



**Universidade de Brasília**

Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas  
Programa de Pós-Graduação em Administração

**Big Data Analytics e Data Mining para Avaliação e  
Previsão da Eficiência Financeira do Setor Bancário  
Brasileiro: Um Estudo de 2011 a 2022**

Paulo Emílio Alcantara Pereira

Dissertação apresentada como  
requisito parcial para a obtenção do  
título de Mestre em Administração

Prof. Dr. Ivan Ricardo Gartner (Orientador)  
PPGA/UnB

Brasília, 30 de junho de 2023



**Universidade de Brasília**

Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas  
Programa de Pós-Graduação em Administração

**Big Data Analytics e Data Mining para Avaliação e  
Previsão da Eficiência Financeira do Setor Bancário  
Brasileiro: Um Estudo de 2011 a 2022**

Paulo Emílio Alcantara Pereira

Dissertação apresentada como  
requisito parcial para a obtenção do  
título de Mestre em Administração

Prof. Dr. Ivan Ricardo Gartner (Orientador)  
PPGA/UnB

Prof. Dr. José Alves Dantas  
PPGCont/UnB/BACEN

Prof. Dr. Carlos Rosano Peña  
PPGA/UnB

Prof. Dra. Solange Alfinito  
Coordenadora do Programa de Pós-graduação em Administração

Brasília, 30 de junho de 2023

# Agradecimentos

Esta dissertação tornou meu mestrado ainda mais especial devido ao assunto em si e às pessoas envolvidas no processo.

Em primeiro lugar, gostaria de agradecer imensamente ao meu orientador acadêmico, Prof. Dr. Ivan Ricardo Gartner, que, desde nosso primeiro contato, se mostrou disponível, entusiasmado e paciente, com críticas sempre construtivas sobre o desenvolvimento dos trabalhos.

Também sou profundamente grato a João Pacheco, meu orientador institucional, que forneceu ideias essenciais e ouviu atentamente minhas explicações, sempre contribuindo decisivamente para o desenvolvimento da dissertação.

Igualmente, agradeço ao Banco Central do Brasil (BCB), que me concedeu a licença, e aos colegas de trabalho que a fizeram acontecer. Obrigado, Emmanuel, Laerte e tantos outros.

Agradeço também aos meus amigos de classe pelos valiosos conselhos durante todo o processo e a todos os novos amigos que tornaram o caminho ainda mais afortunado.

Meu mestrado teria sido impossível sem o apoio da minha esposa, Thaís; durante todo o período de licença, ela me manteve confiante e feliz na caminhada. Além disso, gostaria de agradecer às minhas filhas, Maitê e Alice, por serem a minha felicidade do dia a dia, e ao restante de toda a minha família pelo carinho e amor incondicionais.

Obrigado, Deus, por tudo o que é necessário para completar esta jornada privilegiada.

# Resumo

O setor bancário como um tópico abrangente de dados está em constante evolução sob as influências publicitárias da era do big data. Verificar os recursos analíticos inovadores de detalhes significativos como métodos de mineração de dados é crucial para o setor bancário, que se esforça para revelar informações valiosas da enorme quantidade de informações e obter uma gestão estratégica e uma satisfação do cliente muito melhores. Em virtude da inexistência de estudos tratando essas modelagens tecnológicas em forma conjunta relacionadas ao desempenho financeiro bancário no cenário brasileiro, este trabalho desenvolveu dois estudos teórico-empíricos completos e complementares entre si, tendo o primeiro utilizado dados primários advindos de entrevistas diretas a vários atores envolvidos com o objeto do estudo, e o segundo se utilizando de dados secundários oficiais e públicos fornecidos pelo órgão regulador dos bancos nacionais, o Banco Central do Brasil. Ao reunir e analisar os modismos da concentração de estudos, informações de recursos online, ajudas técnicas e recursos analíticos de informações, esta dissertação em particular contribui para obter insights importantes sobre os avanços sucessivos da análise de big data (BDA) e da mineração de dados (DM) no setor bancário. Dessa forma, o estudo possui como objetivos gerais verificar, num primeiro momento, a influência da BDA no desempenho financeiro bancário e, num segundo tempo, analisar qual a melhor técnica de DM para prever o desempenho financeiro (FP) dos bancos. A proposição em teste nesse estudo é de que essas tecnologias desempenham um papel fundamental na indústria bancária, tornando-se soluções fundamentais para estudiosos de desempenho financeiro corporativo e analistas financeiros. Assim, como conclusão geral, o trabalho sustenta o argumento de que a BDA e a DM exercem, no mercado bancário contemporâneo, influência determinante e finalística para a tomada de decisões da direção corporativa, visto que contribuem de forma fundamental para o desempenho financeiro, podendo até mesmo determinar seu sucesso ou sua falha, a depender da gestão empreendida na análise dos dados. Além disso, o trabalho contribui para as discussões acerca de análise de grandes dados, gestão de ativos, gerenciamento de recursos, entre outros assuntos, sobretudo sob a ótica dos agentes supervisores bancários com ações voltadas para a estabilidade financeira.

**Palavras-chave:** big data analytics; mineração de dados, instituições financeiras; desempenho financeiro corporativo, análise financeira.

**Classificação JEL:** O32, G21, L25, G17.

# *Abstract*

*The banking sector as a comprehensive data topic is constantly evolving under the advertising influences of the big data era. Verifying the innovative analytical resources of significant details such as data mining methods is crucial for the banking sector, which strives to reveal valuable information from the enormous amount of data and achieve better strategic management and customer satisfaction. Due to the lack of studies dealing with these technological models in a joint form related to banking financial performance in the Brazilian scenario, this work developed two complete and complementary theoretical-empirical studies, the first using primary data from direct interviews with various actors involved with the study object, and the second using official and public secondary data provided by the national bank regulator, the Central Bank of Brazil. By gathering and analyzing trends in concentration studies, online resource information, technical aids, and analytical resources for information, this particular dissertation contributes to gaining important insights into the successive advances in big data analysis (BDA) and data mining (DM) in the banking sector. Thus, the study aims to verify, in a first moment, the influence of BDA on banking financial performance, and in a second moment, to analyze which DM technique is best for predicting financial performance (FP) of banks. The proposition being tested in this study is that these technologies play a fundamental role in the banking industry, becoming fundamental solutions for corporate financial performance scholars and financial analysts. Therefore, as a general conclusion, the work supports the argument that BDA and DM play a determining and finalistic influence on decision-making in corporate management, contributing fundamentally to financial performance, and can even determine its success or failure, depending on the management undertaken in data analysis. In addition, the work contributes to discussions about big data analysis, asset management, resource management, among other subjects, especially from the perspective of banking supervisory agents with actions focused on financial stability.*

**Keywords:** *big data analytics, data mining, financial institutions, corporate financial performance, financial analysis.*

**JEL Classification:** *O32, G21, L25, G17.*

# Lista de Figuras

1.1	Modelo Conceitual .....	12
1.2	Análise do modelo de medição .....	32
1.3	Análise de caminho PLS (n = 5000 amostras bootstrap) .....	36
2.1	Determinantes do desempenho financeiro .....	55
2.2	Estrutura de rede neural para prever o desempenho financeiro de bancos no Brasil.....	58
2.3	Resumo da previsão do desempenho financeiro do sistema bancário no Brasil por meio de técnicas de mineração de dados .....	60
2.4	Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica de regressão LASSO .....	63
2.5	Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica RNA.....	63
2.6	Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica RF .....	64
2.7	Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica KNN .....	64

# Lista de Tabelas

1.1	Dados demográficos.....	28
1.2	Modelo de medição .....	32
1.3	Pesos externos e valores do fator de inflação de variância (VIF) .....	34
1.4	Relação heterotraço-monotraço (HTMT) .....	35
1.5	Coeficiente de determinação .....	35
1.6	Resultados do modelo de equações estruturais .....	36
1.7	Resumo das hipóteses (Não Rejeitadas e Rejeitada).....	37
2.1	Percentual do Ativo Total dos 30 maiores bancos no Sistema Financeiro do Brasil .....	52
2.2	Resultados da avaliação de testes e treinamento .....	62
A2.1	Dados trimestrais de CAR, AQ, LM e ROA no sistema bancário brasileiro.....	97
A2.2	Previsão do desempenho financeiro do sistema bancário no Brasil usando técnicas de mineração de dados .....	98

# Lista de Abreviaturas e Siglas

<b>ACE</b>	<i>Audit Committee Effectiveness</i>
<b>AQ</b>	Qualidade de Ativos
<b>BDA</b>	<i>Big Data Analytics</i>
<b>BDAC</b>	<i>Big Data Analytics Competency</i>
<b>CAR</b>	Rácio de Adequação de Capital
<b>DCV</b>	<i>Dynamic Capabilities View</i>
<b>DEA</b>	Análise Envoltória de Dados
<b>DT</b>	Árvore de Decisão
<b>FP</b>	<i>Financial Performance</i>
<b>HTMT</b>	<i>Heterotrait–monotrait</i>
<b>IA</b>	Inteligência Artificial
<b>KNN</b>	<i>K-Nearest Neighbor</i>
<b>LLCI</b>	<i>Lower Limit Confidence Interval</i>
<b>LM</b>	Gerenciamento de Liquidez
<b>MAE</b>	Erro Absoluto Médio
<b>ML</b>	Aprendizado de Máquina
<b>MENA</b>	Oriente Médio e Norte da África
<b>MSE</b>	Erro Quadrático Médio
<b>PLS</b>	<i>Partial Least Squares</i>
<b>PODC</b>	<i>Process Oriented Dynamic Capabilities</i>
<b>RF</b>	Floresta Aleatória
<b>RMSE</b>	Raiz do Erro Quadrático Médio
<b>RNA</b>	Rede Neural Artificial
<b>ROA</b>	Retorno Sobre Ativos
<b>RSS</b>	Soma Residual dos Quadrados
<b>SEM</b>	<i>Structural Equation Modelling</i>
<b>SFN</b>	Sistema Financeiro Nacional
<b>TBR</b>	Teoria Baseada em Recursos
<b>TI</b>	Tecnologia da Informação
<b>TOPSIS</b>	Técnica de Preferência de Ordem por Semelhança com a Solução Ideal
<b>ULCI</b>	<i>Upper Limit Confidence Interval</i>
<b>VBR</b>	Visão Baseada em Recursos
<b>VIF</b>	<i>Variance Inflation Factor</i>

# Sumário

<b>Introdução à pesquisa</b>	<b>1</b>
------------------------------	----------

## **Capítulo 1**

### **A Relação Entre Big Data Analytics e Desempenho Financeiro Bancário: Uma Investigação Com Dados Brasileiros**

<b>1.1 Introdução</b>	<b>5</b>
1.1.1 Antecedentes do Estudo .....	5
1.1.2 Análise de lacunas .....	7
1.1.3 Declaração do problema.....	7
1.1.4 Importância do Estudo .....	8
1.1.5 Teorias de Apoio à Pesquisa .....	8
1.1.6 Definições de Variáveis de Estudo.....	9
1.1.6.1 Desempenho da Empresa .....	9
1.1.6.2 Competência Big Data Analytics .....	10
1.1.6.3 Eficácia do comitê de auditoria.....	10
1.1.6.4 Capacidades dinâmicas orientadas a processos .....	10
1.1.6.5 Objetivos de Pesquisa .....	11
1.1.6.6 Questões de Pesquisa .....	11
1.1.6.7 Hipóteses .....	12
<b>1.2 Revisão da literatura</b>	<b>12</b>
1.2.1 Recursos Tangíveis .....	12
1.2.1.1 Dados .....	12
1.2.1.2 Tecnologia .....	13
1.2.1.3 Recursos Básicos.....	14
1.2.1.4 Recursos Humanos .....	14
1.2.1.5 Habilidades Técnicas.....	14
1.2.1.6 Habilidades Gerenciais .....	15
1.2.1.7 Domínio do Conhecimento .....	16
1.2.2 Recursos Intangíveis .....	16
1.2.2.1 Grandeza dos Dados .....	16
1.2.2.2 Qualidade de dados.....	17
1.2.2.3 Cultura Orientada a Dados.....	17

1.2.3	Competência de análise de big data e desempenho financeiro da empresa .....	18
1.2.4	Competência de análise de big data e eficácia do comitê de auditoria .....	19
1.2.5	Eficácia do comitê de auditoria e desempenho da empresa .....	21
1.2.6	Eficácia do comitê de auditoria e desempenho financeiro .....	22
1.2.7	Teoria baseada em recursos (TBR) .....	23
1.2.8	Recursos dinâmicos orientados a processos como moderador na relação entre competência de análise de big data e desempenho financeiro da empresa ...	23
1.2.9	O efeito moderador das capacidades dinâmicas orientadas a processos.....	25
<b>1.3</b>	<b>Metodologia</b>	<b>26</b>
1.3.1	Projeto de Pesquisa .....	27
1.3.1.1	Horizonte de tempo .....	27
1.3.1.2	População e Amostragem .....	28
1.3.1.3	Medidas .....	29
1.3.1.4	Modelagem de Equações Estruturais (SEM).....	29
<b>1.4</b>	<b>Resultados</b>	<b>31</b>
1.4.1	Modelo de Medição .....	32
1.4.1.1	Validade Discriminatória .....	34
1.4.1.2	Avaliação do Modelo Estrutural .....	34
1.4.1.3	Resumo das Hipóteses .....	37
<b>1.5</b>	<b>Discussão</b>	<b>37</b>
1.5.1	Implicações da Pesquisa .....	41
1.5.2	Implicações Teóricas .....	42
<b>1.6</b>	<b>Conclusões</b>	<b>43</b>
<b>1.7</b>	<b>Limitações e recomendações para trabalhos futuros</b>	<b>44</b>

## Capítulo 2

### Previsão do Desempenho Financeiro Bancário no Brasil: Técnicas Avançadas de Mineração de Dados

<b>2.1</b>	<b>Introdução</b>	<b>47</b>
2.1.1	Formulação do Problema .....	47
2.1.2	Objetivo Geral .....	48
2.1.3	Justificativa .....	48
2.1.4	Estrutura e Organização do Trabalho .....	48
<b>2.2</b>	<b>Revisão da Literatura</b>	<b>48</b>
2.2.1	Técnicas de mineração de dados .....	49

2.2.2	Previsão do desempenho financeiro.....	49
<b>2.3</b>	<b>Metodologia de Pesquisa</b>	<b>51</b>
2.3.1	Amostra e coleta de dados .....	51
2.3.2	Variáveis experimentais .....	53
2.3.2.1	Variáveis de entrada.....	53
2.3.2.2	Variáveis de saída .....	55
2.3.3	Metodologia de mineração de dados.....	56
2.3.3.1	Regressão LASSO.....	56
2.3.3.2	Floresta aleatória .....	57
2.3.3.3	Redes neurais artificiais .....	57
2.3.3.4	<i>K-nearest neighbor</i> .....	59
<b>2.4</b>	<b>Discussão dos Resultados</b>	<b>59</b>
2.4.1	Teste e precisão .....	60
2.4.2	Previsões.....	62
<b>2.5</b>	<b>Conclusão</b>	<b>64</b>
2.5.1	Implicações práticas para pesquisa e trabalho .....	65
2.5.2	Limitações e pesquisas futuras .....	65
	<b>Considerações finais da pesquisa</b>	<b>67</b>
	<b>Referências da Introdução à pesquisa</b>	<b>69</b>
	<b>Referências do Capítulo 1</b>	<b>70</b>
	<b>Referências do Capítulo 2</b>	<b>85</b>
	<b>Apêndice do Capítulo 1</b>	<b>91</b>
	<b>Apêndice do Capítulo 2</b>	<b>97</b>

# Introdução à pesquisa

A era do big data veio junto com grandes desafios e oportunidades. Quase todas as disciplinas científicas estão experimentando informações transbordantes em volumes e velocidades imprevisíveis. Como resultado, revelar as informações ocultas em análise de big data (BDA) por estratégias de mineração de dados (DM) tornou-se um padrão emergente e objetivo supremo para uma ampla gama de estudos (Batoool & Khan, 2022). Como uma questão abrangente de dados, o setor bancário é uma área de implementação favorita para cientistas com habilidades de DM nos últimos anos da revolução da ciência da informação. Os bancos reconheceram que a expertise, e não os recursos financeiros, é o novo maior ativo. Além disso, o crescimento e a popularidade dos serviços bancários móveis e eletrônicos aumentam o crescimento exponencial de informações bancárias em tempo real. Esses desenvolvimentos contínuos bem como a acessibilidade crescente de big data dominam as ferramentas de análise de big data relacionadas, fazendo-as as coisas mais essenciais para o setor bancário (Broby, 2022).

Há um extenso estudo da literatura atual, com vários documentos de pesquisa focados nos usos da DM no setor bancário, com discussões sobre as implementações da DM já antes de 2013. Para um tópico em rápido desenvolvimento, que avança a cada dia, é essencial oferecer às partes interessadas e pesquisadores a vertente mais atual de DM, bem como colaborações bancárias (Doğuç, 2022).

Tendo resumido os usos mais recentes, o big data no setor bancário foi explorado para aumentar a satisfação do consumidor, a publicidade e o controle estratégico. Em particular, as aplicações mais recentes reunidas nesta monografia visam principalmente a investigar a influência do big data no desempenho financeiro e também contribui reunindo e analisando as informações de recursos de dados e técnicas de DM. Portanto, os itens avaliados neste estudo garantem a novidade de suas contribuições mais importantes para a lacuna atual na literatura acadêmica (Jahan & Candidate, 2022).

Desta forma, a razão que motiva esta pesquisa é a de que o uso de ferramentas para otimizar o desempenho financeiro e fazerem-se previsões desse desempenho é algo relevante para a indústria bancária, especificamente no tocante aos processos de supervisão bancária, e tende a ganhar eficácia ao considerar também medidas de gestão estratégica nas organizações, aplicando técnicas de modelagem de problemas complexos.

O objetivo geral de toda a pesquisa é avaliar a potencial eficácia da aplicação de tecnologias de informação recentes no intuito de otimizar e de prever o desempenho financeiro de instituições bancárias brasileiras, tanto em sua abordagem tradicional, com dados sintéticos do desempenho econômico-financeiro das organizações, quanto em

perspectivas que acrescentem outras dimensões de análise.

A partir disso, os seguintes objetivos específicos são derivados: a) analisar como a competência de análise de big data influencia os resultados financeiros dos bancos, tendo a eficácia do comitê de auditoria com um papel mediador e as capacidades dinâmicas orientadas a processos com um papel moderador na competência do setor bancário brasileiro em análise de big data; e b) fornecer ferramentas eficientes de previsão do desempenho financeiro dos bancos para pesquisadores financeiros e tomadores de decisão no setor bancário brasileiro.

Além desta parte introdutória, este documento é composto por dois capítulos destinados às investigações empíricas relacionadas aos objetivos específicos delineados. A última seção traz as considerações finais.

O primeiro trabalho é apresentado no Capítulo 1 e tem como proposição de pesquisa a investigação sobre as relações entre as variáveis, a fim de determinar o efeito do BDA no desempenho financeiro e, com isso, aprimorar a literatura atual sobre o BDA e o desempenho financeiro.

O Capítulo 2 traz o segundo estudo, que possui como problemática a seguinte questão: qual é a técnica de mineração de dados mais adequada para prever o desempenho financeiro dos bancos no Brasil?

Em relação às abordagens metodológicas, foram utilizadas modelagem de equações estruturais com mínimos quadrados parciais aplicadas a dados coletados por meio de entrevistas baseadas em escala Likert de cinco pontos, para prever e explicar a variância das variáveis de critério, no primeiro estudo; no segundo estudo, foram utilizadas técnicas de mineração de dados, usando a biblioteca *Orange*, que contém algoritmos de análise de dados da linguagem *Python*, como regressão LASSO, Floresta aleatória (RF), Redes Neurais Artificiais (RNAs) e *K-Nearest Neighbor* (KNN), com base em dados secundários do desempenho bancário coletados do sistema IF. Data do Banco Central do Brasil.

Quanto à classificação meta-teórica dos estudos que compõem esta pesquisa, o primeiro trabalho está centrado no princípio da sociomaterialidade e da teoria baseada em recursos, que se relaciona com a unificação ontológica da sociedade e dos materiais. O segundo estudo adota uma abordagem mais funcionalista, examinando o impacto do ambiente organizacional e dos administradores nos resultados, com a crença ontológica de que a realidade é um campo de informação no qual os agentes agem de maneiras que possuem significado pessoal, segundo a perspectiva de Morgan e Smircich (1980).

O campo da teoria organizacional é frequentemente descrito por meio de metáforas representativas, como as do organismo, no primeiro estudo; e da máquina, no segundo trabalho. Essas imagens são usadas para acessar a realidade de várias maneiras. A primeira é por meio de dados de gestão estratégica, que incluem preocupações organizacionais com o meio ambiente e estratégias setoriais únicas que as distinguem dos concorrentes. A

segunda é por meio de dados gerados por processos internos dentro de organizações bancárias, como medições ou sensores de desempenho financeiro. Essas metáforas são exemplificadas em diferentes estudos, e seu uso varia de acordo com o contexto (Morgan, 2005).

A perspectiva do organismo reconhece a influência do ambiente em que as organizações existem e as vê como parte de um sistema aberto que desafia sua evolução e sobrevivência. Por outro lado, a perspectiva maquinária das organizações prioriza a precisão e a eficiência, com foco no planejamento, na racionalização e no organograma para atingir as metas. Essa visão considera as organizações como participantes de um sistema fechado.

## Capítulo 1

### A Relação Entre Big Data Analytics e Desempenho Financeiro Bancário: Uma Investigação Com Dados Brasileiros

**Resumo:** Nos últimos anos, devido a avanços tecnológicos, internet, máquinas portáteis, sistemas de apoio à decisão, transmissão e computação, a ideia de big data foi impulsionada pelo avanço da velocidade no processamento e no armazenamento de dados. O objetivo do nosso estudo é determinar como as capacidades de big data analytics dos bancos brasileiros os ajudarão a melhorar o seu desempenho financeiro, tendo a eficácia da comissão de auditoria (ACE) como mediadora dessa relação e as capacidades dinâmicas orientadas a processos (PODC) como moderadoras dessa mesma relação. As empresas têm se deparado com fatores e mudanças ambientais imprevisíveis e tendências cada vez mais competitivas. Para obter vantagem competitiva, desempenho e sucesso superiores e aumentar a sobrevivência e a longevidade a longo prazo, faz-se necessário introduzir e aplicar técnicas e estratégias adequadas. Nesse caminho, o comitê de auditoria é uma das ferramentas e procedimentos mais eficazes usados pelas empresas e é um dos principais impulsionadores de mecanismos organizacionais bem-sucedidos. O estudo adotou um desenho de pesquisa de levantamento. A amostragem aleatória simples foi usada para que os funcionários fossem encontrados em seus locais de trabalho, locais estes que foram usados para o estudo. Este desenho foi quantitativo para permitir a análise descritiva e inferencial. Os dados foram coletados no prazo de um mês (ou seja, dezembro de 2022 a janeiro de 2023). Como foi coletado de uma só vez, o design é de natureza transversal. Foram recebidos dados de 330 respondentes-alvo, nos diferentes bancos públicos e privados. O estudo atual faz uma adição substancial ao gerenciamento de análise de big data e ao desempenho da empresa, e possui uma ampla gama de aplicações. Como não foram encontrados estudos anteriores abordando diretamente o comportamento moderador da capacidade dinâmica orientada a processos e o papel mediador da eficácia do comitê de auditoria nesse cenário, a função desempenhada pelo estudo atual é bastante importante. Além disso, a convergência dos recursos organizacionais examinados neste estudo para avaliar a competência analítica de big data de uma empresa pode ser considerada relevante. Usando a teoria baseada em recursos (TBR) e a visão de capacidades dinâmicas (DCV) como lentes teóricas, o estudo investiga a relação entre as capacidades de BDA e o desempenho da empresa. A principal justificativa para a adoção desses dois pontos de vista é que a capacidade tecnológica de utilizar o BDA requer vários recursos adicionais específicos da empresa que podem eventualmente contribuir para o aumento do desempenho.

**Palavras-chave:** Big data analytics, Instituições Financeiras, Desempenho Financeiro Corporativo.

**Classificação JEL:** O32, G21, L25

# 1.1 Introdução

Serão explicados os antecedentes do estudo, as análises de lacunas, a declaração do problema, o significado do estudo, as questões de pesquisa, os objetivos de pesquisa, a teoria de apoio e as definições de variáveis de estudo usadas nesta pesquisa.

## 1.1.1 Antecedentes do estudo

BDA (Big Data Analytics) está se tornando um assunto comum entre acadêmicos e profissionais. BDA é um método sistemático para manipular, coletar e avaliar cinco “V’s” de dados (ou seja, volume, variedade, velocidade, veracidade e valor) para gerar ideias práticas para valor sustentável, medição de sucesso e vantagens competitivas (A. K. Gupta & Goyal, 2021; Morimura & Sakagawa, 2023). Inúmeros profissionais e estudiosos sugeriram que o BDA é o “quarto paradigma científico” (Rana et al., 2022), “O mais recente modelo de recursos de informação” ou “o próximo limite para criatividade, competitividade e produtividade” (Elia et al., 2022). Isso é motivado principalmente pelo amplo uso e adoção de ferramentas, tecnologia e redes habilitadas para BDA, incluindo mídias sociais, dispositivos inteligentes, tecnologias automatizadas de reconhecimento da Internet e serviços baseados em nuvem para as empresas obterem benefícios competitivos e mantê-los (Maheshwari et al., 2021). Big data é uma das indústrias e desenvolvimentos tecnológicos mais recentes e é extremamente importante em muitas áreas de estudo, de modo especial, no setor financeiro (Hasan et al., 2020). Neste contexto, os desenvolvimentos mais recentes obrigam as empresas a recrutar especialistas em finanças, nomeadamente contabilistas, que também possuam competências em Tecnologias de Informação (TI), sobretudo porque esta é uma exigência da dinâmica do mercado moderno. Ao mesmo tempo, especialistas em finanças e contadores com conhecimento relevante em seu campo se aproveitarão disso (antes apenas profissionais altamente qualificados) para integrar big data nos processos de tomada de decisão.

Nos últimos anos, big data tem sido um avanço tecnológico popular em jogos e em mudanças (Corvello et al., 2023) como uma 'fronteira' virtual de uma ampla variedade de tecnologias orientadas por TI e possibilidades de habilitação de informações (Niu et al., 2021). Devido aos avanços tecnológicos, a internet, dispositivos portáteis, sistemas de suporte à decisão, transmissão e cálculos, a ideia de big data foi impulsionada por avanços rápidos no processamento e armazenamento de dados (Y. Cui et al., 2023). Adicionando mais a isso, Morimura & Sakagawa (2023) mostraram que esse aumento de dados é motivado por diferentes fontes que geralmente são classificadas como dados produzidos

pela máquina e dados humanos (Börsch-Supan et al., 2021); onde os dados gerados por máquinas são usados para criar máquinas digitais automaticamente para big data, a fim de que não haja intervenção humana direta (por exemplo, dados de áudio, de vídeo, de imagem e de fala, dados de sensores, de câmeras de segurança, dados de dispositivos médicos etc.). Em contraste, os dados gerados por humanos envolvem a criação de big data por interações diretas homem-computador (por exemplo, dados gerados por postagens de mídia social, dados de fluxos de cliques, conteúdo da web etc.). Big data é agora conhecido por ser um dos desenvolvimentos futuros de mudança mais rápida, não apenas devido à sua existência distinta em escala e em velocidade, mas também devido a cinco 'V's básicos que o tornam 'grande' em um sentido real (Morimura & Sakagawa, 2023). Além disso, apenas o uso de dados de alto volume não garantirá a eficiência, mas também a alta qualidade e a relevância dos dados são as principais características para melhorar o processo de tomada de decisão (Meiyou & Ye, 2022).

Por razões específicas, os clientes escolhem um ou mais bancos para investirem. Os bancos querem conhecer seus consumidores, pois eles têm as habilidades e precisam fazer melhores escolhas em sua organização (por exemplo, se um cliente está ou não apto a receber um empréstimo de alto valor) (Banna et al., 2021; Morimura & Sakagawa, 2023). A recente crise econômica também afetou a forma como os consumidores, especialmente no Brasil, se relacionam com os bancos (da Rosa München, 2022; Scalco et al., 2021). Os clientes gostariam de saber mais sobre a instituição em que acreditam, na qual querem ter certeza de que o dinheiro não desaparecerá de repente, mas também querem usar seus investimentos de forma a maximizar o lucro.

A literatura existente indica que habilidades analíticas, expertise de domínio (Bag et al., 2022) e ferramentas de análise estão entre as variáveis mais essenciais que podem anular, se não usadas de forma correta e eficiente, o efeito benéfico do big data nos resultados da empresa (Salim et al., 2022). As organizações então correm para obter analistas de big data competentes (Buhalis & Volchek, 2021) para validar e analisar os dados e intensificar sua competência em big data analytics (Morimura & Sakagawa, 2023). No estudo de big data, é benéfico que as empresas estejam atentas às perspectivas crescentes e decisões comerciais (Yu et al., 2021). Isso promove o interesse da pesquisa atual em preencher essa lacuna, levando em consideração as indústrias de telecomunicações e a bancária brasileira para investigar a experiência da análise de big data e seu efeito na eficiência dos negócios.

O objetivo do nosso estudo é determinar como as capacidades de big data analytics dos bancos brasileiros os ajudarão a melhorar o seu desempenho financeiro: a eficácia do comitê de auditoria (ACE) como mediadora e as capacidades dinâmicas orientadas para o processo (PODC) como moderadora. As empresas se depararam recentemente com fatores e mudanças ambientais imprevisíveis e com tendências cada vez mais competitivas. Para obter uma vantagem competitiva, obter desempenho e sucesso superiores e aumentar a

sobrevivência e a longevidade, elas devem introduzir e aplicar técnicas e estratégias adequadas. Técnicas e métodos de uma variedade de campos foram examinados rigorosamente e analiticamente. Neste estudo, o comitê de auditoria é uma das ferramentas e procedimentos mais eficazes utilizados pelas empresas e é um dos principais impulsores de mecanismos organizacionais bem-sucedidos. Em comparação, as funções de monitoramento do comitê de auditoria têm sido consideravelmente maiores e mais eficazes.

### **1.1.2 Análise de lacunas**

O estudo atual contribui para a literatura de competência de análise de big data de várias maneiras. Por exemplo, analisa o efeito do sucesso da empresa na análise de big data, que foi negligenciado em estudos anteriores. O estudo também reconhece possíveis mediadores e moderadores na análise de big data e na relação de desempenho da empresa (eficácia do comitê de auditoria e capacidades dinâmicas orientadas a processos, respectivamente).

As capacidades dinâmicas orientadas a processos são descritas como a possibilidade de uma empresa de ajustar (melhorar, adaptar ou reconfigurar) um processo de negócios de forma mais eficaz do que a concorrência na integração de operações, na redução de custos e na alavancagem de inteligência/aprendizado de negócios. As corporações têm uma ampla gama de mudanças nos procedimentos de negócios, variando de modificações e aprimoramentos contínuos a mudanças dramáticas únicas. Embora a maioria das melhorias possa ser gradual, a capacidade de uma empresa se adaptar rapidamente geralmente significa que ela está disposta a fazer mudanças drásticas, se necessário (Bieńkowska & Tworek, 2022).

Além disso, em pesquisas anteriores, o desempenho da empresa alcança um papel vital como variável dependente ou independente (Ulbert et al., 2022; E. G. Xu et al., 2022; Yasmin et al., 2020). Da mesma forma, muitos pesquisadores usaram a capacidade dinâmica orientada a processos como mediadora (Kindermann et al., 2022; Yasmin et al., 2020), mas não a utilizaram como moderadora (Awan et al., 2021). Assim, o objetivo desta pesquisa é usar a capacidade dinâmica orientada a processos como uma moderadora. Este estudo usa um conjunto de dados inéditos para testar essa estrutura, com metodologia quantitativa, e fornece informações importantes seguindo uma lógica explicativa.

### **1.1.3 Declaração do problema**

No contexto do BDA, um estudo recente constatou que vários benefícios do BDA,

particularmente sua capacidade analítica avançada, criam diferentes capacidades dinâmicas na organização, que, em última análise, ajudam-na no sucesso (Choi & Park, 2022). As organizações em países em desenvolvimento, especialmente no Brasil, podem vir a enfrentar uma situação financeira difícil, e o BDA poderá ser uma das possíveis soluções para esse problema. "Uma distinção conceitual entre capacidades comuns e a grande base de recursos da empresa, por um lado, e capacidades dinâmicas, por outro", de acordo com a literatura sobre capacidades dinâmicas (Choi & Park, 2022). "As capacidades comuns (também conhecidas como capacidades operacionais ou de ordem zero) decidem como uma empresa ganha dinheiro no presente, enquanto as capacidades dinâmicas permitem que a empresa mude" (S. Li et al., 2019; Murray et al., 2022).

A declaração do problema do estudo é: como a competência de análise de big data influencia, por meio da capacidade dinâmica orientada a processos e por meio da eficácia do comitê de auditoria, o desempenho financeiro bancário? Ao examinar as relações entre as variáveis, o efeito do BDA no desempenho financeiro pode ser determinado e a literatura atual sobre o BDA e o desempenho financeiro pode ser aprimorada.

Embora todas essas variáveis sejam levadas em consideração na literatura, os pesquisadores não as estão modelando todas juntas em um único modelo lógico. O presente estudo constrói um modelo dessas variáveis, analisando como a competência de análise de big data causa resultados firmes com a eficácia do comitê de auditoria como um papel mediador e as capacidades dinâmicas orientadas a processos como um papel moderador na competência do setor bancário brasileiro em análise de big data, análise essa utilizada para otimizar o desempenho financeiro da indústria bancária nacional.

### **1.1.4 Importância do estudo**

Nossa pesquisa também contribui para o BDA existente e para a literatura de desempenho da empresa, podendo ajudar as organizações baseadas em projetos a integrar com confiança a imaginação e a inventividade em suas propostas. Isso pode ser feito buscando oportunidades de mercado e arriscando (indústria financeira, por exemplo) para colocar certas ideias em prática. Isso não apenas os ajudará a aumentar a lucratividade executando com sucesso suas inovações, mas também lhes dará uma vantagem competitiva.

### **1.1.5 Teorias de apoio à pesquisa**

A teoria baseada em recursos (TBR) foi desenvolvida por Penrose, em 1959. A TBR

propõe que a empresa pode adquirir uma vantagem competitiva sobre seus concorrentes por possuir recursos estratégicos (Penrose, 1959). A teoria dos recursos examina os recursos disponíveis para as empresas como um conjunto de capacidades dinâmicas e competências de vantagem competitiva. Mishra & Yadav (2021) afirmaram que as empresas operam em um ambiente competitivo e precisam usar seus recursos, suas habilidades e suas capacidades únicas, aproveitar novas oportunidades, expulsar ameaças e atender aos desejos dos clientes.

O paradigma teórico divide os recursos da organização em dois: recursos tangíveis e intangíveis. Os recursos tangíveis compreendem materiais, *commodities*, recursos físicos e reservas financeiras. Recursos de negócios intangíveis incluem reputação, tecnologia e pessoal. Segundo Garcia & Meurer (2022), se a empresa possui recursos raros, únicos, valiosos e únicos, ela alcança uma vantagem competitiva. A TBR observa que a vantagem competitiva está no ambiente da empresa, visível por meio de seus recursos e de suas competências. Essa estratégia foi então criticada principalmente por suas características estáticas e pelo baixo poder explicativo das condições sob as quais empresas em ambientes altamente dinâmicos podem ter uma vantagem competitiva (Rehman et al., 2022).

