



Universidade de Brasília

**MODELO DE ISING EM REDES COMPLEXAS –
CARACTERIZAÇÃO E PREDIÇÃO EM SÉRIES TEMPORAIS COM
RUÍDO**

ISRAEL DA SILVA OLIVEIRA

**DISSERTAÇÃO DE MESTRADO EM FÍSICA
INSTITUTO DE FÍSICA**



**INSTITUTO
DE FÍSICA**

DA UNIVERSIDADE DE BRASÍLIA

**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA**

**ISING MODEL IN COMPLEX NETWORKS – CHARACTERIZATION
AND PREDICTION IN NOISY TIME SERIES**

**MODELO DE ISING EM REDES COMPLEXAS –
CARACTERIZAÇÃO E PREDIÇÃO EM SÉRIES TEMPORAIS COM
RUÍDO**

ISRAEL DA SILVA OLIVEIRA

ORIENTADOR: TARCÍSIO MARCIANO DA ROCHA FILHO

DISSERTAÇÃO DE MESTRADO EM FÍSICA

PUBLICAÇÃO: PPGFIS.TD-001/25

BRASÍLIA/DF: MARÇO - 2025

**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA**

**MODELO DE ISING EM REDES COMPLEXAS –
CARACTERIZAÇÃO E PREDIÇÃO EM SÉRIES TEMPORAIS COM
RUÍDO**

ISRAEL DA SILVA OLIVEIRA

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO INSTITUTO DE FÍSICA DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

APROVADA POR:

**Prof. Dr. Tarcísio Marciano da Rocha Filho – IF/UnB
Orientador**

**Prof. Dr. Ismael Segundo da Silva Carrasco – IF/UnB
Membro Interno**

**Dr. Hernane Borges de Barros Pereira – SENAI CIMATEC
Membro Externo**

BRASÍLIA, 05 DE MARÇO DE 2025.

FICHA CATALOGRÁFICA

OLIVEIRA, ISRAEL S.

Modelo de Ising em Redes Complexas – Caracterização e Predição em Séries Temporais com Ruído [Distrito Federal] 2025.

xiii, 67p., 210 x 297 mm (IF/UnB, Mestre, Física, 2025).

Dissertação de mestrado – Universidade de Brasília, Instituto de Física.

Instituto de Física

1. Modelo de Ising

2. Aprendizado por Reforço

3. Mercado Financeiro

4. Redes Complexas

I. IF/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

OLIVEIRA, I.S. (2025). Modelo de Ising em Redes Complexas – Caracterização e Predição em Séries Temporais com Ruído. Dissertação de mestrado em Física, Publicação PPG-FIS.TD-001/25, Instituto de Física, Universidade de Brasília, Brasília, DF, 67p.

CESSÃO DE DIREITOS

AUTOR: Israel da Silva Oliveira

TÍTULO: Modelo de Ising em Redes Complexas – Caracterização e Predição em Séries Temporais com Ruído.

GRAU: Mestre ANO: 2025

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Israel da Silva Oliveira

Instituto de Física - IF

Universidade de Brasília - UnB

Campus Darcy Ribeiro

CEP 70919-970 - Brasília - DF - Brasil

*“Não hesitarei nem contemporizarei
no trabalho pela física e pela ciência
no Brasil”*

(LOPES, 1998, p. 44)

AGRADECIMENTOS

Ao longo desta jornada, contei com o apoio e a colaboração de muitas pessoas sem as quais este trabalho não teria sido possível.

Em primeiro lugar, agradeço ao meu orientador, Tarcísio Marciano, por sua orientação, paciência e dedicação ao longo desta pesquisa. Sua experiência e visão foram essenciais para o desenvolvimento deste trabalho, e sou grato pelo aprendizado e pelas discussões engrandecedoras. Também expresso minha gratidão ao meu colega de orientação e xará, Israel Garcia, cuja parceria foi substancial. As trocas de ideias, os desafios compartilhados e o apoio mútuo tornaram essa caminhada muito mais produtiva e motivadora.

Aos meus amigos, André Batista, Marcos de Paula e Pedro Silveira, que compartilham comigo não apenas a rotina acadêmica, mesmo que em diferentes frentes e sob diferentes orientações, deixo meu sincero agradecimento. A convivência, o companheirismo e as conversas dentro e fora do ambiente acadêmico foram importantes para tornar este período mais leve e significativo.

Agradeço especialmente a Gabriel Fiuza, cuja amizade transcende qualquer contexto, acadêmico ou não. Suas reflexões e conversas foram valiosas não apenas para minha formação intelectual, mas também para minha compreensão de mundo.

Por fim, à minha companheira de vida, Laura Sousa, minha maior gratidão. Seu apoio incondicional, paciência (e até mesmo a falta dela) e incentivo foram indispensáveis em cada etapa deste percurso. Seu carinho e presença tornaram tudo mais fácil de lidar e deram ainda mais significado a essa conquista.

A todos vocês, meu mais sincero obrigado.

RESUMO

Título: Modelo de Ising em Redes Complexas – Caracterização e Predição em Séries Temporais com Ruído

Autor: Israel da Silva Oliveira

Orientador: Tarcísio Marciano da Rocha Filho

Programa de Pós-Graduação em Física

Brasília, 05 de março de 2025

Este trabalho propõe a integração do modelo de Ising, amplamente utilizado em física estatística para descrever transições de fase, com aprendizado por reforço, técnica de inteligência artificial, para modelar o comportamento de séries temporais. O objetivo é explorar a dinâmica complexa desses sistemas, onde decisões individuais de agentes interagem para formar padrões coletivos. O modelo de Ising é aplicado em uma rede livre de escala, onde os *spins* são representados pelos nós e as interações entre vizinhos pelas arestas, que reflete a heterogeneidade e a conectividade assimétrica observada em diversos contextos, enquanto o aprendizado por reforço, especificamente o algoritmo *Q-learning*, é utilizado para prever tendências em séries temporais. A pesquisa desenvolve um modelo híbrido, denominado *Q-Ising*, que combina a capacidade do modelo de Ising de capturar interações locais entre agentes (representados como *spins*) com a habilidade do *Q-learning* de otimizar decisões com base em recompensas. O campo externo no modelo de Ising é influenciado pelas previsões do *Q-learning*, permitindo que o sistema simule padrões dinâmicos em séries temporais. Em particular, este trabalho aplica o modelo às séries temporais do mercado financeiro, analisando o comportamento de diferentes tipos de ativos financeiros. Os resultados indicam que o modelo *Q-Ising* supera o *Q-learning* em ativos com maior risco associado, enquanto o *Q-learning* tem melhor desempenho em ativos com melhores retornos ajustados ao risco. A análise dos parâmetros físicos do modelo de Ising, como temperatura e interação entre *spins*, sugere uma correlação com a dispersão de opiniões e a coesão entre investidores no mercado. Conclui-se que a abordagem híbrida proposta oferece uma nova perspectiva para a caracterização de séries temporais, especialmente no contexto dos mercados financeiros, com potenciais aplicações em análise de risco e tomada de decisões estratégicas.

Palavras-chave: Modelo de Ising. Aprendizado por Reforço. Mercado Financeiro. Redes Complexas.

ABSTRACT

Title: Ising Model in Complex Networks – Characterization and Prediction in Noisy Time Series

Author: Israel da Silva Oliveira

Supervisor: Tarcísio Marciano da Rocha Filho

Graduate Program in Physics

Brasília, March 05, 2025

This work proposes the integration of the Ising model, widely used in statistical physics to describe phase transitions, with reinforcement learning, an artificial intelligence technique, to model the behavior of time series. The objective is to explore the complex dynamics of these systems, where individual agent decisions interact to form collective patterns. The Ising model is applied to a scale-free network, where the spins are represented by the nodes and the interactions between neighbors by the edges, reflecting the heterogeneity and asymmetric connectivity observed in various contexts. Meanwhile, reinforcement learning, specifically the Q -learning algorithm, is employed to predict trends in time series. This research develops a hybrid model, termed Q -Ising, which combines the ability of the Ising model to capture local interactions among agents (represented as spins) with the capability of Q -learning to optimize decisions based on rewards. The external field in the Ising model is influenced by the predictions of Q -learning, allowing the system to simulate dynamic patterns in time series. In particular, this study applies the model to financial market time series, analyzing the behavior of different types of financial assets. The results indicate that the Q -Ising model outperforms Q -learning in assets with higher associated risk, whereas Q -learning performs better in assets with higher risk-adjusted returns. The analysis of the physical parameters of the Ising model, such as temperature and spin interactions, suggests a correlation with opinion dispersion and investor cohesion in the market. It is concluded that the proposed hybrid approach offers a novel perspective for characterizing time series, particularly in the context of financial markets, with potential applications in risk analysis and strategic decision-making.

Keywords: Ising Model. Reinforcement Learning. Financial Market. Complex Networks.

SUMÁRIO

1	INTRODUÇÃO	1
2	MODELOS DE <i>SPIN</i> E ESTRUTURAS COMPLEXAS	4
2.1	MODELOS DE ISING E DIFERENTES ESTRUTURAS	4
2.1.1	ESTRUTURA QUADRADA COM PARTÍCULAS ENUMERADAS	5
2.1.2	ENERGIA LIVRE E CALOR ESPECÍFICO	6
2.1.3	MAGNETIZAÇÃO	7
2.1.4	TRANSIÇÕES DE FASE E FENÔMENOS CRÍTICOS	9
2.1.5	MODELO DE ISING DE VIZINHOS PRÓXIMOS	11
2.2	REDES COMPLEXAS	11
2.2.1	REDES COMPLEXAS E TEORIA DOS GRAFOS	11
2.2.2	PROPRIEDADES DAS REDES COMPLEXAS	13
2.2.2.1	GRAU TOTAL DA REDE	13
2.2.2.2	GRAU MÉDIO DA REDE	13
2.2.2.3	DISTRIBUIÇÃO DOS GRAUS	14
2.2.3	GRAFOS ALEATÓRIOS	14
2.2.4	REDES LIVRES DE ESCALA	15
2.2.5	ALGORITMO DE BARABÁSI-ALBERT	16
2.3	INTERPRETAÇÃO DO MODELO DE ISING EM REDES COMPLEXAS	17
2.4	CONSIDERAÇÕES FINAIS	18
3	ESTRUTURA DO MERCADO FINANCEIRO E INSTRUMENTOS DE NEGOCIAÇÃO	19
3.1	ATIVOS FINANCEIROS	19
3.1.1	PREÇO DE ABERTURA, FECHAMENTO E OUTROS PARÂMETROS RELEVANTES	20
3.2	ESTRUTURA DO MERCADO FINANCEIRO	20
3.2.1	BOLSAS DE VALORES	21
3.2.2	MERCADO DE BALCÃO	21
3.2.3	LIQUIDEZ E CAPTAÇÃO DE RECURSOS	21
3.3	CONSIDERAÇÕES FINAIS	21
4	MODELAGEM BASEADA EM AGENTES E APRENDIZADO POR REFORÇO	22
4.1	MODELO DE AGENTES	22
4.2	INTRODUÇÃO AO APRENDIZADO POR REFORÇO	24

4.3	<i>Q-learning</i> : DEFINIÇÃO E FUNDAMENTOS	25
4.3.1	TABELA Q	26
4.4	APLICAÇÃO DO <i>Q-learning</i> NO PROBLEMA DO LABIRINTO: UM ESTUDO DE CASO	27
4.4.1	ENVIESAMENTO DO MODELO	29
4.5	CONSIDERAÇÕES FINAIS	30
5	IMPLEMENTAÇÃO DO MODELO DE ISING E DE APRENDIZADO POR REFORÇO	32
5.1	MODELO DE ISING EM UMA REDE LIVRE DE ESCALA	32
5.1.1	PROCEDIMENTOS DAS SIMULAÇÕES.....	32
5.1.2	ASSUNÇÃO DA CONSTANTE DE BOLTZMANN	34
5.1.3	PROPRIEDADES FÍSICAS DO MODELO DE ISING NA TOPOLOGIA COMPLEXA	34
5.1.3.1	TEMPERATURA CRÍTICA E O LIMITE TERMODINÂMICO EM FUNÇÃO DO TAMANHO DO SISTEMA	35
5.1.3.2	INFLUÊNCIA DO GRAU MÍNIMO NA CONVERGÊNCIA DO LIMITE TERMODINÂMICO	37
5.1.3.3	DIAGRAMA DE FASES	39
5.2	MODELO DE APRENDIZADO POR REFORÇO	41
5.2.1	ESTADOS, AÇÕES E RECOMPENSAS	41
5.2.1.1	ESTADOS	42
5.2.1.2	AÇÕES E RECOMPENSAS	43
5.2.2	TREINAMENTO E VALIDAÇÃO	44
5.3	SIMPLIFICAÇÕES NO MODELO DE <i>Q-learning</i>	46
5.3.1	CONFIGURAÇÃO INICIAL	47
5.3.1.1	ANÁLISE DE RISCO	49
5.3.2	RESULTADOS DAS PREVISÕES: <i>Q-learning</i>	50
5.3.2.1	ANÁLISE DOS PARÂMETROS OTIMIZADOS.....	52
5.4	CONSIDERAÇÕES FINAIS	54
6	IMPLEMENTAÇÃO DO MODELO DE ISING NO MERCADO FINANCEIRO	56
6.1	ACOPLAMENTO DO MODELO DE <i>Q-learning</i> AO MODELO DE ISING	56
6.2	PROCEDIMENTOS DAS SIMULAÇÕES	57
6.2.1	CONFIGURAÇÃO INICIAL	57
6.3	RESULTADOS DAS PREVISÕES: <i>Q-ISING</i>	58
6.3.1	ANÁLISE DOS PARÂMETROS OTIMIZADOS	59
6.3.1.1	INTERPRETAÇÃO CONJUNTA DOS PARÂMETROS	61
6.4	CONSIDERAÇÕES FINAIS	61