As capacidades dinâmicas geram novos métodos para gerar valor ao alterar as capacidades comuns (Heubeck, 2023). Alguns estudiosos, no entanto, argumentam que "embora as capacidades dinâmicas e comuns sejam especificadas localmente, a linha entre elas é desnecessariamente tênue" (S. Li et al., 2019). Seguindo o estudo da TBR, o potencial de uma empresa para atingir os objetivos visados, ou seja, maior integração e exploração de recursos no desenvolvimento de habilidades de análise de big data, seria maior. Além disso, a eficácia de um comitê de auditoria que atua como mediador no presente estudo é outra característica baseada em recursos de uma empresa que desempenha um papel importante no processo dos negócios (Al-Okaily & BenYoussef, 2020).

Ademais, o uso neste estudo das capacidades dinâmicas orientadas a processos de uma empresa como moderadora impactam as dimensões orientadas a recursos, pois os recursos estão entre os principais desafios sobre os quais recaem as decisões organizacionais (Ye et al., 2022). Além de a Teoria Baseada em Recursos ter em seu círculo de impacto todas as dimensões examinadas neste relatório, ela ainda justifica sua implementação nas pesquisas atuais.

## **1.1.6 Definições das Variáveis de Estudo**

### **1.1.6.1 Desempenho da empresa**

O nível em que uma empresa explora seus recursos para entrar em novos mercados e fornecer novos produtos e serviços é chamado de desempenho de mercado. Em contraste,

o nível em que uma empresa explora seus recursos para melhorar a produtividade, a lucratividade e o desempenho financeiro é chamado de desempenho operacional. O termo “performance da empresa” possui diversos conceitos (Chet Miller et al., 2013).

### **1.1.6.2 Competência de Big Data Analytics**

A popularidade do big data foi impulsionada pelo avanço tecnológico tanto em software quanto em hardware (Sardi et al., 2020). O big data despertou o interesse de acadêmicos e profissionais como a próxima grande novidade em gestão, sendo alguns também o considerando o próximo movimento de gestão (Alexander & Lyytinen, 2019; D. Li et al., 2020). O big data analytics (BDA), que substituiu alguns instrumentos estatísticos convencionais e valorizou os negócios em termos financeiros e não financeiros, é um tópico emergente entre estudiosos, comunidades empresariais e instituições públicas (Alexander & Lyytinen, 2019; Ashofteh & Bravo, 2021; Saide & Sheng, 2020; Sena et al., 2019). A integração e implementação de recursos específicos de big data pela empresa, a fim de capacitá-la a realizar uma análise sistemática e orientada à ação de dados detalhados, é denominada "Competência em Análise de Dados em Grande Escala" (L. Li et al., 2022).

### **1.1.6.3 Eficácia do comitê de auditoria**

Segundo Moodley et al. (2022), a responsabilidade primária do comitê de auditoria é monitorar e reportar a eficácia da governança corporativa ao conselho de administração. Eles antecipam um possível conflito entre as funções de auditoria interna e as funções da auditoria externa ou independente, o que pode impactar os resultados organizacionais. Já Ghaleb et al. (2020) sugerem que se faz necessário entender como a função do comitê de auditoria interage para fornecer uma governança corporativa eficaz com a auditoria interna, com a administração ou com os auditores externos.

### **1.1.6.4 Capacidades dinâmicas orientadas a processos**

"O poder da organização para reconfigurar, integrar e coordenar talentos internos e externos para lidar com a rápida turbulência em ambientes comerciais" é como a capacidade dinâmica é definida (Suddaby et al., 2020; Liu & Zhao, 2021). A empresa empregará métodos especializados para implementar e expressar suas capacidades dinâmicas para estabelecer, aprimorar ou transformar suas capacidades organizacionais. Eles se relacionam com a capacidade de uma empresa de melhorar seus processos organizacionais para reduzir custos e aumentar a inteligência dos negócios. A empresa deve otimizar a eficácia desses procedimentos operacionais, melhorar a gestão do conhecimento e alinhar melhor seus recursos com a missão da empresa, integrando a

PODC (Lodato et al., 2020; Contreras Pinochet et al., 2021; Wei et al., 2022).

### **1.1.6.5. Questões de pesquisa**

A questão de pesquisa descreve a área mais ampla do problema, que também é descrita em nossa declaração de problema. As questões de pesquisa a seguir são retiradas da definição do problema de nosso estudo.

1. A competência de análise de big data está positivamente associada ao desempenho financeiro da empresa?
2. A competência de análise de big data está associada à eficácia do comitê de auditoria?
3. A eficácia do comitê de auditoria influencia a relação entre a análise de big data e o desempenho financeiro da empresa?
4. A eficácia do comitê de auditoria desempenha um papel significativo como mediador entre a competência de análise de big data e o desempenho financeiro da empresa?
5. Os recursos dinâmicos orientados a processos desempenham um papel significativo como moderador entre a competência de análise de big data e o desempenho financeiro da empresa?
6. Os recursos dinâmicos orientados a processos desempenham um papel significativo como moderador entre a competência de análise de big data e a eficácia do comitê de auditoria para os resultados financeiros do banco?

### **1.1.6.6. Objetivos de pesquisa**

Os objetivos da pesquisa incluem a justificativa para o estudo de correlações específicas. Atingimos os seguintes objetivos em nossa pesquisa com base na tipologia dos objetivos do estudo.

1. Examinar a relação entre a competência de análise de big data e o desempenho financeiro da empresa.
2. Examinar a relação entre a competência de análise de big data e a eficácia do comitê de auditoria.
3. Examinar a influência da eficácia do comitê de auditoria na análise de big data para o desempenho financeiro da empresa.
4. Examinar o efeito mediador da eficácia do comitê de auditoria entre a competência de análise de big data e o desempenho financeiro da empresa.
5. Examinar o efeito moderador das capacidades dinâmicas orientadas a processos na relação entre a competência de análise de big data e o desempenho financeiro da empresa.
6. Examinar o efeito moderador dos recursos dinâmicos orientados a processos na relação entre a competência de análise de big data e a eficácia do comitê de auditoria para

o desempenho financeiro do banco.

### 1.1.6.7. Hipóteses

As hipóteses de pesquisa do estudo são:

H1: A competência de análise de big data está positivamente associada ao desempenho financeiro da empresa.

H2: A competência de análise de big data está positivamente associada à eficácia do comitê de auditoria.

H3: A eficácia do comitê de auditoria está positivamente associada à relação entre análise de big data e ao desempenho financeiro da empresa.

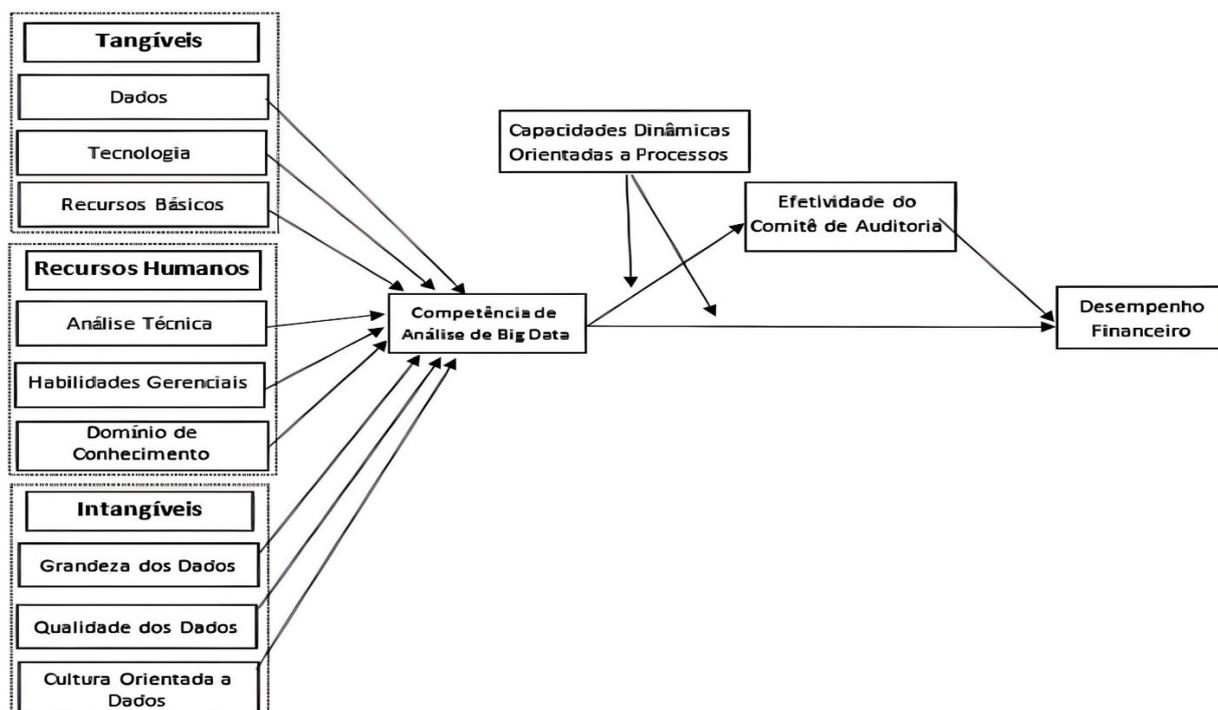
H4: A eficácia do comitê de auditoria como mediador na relação entre a competência de análise de big data e o desempenho financeiro da empresa.

H5: As capacidades dinâmicas orientadas a processos atuam como moderadoras na relação entre a competência de análise de big data e o desempenho financeiro da empresa.

H6: As capacidades dinâmicas orientadas a processos atuam como moderadoras na relação entre a competência de análise de big data, a eficácia do comitê de auditoria e o desempenho financeiro do banco.

## 1.2. Revisão da Literatura.

Figura 1.1: Modelo Conceitual



Fonte: Elaborado pelo autor

## **1.2.1 Recursos Tangíveis**

São os recursos que podem ser mensurados.

### **1.2.1.1 Dados**

As organizações atualmente não estão interessadas apenas em seus dados estruturados exclusivos quando se trata de tomar decisões. Em vez disso, eles visam obter todas as informações disponíveis, independentemente de sua estrutura, de seu tamanho ou de sua taxa de produção (Rosnidah et al., 2022). (i) Dados públicos, (ii) Dados privados, (iii) Esgotamento de dados, (iv) Dados da comunidade e (v) Dados de autoquantificação são as cinco principais fontes de geração de dados de alto volume (Flynn et al., 2020).

Os conjuntos de dados controlados pelo governo e pelas comunidades locais são conhecidos como dados públicos e incluem informações sobre transporte, saúde, uso de energia e mudanças climáticas, entre outros tópicos. Em contraste com os dados públicos, os dados privados referem-se a conjuntos de dados coletados e de propriedade de empresas e de indivíduos, mas não são acessíveis gratuitamente por fontes públicas. Dados de transações do consumidor, dados de uso do telefone celular e outros tipos de dados privados são exemplos de dados pessoais (Cappa et al., 2021). Em seguida, a exaustão de dados refere-se a conjuntos de dados gerados a partir dos comportamentos de um indivíduo, que são coletados passivamente e agregam valor quando combinados com outras fontes de dados para fornecer novos insights. Arquivos de log gerados por um navegador da web são exemplos de exaustão de dados (Bell et al., 2021). A última categoria é a autoquantificação, que são conjuntos de dados formados pela quantificação de comportamentos e ações dos indivíduos, como dados de pulseiras de inteligência em movimento (R. Gupta et al., 2021).

### **1.2.1.2 Tecnologia**

As tecnologias digitais estão transformando substancialmente os mercados, os contextos corporativos e as técnicas de marketing de comunicação (Garrido-Moreno et al., 2020). A digitalização mudou as interações e trocas entre empresas e clientes, gerando novos contatos e caminhos de cooperação no mercado. Embora o aumento da digitalização possa trazer oportunidades corporativas, também é um grande desafio. Os gerentes encontram mercados complicados e em rápida mudança e carecem de conhecimento específico para reconhecer e gerenciar essas mudanças (Saad & Alnuaimi, 2022). Para lidar com essas dificuldades, as empresas precisam entender o uso estratégico da tecnologia digital e a capacidade necessária para responder rapidamente às mudanças do mercado,

alterando o processo de criação de valor (Dutta et al., 2020; Hanelt et al., 2021).

As qualidades fundamentais do big data que o distinguem dos dados tradicionais, de acordo com a estrutura dimensional 3Vs mais comumente usada, são (i) Volume, que enfatiza a quantidade ou número de dados gerados. O tamanho mínimo de geração de dados a ser rotulado como “big” é um terabyte (Zekić-Sušac et al., 2021); (ii) Velocidade, que mostra a rapidez com que os dados são produzidos e processados; a velocidade de big data é alta o suficiente para ser chamada de criação quase em tempo real (Chae & Olson, 2022) e (iii) Variedade, que se refere ao desenvolvimento de várias formas de dados de várias fontes, e big data, que compreende o desenvolvimento de três formas de dados: dados estruturados, não estruturados e semiestruturados de várias fontes, como humanos e robôs (Roh et al., 2021).

Essas três características distinguem big data de dados comuns, tornando necessário o desenvolvimento de novos sistemas capazes de lidar com o gigantismo, diversidade, geração rápida e transporte de big data. As organizações dependem de certas tecnologias para armazenamento e percepção de dados. A tecnologia sempre desempenhou um papel importante na aquisição de uma vantagem competitiva e na formação de uma organização (Muhammad et al., 2020). No entanto, fatores como movimentos de pessoal, reuniões informais e conversas de várias empresas, engenharia reversa etc., dificultam a ocultação das tecnologias exclusivas das empresas (Muhammad et al., 2020).

### **1.2.1.3 Recursos Básicos**

Juntamente com dados e tecnologias avançadas, as organizações são obrigadas a investir adequadamente em recursos básicos para desenvolver suas competências de análise de big data. Levando em consideração a singularidade e a novidade do big data e as tecnologias, empregos e deveres opostos, a maioria das organizações ainda está no caminho de criar estratégias e procedimentos padrões neste domínio. Existe, portanto, uma probabilidade de que as organizações que praticam análise de big data em seu sistema não alcancem os resultados imediatos desejados. O que importa é a firme pertinácia e a determinação em atingir seus objetivos analíticos, dedicando assim recursos suficientes.

Além de dados e de tecnologias inovadoras, as empresas devem investir em recursos essenciais para aprimorar seus conhecimentos em análise de big data. À luz da singularidade e da inovação de big data e de tecnologias, de funções e de obrigações relevantes, a maioria das empresas ainda segue estratégias e procedimentos típicos. Portanto, é provável que as empresas que realizam análises de big data em seus sistemas não consigam obter os resultados pretendidos imediatamente. O que importa é a firme relevância para seus objetivos analíticos e o desejo de dedicar recursos suficientes para eles. Os recursos tangíveis desempenham um papel crucial no estabelecimento de uma

competência adequada de análise de big data ou um ganho subjetivo para o compromisso constante de uma empresa em investir nesses recursos (Banmairuroy et al., 2022; Luo et al., 2022).

#### **1.2.1.4 Recursos Humanos**

O capital humano é visto como um fator de crescimento sustentável de longo prazo (Biedka et al., 2022). A capacidade humana depende de educação, talento, saúde, oportunidades e outros serviços (S. Jain & Singh, 2020). Há uma falta de congruência entre o capital humano e as demandas da empresa (Krammer, 2022), incluindo trabalhadores, renda, realização no trabalho, treinamento e produtividade (Esposito & Scicchitano, 2022; Damelang & Ruf, 2023).

#### **1.2.1.5 Habilidades Técnicas**

As habilidades de análise técnica dos funcionários referem-se ao seu conhecimento e tecnologia avançada para lidar com grandes quantidades de dados. Algumas habilidades técnicas necessárias para examinar dados grandes incluem limpeza e extração de dados, aprendizado de máquina, análise de dados e paradigmas de programação (Börsch-Supan et al., 2021; Ranta et al., 2023). De acordo com a literatura recente, as instituições educacionais têm feito esses cursos para inculcar competências específicas de big data em indivíduos, devido à escalada da análise de big data e de seu uso em empresas; no entanto, a escassez ainda existe (Niu et al., 2021).

Por outro lado, analistas com habilidades de análise técnica insuficientes podem considerar o adiamento de coisas que causam desperdício de tempo e de recursos, cometendo erros e demonstrando que ainda são incapazes de lidar com as dificuldades encontradas (Shamim, Zeng, Khan, et al., 2020). Isso legitima o papel essencial desempenhado pela análise técnica na construção de habilidades e influenciando o desempenho da empresa.

Como resultado, é validada a importância crítica das habilidades de análise técnica no desenvolvimento de competências e na influência no desempenho da empresa. Além disso, existe um conflito nacional e internacional entre o corpo docente e a indústria, necessário para fornecer as pessoas mais competentes para atender às necessidades do setor (Barrera et al., 2023). Smaldone et al. (2022) afirmam que os participantes geralmente não desenvolvem habilidades ou se envolvem em toda a gama de cursos que eles destacam. Os autores acharam desafiador para os alunos lidar com seu tempo, resolver problemas e se envolver na atividade de modelagem.

Para um profissional de negócios, habilidades técnicas e qualificações profissionais são

mais importantes. Na ausência de atividades operacionais, as tarefas não são executadas adequadamente, o que leva a um atraso no crescimento do setor bancário.

### **1.2.1.6 Habilidades Gerenciais**

Embora as habilidades técnicas e analíticas tenham sido consideradas um aspecto importante da competência em big data analytics, as habilidades de gerenciamento também são de igual importância. Enquanto as empresas podem melhorar as habilidades técnicas de sua força de trabalho treinando e recrutando novos especialistas, as habilidades gerenciais são criadas e aprimoradas ao longo do tempo com a experiência organizacional (X. Zhang et al., 2023). Conseqüentemente, é extremamente importante para as empresas que seus gerentes compreendam a aplicação dos valores recém-encontrados, que podem proporcionar ótimos benefícios para uma empresa analisando dados nessas áreas.

Os gerentes de big data devem ter uma compreensão profunda dos requisitos atuais e futuros de outras unidades de negócios, de parceiros e de clientes para exercerem bem seu papel (X. Zhang et al., 2023). Assim, as competências gerenciais são um campo que não pode ser ignorado na análise das competências organizacionais. Uma pesquisa ampla sobre o tema influenciou todas essas dimensões e, no estudo atual, elas estão sendo usadas para explorar o campo das habilidades de análise de big data.

### **1.2.1.7 Domínio de conhecimento**

Além das habilidades técnicas e analíticas dos colaboradores, o domínio do conhecimento é outro aspecto a ser destacado. Um escasso conhecimento de domínio e ineficazes habilidades analíticas podem tornar os dados e a análise de competência no desempenho de tarefas dos funcionários ineficientes (Rajabalee & Santally, 2021). Verificou-se que várias empresas se envolvem extensivamente no desenvolvimento do conhecimento da equipe para descobrir novos insights de negócios a partir dos dados (Bag et al., 2022).

O conhecimento humano é um dos recursos mais difíceis para a organização replicar, de acordo com a visão baseada em recursos (VBR) (Münch et al., 2022). Embora o crescimento da tecnologia moderna não possa reter o conhecimento, novos campos devem sempre ser explorados, e as metas e as capacidades organizacionais devem ser atualizadas para manter um mundo de negócios imprevisível (Ciampi et al., 2021). Faz-se necessário ainda explorar a capacidade dos analistas de identificar os principais pontos fortes e fracos, riscos e oportunidades para que sejam encontradas soluções de negócios eficientes para os problemas (Pan et al., 2022) e tenham um melhor impacto no desempenho da empresa. Portanto, o domínio do conhecimento é considerado um elemento-chave no

desenvolvimento da competência de análise de grandes dados (Y. Yang & Yee, 2022; Saqib Shamim et al., 2020).

## **1.2.2 Recursos Intangíveis**

### **1.2.2.1 Grandeza dos Dados**

O valor dos dados incita a um enorme crescimento na disponibilidade de dados em torno do qual a análise se faz necessária (Aimé et al., 2022). O aumento da geração de conteúdo digital por meio do aumento de dispositivos inteligentes adiciona grandes volumes de dados todos os dias (Marabelli et al., 2021). Em 2011, foi feita uma estimativa do aumento da geração de dados para 50 vezes até 2020 (Jiang et al., 2021). Os aumentos de volume não são desenvolvidos apenas por um tipo de dados, mas sim por diferentes variantes, por exemplo, dados estruturados, dados semiestruturados e dados não estruturados (Tang et al., 2022; Yousuf & Wahid, 2021).

As organizações agora estão interessadas em suas próprias informações internas e externas específicas para derivar novas tendências e padrões que contribuem para o volume e a variedade de dados. A velocidade do big data é a terceira característica mais marcante. O big data é produzido em velocidade quase em tempo real. Ao descobrir padrões ocultos expostos por meio do processamento e da análise de grandes dados, tempos quase reais com grandes volumes e variantes são altamente úteis para o desempenho organizacional (Deb & Fuad, 2021). Esse papel essencial da grandeza de dados o torna um aspecto crítico da análise de dados gigantescos (T. Cui et al., 2022).

### **1.2.2.2 Qualidade de dados**

O ajuste dos dados define sua qualidade, sua limpeza, sua exatidão, sua confiabilidade para uso, sua quantidade de detalhes, sua integridade e seus vários outros recursos. Embora novos métodos analíticos e novas tecnologias tenham progredido suficientemente para detectar dados úteis e importantes (Gu et al., 2021), a qualidade dos dados ainda afeta as descobertas de dados analíticos (Lu et al., 2021). Portanto, as empresas devem empregar apenas dados de qualidade para obter insights valiosos e tomar decisões inteligentes para melhorar o desempenho corporativo (Aimé et al., 2022). T. Cui et al. (2022) o descreveram como um aspecto essencial da competência de análise de grandes dados, dada a importância crucial acerca da qualidade.

### **1.2.2.3 Cultura Orientada a Dados**

Quando se trata de definir a cultura organizacional, há uma variedade de pontos de vista. A pesquisa existente revela que diferentes pesquisadores de gestão possuem diversas perspectivas sobre a cultura corporativa, com alguns acreditando ser um elo que une uma organização (Rass et al., 2023) e outros considerando-o uma relação que define todos os domínios de uma organização (Melo et al., 2022); sendo assim, conseqüentemente, há uma falta de consenso nesta área (Pagda et al., 2021).

Por outro lado, estudos recentes no campo de big data descobriram que a cultura de uma organização desempenha um papel crítico no sucesso ou no fracasso das iniciativas de BDA e que a improdutividade de projetos relacionados a big data está mais ligada à cultura da organização do que à qualidade dos dados e à inadequação tecnológica (Hajiheydari et al., 2021). Além disso, diz-se que a cultura organizacional tem o poder de melhorar a capacidade de uma organização de usar a análise de big data e adquirir uma vantagem competitiva (Shah, 2022).

Com o crescimento da pesquisa no campo de big data, sabe-se agora que, apesar de reunir uma grande quantidade de dados, apenas algumas empresas perceberam o retorno esperado de seu investimento em big data analytics (Shah, 2022). A razão subjacente para isso é que, embora as empresas tenham começado a usar a análise de big data em suas unidades de tomada de decisão, essas entidades ainda dependem da experiência e da intuição de seus principais executivos (Fakhar Manesh et al., 2021). Todos os esforços de uma empresa para reunir grandes quantidades de dados, para adotar tecnologias modernas e para desenvolver capacidades de equipe serão em vão se as escolhas forem feitas com base em designações e não em dados (Fakhar Manesh et al., 2021). A literatura estabelece a cultura orientada a dados como o aspecto mais crítico da especialização em análise de big data.

### **1.2.3 Competência de análise de big data e desempenho financeiro da empresa**

O termo "big data" foi usado para descrever bancos de dados que eram muito grandes para serem manipulados pelos métodos convencionais. O termo "big data" foi cunhado na década de 1990 e tornou-se popular em meados dos anos 2000. O conceito de 'big data', que foi essencialmente generalizado a longo prazo, refere-se tanto aos dados quanto a vários outros elementos, incluindo - mas não limitado a - tendências sociais, diagnósticos ou avanços de habilidades, de períodos e de estruturas (Schoenherr, 2023).

Com a introdução da TI, as atividades financeiras aumentaram, incluindo serviços bancários e de pagamento pela Internet, opções de poupança, seguros e criatividade em serviços bancários e financeiros corporativos. O setor bancário depende principalmente de

tecnologias emergentes para fornecer a maioria de seus serviços. Ele também foi incorporado em sistemas de política chave, como análise de risco organizacional e individual, gerenciamento de transações e relatórios corporativos. A especificação dessa tecnologia de informação e comunicação (TIC) é regulada por leis e regras locais em serviços financeiros (Hasan et al., 2020). Big data é usado em todos os campos da indústria como um fenômeno em evolução. Big data também contribui significativamente para uma melhor interpretação dos mecanismos do mercado financeiro relacionado ao público em geral (Boubaker et al., 2021). Para tomar decisões de curto e de longo prazos, informações infinitas são trocadas regularmente no mercado de ações (Wen et al., 2021). A maior explicação é que, no mercado de serviços financeiros, esse é um novo caminho.

Uma análise mais aprofundada também é essencial para nos ajudar a esclarecer esses problemas no setor bancário. Este estudo se concentra nas maneiras pelas quais os estudiosos não investigaram e examina o efeito do big data nos mercados financeiros como a principal novidade deste relatório. A análise feita aqui sugere a TBR como um forte mecanismo para combinar diferentes aspectos do BDA, seu impacto sinérgico no FP e a continuidade da abordagem vinculada a essa parceria geral de capacidade de desempenho. Parece que a concepção dos requisitos de capacidade, que são cruciais para as previsões de sucesso, pode ser vista em apenas uma pequena parte da pesquisa de big data (Muchenje & Seppänen, 2023; Ma et al., 2022).

Portanto, apoiamos a TBR quanto ao aspecto de que o sucesso da empresa em uma economia de dados pode ser aumentado apenas se as habilidades forem importantes, raras e imperfeitamente replicáveis, e a organização ou o gerenciamento da empresa explorar os recursos potenciais.

Mora et al. (2021) identificaram o sucesso corporativo como a capacidade de uma empresa, em termos de coordenação/integração, redução de custos, inteligência de negócios/aprendizagem, melhorar sua capacidade de ajustar os processos de negócios atuais. Centrada na TBR, a literatura de capacidades de TI reconhece que estas são vantagens estratégicas e diferenciais de sucesso nos negócios (Y. Yang & Yee, 2022; Thompson & Venters, 2021). Como recursos robustos de TI são dimensões essenciais em um conjunto de big data, suas implementações irão distinguir o sucesso da empresa em diferentes funções de negócios (Dieleman et al., 2023). Isso ilumina cada vez mais os estudiosos sobre o papel de distinguir a capacidade de TI em conjunto com o capital organizacional e com as capacidades de influenciar o desempenho das empresas para organizar e implantar recursos baseados em TI.

## **1.2.4 Competência de análise de big data e eficácia do comitê de auditoria**

O conceito de big data na indústria vem da literatura, que inclui muitos artigos conduzidos em periódicos altamente relevantes. Muitos setores da indústria agora estão vinculados ao big data. O big data tem um impacto crucial em vários aspectos do mercado, por exemplo, em termos de gestão, de análise e de crescimento do capital humano (Turulja et al., 2023); de processos de gerenciamento (Ranta et al., 2023; Peng & Ke, 2023); de processos organizacionais, de propaganda e de marketing (Holland et al., 2020; Liu et al., 2020); de montagem mecânica (L. Huang et al., 2019; Dubey et al., 2020); de eficiências organizacionais (S Shamim et al., 2020); de formulação de políticas (Kinra et al., 2020); e de gerenciamento da cadeia de suprimentos (Bag et al., 2022).

Este estudo também examinou três abordagens complexas (análise, proativa e prescrição) para melhorar a medida da análise de conhecimento habitual. A capacidade de uma empresa de coletar e de analisar dados para gerar insights, orquestrando e implantando efetivamente seus dados, tecnologia e talento é conhecida como capacidade de análise de big data (Behl et al., 2022). Como resultado, as empresas devem adquirir e desenvolver uma mistura de dados, de capital técnico, humano e organizacional para desenvolver uma capacidade difícil de copiar e de transferir (Oesterreich, Anton, Teuteberg, et al., 2022).

O comitê de auditoria também apóia sua alta administração, gerentes e auditores na gestão e no tratamento de riscos potenciais e na melhoria dos relatórios financeiros de acordo com as práticas contábeis e outras legislações relevantes sobre transparência correta e total. O comitê de auditoria é uma ferramenta para acompanhar as demonstrações financeiras, os auditores externos e a força de gestão interna (Raimo et al., 2021). Ele também é uma ferramenta crítica de governança corporativa e pode contribuir diretamente para promover a integridade, a transparência, a precisão e o valor agregado dos relatórios financeiros. Ele desempenha um papel importante na redução da assimetria de informações e no aumento da consistência da transparência. Como resultado, uma maior eficácia do comitê de auditoria está favoravelmente ligada a mais desempenho corporativo em ambientes complexos e competitivos (Alzeban, 2020). A atuação do comitê de auditoria torna-se um valioso determinante dos resultados da auditoria interna das empresas, da eficiência dos relatórios financeiros e do progresso corporativo. Além disso, a eficiência do comitê de auditoria auxiliará as operações e as estratégias corporativas das empresas que estão atreladas aos resultados alcançados e à rentabilidade sustentada no longo prazo. As empresas cuja eficácia é melhorada pelo comitê de auditoria tendem a ter mais sucesso corporativo.

A estrutura teórica da presente dissertação está centrada no princípio da sociomaterialidade e da TBR, que se relaciona com a unificação ontológica da sociedade e dos materiais. Essa visão não prova que o material afeta o efeito social (ou seja,

determinismo tecnológico) ou que o efeito social afeta o material ou que há conexão recursiva entre o social e o material (ou seja, construções sociais ou visão sociotécnica). Em vez disso, a pesquisa inclui a ontologia de relacionamento do sociomaterialismo que torna impossível quantificar as contribuições individuais em uma única base de maneira tão entrelaçada como a organização (ou seja, a administração de BDAs), a física (ou seja, infraestrutura de TI) e a humana (por exemplo, capacidade ou habilidade analítica) (Gravett, 2022).

Fomiatti et al. (2021) citam a expressão “inextricavelmente ligado ao social e ao conteúdo”. Discordamos desse princípio de que os aspectos do BDA não atuam isoladamente; em vez disso, percebemos que eles agem em conjunto. Argumentamos que o desempenho da empresa superior no ambiente de big data é produto de recursos organizacionais combinados exclusivos (ou seja, gerenciamento de BDA), físicos (ou seja, infraestrutura de TI) e humanos (ou seja, analíticos ou de conhecimento) consistentes com a TBR (Barney, 1991; Grant, 1991) e ontologias relacionadas do sociomaterialismo (Gravett, 2022; Fomiatti et al., 2021; Priharsari & Abedin, 2021). Esses recursos são divididos em três categorias, de acordo com a TBR: (i) Recursos Tangíveis, (ii) Recursos Intangíveis e (iii) Recursos Humanos. Os recursos intangíveis não são físicos, como exemplo têm-se os ativos baseados no conhecimento, enquanto os recursos humanos incluem treinamento de funcionários, habilidades, relacionamentos, experiências, etc. (Grant, 1996). Os tangíveis são os ativos que podem ser adquiridos ou vendidos.

### **1.2.5 Eficácia do comitê de auditoria e desempenho da empresa**

A palavra auditoria é uma palavra latina que significa inspeção financeira oficial de uma empresa ou de suas contas. O auditor de dados primários compreende facilmente a despesa financeira, o patrimônio líquido e o crescimento líquido dos serviços financeiros. E menos chance de fraude inclui big data porque todas as informações necessárias já são fornecidas em big data (Manita et al., 2020). A principal função do comitê de auditoria é acompanhar o desempenho das funções dos auditores internos e externos.

Nesse sentido, a nomeação, demissão e remuneração dos auditores, o conteúdo e a abrangência do trabalho de auditoria, a integridade do auditor e a resolução de conflitos entre auditores e gestão corporativa são provavelmente fortemente afetados pelos comitês de auditoria. Estes também podem estudar as políticas contábeis selecionadas e tomar decisões sobre elas (Grange et al., 2021). Além disso, há também a persuasão em relação à estratégia, ao nível de informação e ao comprometimento com a prática normal de uma organização em demonstrações financeiras.

Além disso, o comitê de auditoria pode acompanhar a eficácia dos sistemas de contabilidade empresarial para garantir que as práticas organizacionais, incluindo a manutenção de controles preventivos de fraude, sejam cumpridas (Slapničar et al., 2022). “Big data traz medições analíticas adicionais. Ele fornece perspectivas de insights aprimoradas, mas precisa de novos recursos humanos e tecnológicos por causa de seus recursos especiais. Portanto, uma descoberta digna de nota é que poucos estudiosos concordam em incluir BDA e talento em BDA como dimensões centrais da capacidade de gestão do BDA” (Sun et al., 2020).

A auditoria visa confirmar a exatidão das demonstrações financeiras da administração. A supervisão nesta simples declaração pode levantar uma preocupação sobre as habilidades da profissão contábil e, assim, reduzir a confiança do público na profissão (Fakhfakh & Jarboui, 2020; Rodrigues et al., 2022). A posição do auditor tornou-se mais proeminente com a expansão das atividades da empresa internacionalmente. Além disso, o aumento do volume de dados coletados resultou em grande volume de transações (A. Yang et al., 2022). A coleta de dados de transações em tempo real, incluindo localização, hora, quantidade e meio, facilitará a coleta de materiais para uma opinião de auditoria.

Estudos anteriores também implicam que o auditor deve ser qualificado para conduzir uma auditoria de alta qualidade e ter experiência anterior (por exemplo, Brody et al., 1998). O estudo de Boo & Koh (2004) mostrou que as qualidades da equipe de auditoria correspondem às suas habilidades para aconselhar melhorias nos sistemas de controles internos. Para os auditores internos, a experiência anterior é vital, pois muitos julgamentos sobre supervisão são subjetivos e as medidas de gerenciamento podem ter impactos generalizados.

## **1.2.6 Eficácia do comitê de auditoria e desempenho financeiro**

Após a auditoria, faz-se um livro de objeções chamado livro anexo, o principal objetivo do auditor é anotar todos os erros negligenciados sobre o nível dos funcionários e sobre o nível do procedimento operacional padrão (SOPS) das instituições financeiras que resolvem pela gestão de uma organização em um momento específico quando as empresas se exercitam para remover a objeção de seu sinal positivo de organização financeira. Alhababsah & Alhaj-Ismail (2021) apontaram que o comitê de auditoria se originou no final da década de 1930. Após suborno envolvendo a McKesson & Robbin Inc., os membros devem nomear auditores externos (Barney, 1991). Recursos valiosos, raros e inimitáveis são fornecidos pela estrutura e pelos procedimentos organizacionais para melhorar a eficiência corporativa.