<i>SUMÁRIO</i>	x
7 CONCLUSÕES	63
REFERÊNCIAS BIBLIOGRÁFICAS.....	64

LISTA DE FIGURAS

2.1	Arranjo de <i>spins</i> em uma rede quadrada – Os círculos vermelhos denotam <i>spins</i> para cima (positivos) e os círculos azuis denotam <i>spins</i> para baixo (negativos).....	5
2.2	Diagrama de Fases do Modelo de Ising 2D – Gerado computacionalmente	10
2.3	Representação da diferença entre grafos e redes.....	12
2.4	Distribuição de probabilidade dos graus do modelo de Erdős-Rényi	15
2.5	Distribuição de probabilidade dos graus de uma rede livre de escala – Gerado computacionalmente com os seguintes parâmetros: $n = 10^4$, $m_0 = 11$ e $m = 10$	16
2.6	Representação de diferentes modelos com graus indicados nos vértices	16
2.7	Evolução do modelo de Barabási-Albert	17
4.1	Esquema ilustrativo do labirinto com objetivo em $s_M = (4, 4)$ e ponto de partida na célula $(0, 0)$	28
4.2	Trajetória percorrida por agente aleatório, saindo do ponto inicial até o objetivo	28
4.3	Trajetória percorrida por agente treinado, saindo do ponto inicial até o objetivo	29
4.4	Representações visuais do processo do agente enviesado	30
5.1	Variação das quantidades físicas (na ordem: Energia, Magnetização Média, Calor Específico e Suscetibilidade) com a temperatura	35
5.2	Variação da temperatura crítica em função do número de partículas	36
5.3	Variação da temperatura crítica em função do grau mínimo – Valores dos parâmetros ajustados: $a = -2.91 \times 10^{-6}$; $b = -1.66 \times 10^{-3}$; $c = 1.56$; e $d = 9.43$	38
5.4	Diagrama de fases do modelo de Ising em uma rede livre de escala.....	40
5.5	Evolução temporal do preço de fechamento	48
5.6	Simulação do modelo de investimentos com <i>Q-learning</i> para o ativo da Apple Inc. (AAPL) com reescrita da memória – Caso saturado.....	51
5.7	Simulação do modelo de investimentos com <i>Q-learning</i> para o ativo da Apple Inc. (AAPL) – Comparação dos modelos ACM e ARM	51
5.8	Simulação do modelo de investimentos com <i>Q-learning</i> para o Índice Bovespa (\hat{BVSP}) – Comparação dos modelos ACM e ARM	52
5.9	Simulação do modelo de investimentos com <i>Q-learning</i> para o ativo da Vale S.A. (VALE3.SA) – Comparação dos modelos ACM e ARM.....	53
6.1	Simulação do modelo de investimentos com <i>Q-Ising</i> para o ativo da Apple Inc. (AAPL) – Comparação com modelo <i>Q-learning</i>	58

6.2	Simulação do modelo de investimentos com Q -Ising para o ativo do Índice Bovespa (\hat{BVSP}) – Comparação com modelo Q -learning	59
6.3	Simulação do modelo de investimentos com Q -Ising para o ativo da Vale S.A. (VALE3.SA) – Comparação com modelo Q -learning	60

LISTA DE TABELAS

2.1	Comparação entre os termos usados na Ciência de Redes e na Teoria de Grafos	12
4.1	Tabela Q do Q -learning	26
5.1	Parâmetros da Rede e do Modelo de Ising	33
5.2	Índice de Sharpe (\mathcal{S}) e Retorno Anualizado (\mathcal{R}) dos ativos analisados	49
5.3	Eficiência (η) dos modelos ARM e ACM em cada ativo	52
5.4	Parâmetros de aprendizado do Q -learning otimizados (ACM).....	53
6.1	Eficiência (η) dos modelos Q -learning e Q -Ising em cada ativo em porcentagem	58
6.2	Parâmetros do modelo de Ising otimizados	61

1 INTRODUÇÃO

O estudo de sistemas complexos constitui uma das maiores conquistas científicas do século XXI, atravessando disciplinas que vão da biologia à física, da sociologia à economia. Nestes sistemas, séries temporais se destacam pela dinamicidade e imprevisibilidade, nos quais decisões individuais se entrelaçam e formam comportamentos coletivos intrinsecamente não lineares. Apesar de décadas de pesquisa, modelar adequadamente essas interações e prever os padrões emergentes dos sistemas complexos continua sendo um desafio fundamental.

Neste contexto, o presente trabalho propõe combinar o modelo de Ising, reconhecido por sua capacidade de descrever transições de fases em sistemas físicos, com o aprendizado por reforço, ramo da inteligência artificial que permite a adaptação e otimização em ambientes dinâmicos. A integração desses dois métodos visa explorar novas formas de interpretar e prever o comportamento de séries temporais ruidosas de sistemas reais, em especial, dos preços de ativos financeiros. Além de investigar as aplicações tradicionais dos modelos de física estatística no domínio econômico.

Para investigar o comportamento emergente do mercado financeiro, sugerimos que o modelo de Ising seja estruturado em uma rede complexa livre de escala. Este tipo de rede é conhecida por representar hierarquias análogas a redes reais, como redes sociais, a internet e sistemas biológicos e sociais. A interação entre agentes simulada por uma rede livre de escala permite heterogeneização das influências de seus vizinhos, refletindo a complexidade das interações no mercado financeiro. Ao longo deste trabalho, buscamos desenvolver um modelo mínimo como prova de conceito, que servirá como base para futuras melhorias, especialmente no que diz respeito à implementação do modelo de aprendizado de máquina.

Diversos trabalhos aplicam o modelo de Ising para o estudo de ativos financeiros (CIVIDINO; WESTPHAL; SORNETTE, 2023; GUIMARAES; LIMA, 2021; TAKAISHI, 2017; SORNETTE, 2005; HUANG; SHETH, 2021; SORNETTE, 2021; KRAWIECKI, 2009). Cividino, Westphal e Sornette (2023) criam diferentes modelos com tipos diferentes de agentes interagindo, sendo um desses tipos baseados em Ising. As séries temporais analisadas neste trabalho são produzidas pela interação entre a oferta e demanda dos diferentes tipos de agentes. Já Krawiecki (2009) usa somente um tipo agente, baseados em Ising, mas com as conexões feitas seguindo modelos de redes livres de escala. As séries temporais deste trabalho também são simuladas pelo modelo, e não reais. Takaishi (2017) também simula um mercado artificial, porém utiliza o campo externo para influenciar as partículas a comprarem ou venderem o ativo fictício, semelhante ao método que estamos propondo. Os demais trabalhos citados seguem o mesmo tipo de metodologia, variando aspectos como tipos de

agentes, comportamentos específicos e variáveis produzidas pelo modelo, como retorno financeiro ou preço do ativo. No entanto, poucos exploram essa abordagem utilizando séries temporais autênticas. Bury (2013) utiliza ativos reais como partículas no modelo de Ising, enquanto o *spin* indica se o ativo está em alta ou em baixa. Este é o método mais recorrente no uso do modelo de Ising em dados reais. A escassez de pesquisas que integram o modelo de Ising com dados reais de mercado motiva a proposta deste trabalho, que visa realizar essa integração com séries temporais dos preços dos ativos. De maneira que em um dia t , a magnetização indique a opinião média das partículas sobre comprar ou vender o ativo, visando a obtenção de lucros no dia $t + 1$. A integração pode ser feita de várias maneiras, mas a que preserva a implementação do modelo de Ising tradicional é por meio do uso do campo externo que influencia o alinhamento dos *spins*. Considerando que cada *spin* assume valores ± 1 , podemos definir que $+1$ representa a aquisição de um ativo financeiro, enquanto -1 corresponde à venda do ativo. O campo externo, portanto, terá o papel de orientar as partículas a comprar ou vender o ativo, ajustando o comportamento do modelo para simular as dinâmicas reais do mercado. Chamaremos esta combinação do modelo de Ising com aprendizado por reforço de Q -Ising. Para prever o comportamento futuro dos ativos, o campo externo, no dia t , deve indicar a tendência de crescimento do preço no dia $t + 1$. Para isso, empregamos o aprendizado por reforço para determinar, no dia t , o campo externo que melhor representa o crescimento do preço no dia seguinte.

Quanto ao uso da teoria de redes complexas, conforme Krawiecki (2009), os retornos logarítmicos promovidos pela simulação utilizando o modelo de Ising em redes livres de escala podem ser comparáveis aos observados empiricamente. Nesse sentido, optamos por utilizar essa topologia, assumindo que proporcionará interações entre as partículas que são análogas às observadas no mundo real, o que, por sua vez, deve gerar resultados condizentes com as dinâmicas financeiras reais.

O objetivo principal deste trabalho é desenvolver o modelo Q -Ising, que combina o modelo de Ising em uma rede livre de escala com aprendizado por reforço. Para isso, inicialmente, construímos o modelo de Ising nessa topologia para analisarmos suas propriedades físicas, visando compreender seu comportamento sob a influência de um campo externo. Paralelamente, desenvolve-se um modelo de aprendizado por reforço capaz de prever com precisão o comportamento de séries temporais, utilizando informações disponíveis até o instante anterior ao evento predito. Além disso, investiga-se em que medida o modelo Q -Ising pode descrever séries temporais reais e quais aspectos físicos do modelo de Ising, como temperatura e energia de interação, possuem relevância no contexto mercadológico. Este estudo busca avaliar se a integração entre o modelo de Ising e o aprendizado por reforço aprimora a identificação de padrões não lineares e interações complexas em séries temporais, superando as limitações dos modelos isolados. Assim, examina-se a possibilidade de utilizar o modelo Q -Ising para caracterizar séries temporais e auxiliar na quantificação e previsão de riscos de

investimento no contexto financeiro. Por fim, investiga-se a hipótese de que essa abordagem permite a utilização do modelo de Ising para a predição de séries temporais reais, oferecendo uma alternativa aos modelos já conhecidos para a modelagem de séries temporais ruidosas.

2 MODELOS DE *SPIN* E ESTRUTURAS COMPLEXAS

Este capítulo introduz o modelo de Ising, uma ferramenta amplamente utilizada para descrever fenômenos complexos, com aplicações que vão da física (ISING, 1925) e biologia (TORQUATO, 2011) à modelagem de mercados financeiros (SORNETTE, 2014). O modelo, que originalmente surgiu no contexto da física do estado sólido para estudar a magnetização de materiais, oferece estrutura matemática que pode ser aplicada ao comportamento dos investidores em mercados financeiros. Ao considerar o mercado como um sistema de partículas interagindo, onde cada agente representa uma decisão de compra ou venda, o modelo de Ising nos permite explorar como as interações locais influenciam o comportamento coletivo. A Seção 2.1 se dedicará a análise formal das equações que regem o modelo, a discussão das suas implicações e como essas dinâmicas podem ser adaptadas e integradas a dados financeiros reais.

Conforme a discussão sobre o modelo de Ising, este capítulo também explora o conceito de redes complexas, na Seção 2.2, com foco particular em redes livres de escala e sua relevância para a modelagem de séries temporais ruidosas. Nas redes livres de escala, a distribuição de conectividade entre os nós é altamente heterogênea, refletindo a estrutura dos mercados financeiros, onde algumas instituições e grandes investidores possuem maior relevância, conectando-se a muitos agentes, enquanto a maioria dos investidores possui menos conexões. A estrutura dessas redes influencia diretamente o comportamento coletivo no mercado, uma vez que choques em agentes centrais podem ter impactos significativos, disseminando informações ou comportamentos de forma rápida. Por fim, a Seção 2.3 descreve como a analogia entre o sistema de *spins* e a teoria de redes é feita. Este capítulo, portanto, introduz as equações e dinâmicas do modelo de Ising, e também oferece análise de grafos e redes complexas, preparando a base para implementação formal do modelo de Ising em estruturas livres de escala nos capítulos subsequentes.

2.1 MODELOS DE ISING E DIFERENTES ESTRUTURAS

O modelo de Ising representa uma rede de partículas (ordenadas espacialmente) que possuem *spin* e são interagentes, influenciando e sendo influenciadas por partículas próximas (vizinhança). O modelo foi explorado pela primeira vez para investigar a magnetização espontânea em filmes ferromagnéticos (ISING, 1925; SINGH, 2020). Atualmente, encontram-se bastante pesquisas que buscam utilizar-se do modelo para predição e caracterização de

séries temporais de preços de ativos financeiros, como os trabalhos de Guimaraes e Lima (2021), Sornette (2005), Huang e Sheth (2021) e Sornette (2021). O método costuma partir do simples: da magnetização. Isto é, se a magnetização é positiva em um tempo t , o modelo indica comprar o ativo para obter lucros em $t + 1$, e vice-versa.

Discutiremos a seguir dois modelos de Ising para que, futuramente, a implementação do modelo de investimento seja feita em cima destas explicações.

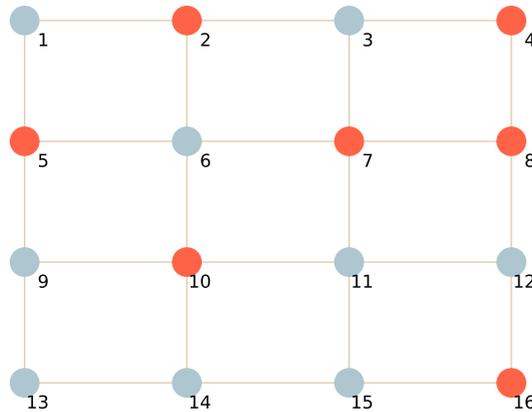
2.1.1 Estrutura quadrada com partículas enumeradas

Segundo Baxter (2016), podemos considerar cada partícula como um ímã microscópico, que aponta ao longo de algum eixo preferencial ou aponta exatamente na direção oposta. Assim, dizemos que cada partícula i tem *spin* σ_i , em que é dito “para cima” quando $\sigma_i = +1$, e “para baixo” quando $\sigma_i = -1$. A indicação de que a estrutura é de partículas enumeradas refere-se a não considerarmos os vizinhos próximos, apenas a ordenação das partículas pelos seus *spins* σ . Por exemplo, na Figura 2.1 podemos ver um arranjo de *spins*

$$\sigma = \{-, +, -, +, +, -, +, +, -, +, -, -, -, -, -, +\}$$

em uma estrutura como a descrita.

Figura 2.1 – Arranjo de *spins* em uma rede quadrada – Os círculos vermelhos denotam *spins* para cima (positivos) e os círculos azuis denotam *spins* para baixo (negativos)



Fonte: Elaboração própria.

Tomemos $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ como o conjunto de todos os n *spins* de um sistema, onde cada uma das 2^n configurações possíveis caracteriza o estado do sistema. O Hamiltoniano é agora uma função $\mathcal{H}(\sigma)$, composta de duas partes: $\mathcal{H}(\sigma) = \mathcal{H}_0(\sigma) + \mathcal{H}_1(\sigma)$, em que \mathcal{H}_0 é a contribuição das forças intermoleculares dentro do ímã e $\mathcal{H}_1(\sigma)$ é a contribuição das interações entre os ímãs (BAXTER, 2016).