Os auditores na prática podem usar o Big Data de forma mais eficiente para coletar

informações importantes e precisas de sua auditoria (Rakipi et al., 2021; Werner et al., 2021). Os dados podem não ser de pouca importância apenas para as empresas; isto é, se mais tentativas de terceiros forem feitas para interpretá-los para uma tomada de decisão sólida, demonstrando assim a necessidade de análises. Os dados são de valor limitado. A análise de dados e os métodos de análise de dados ajudam as equipes de auditoria a revisar os dados do cliente e identificar áreas que requerem análise mais aprofundada no início do processo de auditoria. As equipes de auditoria podem ajustar seus métodos, alterando seus planos de auditoria (F. Huang et al., 2022). A análise de dados explora a transmissão de informações dinâmicas para chegar a decisões importantes. O campo de análise de dados determina quando os dados diferem da indústria comparável e das circunstâncias atuais (Joshi & Marthandan, 2020). Essa correspondência inclui métodos e estratégias informativas, diagnósticas, preditivas e prescritivas. Tais métodos identificam fatores causais que podem prevenir consequências potenciais e, assim, recomendar comportamentos e estratégias apropriadas. Ou seja, por meio da visualização e da interpretação dos dados, eles fornecem insights sobre as principais atividades do mercado (Adomako et al., 2022).

Além disso, os dados permanecem mais significativos para representar a importância, enquanto as análises e as tecnologias para requisitos de dados complexos são usadas (Ghasemaghaei & Calic, 2019). A análise de big data está revolucionando cada vez mais muitas áreas, e é apenas uma questão de tempo para a profissão de auditoria seguir métodos de análise relacionados (Werner et al., 2021). A análise de big data pode aprimorar os dados e a prática dos contadores por sua experiência com conjuntos de dados organizados (Ibrahim et al., 2021).

Há uma necessidade de assumir um papel dominante na tomada de decisões estratégicas para os contadores financeiros, especialmente os auditores. O big data permite que os auditores analisem os processos de geração de dados, incluindo testes populacionais completos, que agregam valor e beneficiam a profissão de auditoria e a de contabilidade (De Santis & D'Onza, 2020). Portanto, os especialistas devem obter conhecimento das estratégias padrões e das estratégias subjacentes da empresa para que os benefícios infinitos e o potencial do big data sejam usados com sucesso.

### **1.2.7 Teoria baseada em recursos (TBR)**

Duas conclusões principais são feitas na TBR sobre o capital corporativo para demonstrar que algumas empresas se saem melhor do que outras e melhoram a eficiência corporativa. A complementaridade de recursos dentro de uma empresa significa que os rivais têm dificuldade em duplicar os recursos (Mora Cortez & Hidalgo, 2022). A complementaridade de recursos ocorre quando um recurso permite que outro faça uso da eficiência da empresa. Em conclusão, o aspecto organizacional se concentra em gerenciar adequadamente o capital útil, incomum e de imitação para aproveitar ao máximo seu potencial de forma competitiva (Y. Zhang et al., 2021).

Os principais componentes da TBR são recursos e habilidades. Considerando que 'capital' significa ativos tangíveis e intangíveis (por exemplo, tecnologia, pessoas e recursos organizacionais), 'capacidade' são subconjuntos de recursos não transferíveis de uma empresa que servem para aumentar a eficiência de outros recursos (T. Xu et al., 2021). A capacidade de usar outras ferramentas e de aumentar a produção total é frequentemente descrita como tangível ou intangível. Em geral, as competências são uma categoria particular de recurso que visa a eficiência de outros capitais da empresa (Mora Cortez & Hidalgo, 2022). A TBR diz que a habilidade de uma empresa é baseada na capacidade de lidar com seu capital chave (tanto recursos humanos quanto outros) com sucesso para alcançar a eficiência da empresa (Grant, 1996).

### **1.2.8 Recursos dinâmicos orientados a processos como moderador na relação entre competência de análise de big data e desempenho financeiro da empresa**

Os recursos de alinhamento com a estratégia de negócios são o caminho crítico para a BDA permitir que as organizações alcancem o FP. A função da capacidade dinâmica orientada para o processo pode ter um efeito moderador no sucesso organizacional. A grande capacidade de análise de dados (BDA) é geralmente reconhecida como um fator importante na melhoria do sucesso de negócios e empresas (FP) (Oesterreich, Anton, & Teuteberg, 2022). Por exemplo, otimização de preços e maximização de benefícios (Chin et al., 2020; Khalajzadeh et al., 2020); receita, lucratividade e participação de mercado são mostrados pela conexão entre BDA e FP nestes relatórios de literatura (Bonnard et al., 2021).

Argumentamos que o FP superior no ambiente de big data é o produto de recursos organizacionais combinados exclusivos (ou seja, gerenciamento de BDA), físicos (ou seja, infraestrutura de TI) e humanos (ou seja, analíticos ou de conhecimento) consistentes com a TBR (Barney, 1991; Grant, 1991) e ontologias relacionadas ao sociomaterialismo (Priharsari & Abedin, 2021; A. Jain & Srinivasan, 2022).

As empresas devem usar processos eficazes para ativar os benefícios potenciais e diversificados da BDA em perspectivas concretas com base nas dimensões dos dados. Este procedimento é considerado um método de análise. Os processos analíticos podem explicar o que aconteceu antes (análise descritiva) e prever o que vai acontecer (análise preditiva). O BDA considera a análise preditiva e a prescritiva como a função chave (Oesterreich, Anton, Teuteberg, et al., 2022).

A literatura da BDA revela que os líderes empresariais gradualmente adotam evidências em vez de decisões baseadas na intuição (Davenport, 2006; Lavallo et al., 2011). Além disso, a abordagem analítica transforma a forma como as empresas operam e competem (Du et al., 2021). Em Desenvolvimento de Valor (BDA), Talaoui et al. (2023) consideram a posição da BDA para fornecer responsabilidade, melhorar a tomada de decisão, a criatividade e segmentar pessoas para ação. A capacidade dinâmica é um

mecanismo organizacional especial para integrar, reconfigurar, ganhar e distribuir capital de dinamismo ambiental (Chin et al., 2020; Ghobakhloo & Fathi, 2020; Rahman et al., 2023). De acordo com os resultados deste estudo, a importância de mercado das ferramentas de análise de big data impacta diretamente os resultados de uma empresa, com um efeito mediador ou moderador (Oesterreich, Anton, & Teuteberg, 2022).

Além disso, ter capacidade de BDA altera os processos corporativos e a maneira como as empresas operam. Como resultado dessas capacidades, o sucesso da empresa é influenciado pela evolução dos processos operacionais e pelo desempenho financeiro do setor. A. Jain & Srinivasan (2022) usaram uma conceituação sociomaterialista hierárquica da capacidade de TI para descobrir uma conexão construtiva e importante entre a capacidade de TI e o desempenho da empresa.

Com base nessa descoberta, nossa pesquisa sugere examinar os efeitos diretos da BDA no FP, e os efeitos moderadores das PODC na relação BDA-FP. Uma associação positiva entre a implantação da análise do consumidor e o desempenho da empresa foi observada na crescente literatura sobre BDA (Panja & Mondal, 2023). Por exemplo, a BDA permite que as empresas examinem e gerenciem estratégias por meio de lentes de dados (Munir et al., 2022). Prevê-se que os efeitos da BDA em várias indústrias sejam enormes. Por exemplo, grandes organizações de varejo usam Big Data para melhorar a experiência do cliente, diminuir a fraude e oferecer recomendações *just-in-time* (Bag et al., 2021). A BDA visa reduzir os custos operacionais e melhorar a qualidade de vida no setor de saúde (Oesterreich, Anton, Teuteberg, et al., 2022). A BDA é considerada um facilitador para monitoramento de ativos e processos de negócios em manufatura e gerenciamento operacional (Ivanov et al., 2021), visibilidade da cadeia de suprimentos, melhoria da manufatura e automação industrial (Wang et al., 2022) e melhoria da transformação dos negócios (Tian et al., 2022).

## **1.2.9 O efeito moderador das capacidades dinâmicas orientadas a processos**

As capacidades dinâmicas orientadas para processos geralmente enfatizam a descrição dos procedimentos pelos quais a vantagem competitiva pode ser alcançada em ambientes diversos. Isso inclui habilidades de identificar e moldar as oportunidades à medida que surgem, bem como as ameaças que possam existir, capturar essas oportunidades e também tentar manter a competitividade da organização melhorando, protegendo, combinando e, quando necessário, reconfigurando os recursos intangíveis e tangíveis da organização (Pisano, 2015).

Como resultado, essa abordagem indica as razões pelas quais algumas organizações conseguem se concentrar na identificação e incorporação de oportunidades provenientes do ambiente externo em seus procedimentos regulares, por meio da gestão e combinação de diferentes recursos disponíveis, a fim de alcançar os resultados necessários, enquanto outras organizações são incapazes de desenvolver e implementar essas capacidades

(Ambrosini et al., 2009).

Teece (2007) considerou que as capacidades dinâmicas são necessárias para se concentrar nas mudanças dos clientes e nos requisitos tecnológicos. Como resultado disso, pode-se desenvolver uma compreensão de como aplicar as capacidades de inovação, com base na compreensão da possível relação que pode existir entre as capacidades dinâmicas e a inovação. Primariamente, a inovação envolve as maneiras pelas quais todo o processo, bem como os recursos, rotinas e capacidades de gestão da organização, são realmente realizados (Soares de Almeida et al., 2019). Devido a isso, as capacidades dinâmicas sólidas permitem que as organizações conceituem, avaliem, testem e implementem rapidamente novas inovações. Isso leva à organização da base de recursos da organização, bem como seus processos, que é o PODC (Teece, 2016).

Existem poucos estudos teóricos e empíricos (Mikalef et al., 2019; Mikalef & Pateli, 2017; Wamba et al., 2017) que também encontraram relações diretas e indiretas entre PODC e Desempenho Organizacional.

## 1.3 Metodologia

O objetivo deste estudo é investigar os efeitos da competência de análise de big data no desempenho financeiro através da eficácia do comitê de auditoria como mediador e das capacidades dinâmicas orientadas a processos atuam como moderadoras no setor bancário brasileiro. A pesquisa de levantamento é utilizada neste estudo para justificar os resultados e descrever a população-alvo. Esta seção conta como selecionamos a população-alvo e como a coleta de dados foi possibilitada para interpretação. Esta pesquisa também determina as variáveis demográficas, ou seja, sexo, idade, designação, estado civil e experiência no setor bancário. As variáveis estudadas são derivadas da revisão da literatura e o aprofundamento da análise dessas variáveis também são executadas neste estudo.

### 1.3.1 Projeto de pesquisa

O estudo adotou um design de pesquisa de levantamento. Este design de pesquisa de levantamento, de acordo com Anjaria (2022), é importante, pois ajudou o pesquisador a obter dados sistemáticos de diferentes respondentes ao mesmo tempo. Foi utilizada uma amostragem aleatória simples para garantir que os funcionários que foram usados no estudo fossem encontrados em seus locais de trabalho. Este design foi quantitativo para permitir a análise descritiva e inferencial (Goldschmidt & Matthews, 2022).

Esta análise é descritiva, já que tenta articular, esclarecer e traduzir as respostas atuais dos representantes sobre a competência em análise de big data, desempenho financeiro, eficácia do comitê de auditoria e capacidades dinâmicas orientadas a processos. Um componente essencial do design de pesquisa é decidir sobre o uso de abordagens de

pesquisa quantitativa, qualitativa ou ambas. Goldschmidt & Matthews (2022) relacionam vários aspectos sobre abordagens de pesquisa quantitativa, quais sejam:

- O teste e a autenticação são ponderados;
- Evidências e causas relacionadas ao caso social são enfatizadas;
- A abordagem é analítica e particular;
- Quantidades organizadas;
- Análise hipotética;
- Preocupação com a consequência;
- Justificação circunstancial e lógica através da participação da população.

O presente estudo trata da abordagem quantitativa de pesquisa. Os procedimentos em direção à abordagem quantitativa envolvem protocolos bem organizados, questionários verbais ou escritos controlados que consistem em respostas pré-arranjadas restritas.

### 1.3.1.1 Horizonte de tempo

Os dados foram coletados dentro de um mês (ou seja, dezembro de 2022 a janeiro de 2023). Como foram coletados de uma só vez, o desenho é de natureza transversal.

### 1.3.1.2 População e amostragem

Uma população é uma coleção de eventos, coisas e pessoas que estão associadas a um interesse que o pesquisador deseja analisar. Trabalhadores do setor bancário foram contatados e questionários foram preenchidos durante o horário de trabalho em seu ambiente natural de trabalho, portanto, é um estudo de campo. A amostra é uma configuração da população que representa toda a população. Para o estudo que utilizamos, foi realizada uma técnica de amostragem aleatória simples.

Foram distribuídos 450 questionários digitais nos diferentes bancos públicos e privados do Brasil. Uma carta introdutória ao lado do objetivo do estudo foi entregue aos administradores e representantes dessas organizações. Uma medida diferente de pesquisa foi dada a eles. Recebemos dados de 330 respondentes-alvo, desses 330 questionários, 30 foram excluídos com base em informações ausentes e em respostas incorretas, o que torna esses questionários inconclusivos e, portanto, excluídos. Dado isso, apenas 300 pesquisas utilizáveis foram recebidas, resultando em uma taxa de resposta de 66,66%.

Durante o mês em que o estudo estava sendo conduzido, os participantes responderam ao questionário, que também descrevia o objetivo do estudo. Garantiu-se a eles que, se precisassem dos resultados, estes lhes seriam fornecidos. Para esta investigação, utilizamos informações de corte transversal. A Tabela 1.1 resume os dados demográficos da amostra.

**Tabela 1.1: Dados demográficos**

Demografia	Categorias	Frequência	Porcentagem válida	Porcentagem cumulativa
------------	------------	------------	--------------------	------------------------

<b>Gênero</b>	Feminino	71	23.7	23.7
	Masculino	229	76.3	100.0
<b>Idade</b>	20-30	106	35.3	35.3
	31-40	135	45.0	80.3
	41-50	55	18.3	98.7
	>50	4	1.3	100.0
<b>Educação</b>	Bacharelado	77	25.7	25.7
	Mestrado	143	47.7	73.3
	Doutorado	69	23.0	96.3
	Outros	11	3.7	100.0
<b>Tempo na organização atual</b>	<3	26	8.7	8.7
	3-5	124	41.3	50.0
	6-10	96	32.0	82.0
	>11-15	54	18.0	100.0
<b>Experiência total</b>	<3	22	7.3	7.3
	3-5	100	33.3	40.7
	6-10	99	33.0	73.7
	>11-15	79	26.3	100.0
<b>Tipo de Banco</b>	Público	67	22.3	22.3
	Privado	233	77.7	100.0
<b>Cargo</b>	Auditor/contador	70	23.3	23.3
	Direção	35	11.7	35.0
	Vice-presidente	91	30.3	65.3
	Diretor de RH	77	25.7	91.0
	Presidente	27	9.0	100.0

Fonte: Elaborado pelo autor

### 1.3.1.3 Medidas

Os dados foram apurados por meio de questionários de diversas fontes. O questionário consistia em um total de 48 itens, incluindo: uma escala de 28 itens desenvolvida por A. K. Gupta & Goyal (2021) para medir a competência de análise de big data à variável independente e uma escala de 04 itens para a variável dependente “Desempenho Financeiro”, a qual foi adotada de Shashi et al. (2019). Para o mediador “eficácia do comitê de auditoria” foi usada a escala de 12 itens desenvolvida por Phornlaphatrachakorn (2020), e as capacidades dinâmicas orientadas ao processo (moderador), com base na escala de 4 itens desenvolvida por Kim et al. (2011). O questionário foi avaliado com base na escala Likert de cinco pontos, compreendendo 1 (discordo totalmente), 2 (discordo), 3 (neutro), 4 (concordo) e 5 (concordo totalmente).

### 1.3.2 Modelagem de Equações Estruturais (SEM)

SEM é chamado de técnica de "segunda geração", que é usada simultaneamente para analisar múltiplas variáveis (Hair & Alamer, 2022). O termo "modulação suave" (Henseler et al., 2015) também foi desenvolvido por Herman Wold na década de 1970 por causa de suas suposições suaves. Em muitas áreas, incluindo gerenciamento organizacional, administração de recursos humanos, marketing etc., ele recebeu grande atenção (Hair Joe F. et al., 2016). O SEM é dividido em modelagem baseada em covariância e em modelagem de equações estruturais com Mínimos Quadrados Parciais (PLS).

PLS é usado como a abordagem baseada em componentes para SEM em estudos exploratórios para prever e explicar a variação das variáveis de critério (Hair Joe F. et al., 2016).

PLS-SEM tem como características mais importantes:

- a) Analisar com robustez pequenas amostras;
- b) Gerenciamento eficiente de modelos complexos;
- c) Nenhuma suposição de normalidade dos dados;
- d) Gerenciamento de edifícios de um ponto;
- e) Medição de estruturas de treinamento.

Existem dois componentes no PLS-SEM. O primeiro componente é chamado de modelo de medição (ou modelo externo), que demonstra a relação entre construtos e seus indicadores. O segundo componente é o modelo estrutural (ou modelo interno), que mostra a conexão entre uma construção e outra (Hair Joe F. et al., 2016). O edifício pode ser exógeno ou endógeno.

Edificações exógenas são variáveis separadas e não há seta apontando para elas. Edifícios endógenos são aqueles explicados por outras variáveis (ou seja, as setas estão apontando para eles). A variável se torna independente se a construção endógena é colocada entre duas variáveis (F. Hair Jr et al., 2014). Dois estágios de resultados de PLS são avaliados. A primeira fase examina o modelo de medição e, após essa fase, avalia-se o modelo estrutural, se os resultados forem satisfatórios.

Os dados coletados são analisados por meio da modelagem de equações estruturais de mínimos quadrados parciais (PLS-SEM) usando o Smart PLS 4.0 para testar as hipóteses deste estudo. A técnica mencionada acima é mais confiável em comparação com o SEM tradicional baseado em covariância, pois este requer um grande tamanho de amostra (Cheah et al., 2023; Gotthardt & Mezhuyev, 2022).

As variáveis latentes endogenéticas são representadas por  $\mu$ , ou seja, as variáveis latentes ou fatores determinados pelas variáveis internas do modelo. A variável explícita endógena é denotada por  $y$ .  $\xi$  refere-se à variável potencial exógena, que é determinada por variáveis fora do modelo. As variáveis explícitas exógenas são indicadas por  $x$ . As variáveis endofíticas latentes foram FP ( $\mu_1$ ) e ACE ( $\mu_2$ ). As variáveis latentes exógenas são BDAC ( $\xi_1$ ) e PODC ( $\xi_2$ ). Com isso, a seguinte equação estrutural é formada com base nas variáveis latentes:

$$\mu = B\mu + \Gamma\xi + \zeta \quad (1.1)$$

A equação 1.1 reflete a relação entre as variáveis latentes, sendo  $B$  e  $\Gamma$  para a matriz de coeficientes das variáveis latentes endógenas e exógenas, e  $\zeta$  para os resíduos da regressão do vetor de erro.

$$Y = \Lambda_Y\mu + \varepsilon, \quad (1.2)$$

$$X = \Lambda_X\xi + \delta. \quad (1.3)$$

A equação 1.2 representa a relação entre as variáveis endógenas explícitas  $Y$  (FP1 a FP4, ACE1 a ACE12) e as variáveis endógenas latentes. A equação 1.3 expressa a relação entre variáveis exógenas explícitas  $X$  (BDAC [MS1 a MS5, BR1 a BR2, TAS1 a TAS3, T1 a T4, BD1 a BD2, D1 a D3, DDC1 a DDC2, DQ1 a DQ3 e DK1 a DK4], PODC1 a PODC4) e variáveis latentes exógenas, em que  $\Lambda_X$  e  $\Lambda_Y$  são a carga fatorial e as variáveis  $\delta$  e  $\varepsilon$  representam o erro de medição.

A seguinte matriz de covariância das variáveis explícitas  $(Y',X)'$  pode ser derivada dos parâmetros acima

$$\Sigma(\theta) = \begin{bmatrix} \sum_{YY}(\theta) & \sum_{YX}(\theta) \\ \sum_{XY}(\theta) & \sum_{XX}(\theta) \end{bmatrix} = \begin{bmatrix} \Lambda_Y B(\Gamma\Phi\Gamma' + \Psi)B'\Lambda_Y' + \Theta_\varepsilon & \Lambda_Y B\Gamma\Phi\Lambda_X' \\ \Lambda_X\Phi\Gamma'B'\Lambda_Y' & \Lambda_X\Phi\Lambda_X' + \Theta_\delta \end{bmatrix} \quad (1.4)$$

em que  $\Phi$  é a matriz de covariância da variável latente  $\xi$ ;  $\Psi$  é a matriz de covariância do termo residual  $\zeta$ ;  $\Theta_\varepsilon$  é a matriz de covariância de  $\varepsilon$ , e  $\Theta_\delta$  é a matriz de covariância de  $\delta$ . Se o modelo teórico for verdadeiro, então  $\Sigma(\theta)$  é igual à matriz de covariância da população  $\Sigma$ ; ou seja,  $\Sigma = \Sigma(\theta)$ , então a variância e a covariância das variáveis explícitas são funções dos parâmetros do modelo.

Como já informado, o PLS-SEM é um processo de dois estágios, sendo o primeiro a validação do modelo externo e, em seguida, a estimativa dos coeficientes do caminho do modelo interno. A validação do modelo externo é desenvolvida estimando a confiabilidade dos construtos dos coeficientes juntamente com sua validade discriminante convergente. Depois de medir a validação do modelo externo, inserimos o modelo interno para medir os coeficientes de caminho. O método *bootstrap* foi utilizado para analisar a significância dos achados.

## 1.4 Resultados

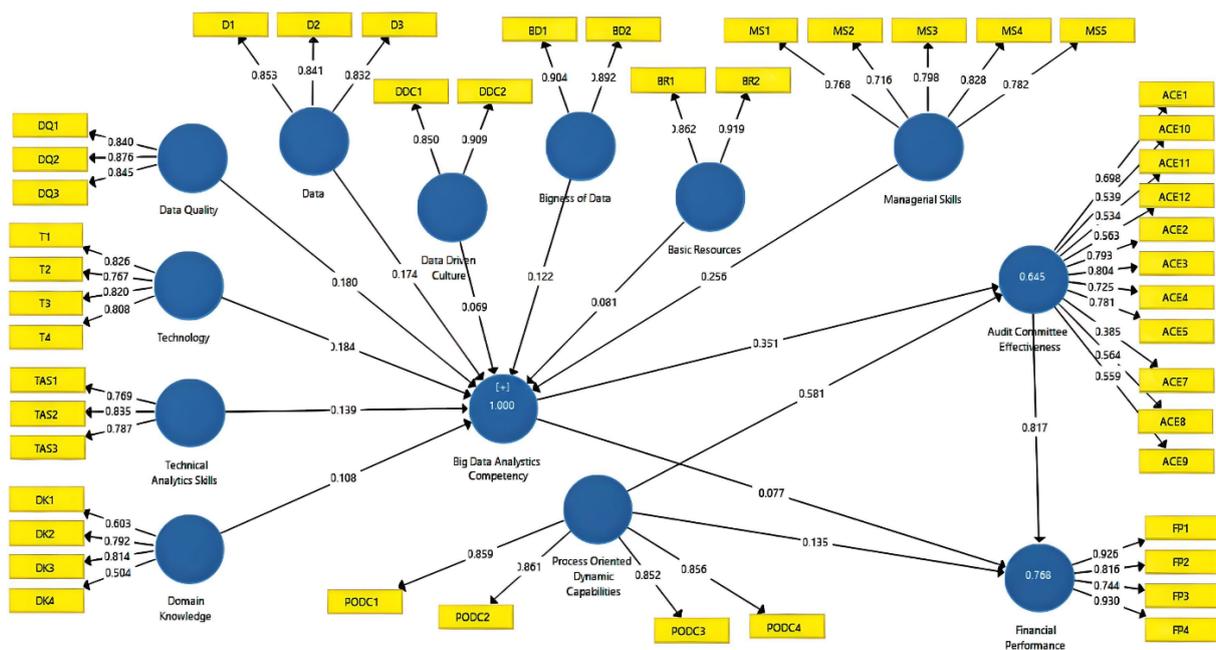
Nesta seção, serão discutidos os métodos estatísticos utilizados para analisar os dados obtidos. Dois sistemas de software diferentes foram usados para análise de dados; STATA versão 17 e Smart PLS versão 4. Será apresentada uma breve introdução ao PLS-SEM assim

como será descrito como a análise descritiva, a análise de confiabilidade, a análise de validade, a mediação e a análise de medição avaliam a medição e o modelo estrutural. As tabelas e os gráficos ilustrarão os resultados.

### 1.4.1 Modelo de medição

O modelo de mensuração é analisado por meio da validade de construtos, convergente e discriminante (confiabilidade de consistência interna). Como a estrutura de orientação empreendedora é reflexiva-formativa, foram implementadas duas etapas. A carga externa dos indicadores de construção reflexiva foi examinada primeiro (ou seja, construção de primeira ordem). Apenas os itens que atenderam aos critérios foram mantidos. Geralmente, o valor da carga externa deve exceder 0,70 (Hair & Alamer, 2022). Aqueles itens cujas cargas externas são 0,40-0,70 só devem ser removidos se a remoção aumentar a confiabilidade dos compósitos ou a variância média extraída (AVE) (Hair et al., 2016). Além disso, a quantidade de itens em uma área que pode reduzir a credibilidade da escala depende do alfa de Cronbach (Hair et al., 2017). A confiabilidade composta, portanto, fornece uma medida de confiabilidade adequada e varia entre 0 e 1. Como limiar, valores superiores a 0,70 são recomendados (Hair & Alamer, 2022). AVE é outra maneira de determinar a validade da convergência. Isso explica o quanto o elemento e sua estrutura correspondente são variáveis (Hair et al., 2017). O limiar de AVE é de 0,50 ou superior, com base nos critérios de Fornell & Larcker (1981). A estimativa inicial do *Path Model* para cargas externas é mostrada na Figura 1.2.

Figura 1.2: Análise do modelo de medição



Fonte: Elaborado pelo autor

Para determinar a validade dos construtos formativos, várias variáveis latentes de todos os edifícios de ordem inferior foram derivadas (ou seja, construto de segunda ordem). As cargas externas para cada item, alfa de Cronbach, confiabilidade composta e variância média estão apresentados na Tabela 1.2.

A confiabilidade composta é usada para testar a consistência interna. O critério implica que

$$P_c = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum_i \text{var}(\varepsilon_i)} \quad (1.5)$$

deve maior que 0,7, onde  $\lambda_i$  é a carga padronizada do indicador  $i$ ;  $\varepsilon_i$  é o erro de medição e  $\text{var}(\varepsilon_i) = 1 - \lambda_i^2$ . No caso desta pesquisa, os valores obtidos são apresentados na Tabela 1.2.

**Tabela 1.2: Modelo de Medição**

Construtos	Itens	Factor Loadings	Cronbach's Alpha	Composite Reliability	Average Variance Extracted
<b>Competência analítica de Big Data</b>	BDAC				
<i>Habilidades gerenciais</i>	MS1	0.768	0.838	0.885	0.608
	MS2	0.716			
	MS3	0.798			
	MS4	0.828			
	MS5	0.782			
<i>Recursos básicos</i>	BR1	0.862	0.743	0.885	0.793
	BR2	0.919			
<i>Habilidade de análise técnica</i>	TAS1	0.769	0.713	0.839	0.636
	TAS2	0.835			
	TAS3	0.787			
<i>Tecnologia</i>	T1	0.826	0.819	0.881	0.649
	T2	0.767			
	T3	0.820			
	T4	0.808			
<i>Grandeza dos dados</i>	BD1	0.710	0.760	0.893	0.806
	BD2	0.892			
<i>Dados</i>	D1	0.853	0.795	0.880	0.709
	D2	0.841			
	D3	0.832			
<i>Cultura orientada a dados</i>	DDC1	0.850	0.711	0.872	0.774
	DDC2	0.909			
<i>Qualidade dos dados</i>	DQ1	0.840	0.814	0.890	0.729
	DQ2	0.876			
	DQ3	0.845			
<i>Conhecimento de domínio</i>	DK1	0.603	0.618	0.779	0.477
	DK2	0.792			
	DK3	0.814			
	DK4	0.504			

<b>Desempenho Financeiro</b>	FP1	0.926	0.876	0.917	0.735
	FP2	0.816			
	FP3	0.744			
	FP4	0.930			
<b>Efetividade do Comitê de Auditoria</b>	ACE1	0.698	0.860	0.882	0.415
	ACE2	0.793			
	ACE3	0.804			
	ACE4	0.725			
	ACE5	0.781			
	ACE7	0.385			
	ACE8	0.564			
	ACE9	0.559			
	ACE10	0.539			
	ACE11	0.534			
	ACE12	0.563			
	<b>Capacidades Dinâmicas Orientadas a Processos</b>	PODC1			
PODC2		0.861			
PODC3		0.852			
PODC4		0.856			

Fonte: Elaborado pelo autor

A validade convergente é analisada por meio da variância média extraída (AVE) proposta por Fornell & Larcker (1981). O valor da expressão

$$AVE = \frac{(\sum \mu_i)^2}{(\sum \mu_i)^2 + \sum_i \text{var}(\varepsilon_i)} \quad (1.6)$$

deve ser maior que 0,5, pois mais de 50% da variância do construto devem ser explicados por suas variáveis. Em nosso estudo, o valor é 0,839.

A multicolinearidade é verificada quando dois ou mais componentes estão intimamente ligados e medidos pelo "Variance Inflation Factor" (VIF) (Hair et al., 2016). Na multicolinearidade, a estrutura formativa foi examinada. Problemas de multicolinearidade ocorrem quando um valor é maior que 5. Os valores VIF de segunda ordem mostrados na Tabela 1.3 abaixo mostram que na pesquisa atual não há problema de multicolinearidade. Os pesos externos dos indicadores formativos foram avaliados. Além disso, o bootstrapping testou a significância dos pesos. Os pesos dos indicadores são significativos conforme indicado na Tabela 1.3.

A Tabela 1.3 mostra pesos externos significativos de recursos básicos, volume de dados, dados, cultura orientada a dados, qualidade de dados, conhecimento de domínio, habilidades gerenciais, habilidades de análise técnica e tecnologia, todos significativos e ambos com seu intervalo de confiança de limite inferior (LLCI) e seu intervalo de confiança de limite superior (ULCI) possuindo o mesmo sinal. Se o peso for pequeno, mas a carga for superior a 0,50, o item deve ser retido (Hair et al., 2016). Assim, de acordo com a literatura, o item deve ser removido se a carga externa também for menor que 0,5.

**Tabela 1.3. Pesos externos e valores da variância do fator de inflação**

Relação entre Construtos	Amostra original	Média da amostra	VIF	LLCI 2.5%	ULCI 97.5%
Recursos Básicos -> BDAC	0.091	0.091	2.367	0.069	0.111
Grandeza dos Dados -> BDAC	0.123	0.123	2.270	0.108	0.139
Dados -> BDAC	0.175	0.175	2.372	0.157	0.194
Cult. Orient. a Dados -> BDAC	0.059	0.058	2.772	0.033	0.081
Qualidade dos Dados -> BDAC	0.172	0.171	2.302	0.156	0.188
Conhec. de Domínio -> BDAC	0.138	0.138	2.270	0.110	0.164
Habilidades Gerenciais -> BDAC	0.257	0.257	2.659	0.236	0.279
Hab. de Análise Téc. -> BDAC	0.132	0.132	2.584	0.114	0.149
Tecnologia -> BDAC	0.173	0.172	2.444	0.148	0.197

Fonte: Elaborado pelo autor

#### 1.4.1.1. Validade discriminatória

Para validade discriminatória, foram utilizadas as técnicas Heterotraço-monotraço (HTMT) apresentadas por Henseler et al. (2015). Essa é a proporção de correlação média dos indicadores entre as várias estruturas e a correlação média dos indicadores da estrutura associada. Segundo Hair et al. (2017), modelo com estruturas semelhantes tem um limite de 0,90, enquanto os construtos com um limite de 0,85 ou inferior não são relacionados. Na Tabela 1.4, pode-se observar que nenhum valor é superior a 0,85. Portanto, a validade discriminatória foi estabelecida.

**Tabela 1.4: Proporção Heterotraço-Monotraço (HTMT)**

Construtos	1	2	3	4	5	6
1. Eficácia do Comitê de Auditoria						
2. Competência da Análise de Big Data	0.810					
3. Desempenho Financeiro	0.837	0.564				
4. Efeito Moderador 1	0.173	0.106	0.154			
5. Efeito Moderador 2	0.173	0.106	0.154	0.137		
6. Capacidades Dinâmicas Orientadas a Processos	0.738	0.507	0.802	0.174	0.264	

Fonte: Elaborado pelo autor

#### 1.4.1.2. Avaliação do Modelo Estrutural

O método padrão de *bootstrap* com subamostra 5000 foi utilizado neste estudo para

adquirir associações significativas entre a estrutura. Seguimos os métodos recomendados por Henseler et al. (2015) para revisar os resultados mediadores da eficácia do comitê de auditoria. Quatro critérios específicos foram utilizados para analisar os efeitos diretos e indiretos dos modelos de equações estruturais. Em primeiro lugar, para calcular a quantidade de variação elucidada por todos os construtos, ou seja, estimar  $R^2$  nos endógenos das variáveis latentes (Hult et al., 2018). Porém, a avaliação satisfatória para  $R^2$ , dependendo do cenário do estudo (Cohen, 1977), mostra uma avaliação alta, moderada e baixa de 0,26, de 0,13 e de 0,09, respectivamente. O valor de  $R^2$  para Desempenho Financeiro foi de 0,759, o que é grande e mostra que BDAC, ACE e PODC elucidaram 75,9% da variação do Desempenho Financeiro.

Da mesma forma, a Tabela 1.5 e a Figura 1.3 indicam que o valor de  $R^2$  para a Eficácia do Comitê de Auditoria (ACE) também foi alto em 0,688, o que significa que 68,8% de mudança ocorreram na eficácia do comitê de auditoria devido ao BDAC e PODC.