Como σ_i é o momento magnético da partícula i , $\mathcal{H}_1(\sigma)$ pode ser escrito como

$$\mathcal{H}_1(\sigma) = -h \sum_i^n \sigma_i, \quad (2.1)$$

onde h é proporcional ao componente do campo na direção do eixo preferencial. Em um sistema físico, esperamos que as interações sejam invariantes sob a inversão do tempo, o que significa que \mathcal{H} não se altera com a inversão de todos os campos e magnetizações, ou seja, com a inversão de h e cada σ_i . Segue-se que \mathcal{H}_0 deve ser função par de σ

$$\mathcal{H}_0(\sigma_1, \sigma_2, \dots, \sigma_n) = \mathcal{H}_0(-\sigma_1, -\sigma_2, \dots, -\sigma_n). \quad (2.2)$$

Para derivarmos as quantidades termodinâmicas deste sistema, devemos considerar a função partição \mathcal{Z} , dada pela soma das probabilidades relativas de cada estado do sistema, ponderada pela sua energia em relação à temperatura (fatores de Boltzmann):

$$\mathcal{Z}_n(h, T) = \sum_{\{\sigma\}} e^{-\frac{\mathcal{H}}{k_B T}} = \sum_{\{\sigma\}} e^{-\frac{\mathcal{H}_0(\sigma) - h \sum_i \sigma_i}{k_B T}}, \quad (2.3)$$

em que k_B é a constante de Boltzmann, T representa a temperatura, e as chaves na soma sobre σ denotam o conjunto de *spins* considerados.

Além disso, é necessário definir a energia livre para caracterizar completamente o modelo, através das derivadas termodinâmicas calculadas nas Subseções seguintes. Assim, definimos a energia livre do sistema como

$$F(h, T) := -k_B T \ln(\mathcal{Z}_n(h, T)) = \mathcal{H}. \quad (2.4)$$

2.1.2 Energia livre e calor específico

Fisicamente, esperamos que a energia livre de um grande sistema seja proporcional ao tamanho do sistema em termos do número de partículas, ou seja, esperamos que exista o limite termodinâmico

$$\lim_{n \rightarrow \infty} f(h, T) = -k_B T \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathcal{Z}_n(h, T), \quad (2.5)$$

em que f é a energia livre por partícula.

Entendendo que a energia interna do sistema U é o valor esperado de \mathcal{H} , escrevemos

$$U = \langle \mathcal{H} \rangle = \mathcal{Z}_n^{-1} \sum_{\{\sigma\}} \mathcal{H} e^{-\frac{\mathcal{H}}{k_B T}}, \quad (2.6)$$

justamente por $e^{-\frac{\mathcal{H}}{k_B T}}$ ser proporcional a probabilidade de se encontrar o estado com energia \mathcal{H} . Logo, a formulação acima descreve a soma das energias ponderada pelas suas probabilidades de ocorrência, portanto: o valor esperado da energia. Usando a Equação 2.3 e a Equação 2.4, formulamos

$$U(h, T) = k_B T^2 \frac{\partial}{\partial T} \ln \mathcal{Z}_n \quad (2.7)$$

$$= -T^2 \frac{\partial}{\partial T} \frac{F(h, T)}{T}. \quad (2.8)$$

Por semelhança, a energia interna por partícula é

$$u(h, T) = -T^2 \frac{\partial}{\partial T} \frac{f(h, T)}{T}. \quad (2.9)$$

O calor específico por partícula é definido como

$$c(h, T) = \frac{\partial}{\partial T} u(h, T), \quad (2.10)$$

i.e., o calor específico em relação à temperatura é a taxa de variação da energia média por partícula com a temperatura.

2.1.3 Magnetização

Retomemos a premissa do modelo de Ising aplicado ao mercado financeiro: os *spins* indicam a intenção de compra e venda. Como a magnetização de um sistema representa o grau de alinhamento dos *spins* na rede, também indicará a coesão das opiniões do mercado. Em termos matemáticos, a magnetização $M(h, T)$ é dada por:

$$M(h, T) = \frac{1}{n} \sum_{i=1}^n \langle \sigma_i \rangle, \quad (2.11)$$

onde σ_i é o *spin* da partícula i (assumido como $+1$ ou -1), n é o número total de partículas, e $\langle \sigma_i \rangle$ é o valor esperado do *spin* σ_i , calculado considerando a distribuição de probabilidades do sistema em equilíbrio. Em outras palavras, a magnetização é a média aritmética dos valores dos *spins* em função da temperatura T e do campo externo h .

Em termos de função partição, \mathcal{Z}_n , podemos expressar a magnetização como:

$$M(h, T) = \frac{1}{n} \mathcal{Z}_n^{-1} \sum_{\{\sigma\}} \left(\sum_{i=1}^n \sigma_i \right) e^{-\frac{\mathcal{H}(\sigma)}{k_B T}}. \quad (2.12)$$

Essa expressão indica que a magnetização é a média ponderada dos *spins*, pela probabilidade

de cada configuração $\{\sigma\}$ dada pela distribuição de Boltzmann.

Além disso, a magnetização pode ser obtida a partir da energia livre $F(h, T)$ do sistema, como:

$$M(h, T) = -\frac{\partial}{\partial h} F(h, T). \quad (2.13)$$

Esse resultado surge da relação entre a magnetização e o comportamento termodinâmico do sistema.

Prova que $-\frac{\partial}{\partial h} f(h, T) = n^{-1} \mathcal{Z}_n^{-1} \sum_{\{\sigma\}} (\sigma_1 + \sigma_2 + \dots + \sigma_n) e^{-\frac{\mathcal{H}}{k_B T}}$:

$$\begin{aligned} -\frac{\partial f}{\partial h} &= \frac{k_B T}{n} \frac{\partial}{\partial h} \ln \mathcal{Z}_n \\ &= \frac{k_B T}{n} \frac{1}{\mathcal{Z}_n} \sum_{\{\sigma\}} \frac{\partial}{\partial h} e^{-\frac{\mathcal{H}}{k_B T}} \\ &= \frac{k_B T}{n} \frac{1}{\mathcal{Z}_n} \sum_{\{\sigma\}} e^{-\frac{\mathcal{H}}{k_B T}} \left(-\frac{1}{k_B T}\right) \frac{\partial \mathcal{H}}{\partial h} \\ &= \frac{1}{n \mathcal{Z}_n} \sum_{\{\sigma\}} e^{-\frac{\mathcal{H}}{k_B T}} \sum_i \sigma_i. \end{aligned}$$

Segue-se daí que $-\frac{\partial f}{\partial h} = (n \mathcal{Z}_n)^{-1} \sum_{\{\sigma\}} \sum_i \sigma_i e^{-\frac{\mathcal{H}}{k_B T}}$. ■

Aqui podemos verificar que $-1 \leq M(h, T) \leq +1$ e que, como a função partição é função par de h e de σ e a energia livre por partícula é função ímpar de h , então M é função ímpar de h :

$$M(-h, T) = -M(h, T). \quad (2.14)$$

Por fim, definimos a suscetibilidade como a variação da magnetização em relação à h

$$\chi = \frac{\partial M}{\partial h} = \frac{1}{n k_B T} \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2, \text{ onde } \mathcal{M} = \sum_i \sigma_i. \quad (2.15)$$

Prova que $\frac{\partial M}{\partial h} = \frac{1}{n k_B T} \{ \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2 \}$:

Primeiro, derivando a Equação 2.12:

$$\begin{aligned}\frac{\partial M}{\partial h} &= n^{-1} \frac{\partial \mathcal{Z}_n^{-1}}{\partial h} \sum_{\{\sigma\}} \sum_i \sigma_i \exp\{-\mathcal{H}/k_B T\} + \frac{1}{n \mathcal{Z}_n} \sum_{\{\sigma\}} \sum_i \sigma_i \frac{\partial}{\partial h} \exp\{-\mathcal{H}/k_B T\} \\ &= \frac{1}{nk_B T \mathcal{Z}_n^2} \left(\sum_{\{\sigma\}} \sum_i \sigma_i \exp\{-\mathcal{H}/k_B T\} \right)^2 + \frac{1}{nk_B T \mathcal{Z}_n} \sum_{\{\sigma\}} \left(\sum_i \sigma_i \right)^2 \exp\{-\mathcal{H}/k_B T\}.\end{aligned}$$

Agora, lembremos que $\langle x \rangle = \mathcal{Z}_n^{-1} \sum x \exp -\mathcal{H}/k_B T$, assim, observamos que a equação acima pode ser reescrita como:

$$\frac{\partial M}{\partial h} = \frac{1}{nk_B T} \left[\left\langle \left(\sum_i \sigma_i \right)^2 \right\rangle - \left\langle \sum_i \sigma_i \right\rangle^2 \right]. \quad (2.16)$$

Segue-se daí que $\frac{\partial M}{\partial h} = \frac{1}{nk_B T} \{ \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2 \}$, com $\mathcal{M} = \sum_i \sigma_i$. ■

Usando a seguinte propriedade $\langle [\mathcal{M} - \langle \mathcal{M} \rangle]^2 \rangle = \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2$, reescrevemos a Equação 2.15

$$\chi = \frac{1}{nk_B T} \langle [\mathcal{M} - \langle \mathcal{M} \rangle]^2 \rangle. \quad (2.17)$$

Como χ é a média de uma quantidade não negativa: $\chi = \partial M / \partial h \geq 0$. A magnetização M é, portanto, função monotônica crescente ímpar de h , situada no intervalo $-1 \leq M(h, T) \leq +1$.

Prova que $\langle [\mathcal{M} - \langle \mathcal{M} \rangle]^2 \rangle = \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2$:

$$\langle [\mathcal{M} - \langle \mathcal{M} \rangle]^2 \rangle = \langle \mathcal{M}^2 - 2\mathcal{M}\langle \mathcal{M} \rangle + \langle \mathcal{M} \rangle^2 \rangle.$$

Como a média de uma constante é a própria constante $\langle \langle \mathcal{M} \rangle \rangle = \langle \mathcal{M} \rangle$, então:

$$\begin{aligned}\langle [\mathcal{M} - \langle \mathcal{M} \rangle]^2 \rangle &= \langle \mathcal{M}^2 \rangle - 2\langle \mathcal{M} \rangle \langle \mathcal{M} \rangle + \langle \mathcal{M} \rangle^2 \\ &= \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2.\end{aligned}$$

Segue-se daí que $\langle [\mathcal{M} - \langle \mathcal{M} \rangle]^2 \rangle = \langle \mathcal{M}^2 \rangle - \langle \mathcal{M} \rangle^2$. ■

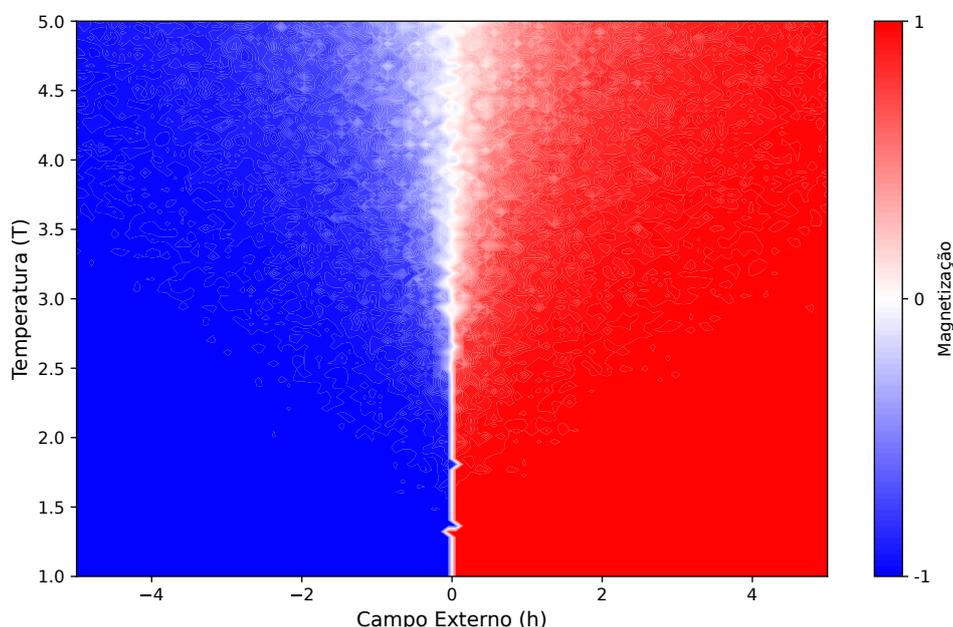
2.1.4 Transições de fase e fenômenos críticos

Devido à sua magnetização, o modelo de Ising exhibe transições de fase, que, no contexto de opinião de compra e venda, correspondem a mudanças abruptas no consenso do mercado.

Para identificar essas transições, apresentamos as derivadas termodinâmicas do modelo por fornecerem informações sobre seus pontos críticos.

Isto é, transições de fase ocorrem em diferentes sistemas como fluidos e materiais magnéticos (como no modelo de Ising). Diversas grandezas termodinâmicas (os calores específicos, a compressibilidade ou a suscetibilidade magnética) apresentam comportamento peculiar na região crítica (SALINAS, 1999, p. 291). Ou seja, nas vizinhanças do ponto crítico, determinadas derivadas termodinâmicas, como a suscetibilidade ou o calor específico, podem apresentar comportamento singular ou anômalo, caracterizando estados críticos da matéria. O estado crítico ocorre quando duas ou mais fases da matéria tornam-se indistinguíveis. A temperatura crítica, correspondente ao ponto crítico, marca essa transição de fase.

Figura 2.2 – Diagrama de Fases do Modelo de Ising 2D – Gerado computacionalmente



Fonte: Elaboração própria.

A Figura 2.2 exibe o diagrama de fases do modelo de Ising. No eixo horizontal, está representada a variação do campo externo h , enquanto o eixo vertical mostra a variação da temperatura T . A coloração da imagem reflete a magnetização média, com a escala cromática indicada na lateral direita. Observa-se uma transição de fase clara em $h = 0$, sendo o ponto crítico em $T \approx 2.25$ ¹. As áreas em vermelho representam a fase ferromagnética ($M > 0$), enquanto as regiões em azul correspondem à fase antiferromagnética ($M < 0$). A região branca, centralizada em torno de $h = 0$, caracteriza a fase paramagnética ($M \approx 0$). Note que, por mais que o diagrama apresente flutuações, estas se mostram homogêneas em torno do eixo $h = 0$.

¹Como a determinação da temperatura crítica foi realizada computacionalmente, assumimos a constante de Boltzmann como $k_b = 1$ e usamos $J = 1$ (ambos em unidades adimensionais), o que torna $k_b T / J = T$.