**Tabela 1.5: Coeficiente de determinação**

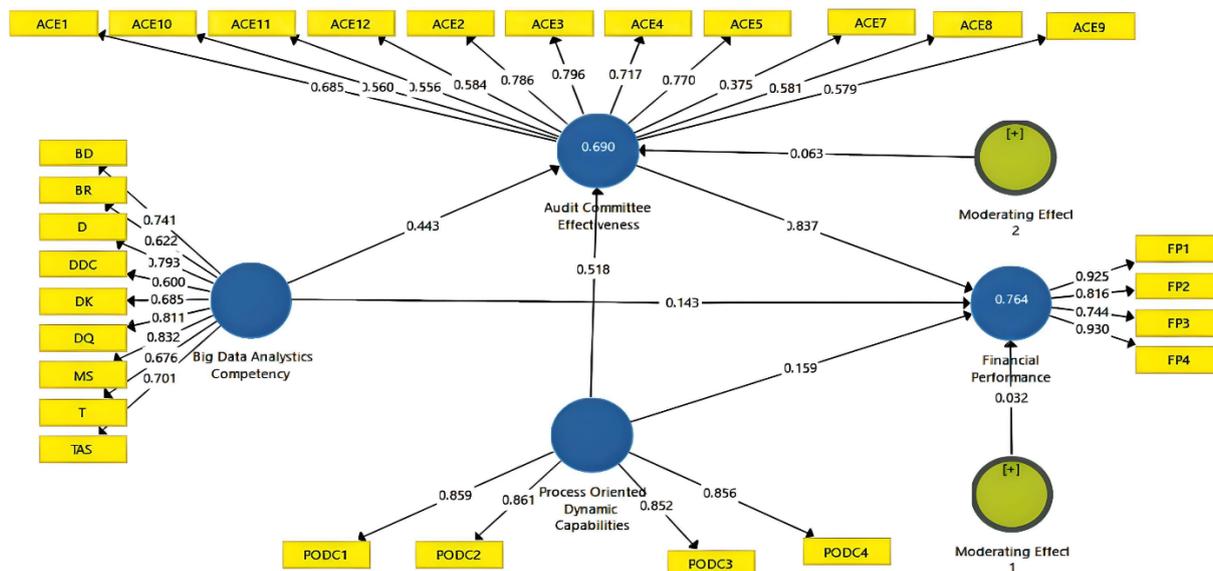
Construtos	$R^2$	$R^2$ ajustado	$Q^2 (=1-SSE/SSO)$
Eficácia do Comitê de Auditoria	0.688	0.685	0.254
Desempenho Financeiro	0.759	0.756	0.549

Fonte: Elaborado pelo autor

Em segundo lugar, a medida de relevância preditiva ( $Q^2$ ) também foi usada para medir o significado avaliado do modelo de investigação usando uma medida de redundância de validação cruzada (Hair & Alamer, 2022). A Tabela 1.5 mostra as estimativas apropriadas do valor do modelo de efeito direto, uma vez que o valor da variável latente endógena de  $Q^2$  é maior que zero, assim, tanto os construtos endógenos diretos quanto os indiretos, Desempenho Financeiro ( $Q^2 = 0,549$ ) (FP) e Eficácia do Comitê de Auditoria ( $Q^2 = 0,254$ ) (ACE), respectivamente, mostram valores superiores a zero e podem ser considerados como uma relevância preditiva aceitável do modelo (Henseler et al., 2015).

A Figura 1.3 e a Tabela 1.6 mostram que o efeito direto do BDAC no FP foi positivo e significativo ( $\beta = 0,144$ ,  $p < 0,001$ ). O coeficiente de caminho (*path coefficient*) indica que a mudança de uma unidade em BDAC provoca 14,4% de mudança em FP. Além disso, o efeito direto de BDAC em ACE ( $\beta = 0,433$ ,  $p < 0,000$ ) e de ACE em FP ( $\beta = 0,853$ ,  $p < 0,000$ ) também foi positivo e significativo. Portanto, todas as hipóteses diretas H1, H2 e H3 não foram rejeitadas.

**Figura 1.3: Análise de caminho PLS (n = 5.000 amostras autoinicializadas [bootstrapped])**



Fonte: Elaborado pelo autor

A análise da mediação é o fenômeno que explica a relação entre construtos independentes e dependentes. No presente estudo, a Efetividade do Comitê de Auditoria (ACE) mede a relação entre o BDAC e o FP. A Tabela 1.6 ilustra que os resultados do efeito indireto de ACE em BDAC e em FP também são positivos e significativos ( $\beta = 0,369$ ,  $p < 0,000$ ). Por fim, o moderador Capacidades Dinâmicas Orientadas a Processos modera positiva e significativamente a relação entre a BDAC e a ACE ( $\beta = 0,063$ ,  $p < 0,021$ ), no entanto, mostra-se insignificante na relação entre a BDAC e o FP ( $\beta = 0,034$ ,  $p = 0,109$ ) porque p-valor é maior que 0,05. Portanto, as hipóteses H4 e H6 não foram rejeitadas e a H5 foi rejeitada.

**Tabela 1.6: Resultados do Modelo de Equações Estruturais**

	Hipóteses/ Relacionamento entre Variáveis	B	Média da Amostra	Desvio Padrão	T- Valores	P- Valores	Obs.
<i>Efeito Direto</i>							
H1	BDAC -> FP	0.144	0.142	0.043	3.325	0.001	Supported
H2	BDAC -> ACE	0.433	0.434	0.052	8.344	0.000	Supported
H3	ACE -> FP	0.853	0.856	0.045	19.105	0.000	Supported
<i>Efeito Indireto</i>							
H4	BDAC -> ACE-> FP	0.369	0.372	0.048	7.645	0.000	Supported
H5	Efeito Moderador 1 -> FP	0.034	0.032	0.021	1.602	0.109	Not Supported
H6	Efeito Moderador 2 -> ACE	0.063	0.058	0.026	2.310	0.021	Supported

Fonte: Elaborado pelo autor

### 1.4.1.3. Resumo das Hipóteses

As hipóteses propostas para o estudo estão resumidas na Tabela 1.7 abaixo.

**Tabela 1.7: Resumo das Hipóteses (Não Rejeitadas e Rejeitada)**

Hipótese	Declaração	Resultado
H1	A competência de análise de big data está positivamente associada ao desempenho financeiro da empresa.	Não Rejeitada
H2	A competência de análise de big data está positivamente associada à eficácia do comitê de auditoria.	Não Rejeitada
H3	A eficácia do comitê de auditoria está positivamente associada à relação entre análise de big data e ao desempenho financeiro da empresa.	Não Rejeitada
H4	A eficácia do comitê de auditoria como mediador na relação entre a competência de análise de big data e o desempenho financeiro da empresa.	Não Rejeitada
H5	As capacidades dinâmicas orientadas a processos atuam como moderadoras na relação entre a competência de análise de big data e o desempenho financeiro da empresa.	Rejeitada
H6	As capacidades dinâmicas orientadas a processos atuam como moderadoras na relação entre a competência de análise de big data, a eficácia do comitê de auditoria e o desempenho financeiro do banco	Não Rejeitada

Fonte: Elaborado pelo autor

## 1.5 Discussão

A literatura há tempos vem demonstrando estudos importantes sobre análise de big data e desempenho da empresa (Collymore et al., 2017; Ghasemaghaei et al., 2015; Raguseo & Vitari, 2018; Song et al., 2018). Além disso, a pesquisa mostrou que o desempenho da empresa, a eficácia do comitê de auditoria e as capacidades dinâmicas orientadas a processos são essenciais para investigar e impactar o desempenho da empresa (Cao & Duan, 2015; López-Cabarcos et al., 2015; Ringov, 2017). Na sequência de pesquisas anteriores e conclusões, o objetivo principal deste estudo foi investigar a relação entre a competência de análise de big data e o desempenho financeiro no contexto da indústria bancária no Brasil. Além disso, na relação entre a competência de análise de big data e o desempenho financeiro, investigou-se o papel mediador da eficácia do comitê de auditoria e o papel moderador da capacidade dinâmica orientada a processos.

### **H1: A competência de análise de big data está positivamente associada ao desempenho financeiro da empresa**

A hipótese 1 constatou que a competência de análise de big data está positivamente associada ao desempenho e aos resultados financeiros ( $\beta = 0,144$ ,  $p < 0,001$ ) também

ênfatiou um vínculo significativo entre as variáveis. Nas últimas décadas, grandes dados surgiram como um novo campo de pesquisa. As empresas por todo o mundo estão extremamente interessadas em fazer pleno uso da capacidade do BDA (Ross et al., 2013). A literatura descobriu que a análise de big data continua a evoluir e as empresas desejam ter a capacidade de análise de big data, o que pode impactar substancialmente o desempenho corporativo (Akter et al., 2016; M. Gupta & George, 2016). Além disso, estudos recentes destacaram a competência da análise de big data como geradora de valor para as empresas, ajudando a obter insights a partir dos dados coletados e fornecendo-lhes informações sobre padrões atuais e futuros para mantê-los atualizados e tomar suas ações de acordo com os dados obtidos (Saggi & Jain, 2018). Além disso, numerosos estudos examinaram e determinaram que a análise de big data desempenhou um papel importante na obtenção de vantagens competitivas (Kubina et al., 2015; Morabito, 2015). Por conseguinte, a vantagem competitiva leva a um melhor desempenho, mantendo-os à frente. A competência em big data analytics é muito importante no setor bancário brasileiro, e o presente estudo provou empiricamente que a BDA permite às empresas operar a um nível elevado. No entanto, as organizações precisam se concentrar em várias dimensões para construir um alto nível de BDA.

## **H2: A competência de análise de big data está positivamente associada à eficácia do comitê de auditoria**

A hipótese 2 antecipou que a competência de análise de big data está positivamente associada à eficácia e aos resultados do comitê de auditoria ( $\beta = 0,433$ ,  $p < 0,001$ ), além de ênfatizar uma ligação significativa entre as variáveis. Todas as dimensões da análise de big data são observadas para aumentar a eficiência da escolha organizacional no contexto da tomada de decisão rápida, exceto a grandeza dos dados (Ghasemaghaei & Calic, 2019). Como afirmado anteriormente, a análise de big data consiste em coletar e em fornecer informações relevantes sobre desejos e ansiedades de clientes, de mercados, de fornecedores entre outros fatores.

## **H3: A eficácia do comitê de auditoria está positivamente associada à relação entre análise de big data e ao desempenho financeiro da empresa.**

A hipótese 3 antecipou que a eficácia do comitê de auditoria está positivamente associada ao desempenho e aos resultados financeiros da empresa ( $\beta = 0,853$ ,  $p < 0,001$ ), além de ênfatizar uma ligação significativa entre as variáveis. O comitê de auditoria dedica

muito tempo e esforço para permitir a separação de itens e permitir que a auditoria interna forneça resultados de alta qualidade (Suryanto, 2014). Por terem competências superiores para descobrir anormalidades e distorções do setor, uma auditoria de alta qualidade executada por uma empresa especializada no setor pode aumentar a aceitação e a credibilidade dos relatórios financeiros de uma empresa (Sirois & Simunic, 2011). O comitê de auditoria deve buscar que a auditoria interna tenha capacidade de fornecer auditorias de alta qualidade, o que decorre de sua experiência de atender a vários clientes no mesmo aprendizado e na mesma indústria e, com isso, compartilhar as melhores práticas em toda a organização. Pesquisas anteriores mostraram que a qualidade da auditoria melhora o desempenho de uma organização (Mandzila et al., 2016). Como resultado deste estudo, a premissa de que existe uma ligação entre a análise de big data, entre a eficácia do comitê de auditoria e entre o desempenho da empresa também se aplica ao setor bancário brasileiro.

#### **H4: A eficácia do comitê de auditoria como mediadora na relação entre a competência de análise de big data e o desempenho financeiro da empresa**

A hipótese 4 antecipou que a eficácia do comitê de auditoria desempenha um papel de mediadora entre a competência de análise de big data, o desempenho e o resultado financeiro da empresa ( $\beta = 0,369$ ,  $p < 0,001$ ) e também enfatizou que há um vínculo significativo entre as variáveis. A qualidade da auditoria é um dos desafios mais fundamentais atualmente nessa profissão. A possibilidade conjunta de que um erro material atual seja reconhecido e relatado por um auditor foi definida como qualidade de auditoria (Jordan et al., 2017). Grandes empresas de contabilidade são normalmente consideradas profissionais, independentes, confiáveis e qualificadas em fornecer qualidade superior de auditoria (Lokatt, 2018). As empresas com mais instabilidade inerente (ou seja, maior assimetria de informações entre a empresa e os clientes) têm um incentivo mais forte para falar sobre sua qualidade distintiva contratando um auditor extra sólido e de primeira linha. De acordo com pesquisas anteriores, não é apenas o tamanho dos dados que causa um melhor desempenho da empresa, mas, sim, como as empresas criam insights de big data e melhoram seu desempenho que impacta no resultado organizacional (Thirathon et al., 2017).

#### **H5: As capacidades dinâmicas orientadas a processos atuam como moderadoras na relação entre a competência de análise de big data e o desempenho financeiro da empresa**

A hipótese 5 antecipou que as capacidades dinâmicas orientadas a processos desempenham um papel de moderador entre a competência de análise de big data e o desempenho financeiro da empresa. No entanto, os resultados indicam que a PODC é insignificante ( $\beta = 0,034$ ,  $p=0,109$ ) entre as variáveis. Em termos de coordenação, integração, redução de custos, inteligência de negócios e aprendizado conectado às iniciativas de BDA, a PODC refere-se à medida em que uma empresa pode criar ou adquirir habilidades essenciais para transformar seus processos de negócios existentes de maneira mais robusta do que seus concorrentes (Kim et al., 2011). O termo "dinâmico" refere-se à capacidade da empresa de renovar suas possibilidades de conformidade, produzindo respostas inventivas para o ambiente de negócios em constante mudança. O termo "capacidade" mostra como habilidades e recursos organizacionais são ajustados, integrados e reconfigurados para atender às necessidades de um ambiente em mudança. O desenvolvimento de capacidades requer o trabalho coordenado de indivíduos e de equipes da organização e é crucial para as empresas nos mercados em mudança atuais (Helfat & Peteraf, 2003). O desempenho aprimorado vem do uso de recursos técnicos e da adição de recursos organizacionais para gerar capacidades (Trainor et al., 2014). Embora o papel moderador da dinâmica orientada ao processo não tenha sido investigado anteriormente neste domínio, a pesquisa existente ofereceu um terreno para investigar esse comportamento na correlação especificada, com resultados favoráveis confirmando a relação prevista.

Os resultados mostram que há um impacto negativo da orientação a processos no desempenho financeiro da empresa e na análise de big data. A capacidade da empresa de recrutar e manter consumidores e melhorar as vendas, a lucratividade e o retorno sobre o investimento (ROI) é medida pelo FP (Mithas et al., 2011; Tippins & Sohi, 2003). As exibições de recursos dinâmicos e RBV e pesquisas recentes de análise de big data são usadas neste estudo. Os resultados empíricos enfatizam a necessidade de investir em todos os recursos complementares de big data (tangíveis, humanos e intangíveis) que contribuem para o desenvolvimento de uma BDAC. As empresas atingem a aptidão evolutiva devido ao conhecimento gerado pela BDAC, que aumenta as capacidades dinâmicas de uma empresa, resultando em maiores capacidades operacionais. Para resumir, este estudo demonstra que a) big data é mais do que apenas os dados em si, b) desenvolver uma capacidade requer a consideração de vários recursos complementares e c) capturar ganhos de desempenho BDAC requer identificar os mecanismos e principais facilitadores/obstáculos que influenciam o desempenho.

## **H6: As capacidades dinâmicas orientadas a processos atuam como**

## **moderadoras na relação entre a competência de análise de big data, a eficácia do comitê de auditoria e o desempenho financeiro do banco**

A hipótese 6 antecipou que as capacidades dinâmicas orientadas a processos desempenham um papel de moderador entre a competência de análise de big data e a eficácia e resultados do comitê de auditoria ( $\beta = 0,433$ ,  $p < 0,001$ ) também enfatizou que há um vínculo significativo entre as variáveis. As capacidades dinâmicas orientadas para o processo são definidas como a capacidade de uma empresa, em termos de integração de operações, de redução de custos e de capitalização em inteligência/aprendizagem de negócios, para alterar (melhorar, adaptar ou reconfigurar) processos de negócios em vez da concorrência. Elas abrangem várias mudanças nos processos da empresa, desde modificações e aprimoramentos constantes até mudanças radicais únicas na vida. Embora a maioria dos ajustes possa ser modesta, a capacidade de uma empresa de se ajustar prontamente também sugere sua disposição de fazer ajustes drásticos sempre que necessário. Embora o papel moderador das capacidades dinâmicas orientadas ao processo não tenha sido investigado anteriormente neste domínio, a pesquisa existente ofereceu terreno para investigar esse comportamento na correlação especificada, com resultados favoráveis, confirmando a relação prevista.

### **1.5.1 Implicações da pesquisa**

O estudo atual faz uma adição substancial ao gerenciamento de análise de big data e ao desempenho da empresa, e possui uma ampla gama de aplicações. Por isso, seria relevante para a literatura abordar diretamente o comportamento moderador da capacidade dinâmica orientada a processos e o papel mediador da eficácia do comitê de auditoria nesse cenário. Além disso, a convergência de recursos organizacionais examinados neste estudo para avaliar a competência de análise de big data de uma empresa pode ser considerada relevante. Soma-se a isso o fato de que pesquisas anteriores revelaram uma escassez de literatura em análise de big data no ambiente bancário do Brasil; portanto, o presente estudo contribuiu para preencher essa lacuna na literatura.

Os resultados da pesquisa mostraram que a ligação direta e indireta entre a análise de big data e o desempenho da empresa é influenciada pelo desempenho do conselho fiscal e pela dinâmica orientada a processos e, isoladamente, a BDA não é suficiente para gerar um desempenho forte e igualmente importante para a pesquisa. A BDA torna o trabalho simples e reduz o efeito adverso. Existe uma tendência em todo o mundo de usar a BDA em vez de confiar na intuição e na experiência para tomar decisões eficazes. Para se beneficiar do Big Data, os pesquisadores fornecem orientação estratégica e prática. No entanto, as perspectivas sobre a aplicação de big data continuam a se expandir por meio de pesquisas e teorias acadêmicas rigorosas. Além disso, os mecanismos internos baseados em BDA não foram completamente estudados. Muitas investigações até o momento

exploraram as vantagens relacionadas ao desempenho da BDA. Esta pesquisa contribui de várias maneiras para a pesquisa em BDA. Embora a literatura e a prática da BDA tenham se concentrado principalmente em países industrializados, parece faltar estudo empírico em contextos de outros países, particularmente mercados emergentes. O Brasil oferece um cenário rico para expandir esta linha de pesquisa, pois tornou-se crítico para empresas de todos os setores desenvolver capacidades de BDA para competir com outras economias. Embora a maioria das economias emergentes tenha lacunas institucionais em infraestrutura econômica, jurídica e financeira, os especialistas da BDA em mercados emergentes descobriram que esta pode ajudar empresas públicas e privadas a lidar com as demandas da competitividade global.

## 1.5.2 Implicações teóricas

Usando a Teoria Baseada em Recursos (TBR) e a Visão de Capacidades Dinâmicas (DCV) como lentes teóricas, o estudo investiga a relação entre as capacidades de BDA e o desempenho da empresa. A principal justificativa para a adoção desses dois pontos de vista é que a capacidade tecnológica de utilização do BDA requer vários recursos adicionais específicos da empresa que podem eventualmente contribuir para o aumento do desempenho. Outra característica importante desta pesquisa é que ela é realizada em um importante mercado emergente, o brasileiro, que é inegavelmente um mercado em rápido crescimento com características semelhantes a outros mercados emergentes de países significativos como Rússia, África do Sul e Índia. Embora os comitês de auditoria também façam parte do sistema de controle interno de uma empresa, este estudo visa verificar se o comitê de auditoria influencia a relação entre a capacidade dinâmica orientada a processos, a BDA e o FP crescente da empresa.

A VBR da empresa identifica duas fontes potenciais de vantagem competitiva: a variedade de seus recursos estratégicos e a total imobilidade da empresa (Barney, 1991). Uma combinação eficiente de capital físico, humano e empresarial, conforme proposto por Hambrick (1987), é um recurso raro e imperfeitamente imitativo. O relacionamento interpessoal entre os gerentes, a cultura da empresa e a reputação de uma empresa entre fornecedores e clientes e sistemas de processamento de informações podem evitar a imitação dos recursos de uma empresa (Barney, 1991). Ao desenvolver os recursos de BDA, as empresas podem, assim, obter imitabilidade imperfeita, criando um ambiente de decisão probatório e uma cultura orientada por dados.

As desvantagens da VBR são destacadas por Teece & Pisano (2003), que afirmam que ela é bastante estática e incompatível com o ambiente corporativo em constante mudança. Eles argumentam que as capacidades dinâmicas são a fonte da vantagem competitiva e que o sucesso competitivo requer tanto a exploração de talentos internos e externos específicos

da empresa quanto o desenvolvimento de novos talentos. Como o comportamento, os métodos e as operações de uma empresa são difíceis de replicar, a premissa central da capacidade dinâmica é o conhecimento tácito; este é o nexos dos processos de gestão coordenada. Outra falha da VBR é que ela não leva em conta a evolução dos recursos ao longo do tempo.

## 1.6 Conclusões

A pesquisa é suportada pela Teoria Baseada em Recursos, que afirma que uma melhor integração e implantação de recursos leva a melhores resultados organizacionais. Primeiro, procurou-se integrar e implementar recursos específicos de big data analytics, que são a força motriz por trás desse tipo de análise, e, depois, o seu impacto no desempenho dos bancos brasileiros. Testes estatísticos foram usados para avaliar a confiabilidade e validade do modelo proposto. Além disso, outros testes, como análise de correlação, análise de regressão e análise de mediação, foram usados para determinar se os relacionamentos reivindicados são aceitos ou rejeitados. O estudo atual levantou a hipótese de que a BDA e o FP estão positivamente relacionados, e os resultados confirmaram isso, implicando que as empresas com alto nível de competência em análise de big data também detêm um alto nível de desempenho financeiro.

Os resultados são compatíveis com um recente estudo (Thottoli & K.V, 2022), que afirma que os auditores devem ser mais formados e a análise de dados aplicada para melhorar a sua capacidade de investigação. Em seguida, esta pesquisa afirmou que a competência de análise de big data está relacionada positivamente à eficácia do comitê de auditoria, e os resultados apoiaram essa afirmação, demonstrando que, se a competência de uma organização em análise de big data é alta, seu desempenho também o é. Além disso, os resultados da pesquisa ofereceram suporte para outra teoria, que prevê que a eficácia do comitê de auditoria está favoravelmente correlacionada com o desempenho da empresa. Ademais, as hipóteses do estudo sugeriram que a eficácia do comitê de auditoria medeia a relação entre a BDAC e o FP, e os resultados mostraram que a eficácia do comitê de auditoria mediou a relação. Ao examinar o comportamento mediador da eficácia do comitê de auditoria e moderador da capacidade dinâmica orientada a processos no setor bancário brasileiro, este estudo tentou apresentar uma compreensão holística do impacto da competência de big data analytics no desempenho da empresa.

A análise, no entanto, mostrou que, na pesquisa de auditoria, os auditores em países mais desenvolvidos estão pelo menos razoavelmente mais confortáveis com o big data analytics (Sanoran & Ruangprapun, 2023). Também foi enfatizado que o progresso técnico geral (incluindo as habilidades profissionais dos auditores) na aplicação da análise de dados de auditoria seria guiado pelo progresso no país e pelas normas internacionais de auditoria (F. Huang et al., 2022). Como são responsáveis por garantir a revisão dos

relatórios financeiros, o comitê de auditoria do conselho tem uma função a desempenhar nos negócios, fornecendo informações ou resultados abertos, direta ou indiretamente. Nesse ínterim, isso tem um impacto substancial no sucesso da empresa. O comitê de auditoria e as empresas de contabilidade desempenham um papel importante na determinação da validade, aceitação e confiabilidade da alta qualidade. Estas constatações têm consequências para a prática dos auditores no Brasil.

Ao auditar os registros dinâmicos de atividades de negócios de seus clientes ou de empresas, os auditores terão que fazer novas tentativas para sumeter suas análises de risco aprofundadas usando a big data analytics. Se a empresa melhorar seus resultados, dependerá principalmente de como a empresa e seus analistas usam o incentivo crítico. As vendas da empresa podem ser melhoradas drasticamente usando processos automatizados e as mudanças certas no momento certo, juntamente com a integração do processamento de big data (Mcbride & Philippou, 2022). Big data é usado para digitalizar o setor financeiro. A revolução digital aumenta a competitividade da empresa e usa essas inovações para satisfazer as necessidades dos consumidores nas grandes empresas. A maioria das empresas detém os registros atuais e vitais, e os efeitos e as consequências desse conhecimento coletado no setor financeiro são uma questão cada vez mais importante. Os serviços financeiros são tecnologicamente inovadores em retrospectiva e tratam os dados como um elemento central da empresa. Portanto, os resultados do estudo indicam que o setor financeiro foi revolucionado pelo uso de big data, principalmente pela mudança no comércio e nos investimentos devido ao real insight em ações e à detecção e prevenção de fraudes, bem como pelo uso de máquinas - aprender algoritmos para analisar os riscos com mais precisão. Os serviços fornecidos pela análise de big data mudam e melhoram a fidelidade do cliente, entregam vendas, minimizam o tempo de resposta para operações manuais, garantem o processamento eficiente do dispositivo, aprimoram a direção da aquisição, analisam o sucesso financeiro da empresa e melhoram e monitoram os produtos, contribuindo para o desenvolvimento dos negócios.

## **1.7 Limitações e recomendações para trabalhos futuros**

Como qualquer outra pesquisa, este estudo tem algumas limitações:

1. Esta pesquisa é limitada ao setor bancário brasileiro. Estudos futuros podem testar este modelo em diferentes indústrias, ou seja, manufatureira, hospitalar e setor de construção.
2. O presente estudo é transversal em termos de intervalo de tempo; no entanto, por se tratar de um estudo baseado em capacidades, os pesquisadores devem considerar dados longitudinais no futuro para investigar essa relação, o que pode gerar resultados diferentes.

3. O estudo atual incluiu apenas algumas das dimensões disponíveis para investigar a competência de análise de big data de uma organização; no entanto, dimensões adicionais podem ter um impacto significativo e os pesquisadores devem usar outras variáveis no futuro.

4. Embora o uso de medidas objetivas do desempenho da empresa seja benéfico, adquirir acesso a essas medidas que incluem múltiplas dimensões do desempenho da empresa é um grande desafio. Como resultado, pesquisas futuras devem combinar medidas rígidas e suaves para avaliar melhor o desempenho da empresa. Além disso, novas pesquisas são necessárias, incluindo recursos adicionais ou fatores externos que influenciam o desenvolvimento da BDA. Sendo assim, apesar de o impacto da BDAC no desempenho corporativo da empresa ter sido examinado, os pesquisadores podem, ainda, no entanto, explorar a influência da BDAC em determinados departamentos ou em determinados campos, como o gerenciamento da cadeia de suprimentos.

## Capítulo 2

### Previsão do Desempenho Financeiro Bancário no Brasil: Técnicas Avançadas de Mineração de Dados

**Resumo:** O objetivo deste artigo é aplicar várias técnicas de mineração de dados para prever o desempenho financeiro do sistema bancário brasileiro por meio dos principais determinantes exógenos de lucratividade, escolhendo a melhor técnica de mineração de dados com base nos critérios de maior pontuação de precisão de teste e de treinamento. Este estudo usou técnicas de mineração de dados para prever o desempenho financeiro do sistema bancário nacional, aplicando regressão LASSO, floresta aleatória (RF), redes neurais artificiais (RNAs) e *K-Nearest Neighbors* (KNN) em conjuntos de dados trimestrais dos 30 maiores bancos em atividade no Brasil do primeiro trimestre de 2011 ao primeiro trimestre de 2020. Este estudo usou o retorno sobre ativos (ROA) como uma medida real de desempenho financeiro, enquanto os índices de adequação de capital, de qualidade de ativos e de gerenciamento de liquidez foram usados como determinantes exógenos de desempenho financeiro. Os resultados experimentais mostraram que o modelo ideal para prever o desempenho financeiro dos bancos no Brasil é a técnica KNN, que oferece a melhor precisão de previsão e fornece o conhecimento ideal do desempenho financeiro dos determinantes bancários brasileiros. Além disso, a RF fornece valores mais próximos da precisão ideal do KNN, o que a torna outra técnica robusta para prever o desempenho financeiro do sistema bancário nacional. Esta pesquisa restringiu a modelagem do desempenho financeiro do sistema bancário brasileiro à lucratividade por meio dos principais determinantes do retorno sobre ativos no Brasil. Pesquisas futuras podem considerar a ampliação da modelagem do desempenho financeiro usando outros modelos, como CAMELS e Z-Score, para prever o desempenho financeiro do sistema bancário brasileiro, sob técnicas de mineração de dados. Devido à falta de uso de técnicas de mineração de dados no setor bancário brasileiro, este estudo preencheria a lacuna da literatura, fornecendo novas técnicas eficazes para prever o desempenho financeiro nesse setor, usando abordagens de mineração de dados, que podem ser ferramentas eficientes em modelagem de negócios e de gerenciamento, para pesquisadores financeiros e para tomadores de decisão na indústria bancária nacional. De acordo com as pesquisas feitas, este artigo pode ser relevante para fornecer técnicas de mineração de dados para prever o desempenho financeiro do setor bancário no Brasil.

**Palavras-chave:** Mineração de Dados, Bancos, Análise Financeira.

**Classificação JEL:** O32, G21, L2.

## 2.1 Introdução

Mudanças radicais no negócio bancário foram feitas por tecnologias recentes. Os novos bancos com tecnologia bancária moderna forçaram outros bancos tradicionais a adotar novas tecnologias. Para ampliar o desempenho financeiro e o relacionamento com os clientes, as instituições financeiras passaram a utilizar novos métodos de previsão para processar pedidos de diversos produtos, o que diminui tempo e custo e aumenta as receitas (Bany Mohammad et al., 2022; Sener et al., 2019). A mineração de dados oferece muitas funções efetivas, como associação, classificação, previsão, análise de cluster, análise de outliers e outras análises estatísticas (Bany Mohammad et al., 2022; Sener et al., 2019).

A mineração de dados é um método de múltiplos procedimentos de descobrir padrões em dados, por meios automáticos ou semiautomáticos, de grandes quantidades de dados qualquer que seja o volume (séries temporais ou grandes conjuntos de dados de big data) para descobrir padrões e regras significativos, em que os padrões úteis nos permitem fazer previsões não triviais em novos dados (Dogan & Birant, 2021; Wang et al., 2022; Zhang et al., 2023). A mineração de dados pode ser bem definida como o processo de extração e mineração de informações desconhecidas, ocultas e significativas de grandes bancos de dados, que são os big data. Em seguida, a mineração de dados usou essas informações valiosas e significativas para obter conhecimento para a tomada de importantes decisões financeiras e de negócios (Dogan & Birant, 2021; Wang et al., 2022; Zhang et al., 2023).

### 2.1.1 Formulação do Problema

Devido à eficácia da mineração de dados no setor bancário (Al-Hashedi & Magalingam, 2021), este artigo aplicou técnicas de mineração de dados para prever o desempenho financeiro do sistema bancário brasileiro, por causa de sua tecnologia moderna no setor financeiro, com fintechs já fazendo o uso de big data, inteligência artificial (IA) e aprendizado de máquina (ML) (Ahmed et al., 2022). De acordo com Bejar et al. (2022), o Brasil é o maior país em fintechs da América Latina; a maior fintech latinoamericana é a Nubank. O Brasil tem uma das maiores indústrias financeiras do mundo, com 159 bancos (Banco Central do Brasil, 2023a). A estrutura dos bancos no Brasil é grande e moderna, o que requer técnicas atualizadas e adequadas, como técnicas de mineração de dados, para analisar e sobreviver às competições do mercado financeiro, já que atualmente o Brasil está entre as 10 principais economias do mundo (World Economic Forum, 2023). Tudo isso leva à necessidade de avaliar o desempenho dos bancos modernos no Brasil, usando técnicas de mineração de dados, nas quais essas técnicas podem extrair o conhecimento dos dados dos bancos, não importando se os dados são grandes como big data ou pequenos como um tipo de série temporal.

Além disso, este estudo investigou a técnica mais adequada para prever o desempenho financeiro com base nos critérios de maior pontuação de precisão do processo de teste e treinamento em mineração de dados. Portanto, este artigo

respondeu à seguinte pergunta: “Qual é a técnica de mineração de dados mais adequada para prever o desempenho financeiro dos bancos no Brasil?”

## **2.1.2 Objetivo Geral**

Consequentemente, o objetivo deste estudo é fornecer, dentre várias ferramentas eficientes em modelagem de negócios e de gerenciamento para pesquisadores financeiros e tomadores de decisão no setor bancário brasileiro, qual a técnica de mineração de dados mais adequada para prever o desempenho financeiro da indústria bancária.

## **2.1.3 Justificativa**

Uma das tecnologias mais revolucionárias que têm sido utilizadas pelo setor bancário é a mineração de dados; ela mudou completamente os conceitos básicos de muitas áreas bancárias, como negócios no mercado monetário, detecção de operações fraudulentas, previsão de risco de crédito, classificação de investimentos, gerenciamento de caixa e previsão de operações e, principalmente, na avaliação do desempenho financeiro dos bancos (Cai & Zhang, 2020; Dogan & Birant, 2021; Doğuç, 2022; Gao, 2022; Liébana-Cabanillas et al., 2013; Luo et al., 2022).

A contribuição deste artigo é aplicar técnicas de mineração de dados no setor bancário brasileiro, no intuito de ser um importante estudo que usou técnicas de mineração de dados para prever o desempenho financeiro do sistema bancário no Brasil. Além disso, neste artigo, o processo de previsão de desempenho financeiro usando mineração de dados é diferente num banco moderno em relação a um banco convencional, devido à diferença na estrutura de cada setor em termos de capital, retornos, depósitos e custos operacionais. Isso é o que Iqbal et al. (2023) confirmaram quando compararam e previram o desempenho financeiro de bancos modernos e bancos convencionais em países asiáticos com base nas redes neurais artificiais (RNAs).

## **2.1.4 Estrutura e Organização do Trabalho**

O restante deste estudo está enquadrado da seguinte forma: a revisão da literatura e a lacuna do estudo são discutidas na “Revisão da literatura” (seção 2.2). Dados amostrais e coletados e o método de análise usando mineração de dados são apresentados em “Metodologia de pesquisa” (seção 2.3). As descobertas são então discutidas e deliberadas em “Discussão de resultados experimentais” (seção 2.4). Como ponto final, a conclusão e os esclarecimentos com algumas implicações são dados em “Conclusão” (seção 2.5).

## **2.2 Revisão da literatura**

Esta seção tem o objetivo de apresentar os principais conceitos sobre técnicas de mineração de dados e sobre previsão do desempenho financeiro e os estudos anteriores relacionados aos assuntos.

### **2.2.1 Técnicas de mineração de dados**

Muitos pesquisadores que aplicaram a mineração de dados no setor bancário, especialmente prevendo o desempenho financeiro, concordaram com a utilidade dessas técnicas modernas. No entanto, até onde se sabe, há uma falta de literatura que inclua o uso de mineração de dados no setor bancário brasileiro, na qual se perderam oportunidades para colher informações sobre técnicas de mineração de dados a serem usadas no setor bancário brasileiro. Assim, este artigo preenche uma lacuna na literatura usando técnicas de mineração de dados nos bancos nacionais.

Em relação à literatura sobre o uso de técnicas de mineração de dados na previsão do desempenho financeiro do setor bancário convencional, Lin et al. (2009) previram o desempenho do banco usando abordagens de mineração de dados aprimoradas em uma amostra de bancos comerciais tailandeses; aplicaram a máquina de vetores de suporte usando otimização por enxame de partículas e árvore de decisão (DT); seus resultados experimentais mostraram que os métodos utilizados obtiveram uma configuração de parâmetros eficiente e melhoraram a precisão da classificação significativa para prever o desempenho bancário.

Na mesma linha, Moradi et al. (2013) previram dificuldades financeiras usando técnicas de mineração de dados em uma amostra de fonte de dados de empresas iranianas entre 2000 e 2009; eles aplicaram C-means fuzzy e vetor de suporte; os resultados de seus experimentos mostraram que o vetor de suporte supera o C-means fuzzy.