2.1.5 Modelo de Ising de vizinhos próximos

A subseção anterior foi colocada para introduzir o modelo mais geral de Ising, de modo que, agora, torna-se natural discutirmos as possibilidades de alteração da estrutura do modelo. Para introduzirmos, por exemplo, a interação de vizinhos próximos (aqueles imediatamente espacialmente ao lado da partícula i), basta alterar a hamiltoniana \mathcal{H}_0 , responsável por contabilizar as forças intermoleculares do sistema:

$$\mathcal{H}_0 = - \sum_{i,j \in \mathcal{V}_i} J_{ij} \sigma_i \sigma_j. \quad (2.18)$$

Na hamiltoniana acima, a soma é feita sobre i e sobre todas as partículas j que são vizinhas próximas de i ($j \in \mathcal{V}_i$). J_{ij} é a energia de interação entre os vizinhos e, em geral, $J_{ij} \neq J_{ji}$. A maneira como preferimos escrever este somatório é relevante para que este modelo não tenha uma estrutura específica.

2.2 REDES COMPLEXAS

Podemos entender uma rede como a topologia das conexões feitas em um sistema. Isto é, como cada parte se conecta umas nas outras. O estudo de topologias complexas é recente no contexto da pesquisa acadêmica. Dorogovtsev (2010, p. 19) diz que no final da década de 1990 o único modelo de referência amplamente utilizado era um grafo aleatório clássico. Foi somente nesta época que o planeta obteve escala o suficiente para o surgimento de falhas significativas nas redes reais. Um trabalho revolucionário que fomentou as pesquisas em redes complexas foi o Diâmetro da internet (tradução livre de “Diameter of the web”) publicado por Albert, Jeong e Barabási (1999).

Para Strogatz (2001, p. 268), estudamos redes complexas porque a estrutura sempre afeta a função. Isto significa que, a forma que aplicamos a estrutura das conexões feitas em um sistema altera fundamentalmente seu funcionamento. Desta forma, é preciso conhecer a topologia das ligações entre as partes do sistema em questão para podermos modelá-la ou, ao menos, nos aproximarmos da lógica de como estas são feitas. Além disso, entender a estrutura nos auxiliará a interpretar os resultados advindos da implementação do modelo de Ising.

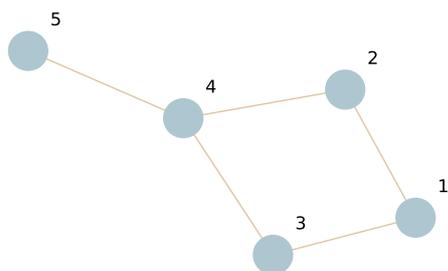
2.2.1 Redes complexas e teoria dos grafos

Os grafos são assim chamados porque podem ser representados graficamente, e é essa representação gráfica que nos ajuda a compreender muitas de suas propriedades. Cada vértice é indicado por um ponto, e cada aresta por uma linha que

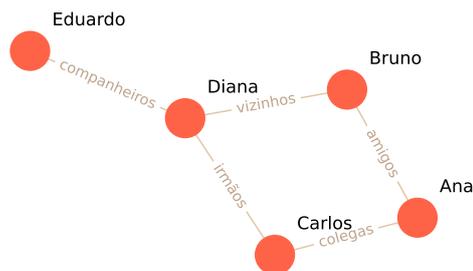
une os pontos, representando seus fins (BONDY; MURTY, 2008, p. 2, tradução nossa).

Figura 2.3 – Representação da diferença entre grafos e redes

(a) Grafo



(b) Rede Social representada em grafo



Fonte: Elaboração própria.

Barabási (2016, p. 4) discorre que a diferença entre grafos e redes está na abordagem da aplicação. Entende-se a rede como a aplicação do grafo: quem ou o que são os vértices e como as arestas realizam a interação entre os vértices. Assim, diferentes redes podem ter um mesmo grafo, enquanto a estrutura for a mesma, e alterando apenas a interpretação dos vértices e das arestas. Na Figura 2.3a encontramos a representação de um grafo dada por vértices e arestas. Já na Figura 2.3b temos o mesmo grafo, contudo aplicado a uma rede social.

Para a diferença ser mais direta, convencionam-se diferentes nomenclaturas para cada caso, especificando-as na Tabela 2.1.

Tabela 2.1 – Comparação entre os termos usados na Ciência de Redes e na Teoria de Grafos

Ciência de redes	Teoria de Grafos
Rede	Grafo
Nó	Vértice
Ligação	Aresta

Fonte: Adaptado de Barabási (2016, p. 6)

2.2.2 Propriedades das redes complexas

O grau do i -ésimo nó é definido como número de suas ligações k_i . Em uma rede direcionada, um nó pode ter dois tipos de graus: de entrada k_i^{in} e de saída k_i^{out} . Os sobrescritos “in” e “out” referem-se a entrada e saída. As ligações de entrada são aquelas que chegam ao nó, enquanto as de saída são as que partem do nó. Desta forma, o grau do nó será $k_i = k_i^{in} + k_i^{out}$. Em redes não direcionadas não se faz a distinção das ligações, e o grau do i -ésimo nó é a soma de todas as suas ligações k_i .

2.2.2.1 Grau total da rede

O grau total da rede direcionada é determinado como a soma de todas as ligações:

$$L = \sum_i^n k_i^{in} = \sum_i^n k_i^{out}. \quad (2.19)$$

Enquanto em uma rede não direcionada:

$$L = \frac{1}{2} \sum_i^n k_i. \quad (2.20)$$

O fator $\frac{1}{2}$ ocorre porque, ao contabilizarmos a ligação de um determinado nó, estamos simultaneamente contando a mesma ligação quando consideramos o grau do nó conectado a ele.

2.2.2.2 Grau médio da rede

O grau médio é uma característica definidora da rede. Com esta informação, temos o número médio das ligações feitas por cada nó. Este é definido como a média dos graus de cada nó (rede não direcionada):

$$\langle k \rangle = \frac{1}{n} \sum_i^n k_i = \frac{2L}{n}. \quad (2.21)$$

Analogamente, para uma rede direcionada:

$$\langle k^{in} \rangle = \frac{1}{n} \sum_i^n k_i^{in} = \langle k^{out} \rangle = \frac{1}{n} \sum_i^n k_i^{out} = \frac{L}{n}. \quad (2.22)$$

²Esta relação é válida para redes direcionadas simples, nas quais se admite no máximo uma aresta em cada direção entre dois vértices distintos.

Se o número de ligações que cada nó faz (grau de saída) é fixo para todos os nós e equivale a m , isso implica que $m = L/n$, assim, $\langle k \rangle = 2m$.