Em estudo semelhante de Kassani et al. (2015), eles aplicaram uma mineração de dados e um modelo híbrido de análise envoltória de dados (DEA) para avaliar a eficiência em uma amostra de 589 agências bancárias localizadas no Irã, eles usaram algoritmos de agrupamento, classificador de padrão polinomial como técnicas de mineração com envoltório de análise de dados, um algoritmo de *clustering* hierárquico aglomerativo, K-means e clusters RNA; os resultados do estudo demonstraram que a abordagem modular tem um desempenho significativamente melhor do que a abordagem não modular com base em termos de precisão de classificação.

Da mesma forma, S. Moradi & Mokhatab Rafiei (2019) aplicaram técnicas de mineração de dados integradas ao modelo de avaliação de risco de crédito em uma amostra de bancos iranianos; eles usaram todos os métodos fuzzy e uma rede adaptativa fuzzy; seus resultados mostraram que o modelo construído com base em mineração de dados é útil e eficaz na avaliação do risco de crédito e supera os modelos clássicos.

### **2.2.2 Previsão do desempenho financeiro**

Em outro estudo de Anouze & Bou-Hamad (2019), eles usaram técnicas de mineração

de dados e DEA de segundo estágio para prever o desempenho do banco por meio de fatores ambientais em uma amostra de bancos na região MENA; eles usaram floresta aleatória (RF), ensacamento, regressão logística e RNAs; suas descobertas mostraram que a RF e o método de ensacamento superaram outras técnicas artificiais com base na precisão preditiva.

Com relação à literatura sobre o uso de técnicas de mineração de dados na previsão do desempenho financeiro do setor bancário, até onde se sabe, poucos estudos limitados usaram técnicas modernas de mineração de dados no setor bancário, especialmente na previsão do desempenho financeiro dos bancos modernos. A maioria desses estudos se concentrou no uso de RNAs para prever o desempenho financeiro no setor bancário. Em um estudo de Anwar & Ali (2018), o desempenho financeiro dos bancos islâmicos foi previsto usando a técnica RNA. Esses autores usaram índices financeiros como insumos para prever o desempenho financeiro dos bancos islâmicos indonésios. Seus resultados mostraram que o modelo de rede neural proposto para prever o desempenho financeiro era preciso para bancos islâmicos na Indonésia. Em outro estudo mais recente, Bolívar et al. (2023) previram a sobrevivência e o desempenho do sistema bancário da União Europeia usando a técnica RF para crescer uma floresta de árvores de decisão, calcular as contribuições dos componentes do modelo de negócios para a lucratividade por meio do algoritmo do interpretador de árvores, agrupar os bancos por meio de análise de agrupamento e, finalmente, analisar os recursos característicos dos modelos de negócios; suas descobertas demonstraram que a RF é eficiente em prever o desempenho financeiro e a sobrevivência dos bancos europeus. No mesmo contexto, Wanke et al. (2016) previram a eficiência bancária islâmica da Malásia usando redes neurais e a abordagem TOPSIS; seus resultados indicaram que a técnica de redes neurais fornece uma eficiência preditiva relevante na avaliação de bancos islâmicos na Malásia. Da mesma forma, Bakar & Tahir (2009) previram o desempenho financeiro do sistema bancário islâmico da Malásia usando a técnica de redes neurais e regressão linear múltipla; sua descoberta revelou que as redes neurais podem ser uma ferramenta eficiente para prever o desempenho financeiro do sistema bancário islâmico da Malásia. No mesmo contexto, Al-Osaimy (1998) previu o desempenho bancário islâmico usando métodos de rede neural de Kohonen e *Backpropagation* em uma amostra de 26 bancos islâmicos internacionais; seus resultados indicaram que a técnica de redes neurais forneceu alta capacidade de previsão para o sistema bancário islâmico.

Além disso, geralmente, apenas alguns estudos limitados usaram as técnicas modernas de mineração de dados no sistema bancário fora do propósito de prever o desempenho financeiro; no entanto, o único foco foi o uso da técnica de redes neurais, como Aydin et al. (2022), Hu & Su (2022), Naveed et al. (2023) e Yan (2022) que fizeram previsões usando RNAs como a única melhor técnica.

Em conclusão, as técnicas de mineração de dados podem ser ferramentas eficazes para prever o desempenho financeiro tanto no setor bancário convencional quanto no setor bancário moderno; no entanto, o que torna a mineração de dados na previsão do desempenho financeiro do banco moderno diferente do banco convencional é a estrutura

de cada setor em termos de capital, retornos, depósitos e custos operacionais. Isso é o que Fakhri et al. (2019) confirmaram quando compararam e previram o desempenho financeiro de bancos modernos e bancos convencionais na Indonésia usando uma das técnicas de mineração de dados, que são as RNAs. Além disso, com base na literatura, todos os estudos de Lin et al. (2009), Moradi et al. (2013), Kassani et al. (2015), Moradi & Mokhatab (2019) e Anouze & Bou-Hamad (2019) utilizaram técnicas de mineração de dados no setor bancário convencional para prever e avaliar as operações bancárias e financeiras. No que este artigo revisou, os estudos de Aydin et al. (2022), Kupka et al. (2020), Luo et al. (2022), S. Moradi & Mokhatab Rafiei (2019) e Sener et al. (2019) provaram que a mineração de dados é adequada e eficiente para prever o desempenho financeiro. No entanto, falta o uso de técnicas modernas de mineração de dados no setor bancário brasileiro, que podem ser instrumentos eficazes na previsão do desempenho financeiro; a maioria dos estudos utilizou apenas a técnica de rede neural para prever o desempenho financeiro. Soma-se a isso o fato de que outras técnicas de mineração de dados estavam ausentes, o que torna a falta de uso de técnicas de mineração de dados na previsão do desempenho financeiro no setor bancário brasileiro uma lacuna na literatura. Este artigo aplicou várias técnicas de mineração de dados para prever o desempenho financeiro do setor bancário nacional, escolhendo a melhor técnica de previsão com base nos critérios da pontuação de precisão ideal de muitos processos de teste. Consequentemente, acredita-se que este estudo trará uma contribuição significativa para a literatura, pois é o primeiro que aplicou várias técnicas de mineração de dados no setor bancário brasileiro para prever o desempenho financeiro.

## **2.3 Metodologia de pesquisa**

### **2.3.1 Amostra e coleta de dados**

Foram coletados conjuntos de dados trimestrais ao longo do primeiro trimestre de 2011 até o primeiro trimestre de 2020 dos 30 maiores bancos ativos do Brasil. Tentou-se evitar o período pandêmico, a fim de se garantir resultados estáveis, em momentos de normalidade econômica. Os dados foram extraídos do sistema IF.data do Banco Central do Brasil (Banco Central do Brasil, 2023b), que fornece uma visão geral do desenvolvimento do sistema bancário brasileiro (Tabela A2.1).

Foi escolhida uma amostra das 30 maiores instituições financeiras organizadas em ordem decrescente do ativo total e pertencentes, em sua grande maioria, ao Tipo de Consolidado Bancário “b1” (Banco Comercial, Banco Múltiplo com Carteira Comercial ou Caixa Econômica), tendo ainda, entre as 30 maiores, apenas um banco de desenvolvimento (BNDES) – tipo b4, um banco múltiplo sem carteira comercial (Volkswagen) – tipo b2, uma instituição não bancária de crédito (NU Pagamentos) – tipo n1, e uma instituição de pagamento (Cielo) – tipo n4.

Os bancos selecionados para a pesquisa distinguem-se pela origem do controle de

capital, ou seja, o tipo de controle que, segundo o Banco Central do Brasil (2023b), pode ser público (por exemplo, Banco do Brasil e Caixa Econômica Federal), privado (por exemplo, Banco Itaú Unibanco e Banco Bradesco) e privado com controle estrangeiro (por exemplo, Banco Santander). O critério adotado para a inclusão de apenas os 30 maiores bancos pelo ativo total deve-se, principalmente, a aspectos de sua significativa participação no Sistema Financeiro Nacional (SFN) (o sistema inclui bancos, cooperativas de crédito e instituições não bancárias) e ao fato de que seus ativos totais somados representam, em média, 91,54% de todo o ativo total do sistema financeiro entre o primeiro trimestre de 2011 e o último trimestre de 2022, segundo os dados coletados no portal IF.Data (Tabela 2.1).

**Tabela 2.1 – Percentual do Ativo Total dos 30 maiores bancos no Sistema Financeiro do Brasil.**

Trimestre	Ativo Total do SFN (em R\$ mil)	Ativo Total dos 30 maiores bancos (em R\$ mil)	Percentual do ativo total dos 30 maiores em relação ao SFN
2011T01	4.573.683.845	4.239.710.948	92,70%
2011T02	4.739.208.111	4.402.817.668	92,90%
2011T03	4.992.714.716	4.641.926.777	92,97%
2011T04	5.078.485.742	4.722.098.832	92,98%
2012T01	5.285.650.391	4.910.917.263	92,91%
2012T02	5.463.995.073	5.079.143.965	92,96%
2012T03	5.708.657.512	5.318.006.110	93,16%
2012T04	5.910.508.940	5.525.308.260	93,48%
2013T01	6.067.588.189	5.664.022.604	93,35%
2013T02	6.303.853.869	5.883.878.062	93,34%
2013T03	6.446.675.627	6.019.972.099	93,38%
2013T04	6.521.449.065	6.092.535.899	93,42%
2014T01	6.812.431.051	6.340.183.006	93,07%
2014T02	6.865.301.814	6.398.250.853	93,20%
2014T03	7.144.544.925	6.635.879.981	92,88%
2014T04	7.455.863.006	6.947.255.857	93,18%
2015T01	7.828.665.349	7.287.592.120	93,09%
2015T02	7.718.827.798	7.178.780.353	93,00%
2015T03	8.265.453.388	7.682.705.757	92,95%
2015T04	8.215.758.194	7.644.743.408	93,05%
2016T01	8.157.328.560	7.578.001.311	92,90%
2016T02	8.245.644.963	7.666.384.746	92,97%
2016T03	8.253.989.990	7.671.037.612	92,94%
2016T04	8.110.164.674	7.513.991.927	92,65%
2017T01	8.157.354.750	7.549.429.703	92,55%
2017T02	8.195.990.443	7.582.680.153	92,52%
2017T03	8.224.882.339	7.569.932.225	92,04%
2017T04	8.277.501.914	7.598.129.145	91,79%
2018T01	8.497.985.839	7.802.340.983	91,81%
2018T02	8.632.099.576	7.902.890.851	91,55%
2018T03	8.864.898.100	8.089.193.116	91,25%
2018T04	8.864.093.428	8.054.908.377	90,87%
2019T01	9.176.792.754	8.347.851.342	90,97%
2019T02	9.312.248.831	8.451.247.650	90,75%

2019T03	9.361.665.945	8.454.767.898	90,31%
2019T04	9.279.547.771	8.358.522.190	90,07%
2020T01	10.235.891.782	9.247.314.608	90,34%
2020T02	10.786.994.310	9.722.360.679	90,13%
2020T03	11.275.457.962	10.117.568.398	89,73%
2020T04	11.195.441.927	9.983.332.988	89,17%
2021T01	11.483.358.416	10.251.643.612	89,27%
2021T02	11.397.927.576	10.125.057.862	88,83%
2021T03	11.972.812.382	10.611.013.772	88,63%
2021T04	11.885.465.236	10.480.904.751	88,18%
2022T01	12.272.361.562	10.782.407.156	87,86%
2022T02	12.751.202.849	11.166.457.332	87,57%
2022T03	13.387.302.105	11.689.278.372	87,32%
2022T04	13.319.319.914	11.570.133.882	86,87%
<b>Média</b>	<b>8.478.646.719</b>	<b>7.719.885.676</b>	<b>91,54%</b>

**Fonte: Elaborado pelo autor com os dados do IF.Data do Banco Central do Brasil**

Dado isso, a seleção de apenas 30 dentre os maiores bancos do Brasil se justifica porque estes abrangem a maioria dos ativos e operações do setor bancário do país, refletindo sua importância econômica e estratégica (Garcia & Meurer, 2022). Outro fator importante é que a concentração de bancos no Brasil é relativamente alta, com poucas instituições dominando a maior parte do mercado financeiro (Elfeituri, 2022). Portanto, selecionar os 30 maiores bancos ajudaria a capturar a maioria dos recursos e transações do setor. Além disso, os grandes bancos no Brasil compartilham características semelhantes em termos de estrutura organizacional, operações, estratégias e regulamentação aplicável (Capeleti et al., 2022). Essa homogeneidade pode aumentar a validade interna da pesquisa, permitindo análises comparativas mais robustas.

Pode-se concluir a partir disso, ainda, que os 30 maiores bancos no Brasil representam a maioria dos ativos e passivos do setor bancário. Isso significa que eles são os principais participantes dessa indústria e seu desempenho provavelmente terá um impacto significativo no desempenho geral do setor (da Silva et al., 2022). Esses bancos também possuem mais recursos e expertise para coletar e manter dados financeiros precisos. Isso significa que os dados coletados desses bancos têm mais probabilidade de ser confiáveis e representativos do desempenho geral do setor bancário (Ahamed et al., 2021). Além disso, eles estão mais sujeitos à supervisão regulatória; isso significa que seu desempenho financeiro é mais provável de ser transparente e responsável (Montes et al., 2021).

Considera-se, assim, para a finalidade desta pesquisa, que a amostra contando com apenas os 30 maiores bancos é suficiente para se prever o desempenho financeiro do setor bancário brasileiro.

## 2.3.2 Variáveis experimentais

### 2.3.2.1 Variáveis de entrada

Na sequência de estudos que avaliaram o desempenho financeiro do sistema bancário

através dos seus determinantes exógenos, este artigo utilizou o rácio de adequação de capital (CAR) (Cai & Zhang, 2020; Chen et al., 2022; Lin et al., 2009), a qualidade de ativos (AQ) (Ashiru et al., 2023; Gong & Wei, 2022) e a gestão de liquidez (LM) (Jeanne & Sandri, 2023; Tata, 2023) como os principais insumos que determinaram o ROA.

O CAR, também conhecido como Índice de Basileia ou Índice de Adequação de Capital, é calculado dividindo o capital de um banco por seus ativos ponderados pelo risco. A fórmula geral para calcular o CAR é a seguinte:

$$\text{CAR} = \text{Patrimônio de Referência para Comparação com o RWA} / \text{RWA} \quad (2.1)$$

onde o Patrimônio de Referência para Comparação com o RWA é calculado somando-se o capital Tier 1, que é o capital mais sólido e de melhor qualidade (que inclui o capital social, as reservas e as ações preferenciais não cumulativas), com o capital Tier 2, que é o capital de qualidade inferior (que inclui os empréstimos subordinados e as reservas de reavaliação) (Moyer, 1990; Karim et al., 2022). Os *Risk-Weighted Assets* (RWA) são os ativos ponderados pelo risco, que levam em consideração o risco de crédito, o risco de mercado e o risco operacional dos ativos do banco (Moyer, 1990; Karim et al., 2022). De acordo com as normas estabelecidas pelo Banco Central do Brasil, os bancos devem manter um CAR mínimo de 11% sobre seus ativos ponderados pelo risco.

A qualidade dos ativos (AQ), também conhecido como ativo total ponderado pelo risco de crédito, revela a perda esperada sobre os ativos. Para calcular o AQ dos bancos, é possível usar a seguinte fórmula:

$$\text{AQ} = 1 - (\text{Operações de Crédito Líquidas de Provisão} / \text{Operações de Crédito}) \quad (2.2)$$

onde as operações de crédito líquidas de provisão são as operações de crédito totais descontadas das provisões sobre os empréstimos em atraso (que não foram pagos no prazo), sobre os empréstimos em processo de renegociação ou reestruturação, sobre as garantias executadas e sobre os ativos imobiliários adquiridos por meio de execução de garantias (Kladakis et al., 2020; Gong & Wei, 2022). O índice revela o risco de perda sobre os ativos do banco, portanto, quanto menor é o índice, melhor é a qualidade dos ativos (Kladakis et al., 2020; Gong & Wei, 2022).

A gestão ou gerenciamento de liquidez (LM) é o índice que revela o risco de disponibilidade de recursos líquidos para garantir que a instituição financeira possa atender às suas obrigações quando necessário (Bianchi & Bigio, 2022; Li, 2021). Para calcular o LM dos bancos, é possível usar a seguinte fórmula:

$$\text{LM} = 1 - (\text{Captações} / \text{Passivo Circulante}) \quad (2.3)$$

onde as captações incluem os recursos que podem ser facilmente convertidos em dinheiro, como os depósitos à vista, depósitos a prazo, letras de crédito, debêntures, empréstimos interbancários, captações externas entre outros, e o passivo circulante inclui as obrigações que devem ser pagas em um curto prazo, como as captações, instrumentos derivativos e outras obrigações (Bianchi & Bigio, 2022; Li, 2021).

### 2.3.2.2 Variáveis de saída

Segundo Adhikari et al. (2023), Chen (2020) e Chen et al. (2022), o retorno sobre o ativo (ROA) é a medida de lucratividade ideal do desempenho financeiro no setor bancário. Consequentemente, este artigo usou o ROA como variável-alvo (variável dependente) para prever o desempenho financeiro do sistema bancário no Brasil. Assim, esta pesquisa é semelhante ao estudo de Bakar & Tahir (2009), em que o ROA foi utilizado como variável-alvo na previsão do desempenho financeiro de bancos islâmicos da Malásia com base na técnica de redes neurais.

É possível determinar o ROA de um banco considerando essas três medidas de desempenho financeiro. Para calcular o ROA pelo CAR, é necessário dividir o lucro líquido do banco pelo ativo total, ponderado pelo índice CAR. Essa medida indica a rentabilidade do banco em relação ao seu nível de capitalização e capacidade de absorção de riscos (Ko & Park, 2022; Tariq & Masih 2021).

$$\text{ROA pelo CAR} = (\text{Lucro Líquido} / \text{Ativo Total}) * \text{CAR} \quad (2.4)$$

Para calcular o ROA pelo AQ, é necessário dividir o lucro líquido do banco pelo ativo total, ponderado pelo índice AQ. Essa medida indica a rentabilidade do banco em relação à qualidade de sua carteira de empréstimos e ao controle de risco de crédito (Lin et al., 2022; Hassan & Marimuthu, 2022).

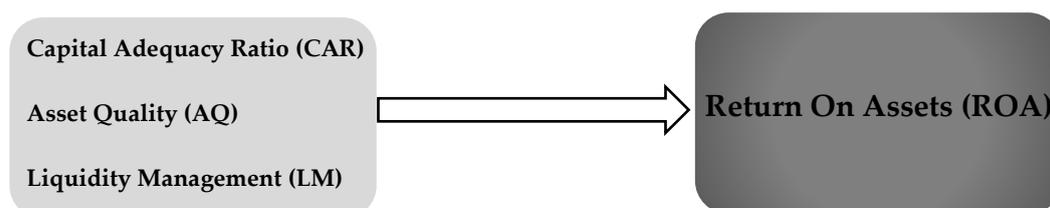
$$\text{ROA pelo AQ} = (\text{Lucro Líquido} / \text{Ativo Total}) * \text{AQ} \quad (2.5)$$

Para calcular o ROA pelo LM, é necessário dividir o lucro líquido do banco pelo ativo total, ponderado pelo índice LM. Essa medida indica a rentabilidade do banco em relação à sua capacidade de gerenciar a liquidez e os fluxos de caixa (Al-Khouri & Al-Otaibi, 2022; Beder, 2021).

$$\text{ROA pelo LM} = (\text{Lucro Líquido} / \text{Ativo Total}) * \text{LM} \quad (2.6)$$

Assim, a previsão do desempenho financeiro dos bancos brasileiros é baseada em variáveis exógenas, em que o alvo previsto é o ROA como uma medida real do desempenho financeiro dos bancos nacionais por meio de seus determinantes de CAR, AQ e LM (Figura 2.1).

Figura 2.1: Determinantes do desempenho financeiro



Fonte: Elaborado pelo autor

## 2.3.3 Metodologia de mineração de dados

Para alcançar a previsão ideal para o alvo do estudo, foram aplicadas várias técnicas de mineração de dados usando a biblioteca *Orange*, que contém algoritmos de análise de dados da linguagem *Python*, como regressão LASSO, RF, RNAs e *K-Nearest Neighbor* (KNN).

### 2.3.3.1 Regressão Lasso

De acordo com Centofanti et al. (2022), Ciner et al. (2022) e Wada (2022), a estimativa LASSO  $\tilde{B}_\lambda$  é a solução subsequente dos problemas de minimização idênticos para observações iguais  $\{A, V\}$ :

$$\| \mathbf{V} - \mathbf{AB} \|^2 + \lambda \| B \|_1 \rightarrow \min, \quad (2.7)$$

Para todo  $\lambda > 0$ , existe um  $t(\lambda) > 0$  tal que:

$$\| \mathbf{V} - \mathbf{AB} \|^2 \rightarrow \min \text{ sujeito a } \| B \|_1 \leq t(\lambda) \quad (2.8)$$

onde  $\| B \|_1 = \sum_{j=1}^{k-1} |\beta_j|$ .

A consequência do vetor fator  $\beta_j, j = 1, \overline{k-1}$  forçado pela regressão LASSO é um pouco diferente da regressão Ridge. No caso de LASSO, a restrição  $\lambda$  é reproduzida pela norma  $l_1$  do vetor de  $(\beta_1, \dots, \beta_{k-1})$ .

A regularização LASSO também realiza a seleção de variáveis, o que facilita a compreensão do modelo devido às suas propriedades positivas. Seus resultados no modelo são tais que algumas estimativas de coeficientes são exatamente iguais a zero quando  $\lambda$  é grande e diferentes valores de  $\lambda$  geram diferentes vetores  $\tilde{B}_\lambda$ . Portanto, é necessário escolher um valor adequado para  $\lambda$ .

Para esse fim, a validação cruzada é usada para encontrar o valor correto para o parâmetro  $\lambda$  que permite a previsão dos valores de resposta com máxima precisão (Patonai et al., 2022).

De acordo com Bedoui & Lazar (2020), Decelle (2022), Frndak et al. (2023) e James et al. (2021), o processo de realização da validação cruzada começa com a divisão dos dados em duas subseções: o primeiro processo é o treinamento e o outro é o teste. O processo de treinamento é usado para calcular as estimativas de coeficientes; em seguida, essas estimativas são validadas no processo de teste.

O algoritmo de validação cruzada é um processo de dividir aleatoriamente o conjunto de dados inicial em Q blocos de dimensão equivalente. Cada bloco é uma parte alocada do processo de teste, enquanto os Q-1 blocos restantes iniciam o processo de treinamento. Geralmente, o número Q de blocos é escolhido como cinco ou dez. Em seguida, essa técnica indica uma rede de valores  $\lambda = [\lambda_s]$  e calcula os coeficientes na regressão para cada valor  $\lambda_s$ . Dados esses coeficientes na regressão, este estudo calcula a soma residual dos quadrados na equação a seguir:

$$RSS_{\lambda_s}^q = \sum_{i=1}^n \left( y_i - \sum_{j=0}^{k-1} \hat{b}_j(q, \lambda_s) x_{ij} \right)^2 \quad (2.9)$$

onde:  $q = \overline{1, Q}$  é o índice do bloco designado no processo de teste, em que o ganho médio dos valores de RSS determinou o total de blocos nesta seguinte equação:

$$MSE_{\lambda_s} = \frac{1}{Q} \sum_{q=1}^Q RSS_{\lambda_s}^q \quad (2.10)$$

Este estudo aplicou a regressão LASSO na força de regularização de 0,0001 Alpha como nível de significância.

### 2.3.3.2 Floresta aleatória

RF é um agrupamento (cluster) de Árvores de Decisão (DTs) (Granitto et al., 2007), que é criado após uma abordagem principal efetiva destinada a aumentar a diversidade entre as árvores. DTs são perturbadas porque são abordagens com apenas uma pequena variação no conjunto de dados que podem causar grandes variações no modelo estabelecido (Breiman, 1996). Então, para aumentar a diversidade entre as associações de cluster, a RF ajusta cada árvore em um processo de *bootstrap* para imitar o conjunto total de amostras. O processo de *bootstrap* é uma subseção aleatória do conjunto de dados existente com a mesma dimensão, reservado com auxílio (Efron & Tibshirani, 1994).

Outro processo de diversidade é introduzido durante o crescimento de cada árvore; em cada nó, o método seleciona uma pequena subclasse aleatória de atributos e usa apenas essa subclasse para procurar a melhor divisão (Cao et al., 2023).

A mistura dos dois processos de diversidade (*bootstrap* com escolha em cada nó apenas de uma subclasse de características) gera uma construção melhor dos conjuntos com desempenho robusto para uma técnica de previsão em mineração de dados (Cao et al., 2023; Granitto et al., 2007).

Este estudo aplicou dois processos concluindo dez árvores com dados classificados em um nível de maioria de 95%.

### 2.3.3.3 Redes neurais artificiais

As técnicas de mineração de dados mais comumente usadas tanto para predição quanto para classificação são RNAs. Uma RNA é estimulada pela construção da rede neural biológica, na qual os neurônios são inter-relacionados e simulados em aprendizagem por experiência. A RNA é composta por neurônios organizados em camadas que são totalmente ligadas à camada anterior por meio de um sistema de pesos. Várias arquiteturas de rede neural diferentes já foram projetadas, no entanto, as redes de alimentação múltipla

em camadas são consideradas as aplicações mais eficazes e já foram testadas em aplicações financeiras, como previsões de desempenho financeiro e previsões de falência. Esses tipos de RNAs são compostos pela camada de entrada, uma ou mais camadas ocultas e uma camada de saída (dos Santos et al., 2019; Wang et al., 2022).

A camada de entrada é onde as características de entrada são alimentadas e encaminhadas para a camada oculta, que é novamente encaminhada para a camada de saída (dos Santos et al., 2019; Wang et al., 2022).

A saída do nó da camada oculta é calculada da seguinte forma: inicialmente, é calculada uma soma ponderada das entradas e, em seguida, um processo é aplicado a essa soma chamado de função de transferência (dos Santos et al., 2019; Wang et al., 2022).

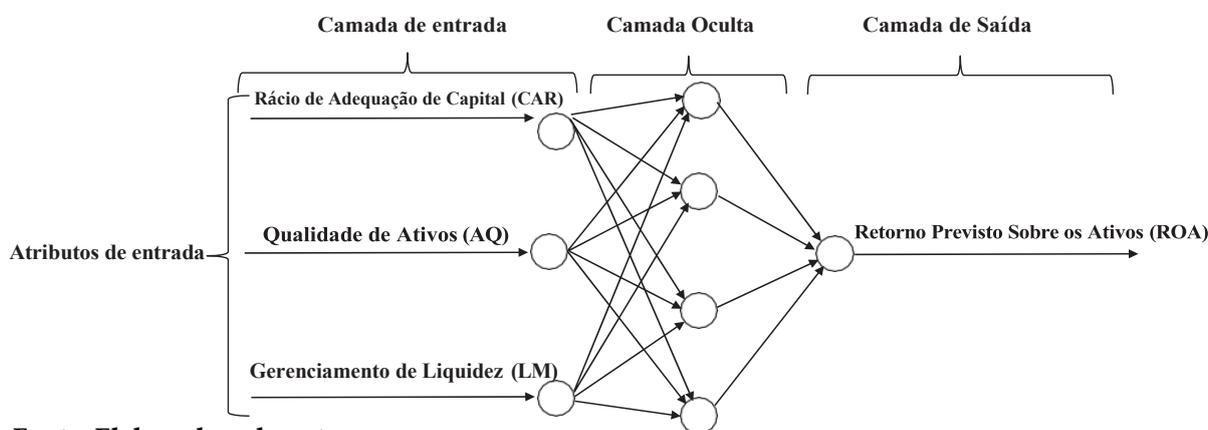
Para os valores de entrada,  $x_1, x_2, \dots, x_m$ , a saída do nó  $j$  é calculada usando a soma ponderada  $\theta_j + \sum_{i=1}^m \omega_{ij}x_{ij}$  onde  $\theta_j, \omega_{1j}, \dots, \omega_{mj}$  são pesos que são estabelecidos principalmente aleatoriamente e regulados à medida que a rede aprende. A etapa seguinte é aplicar a função de transferência a essa soma, em que a função monótona é a função de transferência. A função de ativação mais comum é a função logística:

$$g(s) = \frac{1}{(1 + e^{-s})} \quad (2.11)$$

Finalmente, a camada de saída recebe os valores da camada de entrada da camada oculta e a mesma função de transferência é executada novamente para gerar a saída (dos Santos et al., 2019; Wang et al., 2022).

O algoritmo de retropropagação é o mais comum utilizado para estimar e atualizar os pesos (Rumelhart et al., 1986). No entanto, esse algoritmo comum tem uma baixa velocidade de aprendizado (Castillo et al., 2006). Por esse motivo, este estudo utilizou o algoritmo Broyden-Fletcher-Goldfarb-Shanno para resolução com regularização de alfa = 0,0001, 500 iterações, concluindo com 500 neurônios nas camadas ocultas. Além disso, este estudo ativou a rede neural com a função logística e o treinamento replicável. Por conseguinte, a estrutura de rede neural típica para prever o desempenho financeiro do sistema bancário brasileiro é apresentada na Figura 2.2.

**Figura 2.2 - Estrutura de rede neural para prever o desempenho financeiro de bancos no Brasil.**



Fonte: Elaborado pelo autor

### 2.3.3.4 *K-nearest neighbor*

De acordo com Hastie et al. (2009) e Urso et al. (2018), o KNN funciona como um método de regressão quando aplicado a um conjunto de dados de casos  $(x_i; y_i)$  definidos por características numéricas como um elemento do vetor  $x_i$  e um valor numérico correspondente  $y_i$ . Portanto, permite prever o valor  $\hat{y}$  consistente com um vetor de ocorrência oculta  $X$  com base nas informações do KNN obtidas durante o processo de treinamento, que delibera a instância simples do conjunto de dados pronunciada por uma única característica. O  $i$ -ésimo caso do conjunto de dados é caracterizado pelo par  $(x_i; y_i)$ , onde  $x_i$  é um escalar.

Dado um novo caso  $(x = k; \hat{y} = ?)$ , o KNN fornece como saída um número  $\hat{y}$  decorrente da seguinte equação:

$$\hat{y} = \frac{1}{k} \sum_{x_i \in N_k(x)} y_j \quad (2.12)$$

onde:  $N_k(x)$  é o conjunto  $\{x_i, \dots, x_k\}$  de vizinhos do processo de treinamento do novo valor  $x$ . Evidentemente,  $N_k(x)$  depende do valor selecionado para  $K$ .

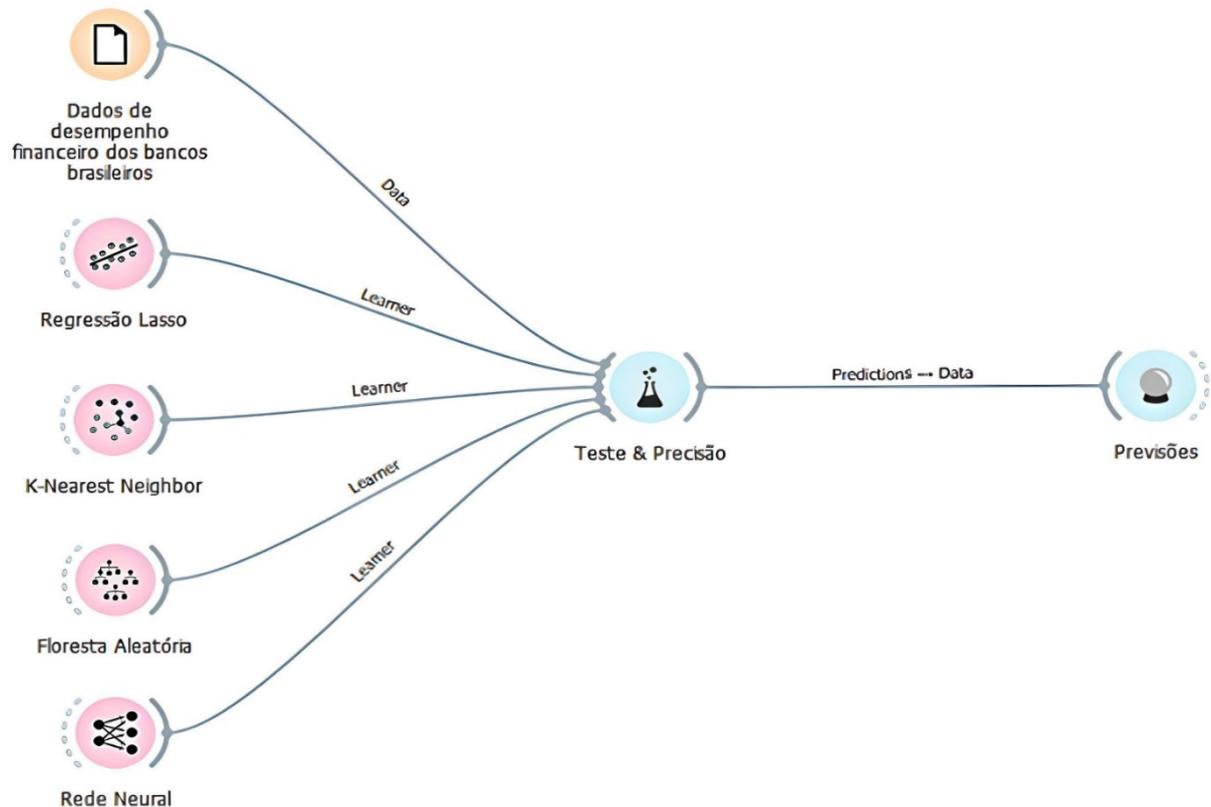
O KNN não adota uma forma óbvia para a função de previsão. Portanto, a regressão KNN é uma técnica não paramétrica.

Seguindo dos Santos et al.(2019) e Wang et al. (2022) ao prever o desempenho financeiro do sistema bancário brasileiro com base nas instâncias de treinamento mais próximas, este estudo usou cinco vizinhos com métrica de Manhattan e ponderamento pela distância.

## 2.4 Discussão dos resultados

Nesta pesquisa, a principal operação antes da previsão é escolher o melhor método de mineração de dados com base na melhor pontuação de precisão (Figura2.3).

Figura 2.3 - Resumo da previsão do desempenho financeiro do sistema bancário no Brasil por meio de técnicas de mineração de dados.



Fonte: Elaborado pelo autor

## 2.4.1 Teste e precisão

A análise de regressão compõe uma grande parte do aprendizado de máquina supervisionado e consiste na previsão de um alvo independente contínuo de um conjunto de outras variáveis de previsão; a diferença entre classificação binária e regressão está no intervalo de destino: na classificação binária, o destino pode ter apenas dois valores (geralmente codificados como 0 e 1), enquanto na regressão o destino pode ter vários valores; mesmo que a análise de regressão tenha sido empregada em um grande número de estudos de aprendizado de máquina, nenhum consenso foi alcançado em uma única métrica padrão unificada para avaliar os resultados da própria regressão (Chicco et al., 2021). Muitos estudos empregam o erro quadrático médio (MSE), ou sua variante enraizada (RMSE), ou o erro absoluto médio (MAE), ou o coeficiente de determinação ( $R^2$ ). Apesar de úteis, essas taxas compartilham uma desvantagem comum: como seus valores podem variar entre zero e +infinito, um único valor delas não diz muito sobre o desempenho da regressão com relação à distribuição dos elementos de base (Chicco et al., 2021). O objetivo da Regressão Linear é encontrar uma linha que minimize o erro de previsão de todos os pontos de dados. A etapa essencial em qualquer modelo de aprendizado de máquina é avaliar a precisão do modelo. As métricas de erro quadrático médio, erro absoluto médio, raiz do erro quadrático médio e R-quadrado ou coeficiente de

determinação são usadas para avaliar o desempenho do modelo na análise de regressão (Chicco et al., 2021).

O MSE é uma medida comum de avaliação da qualidade de um modelo de regressão. Ele é calculado pela média dos quadrados dos erros, que é a diferença entre os valores previstos pelo modelo e os valores observados. O MSE é frequentemente utilizado para avaliar a precisão das previsões de modelos de regressão. Quanto menor for o valor do MSE, melhor será a precisão do modelo (Kim et. al, 2022). Ele mede a variância dos resíduos:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2 \quad (2.13)$$

onde  $y_i$  é a  $i$ -ésima observação,  $\hat{y}$  é o valor estimado (previsão) de  $y_i$  e  $N$  é o número total de observações.

O RMSE é uma medida de avaliação da qualidade de um modelo de regressão que é semelhante ao MSE, mas que leva em conta a raiz quadrada da média dos quadrados dos erros. O RMSE é uma medida de erro absoluto, o que significa que ele expressa o desvio médio em unidades de medida da variável dependente (Kim et. al, 2022). Ele mede o desvio padrão dos resíduos:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2} \quad (2.14)$$

O MAE é uma medida de avaliação da qualidade de um modelo de regressão que mede a diferença média absoluta entre os valores previstos pelo modelo e os valores observados. Ao contrário do MSE e do RMSE, o MAE é uma medida de erro médio absoluto, o que significa que ele expressa o desvio médio em unidades de medida da variável dependente (Pakzad et al., 2023). Ele mede a média dos resíduos no conjunto de dados:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}| \quad (2.15)$$

O  $R^2$  é uma medida de avaliação da qualidade de um modelo de regressão que indica o quanto da variação na variável dependente pode ser explicada pelas variáveis independentes. O  $R^2$  é calculado a partir da soma dos quadrados da diferença entre os valores previstos e os valores observados, dividida pela soma dos quadrados da diferença entre os valores observados e a média dos valores observados. O coeficiente de determinação pode assumir valores no intervalo  $(0, 1]$  de acordo com a relação mútua entre a verdade fundamental e o modelo de previsão (Chicco et al., 2021). Ele pode, ainda, ser interpretado como a proporção da variância na variável dependente que é previsível a

partir das variáveis independentes, sendo o pior valor = 0, e o melhor valor = +1 (Wright, 1921). O coeficiente de determinação é uma pontuação sem escala, ou seja, independentemente de os valores serem pequenos ou grandes, o valor de  $R^2$  será menor que um:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (2.16)$$

onde  $\bar{y}$  é a média dos valores observados.

O RMSE é mais usado do que o MSE para avaliar o desempenho do modelo de regressão com outros modelos aleatórios, pois possui as mesmas unidades da variável dependente (eixo Y). MSE é uma função diferenciável que facilita a realização de operações matemáticas em comparação com uma função não diferenciável como MAE. Portanto, em muitos modelos, o RMSE é usado como uma métrica padrão para calcular a função de perda, apesar de ser mais difícil de interpretar do que o MAE (Kim et. al, 2022). O menor valor de MAE, MSE e RMSE implica maior precisão de um modelo de regressão. No entanto, um valor mais alto de R quadrado é considerado desejável. O  $R^2$  é usado para explicar como as variáveis independentes no modelo de regressão linear explicam a variabilidade na variável dependente (Chicco et al., 2021).

Este estudo aplicou o treinamento repetido e o teste em dez procedimentos com um tamanho de conjunto de treinamento baseado em 75%. Os resultados da avaliação dos testes mostraram que a precisão ideal para prever o desempenho financeiro do sistema bancário no Brasil é aplicar a técnica KNN com um MSE = 0,000, RMSE = 0,001, MAE = 0,000 e  $R^2 = 0,945$ . Como comparação, a técnica de RF é a que fornece valores mais próximos da precisão ideal do KNN com um MSE = 0,000, RMSE = 0,002, MAE = 0,001 e  $R^2 = 0,854$  (Tabela 2.2).

**Tabela 2.2: Resultados da avaliação de testes e treinamento**

Modelo	MSE	RMSE	MAE	R2
kNN	0.000*	0.001*	0.000*	0.945*
Floresta Aleatória	0.000	0.002	0.001	0.854
Neural Network	0.000	0.003	0.002	0.654
Regressão Lasso	0.000	0.004	0.003	0.471

**Nota: \*Precisão ideal para a previsão**

**Fonte: Elaborado pelo autor**

## 2.4.2 Previsões

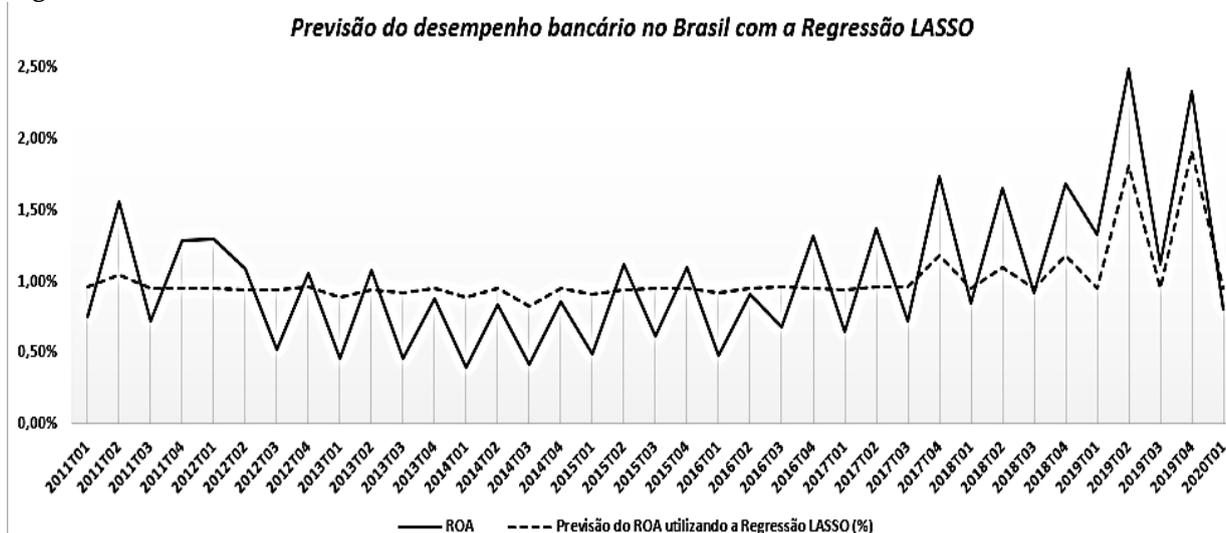
Este estudo realizou quatro técnicas de mineração de dados para prever o desempenho financeiro do sistema bancário brasileiro, conforme mostrado nas Figuras 2.4, 2.5, 2.6 e 2.7 e na Tabela A2.1 (Apêndice).

Além disso, de acordo com a Lei Federal Brasileira nº 4.595, de 1964, as instituições financeiras devem apurar resultados em 30 de junho e 31 de dezembro de cada ano,

obrigatoriamente, com observância das regras contábeis estabelecidas pelo Conselho Monetário Nacional. Desta forma, os gráficos mostram, em março e setembro, os dados contábeis referentes a receitas e despesas correspondem aos saldos acumulados de janeiro a março e de julho a setembro, respectivamente. Os demonstrativos de resultado de junho e dezembro registram os valores acumulados de janeiro a junho e de julho a dezembro, respectivamente (Banco Central do Brasil, 2023c).

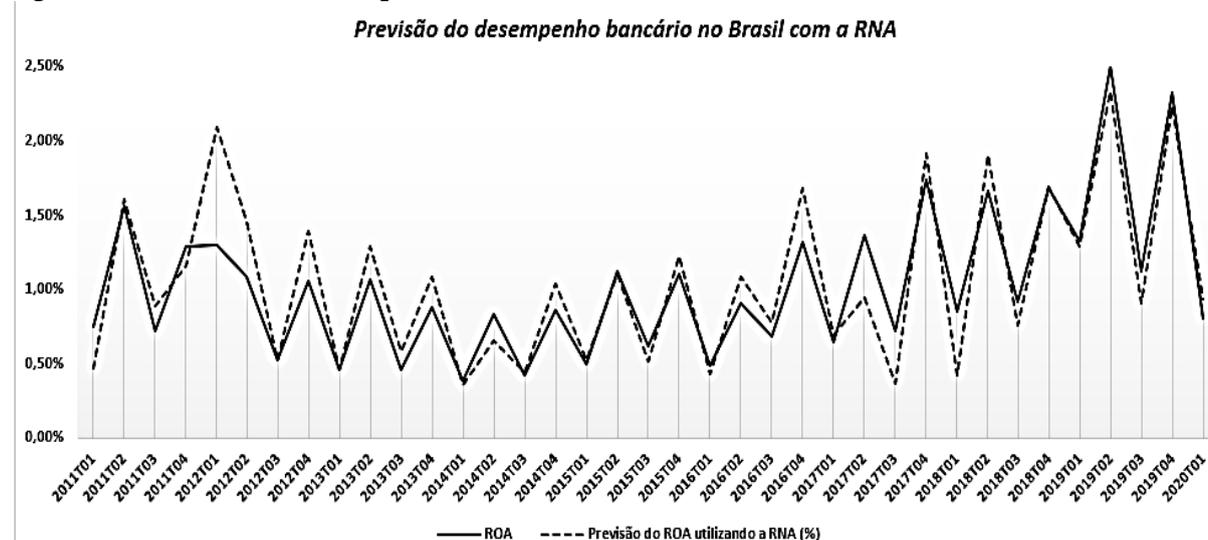
Tendo isso em conta, os resultados experimentais estabeleceram que tanto a técnica KNN quanto a RF são as técnicas mais apropriadas para prever o desempenho financeiro do sistema bancário no Brasil. No entanto, como a técnica KNN fornece os valores de precisão ideais para previsão, é melhor executar o KNN para prever o desempenho financeiro do sistema bancário brasileiro.

**Figura 2.4 - Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica de regressão LASSO.**



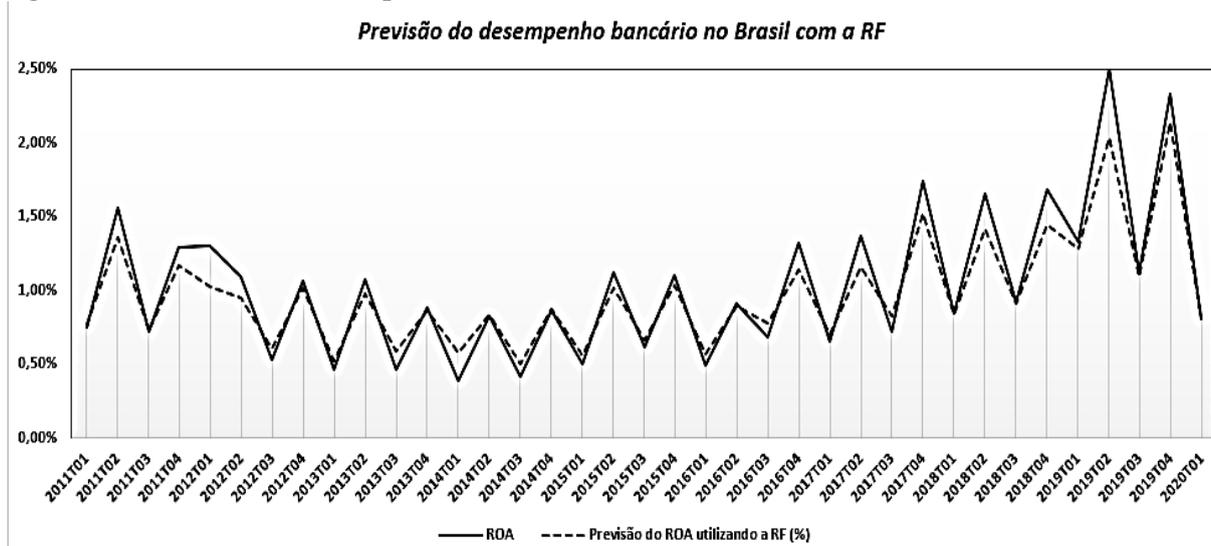
Fonte: Elaborado pelo autor

**Figura 2.5 - Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica RNA.**



Fonte: Elaborado pelo autor

Figura 2.6 - Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica RF.



Fonte: Elaborado pelo autor

Figura 2.7 - Previsões do desempenho financeiro do sistema bancário no Brasil usando a técnica KNN.



Fonte: Elaborado pelo autor

## 2.5 Conclusão

Este estudo enriqueceu a literatura ao fornecer um novo modelo para prever o desempenho financeiro dos bancos brasileiros. Foram aplicadas quatro técnicas de mineração de dados (regressão LASSO, RF, RNAs e KNN) e escolheu-se a melhor técnica entre elas usando testes de treinamento e de precisão, após processos de aprendizado.

Os resultados experimentais demonstraram que tanto a técnica KNN quanto a RF são as melhores técnicas para prever o desempenho financeiro do sistema bancário no Brasil devido à melhor precisão de previsão com base nos critérios MSE, RMSE, MAE

e  $R^2$ . No entanto, embora a RF forneça uma precisão de previsão mais próxima da KNN, a KNN é mais precisa do que a RF. Portanto, como resposta à pergunta principal da pesquisa, a KNN é a técnica de mineração de dados mais adequada para prever o desempenho financeiro do sistema bancário no Brasil.

Este artigo experimental preenche a lacuna da ausência de técnicas de mineração de dados na previsão do desempenho financeiro do sistema bancário brasileiro. Além disso, em consonância com os estudos de Cai & Zhang (2020), Dogan & Birant (2021), Doğuç (2022), Gao (2022), Liébana-Cabanillas et al. (2013) e Luo et al. (2022), este estudo confirma que as técnicas de mineração de dados são ferramentas eficazes na previsão do desempenho financeiro bancário.

Como o setor bancário no Brasil está testemunhando a modernização da tecnologia por meio de fintech que usa big data, AI e ML, a estrutura dos bancos nacionais se tornou grande e moderna, com mais desafios e concorrência no mercado financeiro (World Economic Forum, 2023). Por esse motivo, este estudo pode fornecer insights sobre a necessidade de adotar ferramentas eficazes de mineração de dados para prever o desempenho financeiro dos bancos no Brasil, no intuito de sobreviver à concorrência no mercado financeiro.

### **2.5.1 Implicações práticas para pesquisa e trabalho**

Este estudo tem algumas implicações para a pesquisa e prática da seguinte forma:

- O objetivo deste estudo foi o de aplicar diversas técnicas de mineração de dados para prever o desempenho financeiro dos bancos brasileiros por meio dos principais determinantes de lucratividade, selecionando a melhor técnica de mineração de dados com base nos critérios de pontuação de precisão ótima para a previsão. Assim, este artigo preencheria a lacuna na literatura ao fornecer novas técnicas eficazes para prever o desempenho financeiro no setor bancário usando abordagens de mineração de dados.

- Devido à falta de uso de técnicas de mineração de dados no setor bancário brasileiro, este artigo oferece técnicas de mineração de dados como ferramentas eficientes para pesquisadores financeiros e tomadores de decisão para prever o desempenho financeiro de bancos.

- Em termos de implicação prática, este artigo fornece uma visão importante para pesquisadores e tomadores de decisão em países do setor bancário sulamericano sobre a necessidade de adotar técnicas de mineração de dados como ferramentas importantes para analisar e extrair conhecimentos valiosos, pois essas técnicas podem ser benéficas na modelagem de negócios e na gestão do sistema bancário ao prever o desempenho financeiro.

### **2.5.2 Limitações e pesquisas futuras**

Para prever o desempenho financeiro do sistema bancário usando mineração de

dados, este artigo restringiu a modelagem do desempenho financeiro à rentabilidade através dos principais determinantes do ROA no Brasil. Pesquisas futuras poderiam considerar o uso de outros modelos financeiros, como CAMELS e Z-Score, para avaliar o desempenho financeiro da indústria bancária sob mineração de dados.

# Considerações finais da pesquisa

Este estudo teve como objetivo investigar duas questões usando dados empíricos, técnicas econométricas e simulação computacional. As questões visavam averiguar se o desempenho financeiro dos bancos no Brasil poderia ser mais efetivo quando se lançasse mão de tecnologias rápidas de análise de grandes dados, além de se fazerem previsões sobre essa performance, por meio de técnicas de aprendizado de máquina. Embora a primeira abordagem seja comum na literatura, a última pode ser mais eficaz, visto trabalhar com dados reais oficiais das instituições financeiras.

A teoria organizacional aplica metáforas a vários quadros teóricos. Nos dois estudos, a revisão da literatura prioriza a estabilidade financeira. O primeiro e o segundo trabalhos associam a estabilidade financeira à implantação de medidas tecnológicas para melhorar a saúde financeira organizacional, valendo-se da metáfora do organismo e da máquina, respectivamente. O primeiro estudo utilizou dados coletados de forma transversal, e o segundo dados periódicos para avaliar indicadores financeiros críticos, refletindo dimensões significativas das organizações bancárias. Esses indicadores incluem capitalização, qualidade dos ativos, qualidade da gestão, lucratividade e liquidez.

No primeiro capítulo, foi abordada a teoria baseada em recursos, a qual estabelece que uma integração e uma implantação mais efetivas de recursos conduzem a melhores resultados organizacionais. Inicialmente, o estudo buscou integrar e implantar as capacidades específicas de análise de big data, que impulsionam tal análise, e posteriormente avaliou seu impacto no desempenho dos bancos brasileiros através de testes estatísticos para verificação da confiabilidade e validade dos modelos propostos. Adicionalmente, foram realizados outros testes, como análise de correlação, análise de regressão e análise de intermediação, para determinar se o relacionamento reivindicado era ou não rejeitado. O estudo considerou a hipótese de que existe correlação positiva entre análise de grandes dados e desempenho financeiro bancário, e os resultados confirmaram essa suposição, indicando que empresas com altos níveis de recursos de análise de big data também possuem altos níveis de desempenho financeiro.

O segundo capítulo contribuiu para a literatura ao apresentar um novo modelo de previsão de desempenho financeiro dos bancos brasileiros. Foram aplicadas quatro técnicas de mineração de dados (regressão LASSO, floresta aleatória (RF), redes neurais artificiais (RNAs) e k-vizinhos mais próximos (KNN)), e a melhor técnica foi selecionada por meio de testes de treinamento e de precisão, após os processos de aprendizagem. Os resultados experimentais indicaram que tanto a técnica KNN quanto a RF apresentaram as melhores precisões de previsão com base nos critérios do erro quadrático médio, da raiz do erro quadrático médio, do erro absoluto médio e do coeficiente de correlação, tornando-as as técnicas mais adequadas para prever o desempenho financeiro do sistema bancário no Brasil. No entanto, apesar de a RF

fornecer uma precisão de previsão próxima ao KNN, este teve um desempenho superior. Consequentemente, em resposta à questão principal de pesquisa do capítulo, concluiu-se que o KNN é a técnica de mineração de dados mais indicada para prever o desempenho financeiro do sistema bancário brasileiro.

A partir dos estudos conduzidos na pesquisa, conclui-se que o uso de tecnologias de análise de grandes dados associado a técnicas de mineração de dados e aprendizado de máquina podem contribuir para a otimização do desempenho financeiro presente e futuro. Além da tradicional visão mecanicista, a incorporação de diferentes abordagens teóricas das organizações pode ser benéfica, como as que relacionam as organizações a organismos.

As principais limitações encontradas ao longo do desenvolvimento dos dois capítulos foram, num primeiro momento: uso restrito à indústria bancária, a utilização de apenas algumas dimensões para investigar a competência de análise de big data ao desempenho financeiro corporativo, incluir uma combinação de medidas rígidas e suaves para avaliar melhor o desempenho das firmas. Além dessas limitações encontradas no primeiro capítulo da pesquisa, o segundo estudo sugere outras limitações, como: para prever o desempenho financeiro do sistema bancário brasileiro, a pesquisa restringiu-se à lucratividade e aos principais determinantes do ROA no país; embora essa abordagem seja útil, futuras pesquisas podem explorar outros modelos financeiros, como CAMELS e Z-Score, para avaliar o desempenho do sistema bancário moderno, por meio de técnicas de mineração de dados.

Para expandir a aplicabilidade deste modelo, pesquisas futuras podem testá-lo em diferentes setores, tais como manufatura, saúde e construção. Além disso, é recomendável explorar outras dimensões que possam ter impacto significativo, bem como utilizar variáveis adicionais. Ademais, seria interessante investigar a influência desta tecnologia em áreas específicas, como o gerenciamento da cadeia de suprimentos.

# Referências da Introdução à pesquisa

- Batool, I., & Khan, T. A. (2022). Software fault prediction using data mining, machine learning and deep learning techniques: A systematic literature review. *Computers and Electrical Engineering*, 100, 107886. <https://doi.org/https://doi.org/10.1016/j.compeleceng.2022.107886>
- Broby, D. (2022). The use of predictive analytics in finance. *The Journal of Finance and Data Science*, 8, 145–161. <https://doi.org/https://doi.org/10.1016/j.jfds.2022.05.003>
- Doğuş, Ö. (2022). Data Mining Applications in Banking Sector While Preserving Customer Privacy. *Emerging Science Journal*, 6(6), 1444–1454. <https://doi.org/10.28991/ESJ-2022-06-06-014>
- Jahan, S. A., & Candidate, M. B. A. (2022). Factors Affecting The Adoption Of Financial Technology Among The Banking Customers In Emerging Economies. *Review of Financial Studies*, 39.
- Morgan, G., & Smircich, L. (1980). The Case for Qualitative Research. *The Academy of Management Review*, 5 (4), 491-500.
- Morgan, G. (2005). Paradigmas, Metáforas e Resolução de Quebra-Cabeças na Teoria das Organizações. Tradução do original (Administrative Science Quarterly, 1980). *Revista de Administração de Empresas*, 45(1), 58-71.

# Referências do Capítulo 1

- Adomako, S., Amankwah-Amoah, J., Donbesuur, F., Ahsan, M., Danso, A., & Uddin, M. (2022). Strategic agility of SMEs in emerging economies: Antecedents, consequences and boundary conditions. *International Business Review*, 31(6), 102032. <https://doi.org/https://doi.org/10.1016/j.ibusrev.2022.102032>
- Ahmed, R., Shaheen, S., & Philbin, S. P. (2022). The role of big data analytics and decision-making in achieving project success. *Journal of Engineering and Technology Management*, 65, 101697. <https://doi.org/https://doi.org/10.1016/j.jengtecman.2022.101697>
- Aimé, I., Berger-Remy, F., & Laporte, M.-E. (2022). The brand, the persona and the algorithm: How datafication is reconfiguring marketing work☆. *Journal of Business Research*, 145, 814–827. <https://doi.org/https://doi.org/10.1016/j.jbusres.2022.03.047>
- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131.
- Alexander, D. T., & Lyytinen, K. (2019). Organizing around big data: Organizational analytic capabilities for improved performance. *40th International Conference on Information Systems, ICIS 2019*. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114902363&partnerID=40&md5=20d766285d278073794529dc14bc3988>
- Alhababsah, S., & Alhaj-Ismail, A. (2021). Does shared tenure between audit committee chair and engagement partner affect audit outcomes? Evidence from the UK. *The British Accounting Review*, 101067. <https://doi.org/https://doi.org/10.1016/j.bar.2021.101067>
- Al-Okaily, J., & BenYoussef, N. (2020). Audit committee effectiveness and non-audit service fees: Evidence from UK family firms. *Journal of International Accounting, Auditing and Taxation*, 41, 100356. <https://doi.org/10.1016/j.intaccaudtax.2020.100356>
- Alzeban, A. (2020). The relationship between the audit committee, internal audit and firm performance. *Journal of Applied Accounting Research*, 21(3), 437–454. <https://doi.org/10.1108/JAAR-03-2019-0054>
- Ambrosini, V., Bowman, C., & Collier, N. (2009). Dynamic capabilities: An exploration of how firms renew their resource base. *British journal of management*, 20, S9-S24.
- Anjaria, K. (2022). Knowledge derivation from Likert scale using Z-numbers. *Information Sciences*, 590, 234–252. <https://doi.org/https://doi.org/10.1016/j.ins.2022.01.024>
- Ashofteh, A., & Bravo, J. M. (2021). Data science training for official statistics: A new scientific paradigm of information and knowledge development in national statistical systems. *Statistical Journal of the IAOS*, 37(3), 771–789. <https://doi.org/10.3233/SJI-210841>
- Awan, U., Shamim, S., Khan, Z., Zia, N. U., Shariq, S. M., & Khan, M. N. (2021). Big data analytics capability and decision-making: The role of data-driven insight on circular economy performance. *Technological Forecasting and Social Change*, 168(March). <https://doi.org/10.1016/j.techfore.2021.120766>
- Bag, S., Pretorius, J. H. C., Gupta, S., & Dwivedi, Y. K. (2021). Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. *Technological Forecasting and Social Change*, 163, 120420. <https://doi.org/https://doi.org/10.1016/j.techfore.2020.120420>

- Bag, S., Rahman, M. S., Srivastava, G., Chan, H. L., & Bryde, D. J. (2022). The role of big data and predictive analytics in developing a resilient supply chain network in the South African mining industry against extreme weather events. *International Journal of Production Economics*, 251. <https://doi.org/10.1016/j.ijpe.2022.108541>
- Banmairuroy, W., Kritjaroen, T., & Homsombat, W. (2022). The effect of knowledge-oriented leadership and human resource development on sustainable competitive advantage through organizational innovation's component factors: Evidence from Thailand 's new S-curve industries. *Asia Pacific Management Review*, 27(3), 200–209. <https://doi.org/10.1016/j.apmr.2021.09.001>
- Banna, H., Kabir Hassan, M., & Rashid, M. (2021). Fintech-based financial inclusion and bank risk-taking: Evidence from OIC countries. *Journal of International Financial Markets, Institutions and Money*, 75(October), 101447. <https://doi.org/10.1016/j.intfin.2021.101447>
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120. <https://doi.org/10.1177/014920639101700108>
- Barrera, C., Seguí, L., & Castell, M. L. (2023). Bridging the academia-industry gap in the food sector through collaborative courses and internships. *Education for Chemical Engineers*, 42(November 2022), 33–43. <https://doi.org/10.1016/j.ece.2022.11.003>
- Behl, A., Gaur, J., Pereira, V., Yadav, R., & Laker, B. (2022). Role of big data analytics capabilities to improve sustainable competitive advantage of MSME service firms during COVID-19 – A multi-theoretical approach. *Journal of Business Research*, 148, 378–389. <https://doi.org/https://doi.org/10.1016/j.jbusres.2022.05.009>
- Bell, D., Lycett, M., Marshan, A., & Monaghan, A. (2021). Exploring future challenges for big data in the humanitarian domain. *Journal of Business Research*, 131, 453–468. <https://doi.org/10.1016/j.jbusres.2020.09.035>
- Biedka, W., Herbst, M., Rok, J., & Wójcik, P. (2022). The local-level impact of human capital investment within the EU cohesion policy in Poland. *Papers in Regional Science*, 101(2), 303–325. <https://doi.org/10.1111/pirs.12648>
- Bieńkowska, A., & Tworek, K. (2022). IT dynamic capabilities as a factor influencing Controlling Effectiveness. *Procedia Computer Science*, 207, 24–33. <https://doi.org/10.1016/j.procs.2022.09.034>
- Bonnard, R., Arantes, M. D. S., Lorbiecki, R., Vieira, K. M. M., & Nunes, M. C. (2021). Big data/analytics platform for Industry 4.0 implementation in advanced manufacturing context. *International Journal of Advanced Manufacturing Technology*, 117(5–6), 1959–1973. <https://doi.org/10.1007/s00170-021-07834-5>
- Boo, E., & Koh, H. (2004). Are Client-Perceived Audit Firm Reputation and Audit Team Attributes Associated with Big Audit Fees? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.542282>
- Börsch-Supan, A., Hunkler, C., & Weiss, M. (2021). Big data at work: Age and labor productivity in the service sector. *Journal of the Economics of Ageing*, 19(March). <https://doi.org/10.1016/j.jeoa.2021.100319>
- Boubaker, S., Liu, Z., & Zhai, L. (2021). Big data, news diversity and financial market crash. *Technological Forecasting & Social Change*, 168(59).
- Brody, R. G., Golen, S. P., & Reckers, P. M. J. (1998). An empirical investigation of the interface between internal and external auditors. *Accounting and Business Research*, 28(3), 160–171. <https://doi.org/10.1080/00014788.1998.9728907>
- Buhalis, D., & Volchek, K. (2021). Bridging marketing theory and big data analytics: The

- taxonomy of marketing attribution. *International Journal of Information Management*, 56(October 2019), 102253. <https://doi.org/10.1016/j.ijinfomgt.2020.102253>
- Cao, G., & Duan, Y. (2015). The Affordances of Business Analytics for Strategic Decision-Making and Their Impact on Organisational Performance. *PACIS*, 255.
- Cappa, F., Oriani, R., Peruffo, E., & McCarthy, I. (2021). Big Data for Creating and Capturing Value in the Digitalized Environment: Unpacking the Effects of Volume, Variety, and Veracity on Firm Performance\*. *Journal of Product Innovation Management*, 38(1), 49–67. <https://doi.org/https://doi.org/10.1111/jpim.12545>
- Chae, B., & Olson, D. (2022). Technologies and applications of Industry 4.0: insights from network analytics. *International Journal of Production Research*, 60(12), 3682–3704. <https://doi.org/10.1080/00207543.2021.1931524>
- Chet Miller, C., Washburn, N. T., & Glick, W. H. (2013). The myth of firm performance. *Organization Science*, 24(3), 948–964. <https://doi.org/10.1287/orsc.1120.0762>
- Chin, W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.-J., & Cham, T. H. (2020). Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Industrial Management and Data Systems*, 120(12), 2161–2209. <https://doi.org/10.1108/IMDS-10-2019-0529>
- Choi, H. Y., & Park, J. (2022). Do data-driven CSR initiatives improve CSR performance? The importance of big data analytics capability. *Technological Forecasting and Social Change*, 182(June), 121802. <https://doi.org/10.1016/j.techfore.2022.121802>
- Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research*, 123, 1–13. <https://doi.org/10.1016/j.jbusres.2020.09.023>
- Cohen, J. (1977). CHAPTER 2 - The t Test for Means. Em J. Cohen (Org.), *Statistical Power Analysis for the Behavioral Sciences* (p. 19–74). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-179060-8.50007-4>
- Collimore, A., Rosado-Muñoz, F. J., & Ojeda-Castro, A. (2017). Big data analytics, competitive advantage and firm performance. *International Journal of Information Research and Review*, 4(2), 3599–3603.
- Contreras Pinochet, L. H., Amorim, G. C. B., Lucas Júnior, D., & Souza, C. A. (2021). Consequential factors of Big Data's Analytics Capability: how firms use data in the competitive scenario. *Journal of Enterprise Information Management*, 34(5), 1406–1428. <https://doi.org/10.1108/JEIM-11-2020-0445>
- Corvello, V., Belas, J., Giglio, C., Iazzolino, G., & Troise, C. (2023). The impact of business owners' individual characteristics on patenting in the context of digital innovation. *Journal of Business Research*, 155(PA), 113397. <https://doi.org/10.1016/j.jbusres.2022.113397>
- Cui, T., Hua, J., & Hoo, C. (2022). Information technology in open innovation: A resource orchestration perspective. *Information & Management*, 59(8), 103699. <https://doi.org/10.1016/j.im.2022.103699>
- Cui, Y., Ma, Z., Wang, L., Yang, A., Liu, Q., Kong, S., & Wang, H. (2023). A survey on big data-enabled innovative online education systems during the COVID-19 pandemic. *Journal of Innovation & Knowledge*, 8(1), 262–268. <https://doi.org/10.1016/j.jik.2022.100295>
- da Rosa München, D. (2022). The effect of financial distress on capital structure: The case of Brazilian banks. *Quarterly Review of Economics and Finance*, 86, 296–304. <https://doi.org/10.1016/j.qref.2022.07.010>

- Damelang, A., & Ruf, K. (2023). Once outside, always outside? The link between overeducation persistence and training systems throughout the employment career. *Social Science Research*, 109(August 2022), 102774. <https://doi.org/10.1016/j.ssresearch.2022.102774>
- Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*, 84(1), 98-107+134.
- De Santis, F., & D'Onza, G. (2020). Big data and data analytics in auditing: in search of legitimacy. *Meditari Accountancy Research*, 29(5), 1088–1112. <https://doi.org/10.1108/MEDAR-03-2020-0838>
- Deb, D., & Fuad, M. (2021). Integrating big data and cloud computing topics into the computing curricula: A modular approach. *Journal of Parallel and Distributed Computing*, 157, 303–315. <https://doi.org/10.1016/j.jpdc.2021.07.012>
- Dieleman, N. A., Berkhout, J., & Heidergott, B. (2023). A neural network approach to performance analysis of tandem lines: The value of analytical knowledge. *Computers & Operations Research*, 152, 106124. <https://doi.org/https://doi.org/10.1016/j.cor.2022.106124>
- Du, X., Yang, J., Shelton, B. E., Hung, J.-L., & Zhang, M. (2021). A systematic meta-Review and analysis of learning analytics research. *Behaviour and Information Technology*, 40(1), 49–62. <https://doi.org/10.1080/0144929X.2019.1669712>
- Dubey, R., Gunasekaran, A., Childe, S. J., Bryde, D. J., Giannakis, M., Foropon, C., Roubaud, D., & Hazen, B. T. (2020). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. *International Journal of Production Economics*, 226. <https://doi.org/10.1016/j.ijpe.2019.107599>
- Dutta, G., Kumar, R., Sindhwani, R., & Singh, R. K. (2020). Digital transformation priorities of India's discrete manufacturing SMEs – a conceptual study in perspective of Industry 4.0. *Competitiveness Review*, 289–314. <https://doi.org/10.1108/CR-03-2019-0031>
- Elia, G., Raguseo, E., Solazzo, G., & Pigni, F. (2022). Strategic business value from big data analytics: An empirical analysis of the mediating effects of value creation mechanisms. *Information and Management*, 59(8), 103701. <https://doi.org/10.1016/j.im.2022.103701>
- Esposito, P., & Scicchitano, S. (2022). Educational mismatch and labour market transitions in Italy : Is there an unemployment trap? *Structural Change and Economic Dynamics*, 61, 138–155. <https://doi.org/10.1016/j.strueco.2022.02.011>
- F. Hair Jr, J., Sarstedt, M., Hopkins, L., & G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM). *European Business Review*, 26(2), 106–121. <https://doi.org/10.1108/EBR-10-2013-0128>
- Fakhar Manesh, M., Pellegrini, M. M., Marzi, G., & Dabic, M. (2021). Knowledge Management in the Fourth Industrial Revolution: Mapping the Literature and Scoping Future Avenues. *IEEE Transactions on Engineering Management*, 68(1), 289–300. <https://doi.org/10.1109/TEM.2019.2963489>
- Fakhfakh, I., & Jarboui, A. (2020). Audit certification, earnings management and risk governance: a moderated-mediation analysis. *Journal of Financial Reporting and Accounting*, 18(2), 277–299. <https://doi.org/10.1108/JFRA-01-2019-0015>
- Flynn, B., Pagell, M., & Fugate, B. (2020). From the Editors: Introduction to the Emerging Discourse Incubator on the Topic of Emerging Approaches for Developing Supply Chain Management Theory. *Journal of Supply Chain Management*, 56(2), 3–6. <https://doi.org/10.1111/jscm.12227>
- Fomiatti, R., Moore, D., Fraser, S., & Farrugia, A. (2021). Holding 'new recovery' together: Organising relations and forms of coordination in professional sociomaterial practices of

- addiction recovery. *International Journal of Drug Policy*, 97, 103357. <https://doi.org/https://doi.org/10.1016/j.drugpo.2021.103357>
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Garcia, A. S., & Meurer, R. (2022). Effects of a development bank on the profitability of commercial banks: Evidence for Brazil. *Quarterly Review of Economics and Finance*, 85, 246–259. <https://doi.org/10.1016/j.qref.2022.03.008>
- Garrido-Moreno, A., García-Morales, V., King, S., & Lockett, N. (2020). Social Media use and value creation in the digital landscape: a dynamic-capabilities perspective. *Journal of Service Management*, 31(3), 313–343. <https://doi.org/10.1108/JOSM-09-2018-0286>
- Ghaleb, B. A. A., Kamardin, H., & Al-Qadasi, A. A. (2020). Internal audit function and real earnings management practices in an emerging market. *Meditari Accountancy Research*, 28(6), 1209–1230. <https://doi.org/10.1108/MEDAR-02-2020-0713>
- Ghasemaghaei, M., & Calic, G. (2019). Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research*, 104, 69–84. <https://doi.org/https://doi.org/10.1016/j.jbusres.2019.07.006>
- Ghasemaghaei, M., Hassanein, K., & Turel, O. (2015). Impacts of big data analytics on organizations: a resource fit perspective. *Emergent Research Forum Papers*.
- Ghobakhloo, M., & Fathi, M. (2020). Corporate survival in Industry 4.0 era: the enabling role of lean-digitized manufacturing. *Journal of Manufacturing Technology Management*, 31(1), 1–30. <https://doi.org/10.1108/JMTM-11-2018-0417>
- Goldschmidt, G., & Matthews, B. (2022). Formulating design research questions: A framework. *Design Studies*, 78, 101062. <https://doi.org/https://doi.org/10.1016/j.destud.2021.101062>
- Grange, M., Ackers, B., & Odendaal, E. (2021). The impact of the characteristics of audit committee members on audit committee effectiveness. *Global Business and Economics Review*, 24, 409. <https://doi.org/10.1504/GBER.2021.115812>
- Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review*, 33(3), 114–135. <https://doi.org/10.2307/41166664>
- Grant, R. M. (1996). Prospering in Dynamically-competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science*, 7(4), 375–387. <https://doi.org/10.1287/orsc.7.4.375>
- Gravett, K. (2022). Feedback literacies as sociomaterial practice. *Critical Studies in Education*, 63(2), 261–274. <https://doi.org/10.1080/17508487.2020.1747099>
- Gu, R., Qi, Y., Wu, T., Wang, Z., Xu, X., Yuan, C., & Huang, Y. (2021). SparkDQ: Efficient generic big data quality management on distributed data-parallel computation. *Journal of Parallel and Distributed Computing*, 156, 132–147. <https://doi.org/10.1016/j.jpdc.2021.05.012>
- Gupta, A. K., & Goyal, H. (2021). Framework for implementing big data analytics in Indian manufacturing: ISM-MICMAC and Fuzzy-AHP approach. *Information Technology and Management*, 22(3), 207–229. <https://doi.org/10.1007/s10799-021-00333-9>
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049–1064.
- Gupta, R., Srivastava, P., Sharma, S., & Alrasheedi, M. (2021). Leveraging Big Data to Accelerate Supply Chain Management in Covid-19. *Em Studies in Computational Intelligence* (Vol. 974, p. 1–19). Springer Science and Business Media Deutschland GmbH.

[https://doi.org/10.1007/978-3-030-73057-4\\_1](https://doi.org/10.1007/978-3-030-73057-4_1)

- Hair Joe F., Jr., Sarstedt, M., Matthews, L. M., & Ringle, C. M. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: part I – method. *European Business Review*, 28(1), 63–76. <https://doi.org/10.1108/EBR-09-2015-0094>
- Hair, J., & Alamer, A. (2022). Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics*, 1(3), 100027. <https://doi.org/https://doi.org/10.1016/j.rmal.2022.100027>
- Hair, J., Hollingsworth, C., Randolph, A., & Chong, A. (2017). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management & Data Systems*, 117, 442–458. <https://doi.org/10.1108/IMDS-04-2016-0130>
- Hair, J., Sarstedt, M., Matthews, L., & Ringle, C. (2016). Identifying and treating unobserved heterogeneity with FIMIX-PLS: part I – method. *European Business Review*, 28. <https://doi.org/10.1108/EBR-09-2015-0094>
- Hajiheydari, N., Delgosha, M. S., Wang, Y., & Olya, H. (2021). Exploring the paths to big data analytics implementation success in banking and financial service: an integrated approach. *Industrial Management and Data Systems*, 121(12), 2498–2529. <https://doi.org/10.1108/IMDS-04-2021-0209>
- Hambrick, D. C. (1987). The top management team: Key to strategic success. *California management review*, 30(1), 88–108.
- Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2021). A Systematic Review of the Literature on Digital Transformation: Insights and Implications for Strategy and Organizational Change. *Journal of Management Studies*, 58(5), 1159–1197. <https://doi.org/10.1111/joms.12639>
- Hasan, M. M., Popp, J., & Oláh, J. (2020). Current landscape and influence of big data on finance. *Journal of Big Data*, 7(1). <https://doi.org/10.1186/s40537-020-00291-z>
- Helfat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic management journal*, 24(10), 997–1010.
- Henseler, J., Ringle, C., & Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, 43, 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Heubeck, T. (2023). Managerial capabilities as facilitators of digital transformation? Dynamic managerial capabilities as antecedents to digital business model transformation and firm performance. *Digital Business*, 3(1), 100053. <https://doi.org/10.1016/j.digbus.2023.100053>
- Holland, C. P., Thornton, S. C., & Naudé, P. (2020). B2B analytics in the airline market: Harnessing the power of consumer big data. *Industrial Marketing Management*, 86, 52–64. <https://doi.org/10.1016/j.indmarman.2019.11.002>
- Huang, F., No, W. G., Vasarhelyi, M. A., & Yan, Z. (2022). Audit data analytics, machine learning, and full population testing. *The Journal of Finance and Data Science*, 8, 138–144. <https://doi.org/https://doi.org/10.1016/j.jfds.2022.05.002>
- Huang, L., Wu, C., & Wang, B. (2019). Challenges, opportunities and paradigm of applying big data to production safety management: From a theoretical perspective. *Journal of Cleaner Production*, 231, 592–599. <https://doi.org/10.1016/j.jclepro.2019.05.245>
- Hult, G. T. M., Hair, J. F., Proksch, D., Sarstedt, M., Pinkwart, A., & Ringle, C. M. (2018). Addressing Endogeneity in International Marketing Applications of Partial Least Squares Structural Equation Modeling. *Journal of International Marketing*, 26(3), 1–21.

<https://doi.org/10.1509/jim.17.0151>

- Ibrahim, A. E. A., Elamer, A. A., & Ezat, A. N. (2021). The convergence of big data and accounting: innovative research opportunities. *Technological Forecasting and Social Change*, 173, 121171. <https://doi.org/https://doi.org/10.1016/j.techfore.2021.121171>
- Ivanov, D., Tang, C. S., Dolgui, A., Battini, D., & Das, A. (2021). Researchers' perspectives on Industry 4.0: multi-disciplinary analysis and opportunities for operations management. *International Journal of Production Research*, 59(7), 2055–2078. <https://doi.org/10.1080/00207543.2020.1798035>
- Jain, A., & Srinivasan, V. (2022). What happened to the work I was doing? Sociomateriality and cognitive tensions in technology work. *Organizational Dynamics*, 51(4), 100901. <https://doi.org/https://doi.org/10.1016/j.orgdyn.2022.100901>
- Jain, S., & Singh, P. (2020). Intrinsic human capacities of high and low performers of public sector manufacturing industry in India. *Journal of Critical Reviews*, 7(3), 623–630. <https://doi.org/10.31838/JCR.07.03.110>
- Jiang, J., Liao, L., Lu, X., Wang, Z., & Xiang, H. (2021). Deciphering big data in consumer credit evaluation. *Journal of Empirical Finance*, 62, 28–45. <https://doi.org/https://doi.org/10.1016/j.jempfin.2021.01.009>
- Jordan, C. E., Clark, C. S. J., & Thomas, P. B. (2017). Audit quality differentials for constraining cosmetic earnings management in the pre-sox era: An analysis of audit firm size and brand. *Global Journal of Accounting and Finance Volume*, 1(1).
- Joshi, P. L., & Marthandan, G. (2020). Continuous internal auditing: Can big data analytics help? *International Journal of Accounting, Auditing and Performance Evaluation*, 16(1), 25–42. <https://doi.org/10.1504/IJAAP.2020.106766>
- Khalajzadeh, H., Simmon, A. J., Abdelrazek, M., Grundy, J., Hosking, J., & He, Q. (2020). Visual languages for supporting big data analytics development. Em A. R., K. H., M. L., & M. L. (Orgs.), *ENASE 2020 - Proceedings of the 15th International Conference on Evaluation of Novel Approaches to Software Engineering* (p. 15–26). SciTePress.
- Kim, G., Shin, B., Kim, K., & Lee, H. (2011). IT Capabilities, Process-Oriented Dynamic Capabilities, and Firm Financial Performance. *Journal of the Association for Information Systems*, 12. <https://doi.org/10.17705/1jais.00270>
- Kindermann, B., Salge, T. O., Wentzel, D., Flatten, T. C., & Antons, D. (2022). Dynamic capabilities for orchestrating digital innovation ecosystems: Conceptual integration and research opportunities. *Information and Organization*, 32(3), 100422. <https://doi.org/10.1016/j.infoandorg.2022.100422>
- Kinra, A., Beheshti-Kashi, S., Buch, R., Nielsen, T. A. S., & Pereira, F. (2020). Examining the potential of textual big data analytics for public policy decision-making: A case study with driverless cars in Denmark. *Transport Policy*, 98, 68–78. <https://doi.org/https://doi.org/10.1016/j.tranpol.2020.05.026>
- Krammer, S. M. S. (2022). Technovation Human resource policies and firm innovation : The moderating effects of economic and institutional context. *Technovation*, 110(May 2021), 102366. <https://doi.org/10.1016/j.technovation.2021.102366>
- Kubina, M., Varmus, M., & Kubinova, I. (2015). Use of big data for competitive advantage of company. *Procedia Economics and Finance*, 26, 561–565.
- Lavalle, S., Lesser, E., Shockley, R., Hopkins, M., & Kruschwitz, N. (2011). Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, 52, 21–32.
- Li, D., Gong, Y., Ren, M., & Li, D. (2020). The Research and Design of Trust Business

- Management and Analysis System Based on Big Data Technology. *2020 5th IEEE International Conference on Big data analytics, ICBDA 2020*, 68–72. <https://doi.org/10.1109/ICBDA49040.2020.9101191>
- Li, L., Lin, J., Ouyang, Y., & Luo, X. R. (2022). Evaluating the impact of big data analytics usage on the decision-making quality of organizations. *Technological Forecasting and Social Change*, 175. <https://doi.org/10.1016/j.techfore.2021.121355>
- Li, S., Easterby-Smith, M., & Hong, J. F. L. (2019). Towards an understanding of the nature of dynamic capabilities in high-velocity markets of China. *Journal of Business Research*, 97(June 2017), 212–226. <https://doi.org/10.1016/j.jbusres.2018.08.007>
- Liu, Y., Soroka, A., Han, L., Jian, J., & Tang, M. (2020). Cloud-based big data analytics for customer insight-driven design innovation in SMEs. *International Journal of Information Management*, 51. <https://doi.org/10.1016/j.ijinfomgt.2019.11.002>
- Liu, Y., & Zhao, X. (2021). Successful implementation of value-based selling: a value co-creation and dynamic capabilities perspective. *Journal of Business and Industrial Marketing*, 36(3), 372–389. <https://doi.org/10.1108/JBIM-05-2020-0240>
- Lodato, C., Tonini, D., Damgaard, A., & Fruergaard Astrup, T. (2020). A process-oriented life-cycle assessment (LCA) model for environmental and resource-related technologies (EASETECH). *International Journal of Life Cycle Assessment*, 25(1), 73–88. <https://doi.org/10.1007/s11367-019-01665-z>
- Lokatt, C. (2018). *Auditors' Constitution of Performance: a study on the duality of performance in the auditing profession*. Stockholm Business School, Stockholm University.
- López-Cabarcos, M. Á., Göttling-Oliveira-Monteiro, S., & Vazquez-Rodriguez, P. (2015). Organizational capabilities and profitability: The mediating role of business strategy. *Sage Open*, 5(4), 2158244015616852.
- Lu, J., Feng, G., Shum, S., & Keung, K. (2021). On the value of information sharing in the presence of information errors. *European Journal of Operational Research*, 294(3), 1139–1152. <https://doi.org/10.1016/j.ejor.2021.02.028>
- Luo, J., Xu, J., Aldosari, O., Althubiti, S. A., & Deebani, W. (2022). Design and Implementation of an Efficient Electronic Bank Management Information System Based Data Warehouse and Data Mining Processing. *Information Processing and Management*, 59(6). <https://doi.org/10.1016/j.ipm.2022.103086>
- Ma, S., Ding, W., Liu, Y., Ren, S., & Yang, H. (2022). Digital twin and big data-driven sustainable smart manufacturing based on information management systems for energy-intensive industries. *Applied Energy*, 326, 119986. <https://doi.org/https://doi.org/10.1016/j.apenergy.2022.119986>
- Maheshwari, S., Gautam, P., & Jaggi, C. K. (2021). Role of Big data analytics in supply chain management: current trends and future perspectives. *International Journal of Production Research*, 59(6), 1875–1900. <https://doi.org/10.1080/00207543.2020.1793011>
- Mandzila, E. E. W., Zéghal, D., & others. (2016). Content analysis of board reports on corporate governance, internal controls and risk management: Evidence from France. *Journal of Applied Business Research (JABR)*, 32(3), 637–648.
- Manita, R., Elommal, N., Baudier, P., & Hikkerova, L. (2020). The digital transformation of external audit and its impact on corporate governance. *Technological Forecasting and Social Change*, 150. <https://doi.org/10.1016/j.techfore.2019.119751>
- Marabelli, M., Newell, S., & Handunge, V. (2021). *Journal of Strategic Information Systems*  
The lifecycle of algorithmic decision-making systems : Organizational choices and ethical

- challenges. *Journal of Strategic Information Systems*, 30(3), 101683. <https://doi.org/10.1016/j.jsis.2021.101683>
- Mcbride, K., & Philippou, C. (2022). "Big results require big ambitions": big data, data analytics and accounting in masters courses. *Accounting Research Journal*, 35(1), 71–100. <https://doi.org/10.1108/ARJ-04-2020-0077>
- Meiyou, D., & Ye, Y. (2022). Establishment of big data evaluation model for green and sustainable development of enterprises. *Journal of King Saud University - Science*, 34(5), 102041. <https://doi.org/10.1016/j.jksus.2022.102041>
- Melo, A., Carneiro, M., Mariano, T., Rocha, M., Mariano, A. M., Santos, M. C., Mello, M., & Rocha, M. (2022). ScienceDirect The importance of mobile applications for companies ' brand The importance applications for companies ' brand image : of A mobile study using structural equations image : A study using structural. *Procedia Computer Science*, 214, 1128–1135. <https://doi.org/10.1016/j.procs.2022.11.287>
- Mikalef, P., & Pateli, A. (2017). Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: Findings from PLS-SEM and fsQCA. *Journal of Business Research*, 70, 1-16.
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: the mediating role of dynamic capabilities and moderating effect of the environment. *British journal of management*, 30(2), 272-298.
- Mishra, P., & Yadav, M. (2021). "Environmental capabilities, proactive environmental strategy and competitive advantage: A natural-resource-based view of firms operating in India". *Journal of Cleaner Production*, 291. <https://doi.org/10.1016/j.jclepro.2020.125249>
- Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS quarterly*, 237–256.
- Moodley, A., Ackers, B., & Odendaal, E. (2022). Internal audit's evolving performance role: lessons from the South African public sector. *Journal of Accounting and Organizational Change*, 18(5), 704–726. <https://doi.org/10.1108/JAOC-05-2021-0063>
- Mora Cortez, R., & Hidalgo, P. (2022). Prioritizing B2B marketing capabilities: Crossvergence in advanced and emerging economies. *Industrial Marketing Management*, 105, 422–438. <https://doi.org/https://doi.org/10.1016/j.indmarman.2022.07.002>
- Mora, L., Kummitha, R. K. R., & Esposito, G. (2021). Not everything is as it seems: Digital technology affordance, pandemic control, and the mediating role of sociomaterial arrangements. *Government Information Quarterly*, 38(4), 101599. <https://doi.org/https://doi.org/10.1016/j.giq.2021.101599>
- Morabito, V. (2015). Big data and analytics for competitive advantage. *Big Data and Analytics: Strategic and Organizational Impacts*, 3–22.
- Morimura, F., & Sakagawa, Y. (2023). The intermediating role of big data analytics capability between responsive and proactive market orientations and firm performance in the retail industry. *Journal of Retailing and Consumer Services*, 71, 103193. <https://doi.org/10.1016/j.jretconser.2022.103193>
- Muchenje, G., & Seppänen, M. (2023). Unpacking task-technology fit to explore the business value of big data analytics. *International Journal of Information Management*, 69(December 2022). <https://doi.org/10.1016/j.ijinfomgt.2022.102619>
- Muhammad, R. N., Tasmin, R., & Nor Aziati, A. H. (2020). Sustainable Competitive Advantage of Big data analytics in Higher Education Sector: An Overview. *Journal of Physics: Conference Series*, 1529(4). <https://doi.org/10.1088/1742-6596/1529/4/042100>

- Münch, C., Benz, L. A., & Hartmann, E. (2022). Journal of Purchasing and Supply Management Exploring the circular economy paradigm : A natural resource-based view on supplier selection criteria ☆. *Journal of Purchasing and Supply Management*, 28(4), 100793. <https://doi.org/10.1016/j.pursup.2022.100793>
- Munir, S., Abdul Rasid, S. Z., Aamir, M., Jamil, F., & Ahmed, I. (2022). Big data analytics capabilities and innovation effect of dynamic capabilities, organizational culture and role of management accountants. *Foresight*. <https://doi.org/10.1108/FS-08-2021-0161>
- Murray, N., Lynch, P., & Foley, A. (2022). Strategic nets in tourism destinations: Investigating the learning processes underpinning dynamic management capabilities. *Industrial Marketing Management*, 106(December 2021), 363–375. <https://doi.org/10.1016/j.indmarman.2022.09.004>
- Niu, Y., Ying, L., Yang, J., Bao, M., & Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. *Information Processing and Management*, 58(6), 102725. <https://doi.org/10.1016/j.ipm.2021.102725>
- Oesterreich, T. D., Anton, E., & Teuteberg, F. (2022). What translates big data into business value? A meta-analysis of the impacts of business analytics on firm performance. *Information & Management*, 59(6), 103685. <https://doi.org/https://doi.org/10.1016/j.im.2022.103685>
- Oesterreich, T. D., Anton, E., Teuteberg, F., & Dwivedi, Y. K. (2022). The role of the social and technical factors in creating business value from big data analytics: A meta-analysis. *Journal of Business Research*, 153, 128–149. <https://doi.org/https://doi.org/10.1016/j.jbusres.2022.08.028>
- Pagda, Z., Bayraktar, S., & Jimenez, A. (2021). Exploring culture and leadership after 23 years : A replication of GLOBE project in Turkey. *Journal of International Management*, 27(1), 100822. <https://doi.org/10.1016/j.intman.2021.100822>
- Pan, Q., Luo, W., & Fu, Y. (2022). A csQCA study of value creation in logistics collaboration by big data : A perspective from companies in China. *Technology in Society*, 71(February). <https://doi.org/10.1016/j.techsoc.2022.102114>
- Panja, S., & Mondal, S. K. (2023). Integrating online and offline business of a retailer: A customer utility based inventory model. *Computers & Industrial Engineering*, 175, 108888. <https://doi.org/https://doi.org/10.1016/j.cie.2022.108888>
- Peng, T., & Ke, W. (2023). Urban fire emergency management based on big data intelligent processing system and Internet of Things. *Optik*, 273, 170433. <https://doi.org/https://doi.org/10.1016/j.ijleo.2022.170433>
- Penrose, E. (1959). *The Theory of the Growth of the Firm*. John Wiley. <https://doi.org/10.1093/0198289774.001.0001>
- Phornlaphatrachakorn, K. (2020). Audit Committee Effectiveness, Internal Audit Quality, Financial Reporting Quality, and Organizational Success: An Empirical Investigation of Thai Listed Firms. *International Journal of Business*, 25(4), 344–366.
- Pisano, G. P. (2015). You need an innovation strategy. *Harvard Business Review*, 93(6), 44-54.
- Priharsari, D., & Abedin, B. (2021). What facilitates and constrains value co-creation in online communities: A sociomateriality perspective. *Information & Management*, 58(6), 103433. <https://doi.org/https://doi.org/10.1016/j.im.2021.103433>
- Raguseo, E., & Vitari, C. (2018). Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. *International Journal of Production Research*, 56(15), 5206–5221.

- Rahman, M. S., Bag, S., Gupta, S., & Sivarajah, U. (2023). Technology readiness of B2B firms and AI-based customer relationship management capability for enhancing social sustainability performance. *Journal of Business Research*, 156, 113525. <https://doi.org/https://doi.org/10.1016/j.jbusres.2022.113525>
- Raimo, N., Vitolla, F., Marrone, A., & Rubino, M. (2021). Do audit committee attributes influence integrated reporting quality? An agency theory viewpoint. *Business Strategy and the Environment*, 30(1), 522–534. <https://doi.org/10.1002/bse.2635>
- Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies*, 26(3), 2623–2656. <https://doi.org/10.1007/s10639-020-10375-1>
- Rakipi, R., De Santis, F., & D’Onza, G. (2021). Correlates of the internal audit function’s use of data analytics in the big data era: Global evidence. *Journal of International Accounting, Auditing and Taxation*, 42, 100357. <https://doi.org/https://doi.org/10.1016/j.intaccaudtax.2020.100357>
- Rana, S. K., Dangwal, B., Negi, V. S., & Bhatt, I. D. (2022). Scientific research in the Himalaya: Current state of knowledge, funding paradigm and policy implications. *Environmental Science and Policy*, 136(July), 685–695. <https://doi.org/10.1016/j.envsci.2022.07.030>
- Ranta, M., Paroutis, S., Talaoui, Y., & Kohtam, M. (2023). Recovering the divide : A review of the big data analytics – strategy relationship. *Long Range Planning*, March 2021, 1–20. <https://doi.org/10.1016/j.lrp.2022.102290>
- Rass, L., Treur, J., Kucharska, W., & Wiewiora, A. (2023). Adaptive Dynamical Systems Modelling of Transformational Organizational Change with Focus on Organizational Culture and Organizational Learning. *Cognitive Systems Research*. <https://doi.org/10.1016/j.cogsys.2023.01.004>
- Rehman, W. ul, Degirmen, S., & Waseem, F. (2022). Propensity for and Quality of Intellectual Capital Divulgence Across the BRICS Banking Sector: A Knowledge-Based Perspective from Emerging Economies. *Journal of the Knowledge Economy*, 13(2), 1028–1055. <https://doi.org/10.1007/s13132-021-00730-z>
- Ringov, D. (2017). Dynamic capabilities and firm performance. *Long Range Planning*, 50(5), 653–664.
- Rodrigues, V. F., Policarpo, L. M., da Silveira, D. E., da Rosa Righi, R., da Costa, C. A., Barbosa, J. L. V., Antunes, R. S., Scorsatto, R., & Arcot, T. (2022). Fraud detection and prevention in e-commerce: A systematic literature review. *Electronic Commerce Research and Applications*, 56, 101207. <https://doi.org/https://doi.org/10.1016/j.elerap.2022.101207>
- Roh, Y., Heo, G., & Whang, S. E. (2021). A Survey on Data Collection for Machine Learning: A Big Data-AI Integration Perspective. *IEEE Transactions on Knowledge and Data Engineering*, 33(4), 1328–1347. <https://doi.org/10.1109/TKDE.2019.2946162>
- Rosnidah, I., Johari, R. J., Hairudin, N. A. M., Hussin, S. A. H. S., & Musyaffi, A. M. (2022). Detecting and Preventing Fraud with Big data analytics: Auditing Perspective. *Journal of Governance and Regulation*, 11(4), 8–15. <https://doi.org/10.22495/jgrv11i4art1>
- Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You may not need big data after all. *Harvard business review*, 91(12), 90–+.
- Saad, S., & Alnuaimi, S. (2022). Innovation Framework for Digital Era. *Advances in Transdisciplinary Engineering*, 25, 54–59. <https://doi.org/10.3233/ATDE220565>
- Saggi, M. K., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management*, 54(5), 758–790.

- Saide, S., & Sheng, M. L. (2020). Toward Business Process Innovation in the Big Data Era: A Mediating Roles of Big Data Knowledge Management. *Big Data*, 8(6), 464–477. <https://doi.org/10.1089/big.2020.0140>
- Salim, S., Turnbull, B., & Moustafa, N. (2022). Data analytics of social media 3.0: Privacy protection perspectives for integrating social media and Internet of Things (SM-IoT) systems. *Ad Hoc Networks*, 128. <https://doi.org/10.1016/j.adhoc.2022.102786>
- Sanoran, K. (Lek), & Ruangprapun, J. (2023). Initial Implementation of Data Analytics and Audit Process Management. *Sustainability*, 15(3), 1766. <https://doi.org/10.3390/su15031766>
- Sardi, A., Sorano, E., Cantino, V., & Garengo, P. (2020). Big data and performance measurement research: trends, evolution and future opportunities. *Measuring Business Excellence*, October 2019. <https://doi.org/10.1108/MBE-06-2019-0053>
- Scalco, P. R., Tabak, B. M., & Teixeira, A. M. (2021). Prudential measures and their adverse effects on bank competition: The case of Brazil. *Economic Modelling*, 100(March), 105495. <https://doi.org/10.1016/j.econmod.2021.03.014>
- Schoenherr, T. (2023). Supply chain management professionals ' proficiency in big data analytics: Antecedents and impact on performance. *Transportation Research Part E*, 169(June 2022), 102972. <https://doi.org/10.1016/j.tre.2022.102972>
- Sena, V., Bhaumik, S., Sengupta, A., & Demirbag, M. (2019). Big Data and Performance: What Can Management Research Tell us? *British Journal of Management*, 30(2), 219–228. <https://doi.org/10.1111/1467-8551.12362>
- Shah, T. R. (2022). Can big data analytics help organisations achieve sustainable competitive advantage? A developmental enquiry. *Technology in Society*, 68(October 2021), 101801. <https://doi.org/10.1016/j.techsoc.2021.101801>
- Shamim, S., Zeng, J., Khan, Z., & Zia, N. U. (2020). Big data analytics capability and decision making performance in emerging market firms: The role of contractual and relational governance mechanisms. *Technological Forecasting and Social Change*, 161. <https://doi.org/10.1016/j.techfore.2020.120315>
- Shamim, S., Zeng, J., Shafi Choksy, U., & Shariq, S. M. (2020). Connecting big data management capabilities with employee ambidexterity in Chinese multinational enterprises through the mediation of big data value creation at the employee level. *International Business Review*, 29(6). <https://doi.org/10.1016/j.ibusrev.2019.101604>
- Shashi, Centobelli, P., Cerchione, R., & Singh, R. (2019). The impact of leanness and innovativeness on environmental and financial performance: Insights from Indian SMEs. *International Journal of Production Economics*, 212, 111–124. <https://doi.org/https://doi.org/10.1016/j.ijpe.2019.02.011>
- Sirois, L.-P., & Simunic, D. A. (2011). Auditor size and audit quality revisited: The importance of audit technology. *Available at SSRN 1694613*.
- Slapničar, S., Vuko, T., Čular, M., & Drašček, M. (2022). Effectiveness of cybersecurity audit. *International Journal of Accounting Information Systems*, 44, 100548. <https://doi.org/https://doi.org/10.1016/j.accinf.2021.100548>
- Smaldone, F., Ippolito, A., Lager, J., & Pellicano, M. (2022). Employability skills : Profiling data scientists in the digital labour market. *European Management Journal*, 40(5), 671–684. <https://doi.org/10.1016/j.emj.2022.05.005>
- Soares de Almeida, C. A., Del Corso, J. M., Rocha, L. A., da Silva, W. V., & da Veiga, C. P. (2019). Innovation and Performance: The Impact of Investments in R&D According to the Different Levels of Productivity of Firms. *International Journal of Innovation and Technology*

- Management*, 16(05), 1950036.
- Song, P., Zheng, C., Zhang, C., & Yu, X. (2018). Data analytics and firm performance: An empirical study in an online B2C platform. *Information & management*, 55(5), 633–642.
- Suddaby, R., Coraiola, D., Harvey, C., & Foster, W. (2020). History and the micro-foundations of dynamic capabilities. *Strategic Management Journal*, 41(3), 530–556. <https://doi.org/10.1002/smj.3058>
- Sun, X. S., Habib, A., & Bhuiyan, M. B. U. (2020). Workforce environment and audit fees: International evidence. *Journal of Contemporary Accounting and Economics*, 16(1). <https://doi.org/10.1016/j.jcae.2020.100182>
- Suryanto, T. (2014). *Determinants of audit fee based on client attribute, auditor attribute, and engagement attribute to control risks and prevent fraud: A study on public accounting firms in Sumatra-Indonesia*.
- Talaoui, Y., Kohtamäki, M., Ranta, M., & Paroutis, S. (2023). Recovering the divide: A review of the big data analytics—strategy relationship. *Long Range Planning*, 102290. <https://doi.org/https://doi.org/10.1016/j.lrp.2022.102290>
- Tang, L., Li, J., Du, H., Li, L., Wu, J., & Wang, S. (2022). Big Data in Forecasting Research : A Literature Review. *Big Data Research*, 27, 100289. <https://doi.org/10.1016/j.bdr.2021.100289>
- Teece, D., & Pisano, G. (2003). *The dynamic capabilities of firms*. Springer.
- Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic management journal*, 28(13), 1319-1350.
- Teece, D. J. (2016). Dynamic capabilities and entrepreneurial management in large organizations: Toward a theory of the (entrepreneurial) firm. *European Economic Review*, 86, 202-216.
- Thirathon, U., Wieder, B., Matolcsy, Z., & Ossimitz, M.-L. (2017). Impact of big data analytics on decision making and performance. *International conference on enterprise systems, accounting and logistics*.
- Thompson, M., & Venters, W. (2021). Platform, or technology project? A spectrum of six strategic ‘plays’ from UK government IT initiatives and their implications for policy. *Government Information Quarterly*, 38(4), 101628. <https://doi.org/https://doi.org/10.1016/j.giq.2021.101628>
- Thottoli, M. M., & K.V, T. (2022). Characteristics of information communication technology and audit practices: evidence from India. *VINE Journal of Information and Knowledge Management Systems*, 52(4), 570–593. <https://doi.org/10.1108/VJIKMS-04-2020-0068>
- Tian, H., Li, Y., & Zhang, Y. (2022). Digital and intelligent empowerment: Can big data capability drive green process innovation of manufacturing enterprises? *Journal of Cleaner Production*, 377, 134261. <https://doi.org/https://doi.org/10.1016/j.jclepro.2022.134261>
- Tippins, M. J., & Sohi, R. S. (2003). IT competency and firm performance: is organizational learning a missing link? *Strategic management journal*, 24(8), 745–761.
- Trainor, K. J., Andzulis, J. M., Rapp, A., & Agnihotri, R. (2014). Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM. *Journal of business research*, 67(6), 1201–1208.
- Turulja, L., Vugec, D. S., & Bach, M. P. (2023). Big Data and Labour Markets: A Review of Research Topics. *Procedia Computer Science*, 217, 526–535. <https://doi.org/https://doi.org/10.1016/j.procs.2022.12.248>
- Ulbert, J., Takács, A., & Csapi, V. (2022). Golden ratio-based capital structure as a tool for boosting firm’s financial performance and market acceptance. *Heliyon*, 8(6).

<https://doi.org/10.1016/j.heliyon.2022.e09671>

- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356-365.
- Wang, J., Xu, C., Zhang, J., & Zhong, R. (2022). Big data analytics for intelligent manufacturing systems: A review. *Journal of Manufacturing Systems*, 62, 738-752. <https://doi.org/https://doi.org/10.1016/j.jmsy.2021.03.005>
- Wei, S., Yin, J., & Chen, W. (2022). How big data analytics use improves supply chain performance: considering the role of supply chain and information system strategies. *The International Journal of Logistics Management*, 33(2), 620-643. <https://doi.org/10.1108/IJLM-06-2020-0255>
- Wen, C., Yang, J., Gan, L., & Pan, Y. (2021). Big data driven Internet of Things for credit evaluation and early warning in finance. *Future Generation Computer Systems*, 124, 295-307. <https://doi.org/10.1016/j.future.2021.06.003>
- Werner, M., Wiese, M., & Maas, A. (2021). Embedding process mining into financial statement audits. *International Journal of Accounting Information Systems*, 41, 100514. <https://doi.org/https://doi.org/10.1016/j.accinf.2021.100514>
- Xu, E. G., Graves, C., Shan, Y. G., & Yang, J. W. (2022). The mediating role of corporate social responsibility in corporate governance and firm performance. *Journal of Cleaner Production*, 375(August), 134165. <https://doi.org/10.1016/j.jclepro.2022.134165>
- Xu, T., Yang, J., Zhang, F., & Guo, W. (2021). Interfirm coopetition, interfirm knowledge creation, and collaborative innovation performance: The moderating roles of environmental competitiveness and dysfunctional competition. *Industrial Marketing Management*, 99, 123-135. <https://doi.org/https://doi.org/10.1016/j.indmarman.2021.10.003>
- Yang, A., Lu, C., Li, J., Huang, X., Ji, T., Li, X., & Sheng, Y. (2022). Application of meta-learning in cyberspace security: a survey. *Digital Communications and Networks*. <https://doi.org/https://doi.org/10.1016/j.dcan.2022.03.007>
- Yang, Y., & Yee, R. W. Y. (2022). The effect of process digitalization initiative on firm performance: A dynamic capability development perspective. *International Journal of Production Economics*, 254, 108654. <https://doi.org/https://doi.org/10.1016/j.ijpe.2022.108654>
- Yasmin, M., Tatoglu, E., Kilic, H. S., Zaim, S., & Delen, D. (2020). Big data analytics capabilities and firm performance: An integrated MCDM approach. *Journal of Business Research*, 114(March), 1-15. <https://doi.org/10.1016/j.jbusres.2020.03.028>
- Ye, Y., Yu, Q., Zheng, Y., & Zheng, Y. (2022). Investigating the effect of social media application on firm capabilities and performance: The perspective of dynamic capability view. *Journal of Business Research*, 139, 510-519. <https://doi.org/10.1016/j.jbusres.2021.10.008>
- Yousuf, M., & Wahid, A. (2021). The role of Artificial Intelligence in Education: Current Trends and Future Prospects. *International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2021*. <https://doi.org/10.1109/ICISCT52966.2021.9670009>
- Yu, W., Wong, C. Y., Chavez, R., & Jacobs, M. A. (2021). Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture. *International Journal of Production Economics*, 236(April), 108135. <https://doi.org/10.1016/j.ijpe.2021.108135>
- Zekić-Sušac, M., Mitrović, S., & Has, A. (2021). Machine learning based system for managing

- energy efficiency of public sector as an approach towards smart cities. *International Journal of Information Management*, 58. <https://doi.org/10.1016/j.ijinfomgt.2020.102074>
- Zhang, X., Chu, Z., Ren, L., & Xing, J. (2023). Open innovation and sustainable competitive advantage: The role of organizational learning. *Technological Forecasting & Social Change*, 186(PA), 122114. <https://doi.org/10.1016/j.techfore.2022.122114>
- Zhang, Y., Hou, Z., Yang, F., Yang, M. M., & Wang, Z. (2021). Discovering the evolution of resource-based theory: Science mapping based on bibliometric analysis. *Journal of Business Research*, 137, 500–516. <https://doi.org/https://doi.org/10.1016/j.jbusres.2021.08.055>

## Referências do capítulo 2

- Adhikari, B., Kavanagh, M., & Hampson, B. (2023). Analysis of the pre-post-merger and acquisition financial performance of selected banks in Nepal. *Asia Pacific Management Review*. <https://doi.org/https://doi.org/10.1016/j.apmr.2023.02.001>
- Ahamed, M. M., Ho, S. J., Mallick, S. K., & Matousek, R. (2021). Inclusive banking, financial regulation and bank performance: Cross-country evidence. *Journal of Banking & Finance*, 124, 106055. <https://doi.org/https://doi.org/10.1016/j.jbankfin.2021.106055>
- Ahmed, S., Alshater, M. M., Ammari, A. El, & Hammami, H. (2022). Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance*, 61, 101646. <https://doi.org/https://doi.org/10.1016/j.ribaf.2022.101646>
- Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, 40, 100402. <https://doi.org/https://doi.org/10.1016/j.cosrev.2021.100402>
- Al-Khouri, R., & Al-Otaibi, H. (2022). The relationship between liquidity risk and profitability: Evidence from GCC banks. *Journal of Banking & Finance*, 135, 106302. <https://doi.org/10.1016/j.jbankfin.2022.106302>
- Al-Osaimy, M. H. (1998). A Neural Networks System for Predicting Islamic Banks Performance. *Journal of King Abdul Azis University: Islamic Economic & Administration*, 11(1), 33–46.
- Anouze, A. L. M., & Bou-Hamad, I. (2019). Data envelopment analysis and data mining to efficiency estimation and evaluation. *International Journal of Islamic and Middle Eastern Finance and Management*, 12(2), 169–190. <https://doi.org/10.1108/IMEFM-11-2017-0302>
- Anwar, S., & Ali, A. M. (2018). ANNs-Based Early Warning System for Indonesian Islamic Banks. *Buletin Ekonomi Moneter Dan Perbankan*, 20(3). <https://doi.org/10.21098/bemp.v20i3.856>
- Ashiru, O., Balogun, G., & Paseda, O. (2023). Financial innovation and bank financial performance: Evidence from Nigerian deposit money banks. *Research in Globalization*, 6, 100120. <https://doi.org/https://doi.org/10.1016/j.resglo.2023.100120>
- Aydin, N., Sahin, N., Deveci, M., & Pamucar, D. (2022). Prediction of financial distress of companies with artificial neural networks and decision trees models. *Machine Learning with Applications*, 10, 100432. <https://doi.org/https://doi.org/10.1016/j.mlwa.2022.100432>
- Bakar, N. M. A., & Tahir, I. M. (2009). Applying Multiple Linear Regression and Neural Network to Predict Bank Performance. *International Business Research*, 2(4).
- Banco Central do Brasil. (2023a, March 14). *Quantidade de bancos autorizados a funcionar no Sistema Financeiro Nacional*. <https://Dadosabertos.Bcb.Gov.Br/Dataset/24868-Quantidade-de-Bancos-Autorizados-a-Funcionar-No-Sistema-Financeiro-Nacional>.
- Banco Central do Brasil. (2023b, March 15). *IF.data*. <https://www3.bcb.gov.br/ifdata/>.
- Banco Central do Brasil. (2023c, June 27). *IF.data*. <https://www3.bcb.gov.br/ifdata/legado.html>.
- Bany Mohammad, A., Al-Okaily, M., Al-Majali, M., & Masa'deh, R. (2022). Business Intelligence and Analytics (BIA) Usage in the Banking Industry Sector: An Application of the TOE Framework. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(4), 189. <https://doi.org/https://doi.org/10.3390/joitmc8040189>
- Beder, O. A. (2021). Liquidity management and profitability of Islamic and conventional banks:

- Empirical evidence from Turkey. *Journal of Economic Cooperation and Development*, 42(3), 83-104. <https://doi.org/10.32380/jecd.42303>
- Bedoui, A., & Lazar, N. A. (2020). Bayesian empirical likelihood for ridge and lasso regressions. *Computational Statistics & Data Analysis*, 145, 106917. <https://doi.org/https://doi.org/10.1016/j.csda.2020.106917>
- Bejar, P., Ishi, K., Komatsuzaki, T., Shibata, I., Sin, J., & Tambunlertchai, S. (2022). Can Fintech Foster Competition in the Banking System in Latin America and the Caribbean? *Latin American Journal of Central Banking*, 3(2), 100061. <https://doi.org/https://doi.org/10.1016/j.latcb.2022.100061>
- Bianchi, J., & Bigio, S. (2022). Banks, liquidity management, and monetary policy. *Econometrica*, 90(1), 391-454. doi:10.3982/ECTA16599
- Bolívar, F., Duran, M. A., & Lozano-Vivas, A. (2023). Business model contributions to bank profit performance: A machine learning approach. *Research in International Business and Finance*, 64, 101870. <https://doi.org/https://doi.org/10.1016/j.ribaf.2022.101870>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. <https://doi.org/10.1007/BF00058655>
- Cai, S., & Zhang, J. (2020). Exploration of credit risk of P2P platform based on data mining technology. *Journal of Computational and Applied Mathematics*, 372, 112718. <https://doi.org/https://doi.org/10.1016/j.cam.2020.112718>
- Cao, M., Yin, D., Zhong, Y., Lv, Y., & Lu, L. (2023). Detection of geochemical anomalies related to mineralization using the Random Forest model optimized by the Competitive Mechanism and Beetle Antennae Search. *Journal of Geochemical Exploration*, 107195. <https://doi.org/https://doi.org/10.1016/j.gexplo.2023.107195>
- Capéleti, P., Garcia, M., & Miessi Sanches, F. (2022). Countercyclical credit policies and banking concentration: Evidence from Brazil. *Journal of Banking & Finance*, 143, 106589. <https://doi.org/https://doi.org/10.1016/j.jbankfin.2022.106589>
- Castillo, E., Guijarro-Berdiñas, B., Berdiñas, B., Fontenla-Romero, O., Alonso-Betanzos, A., & Es, C. (2006). A Very Fast Learning Method for Neural Networks Based on Sensitivity Analysis. In *Journal of Machine Learning Research* (Vol. 7).
- Centofanti, F., Fontana, M., Lepore, A., & Vantini, S. (2022). Smooth LASSO estimator for the Function-on-Function linear regression model. *Computational Statistics & Data Analysis*, 176, 107556. <https://doi.org/https://doi.org/10.1016/j.csda.2022.107556>
- Chen, X. (2020). Exploring the sources of financial performance in Chinese banks: A comparative analysis of different types of banks. *The North American Journal of Economics and Finance*, 51, 101076. <https://doi.org/https://doi.org/10.1016/j.najef.2019.101076>
- Chen, X., Wang, Y., & Wu, X. (2022). Exploring the source of the financial performance in Chinese banks: A risk-adjusted decomposition approach. *International Review of Financial Analysis*, 80, 102051. <https://doi.org/https://doi.org/10.1016/j.irfa.2022.102051>
- Chicco D, Warrens MJ, Jurman G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, 623. <https://doi.org/10.7717/peerj-cs.623>
- Ciner, C., Lucey, B., & Yarovaya, L. (2022). Determinants of cryptocurrency returns: A LASSO quantile regression approach. *Finance Research Letters*, 49, 102990. <https://doi.org/https://doi.org/10.1016/j.frl.2022.102990>
- da Silva, T. G., de Carvalho Guillén, O. T., Morcerf, G. A. N., & de Melo Modenesi, A. (2022). Effects of monetary policy news on financial assets: Evidence from Brazil on a bivariate

- VAR-GARCH model (2006–17). *Emerging Markets Review*, 52, 100916. <https://doi.org/https://doi.org/10.1016/j.ememar.2022.100916>
- Decelle, A. (2022). An Introduction to Machine Learning: a perspective from Statistical Physics. *Physica A: Statistical Mechanics and Its Applications*, 128154. <https://doi.org/https://doi.org/10.1016/j.physa.2022.128154>
- Dogan, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166, 114060. <https://doi.org/https://doi.org/10.1016/j.eswa.2020.114060>
- Doğuş, Ö. (2022). Data Mining Applications in Banking Sector While Preserving Customer Privacy. *Emerging Science Journal*, 6(6), 1444–1454. <https://doi.org/10.28991/ESJ-2022-06-06-014>
- dos Santos, B. S., Steiner, M. T. A., Fenerich, A. T., & Lima, R. H. P. (2019). Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. *Computers & Industrial Engineering*, 138, 106120. <https://doi.org/https://doi.org/10.1016/j.cie.2019.106120>
- Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap* (Chapman and Hall/CRC, Ed.; 1<sup>a</sup>). Chapman and Hall/CRC. <https://doi.org/https://doi.org/10.1201/9780429246593>
- Elfeituri, H. (2022). Banking stability, institutional quality, market concentration, competition and political conflict in MENA. *Journal of International Financial Markets, Institutions and Money*, 76, 101476. <https://doi.org/https://doi.org/10.1016/j.intfin.2021.101476>
- Fakhri, U. N., Anwar, S., Ismal, R., & Ascarya, A. (2019). Comparison and predicting financial performance of Islamic and conventional banks in Indonesia to achieve growth sustainability. *Al-Uqud: Journal of Islamic Economics*, 3(2), 174–187. <https://doi.org/10.26740/al-uqud.v3n2.p174-187>
- Frndak, S., Mañay, N., Yu, G., Oulhote, Y., Peregalli, F., Queirolo, E. I., Olson, J. R., Barg, G., Ahmed, Z., Vahter, M., & Kordas, K. (2023). Reducing the complexity of high-dimensional environmental data: An analytical framework using LASSO with considerations of confounding for statistical inference. *International Journal of Hygiene and Environmental Health*, 249.
- Gao, B. (2022). The Use of Machine Learning Combined with Data Mining Technology in Financial Risk Prevention. *Computational Economics*, 59(4), 1385–1405. <https://doi.org/10.1007/s10614-021-10101-0>
- Garcia, A. S., & Meurer, R. (2022). Effects of a development bank on the profitability of commercial banks: Evidence for Brazil. *The Quarterly Review of Economics and Finance*, 85, 246–259. <https://doi.org/https://doi.org/10.1016/j.qref.2022.03.008>
- Gong, Y., & Wei, X. (2022). Asset quality, financing structure, and bank regulations. *International Review of Economics & Finance*, 80, 1061–1075. <https://doi.org/https://doi.org/10.1016/j.iref.2022.02.033>
- Granitto, P. M., Gasperi, F., Biasioli, F., Trainotti, E., & Furlanello, C. (2007). Modern data mining tools in descriptive sensory analysis: A case study with a Random forest approach. *Food Quality and Preference*, 18(4), 681–689. <https://doi.org/10.1016/j.foodqual.2006.11.001>
- Hassan, K., & Marimuthu, M. (2022). Asset quality and bank profitability: A dynamic panel data approach. *Journal of Financial Services Research*, 61, 273–299. <https://doi.org/10.1007/s10693-021-00364-9>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. In Springer (Ed.), *Springer Series in Pharmaceutical*

*Statistics* (2nd ed.). Springer.

- Hu, Y., & Su, J. (2022). Research on Credit Risk Evaluation of Commercial Banks Based on Artificial Neural Network Model. *Procedia Computer Science*, 199, 1168–1176. <https://doi.org/https://doi.org/10.1016/j.procs.2022.01.148>
- Iqbal, J., Saeed, A., & Khan, R. A. (2023). The relative importance of textual indexes in predicting the future performance of banks: A connection weight approach. *Borsa Istanbul Review*, 23(1), 240–253. <https://doi.org/https://doi.org/10.1016/j.bir.2022.10.004>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). Statistical Learning. In G. James, D. Witten, T. Hastie, & R. Tibshirani (Eds.), *An Introduction to Statistical Learning: with Applications in R* (pp. 15–57). Springer US. [https://doi.org/10.1007/978-1-0716-1418-1\\_2](https://doi.org/10.1007/978-1-0716-1418-1_2)
- Jeanne, O., & Sandri, D. (2023). Global financial cycle and liquidity management. *Journal of International Economics*, 103736. <https://doi.org/https://doi.org/10.1016/j.jinteco.2023.103736>
- Karim, S., Akhtar, M. U., Tashfeen, R., Raza Rabbani, M., Rahman, A. A. A., & AlAbbas, A. (2022). Sustainable banking regulations pre and during coronavirus outbreak: The moderating role of financial stability. *Economic Research-Ekonomika Istrazivanja*, 35(1), 3360–3377. doi:10.1080/1331677X.2021.1993951
- Kassani, S. H., Kassani, H., & Najafi, E. (2015). Introducing a hybrid model of DEA and data mining in evaluating efficiency. Case study: Bank Branches. *International Conference on Economics, Management and Social Sciences*, 3(2), 72–80.
- Kim, B., Ryu, K. H., & Heo, S. (2022). Mean squared error criterion for model-based design of experiments with subset selection. *Computers & Chemical Engineering*, 159, 107667. <https://doi.org/10.1016/j.compchemeng.2022.107667>
- Kladakis, G., Chen, L. & Bellos, S. K. (2020). Bank Asset and Informational Quality (January 13, 2020). *Journal of International Financial Markets, Institutions and Money*, 69. <http://dx.doi.org/10.2139/ssrn.3518046>
- Ko, D., & Park, S. (2022). Do bank capital adequacy requirements affect bank profitability? Evidence from Korean banks. *Asia-Pacific Journal of Financial Studies*, 51(1), 44–68. <https://doi.org/10.1111/ajfs.12502>
- Kupka, N., Tolosana-Delgado, R., Schach, E., Bachmann, K., Heinig, T., & Rudolph, M. (2020). R as an environment for data mining of process mineralogy data: A case study of an industrial rougher flotation bank. *Minerals Engineering*, 146, 106111. <https://doi.org/https://doi.org/10.1016/j.mineng.2019.106111>
- Li, W. (2021). The role of accounting quality in corporate liquidity management. *Accounting and Finance*, 61(2), 2631–2670. doi:10.1111/acfi.12678
- Liébana-Cabanillas, F., Nogueras, R., Herrera, L. J., & Guillén, A. (2013). Analysing user trust in electronic banking using data mining methods. *Expert Systems with Applications*, 40(14), 5439–5447. <https://doi.org/https://doi.org/10.1016/j.eswa.2013.03.010>
- Lin, J. J., Hsu, Y. H., & Chang, C. L. (2022). Bank risk-taking, asset quality and profitability in Taiwan. *The North American Journal of Economics and Finance*, 56, 101429. <https://doi.org/10.1016/j.najef.2021.101429>
- Lin, S.-W., Shiue, Y.-R., Chen, S.-C., & Cheng, H.-M. (2009). Applying enhanced data mining approaches in predicting bank performance: A case of Taiwanese commercial banks. *Expert Systems with Applications*, 36(9), 11543–11551. <https://doi.org/https://doi.org/10.1016/j.eswa.2009.03.029>
- Luo, J., Xu, J., Aldosari, O., Althubiti, S. A., & Deebani, W. (2022). Design and Implementation

- of an Efficient Electronic Bank Management Information System Based Data Warehouse and Data Mining Processing. *Information Processing & Management*, 59(6), 103086. <https://doi.org/https://doi.org/10.1016/j.ipm.2022.103086>
- Montes, G. C., Valladares, M., & de Moraes, C. O. (2021). Impacts of the sovereign risk perception on financial stability: Evidence from Brazil. *The Quarterly Review of Economics and Finance*, 81, 358–369. <https://doi.org/https://doi.org/10.1016/j.qref.2021.06.010>
- Moradi, M., Salehi, M., Ghorgani, M. E., & Yazdi, H. S. (2013). Financial Distress Prediction of Iranian Companies Using Data Mining Techniques. *Organizacija*, 46(1), 20–27. <https://doi.org/doi:10.2478/orga-2013-0003>
- Moradi, S., & Mokhatab Rafiei, F. (2019). A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks. *Financial Innovation*, 5(1), 15. <https://doi.org/10.1186/s40854-019-0121-9>
- Moyer, S. E. (1990). Capital adequacy ratio regulations and accounting choices in commercial banks. *Journal of Accounting and Economics*, 13(2), 123–154. doi:10.1016/0165-4101(90)90027-2
- Naveed, H. M., HongXing, Y., Memon, B. A., Ali, S., Alhussam, M. I., & Sohu, J. M. (2023). Artificial neural network (ANN)-based estimation of the influence of COVID-19 pandemic on dynamic and emerging financial markets. *Technological Forecasting and Social Change*, 190, 122470. <https://doi.org/https://doi.org/10.1016/j.techfore.2023.122470>
- Pakzad, S.S., Roshan, N. & Ghalehnovi, M. (2023). Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete. *Sci Rep* 13, 3646. <https://doi.org/10.1038/s41598-023-30606-y>
- Patonai, Z., Kicsiny, R., & Géczi, G. (2022). Multiple linear regression based model for the indoor temperature of mobile containers. *Heliyon*, 8(12), e12098. <https://doi.org/https://doi.org/10.1016/j.heliyon.2022.e12098>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Sener, A., Barut, M., Oztekin, A., Avcilar, M. Y., & Yildirim, M. B. (2019). The role of information usage in a retail supply chain: A causal data mining and analytical modeling approach. *Journal of Business Research*, 99, 87–104. <https://doi.org/https://doi.org/10.1016/j.jbusres.2019.01.070>
- Tariq, R., & Masih, M. (2021). Does capital adequacy ratio matter for bank stability and profitability in Islamic banking? Empirical evidence from Malaysia. *International Journal of Finance & Economics*, 26(2), 2753–2765. <https://doi.org/10.1002/ijfe.1943>
- Tata, F. (2023). Proposing an interval design feature to Central Bank Digital Currencies. *Research in International Business and Finance*, 64, 101898. <https://doi.org/https://doi.org/10.1016/j.ribaf.2023.101898>
- Urso, A., Fiannaca, A., La Rosa, M., Ravi, V., & Rizzo, R. (2018). Data mining: Prediction methods. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma*, 1, 3.
- Wada, T. (2022). Out-of-sample forecasting of foreign exchange rates: The band spectral regression and LASSO. *Journal of International Money and Finance*, 128, 102719. <https://doi.org/https://doi.org/10.1016/j.jimonfin.2022.102719>
- Wang, J., Omar, A. H., Alotaibi, F. M., Daradkeh, Y. I., & Althubiti, S. A. (2022). Business intelligence ability to enhance organizational performance and performance evaluation capabilities by improving data mining systems for competitive advantage. *Information Processing & Management*, 59(6), 103075.

- <https://doi.org/https://doi.org/10.1016/j.ipm.2022.103075>
- Wanke, P., Azad, M. D. A. K., & Barros, C. P. (2016). Predicting efficiency in Malaysian Islamic banks: A two-stage TOPSIS and neural networks approach. *Research in International Business and Finance*, 36, 485–498. <https://doi.org/https://doi.org/10.1016/j.ribaf.2015.10.002>
- World Economic Forum. (2023). *The Global Risks Report 2023 18th Edition*. [www.weforum.org](http://www.weforum.org)
- Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20(7), 557–585.
- Yan, J. (2022). Application of improved BP neural network model in bank financial accounting. *Intelligent Systems with Applications*, 16, 200155. <https://doi.org/https://doi.org/10.1016/j.iswa.2022.200155>
- Zhang, H., Ren, S., Li, X., Baharin, H., Alghamdi, A., & Alghamdi, O. A. (2023). Developing scalable management information system with big financial data using data mart and mining architecture. *Information Processing & Management*, 60(3), 103326. <https://doi.org/https://doi.org/10.1016/j.ipm.2023.103326>

# Apêndice do Capítulo 1

## Questionário

Prezado entrevistado,

Sou estudante de Pós-graduação em Administração na Universidade de Brasília, desejando realizar investigação sobre *A relação entre Big Data Analytics e Desempenho Financeiro Bancário: Uma Investigação Com Dados Brasileiros* para a conclusão do meu relatório de pesquisa.

A esse respeito, preparei o seguinte questionário, observe que sua identidade como entrevistado é ocultada. Você pode expressar livremente quaisquer realidades fundamentais que veja e enfrente. Levará de 10 a 15 minutos para responder as perguntas; todas as informações obtidas para esta pesquisa serão usadas apenas para fins acadêmicos.

Para mais perguntas, envie um e-mail para pauloemilio01@gmail.com.

Agradeço o dispêndio de seu tempo para preencher este questionário.

Atenciosamente

**Paulo Emílio**

## DEMOGRAFIA

### 1. Gênero

Masculino

Feminino

### 2. Idade

21 a 30 anos

31 a 40 anos

41 a 50 anos

Mais de 50 anos

### 3. Experiência

Menos de 3 anos

3 a 5 anos

6 a 10 anos

11 a 15 anos

### 4. Educação Nível

Bacharel

Mestrado

Doutorado

Outros (por favor especifique)

---

### 5. Tipo de Banco

Público

Privado

### 6. Experiência na Organização Atual

Menos de 3 anos

3 a 5 anos

6 a 10 anos

11 a 15 anos

### 7. Posição

1. Auditor/contador

2. Direção

3. Vice-presidente

4. Diretor de recursos humanos

5. Presidente

**Tendo em vista a sua experiência na empresa, indique a extensão do seu acordo e desacordo, inserindo a opção apropriada.**

**Discordo Fortemente = 1, Discordo = 2, Neutro = 3, Concordo = 4, Concordo Fortemente = 5**

<b>Escala de competência de análise de big data</b>	<b>Discordo Fortemente</b>	<b>Discordo</b>	<b>Neutro</b>	<b>Concordo</b>	<b>Concordo Fortemente</b>
<b>Dados</b>					
BDAC1: Tenho acesso à análise de dados muito grandes, não estruturados ou de movimento rápido.	1	2	3	4	5
BDAC2: Integro dados de várias fontes internas em um banco de dados ou mercado de fácil acesso.	1	2	3	4	5
BDAC3: Integro dados externos com dados internos para facilitar a análise de alto valor do nosso ambiente de negócios.	1	2	3	4	5
<b>Grandeza dos dados</b>					
BDAC4: Em nossa organização, processamos alto volume de dados.	1	2	3	4	5
BDAC5: Em nossa organização, processamos dados em tempo real.	1	2	3	4	5
<b>Qualidade dos dados</b>					
BDAC6: Em nossa organização, os dados usados nas análises são confiáveis.	1	2	3	4	5
BDAC7: Em nossa organização, os dados usados nas análises têm um nível adequado de detalhes.	1	2	3	4	5
BDAC8: Em nossa organização, os dados usados nas análises são relevantes para a rotina de atividades.	1	2	3	4	5
<b>Tecnologia</b>					
BDAC9: Exploramos ou adotamos abordagens de computação paralela (por exemplo, Hadoop) para processamento de big data.	1	2	3	4	5
BDAC10: Exploramos ou adotamos diferentes ferramentas de visualização de dados.	1	2	3	4	5
BDAC11: Exploramos ou adotamos o serviço de armazenamento baseado em nuvem para processamento de dados e realização de análises.	1	2	3	4	5
BDAC12: Exploramos ou adotamos novas formas de bancos de dados, como o Not Only SQL (NoSQL) para armazenamento de dados.	1	2	3	4	5

<b>Recursos Básicos</b>					
BDAC13: Nossos projetos de análise de big data são adequadamente financiados.	1	2	3	4	5
BDAC14: Aos nossos projetos de análise de big data é fornecido tempo suficiente para alcançar seus objetivos.	1	2	3	4	5
<b>Habilidades de Análise Técnica</b>					
BDAC15: Fornecemos treinamento em análise de big data para nossos funcionários próprios.	1	2	3	4	5
BDAC16: Contratamos novos funcionários que já possuem habilidades de análise de big data.	1	2	3	4	5
BDAC17: Nossa equipe de análise de big data tem as habilidades certas para realizar seus trabalhos com sucesso.	1	2	3	4	5
<b>Habilidades Gerenciais</b>					
BDAC18: Nossos gerentes de análise de big data entendem e apreciam as necessidades de negócios de outros gerentes funcionais, fornecedores e clientes.	1	2	3	4	5
BDAC19: Nossos gerentes de análise de big data são capazes de trabalhar com outros gerentes funcionais, fornecedores e clientes para determinar as oportunidades que o big data pode trazer para o nosso negócio.	1	2	3	4	5
BDAC20: Nossos gerentes de análise de big data são capazes de coordenar atividades relacionadas a big data de maneira que apoiem outros gerentes funcionais, fornecedores e clientes.	1	2	3	4	5
BDAC21: Nossos gerentes de análise de big data têm um bom senso de onde aplicar o big data.	1	2	3	4	5
BDAC22: Nossos gerentes de análise de big data são capazes de entender e avaliar as saídas de dados extraídas do big data.	1	2	3	4	5
<b>Conhecimento de Domínio</b>					
BDAC23: Em nossa organização, há um alto nível de conhecimento do ambiente externo (por exemplo, governo, concorrentes, fornecedores e clientes).	1	2	3	4	5
BDAC24: Em nossa organização, há um alto nível de conhecimento das metas e objetivos organizacionais.	1	2	3	4	5
BDAC25: Em nossa organização, há um alto nível de conhecimento das principais capacidades da organização	1	2	3	4	5
BDAC26: Em nossa organização, há um alto nível de conhecimento dos principais fatores que devem dar certo para que a organização tenha sucesso.	1	2	3	4	5
<b>Cultura orientada por dados</b>					
BDAC27: Estamos dispostos a anular nossa própria intuição quando os dados contradizem nossos pontos de vista.	1	2	3	4	5
BDAC28: Avaliamos e aprimoramos continuamente as regras de negócios em resposta aos insights extraídos dos dados.	1	2	3	4	5

<b>Desempenho Financeiro</b>	<b>Discordo Fortemente</b>	<b>Discordo</b>	<b>Neutro</b>	<b>Concordo</b>	<b>Concordo Fortemente</b>
FP1: O retorno sobre o investimento da nossa empresa é maior em comparação com os concorrentes.	1	2	3	4	5
FP2: O retorno sobre os ativos da nossa empresa é maior em comparação com os concorrentes.	1	2	3	4	5
FP3: O crescimento das vendas e a rentabilidade da nossa empresa são maiores em comparação com os concorrentes.	1	2	3	4	5
FP4: Os custos operacionais totais da nossa empresa são menores em comparação com os concorrentes.	1	2	3	4	5

<b>Eficácia do comitê de auditoria</b>	<b>Discordo Fortemente</b>	<b>Discordo</b>	<b>Neutro</b>	<b>Concordo</b>	<b>Concordo Fortemente</b>
ACE1: O comitê de auditoria incentiva a empresa a fornecer boas práticas corporativas.	1	2	3	4	5
ACE2: O comitê de auditoria fortalece as funções e a eficácia dos diretores não executivos.	1	2	3	4	5
ACE3: O comitê de auditoria tem auxiliado os diretores no cumprimento de suas responsabilidades estatutárias no que diz respeito aos relatórios financeiros.	1	2	3	4	5
ACE4: O comitê de auditoria preservou e reforçou a independência dos auditores internos.	1	2	3	4	5
ACE5: O Comitê de Auditoria tem assistido os auditores na comunicação de deficiências graves no ambiente de controle ou de fraquezas de gerenciamento.	1	2	3	4	5
ACE6: O comitê de auditoria melhorou a comunicação entre o conselho e os auditores internos.	1	2	3	4	5
ACE7: O comitê de auditoria melhorou a comunicação entre o conselho de administração e os auditores externos.	1	2	3	4	5
ACE8: O comitê de auditoria aumentou a confiança do público na credibilidade e na objetividade das demonstrações financeiras.	1	2	3	4	5
ACE9: O comitê de auditoria tem assistido a gestão a cumprir as suas responsabilidades em relação à prevenção de fraudes, de outras irregularidades e de erros.	1	2	3	4	5
ACE10: O comitê de auditoria aumentou a confiança dos analistas de investimento na credibilidade e objetividade das demonstrações financeiras.	1	2	3	4	5
ACE11: O comitê de auditoria forneceu um fórum para arbitragem entre a administração e os auditores.	1	2	3	4	5
ACE12: O comitê de auditoria tem se preocupado com a possibilidade de pressão legislativa.	1	2	3	4	5

<b>Recursos dinâmicos orientados a processos</b>	<b>Discordo Fortemente</b>	<b>Discordo</b>	<b>Neutro</b>	<b>Concordo</b>	<b>Concordo Fortemente</b>
PODC1: Nossa empresa é melhor do que os concorrentes na conexão (por exemplo, comunicação e compartilhamento de informações) de partes dentro de um processo de negócios.	1	2	3	4	5
PODC2: Nossa empresa é melhor do que os concorrentes em redução de custos dentro de um processo de negócios.	1	2	3	4	5
PODC3: Nossa empresa é melhor do que os concorrentes no tratamento de métodos analíticos complexos para suportar um processo de negócios.	1	2	3	4	5
PODC4: Nossa empresa é melhor do que os concorrentes em trazendo informações detalhadas para um processo de negócios	1	2	3	4	5

# Apêndice do Capítulo 2

**Tabela A2.1** - Dados trimestrais de CAR, AQ, LM e ROA no sistema bancário brasileiro.

Trimestre	CAR (%)	AQ (%)	LM (%)	ROA (%)
2011T01	16,15%	5,41%	23,57%	0,75%
2011T02	16,27%	5,48%	22,94%	1,56%
2011T03	15,45%	5,56%	22,86%	0,72%
2011T04	15,98%	5,58%	21,50%	1,29%
2012T01	15,98%	5,62%	21,50%	1,30%
2012T02	15,72%	5,65%	22,70%	1,09%
2012T03	15,41%	5,69%	22,70%	0,53%
2012T04	15,24%	5,48%	21,64%	1,06%
2013T01	15,95%	5,37%	21,99%	0,46%
2013T02	16,27%	5,12%	22,43%	1,07%
2013T03	16,42%	5,04%	22,26%	0,46%
2013T04	15,29%	4,89%	21,52%	0,88%
2014T01	15,03%	4,77%	21,83%	0,39%
2014T02	15,23%	4,71%	20,92%	0,83%
2014T03	16,23%	4,71%	20,90%	0,42%
2014T04	16,46%	4,72%	21,28%	0,86%
2015T01	15,09%	4,74%	23,74%	0,50%
2015T02	16,20%	4,77%	23,15%	1,12%
2015T03	15,39%	5,10%	24,96%	0,62%
2015T04	16,26%	5,32%	24,07%	1,10%
2016T01	16,20%	5,66%	24,19%	0,49%
2016T02	16,47%	5,95%	24,43%	0,91%
2016T03	16,93%	6,26%	23,20%	0,68%
2016T04	17,43%	6,32%	23,35%	1,32%
2017T01	17,90%	6,47%	23,64%	0,65%
2017T02	18,22%	6,59%	22,57%	1,37%
2017T03	18,07%	6,53%	23,26%	0,72%
2017T04	17,68%	6,49%	23,92%	1,74%
2018T01	17,44%	6,36%	24,70%	0,85%
2018T02	17,20%	6,20%	24,30%	1,66%
2018T03	17,82%	6,30%	25,06%	0,92%
2018T04	17,48%	6,41%	25,21%	1,68%
2019T01	17,25%	6,34%	26,13%	1,33%
2019T02	18,06%	6,12%	26,94%	2,50%
2019T03	16,91%	6,01%	27,63%	1,12%
2019T04	16,88%	6,38%	27,01%	2,33%
2020T01	15,37%	6,50%	28,48%	0,81%

**Fonte:** Elaborado pelo autor com os dados do IF.Data do Banco Central do Brasil

**Tabela A2.2** – Previsão do desempenho financeiro do sistema bancário no Brasil usando técnicas de mineração de dados.

Trimestre	Previsão do ROA com a KNN (%)	Previsão do ROA com a RF (%)	Previsão do ROA com a RNA (%)	Previsão do ROA com a regressão LASSO (%)
2011T01	0,75%	0,77%	0,47%	0,96%
2011T02	1,56%	1,36%	1,61%	1,05%
2011T03	0,72%	0,73%	0,89%	0,95%
2011T04	1,29%	1,17%	1,16%	0,95%
2012T01	1,09%	1,03%	2,09%	0,95%
2012T02	0,93%	0,95%	1,46%	0,94%
2012T03	0,53%	0,61%	0,53%	0,94%
2012T04	1,06%	1,02%	1,39%	0,96%
2013T01	0,46%	0,52%	0,46%	0,89%
2013T02	1,07%	0,98%	1,29%	0,94%
2013T03	0,46%	0,59%	0,58%	0,92%
2013T04	0,88%	0,86%	1,09%	0,95%
2014T01	0,39%	0,58%	0,36%	0,89%
2014T02	0,83%	0,84%	0,66%	0,95%
2014T03	0,42%	0,50%	0,44%	0,83%
2014T04	0,86%	0,87%	1,04%	0,95%
2015T01	0,50%	0,56%	0,53%	0,91%
2015T02	1,12%	1,02%	1,10%	0,94%
2015T03	0,62%	0,66%	0,52%	0,95%
2015T04	1,10%	1,04%	1,23%	0,95%
2016T01	0,49%	0,57%	0,43%	0,92%
2016T02	0,91%	0,89%	1,09%	0,95%
2016T03	1,01%	0,78%	0,78%	0,96%
2016T04	1,29%	1,14%	1,68%	0,95%
2017T01	0,65%	0,70%	0,69%	0,94%
2017T02	1,43%	1,16%	0,95%	0,96%
2017T03	0,90%	0,83%	0,37%	0,96%
2017T04	1,74%	1,52%	1,92%	1,18%
2018T01	0,85%	0,84%	0,42%	0,95%
2018T02	1,66%	1,42%	1,91%	1,10%
2018T03	0,92%	0,91%	0,76%	0,95%
2018T04	1,68%	1,45%	1,69%	1,18%
2019T01	1,33%	1,28%	1,29%	0,95%
2019T02	2,50%	2,04%	2,34%	1,81%
2019T03	1,12%	1,09%	0,91%	0,95%
2019T04	2,33%	2,14%	2,23%	1,92%
2020T01	0,81%	0,81%	0,94%	0,95%

**Fonte:** Elaborado pelo autor com os dados do IF.Data do Banco Central do Brasil