2.2.2.3 Distribuição dos graus

Considerando que p_k seja a probabilidade de se obter um nó com grau k em uma rede, temos a condição de normalização

$$\sum_k^{\infty} p_k = 1. \quad (2.23)$$

E para uma rede com n nós, p_k é dado por

$$p_k = \frac{n_k}{n}, \quad (2.24)$$

em que n_k é o número de nós com grau k . Dessa forma, o grau médio da rede pode ser escrito como

$$\langle k \rangle = \sum_k^{\infty} k p_k. \quad (2.25)$$

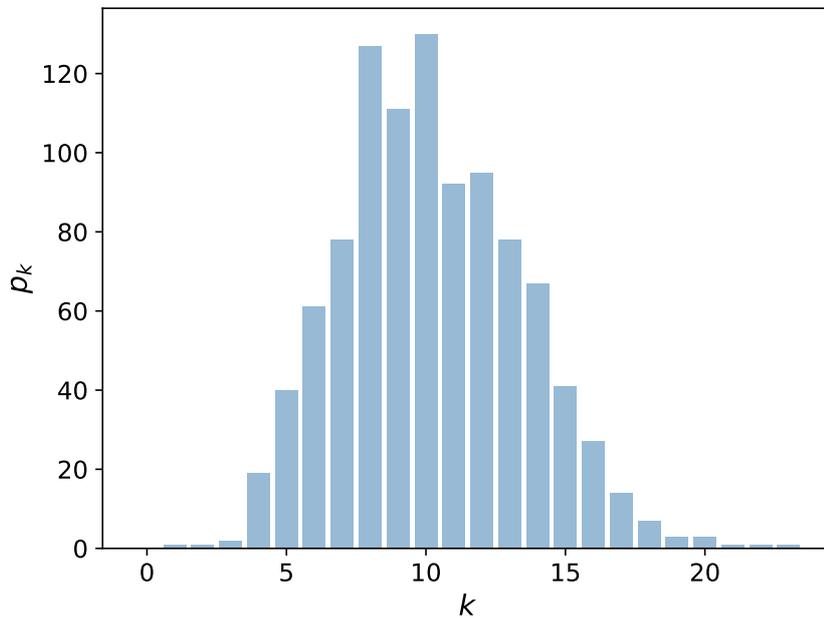
A forma da distribuição de probabilidade p_k depende de como as ligações são distribuídas na rede. Veremos a seguir para os grafos aleatórios e para os livres de escala.

2.2.3 Grafos aleatórios

Segundo Hofstad (2013), o grafo aleatório de Erdős-Rényi é o mais simples possível: assumimos que todas as arestas são possíveis dentre uma coleção de n vértices independentes com igual probabilidade. Os grafos aleatórios são simplificações da realidade, pois não é possível afirmar com certeza que eles representam precisamente qualquer rede do mundo real de maneira quantitativa (HOFSTAD, 2013).

Isso ocorre porque, no grafo, as arestas entre os vértices têm igual probabilidade. Não há distribuição de pesos para os vértices, resultando em uma distribuição de graus proporcional à distribuição de Poisson quando n tende ao infinito, como mostra a Figura 2.4. Este não é o caso de redes reais. Entretanto, essa característica não é observada em diversas redes reais. Em alguns desses sistemas, surge uma classe de vértices denominada *hubs*, caracterizada por possuir graus significativamente superiores à média da rede. Para sanar essa inadequação dos grafos aleatórios, foram desenvolvidos modelos que tenham as características de redes livres de escala.

Figura 2.4 – Distribuição de probabilidade dos graus do modelo de Erdős-Rényi



Fonte: Elaboração própria.

2.2.4 Redes livres de escala

A principal diferença entre o modelo de Erdős-Rényi e o modelo de Barabási-Albert está na distribuição de graus. No segundo caso, a distribuição p_k é dada por uma lei de potências

$$p_k \sim k^{-\gamma}, \quad (2.26)$$

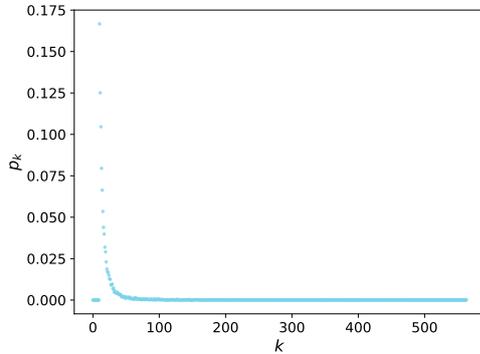
em que γ é dito expoente do grau. Esta distribuição faz com que a maioria dos nós tenha poucas conexões comparados aos *hubs*, que concentram ligações. A Figura 2.5a demonstra a distribuição da Equação 2.26. Perceba que os graus com maior probabilidade de ocorrência são os menores. Quanto maior o grau, mais a probabilidade tende a zero.

Aplicando a função logarítmica na Equação 2.26 obtemos $\log(p_k) \sim -\gamma \log(k)$, a qual é função linear. O gráfico desse resultado pode ser observado na Figura 2.5b. A inclinação da curva é precisamente $-\gamma$ e os *hubs* podem ser caracterizados pela última linha de pontos com menor probabilidade.

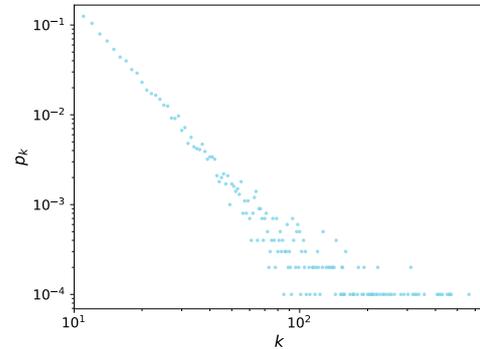
Diante disto, em um grafo aleatório, a maioria dos vértices tem um número médio de conexões, veja na Figura 2.6a. No entanto, em uma rede livre de escala, a distribuição de graus é muito mais variada, e não há valor característico que determine essa distribuição. Em vez disso, a distribuição de graus segue uma lei de potência, sem uma escala definida que distinga o grau dos nós do grau médio, caracterizando, assim, as redes como livres de escala (em tradução livre de “Scale-Free Network”). A Figura 2.6b mostra que, mesmo para

Figura 2.5 – Distribuição de probabilidade dos graus de uma rede livre de escala – Gerado computacionalmente com os seguintes parâmetros: $n = 10^4$, $m_0 = 11$ e $m = 10$

(a) Escala linear



(b) Escala log-log

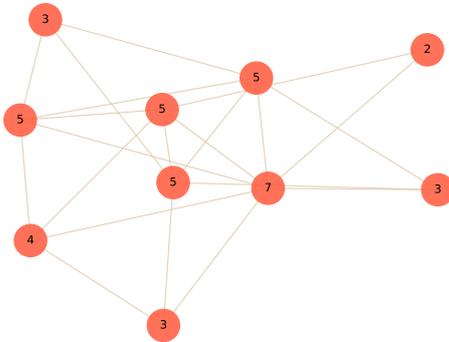


Fonte: Elaboração própria.

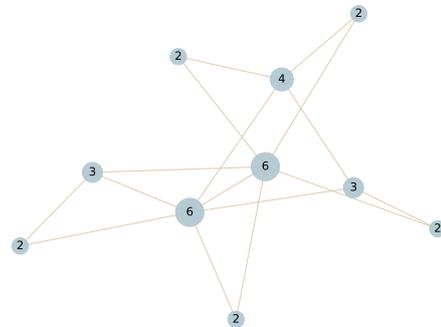
valores pequenos de n , já são observados *hubs* na rede livre de escala.

Figura 2.6 – Representação de diferentes modelos com graus indicados nos vértices

(a) Grafo gerado pelo modelo de Erdős-Rényi



(b) Grafo gerado pelo modelo de Barabási-Albert



Fonte: Elaboração própria.

2.2.5 Algoritmo de Barabási-Albert

O modelo de Barabási-Albert descreve como as redes livres de escala se formam, ou melhor, como fazer para criá-las. O modelo foi proposto pelos físicos Albert, Jeong e Barabási (1999) e descreve como a atração preferencial faz com que nós altamente conectados tenham maior probabilidade de receber novas conexões em comparação com aqueles com poucas conexões.

Barabási (2016, Seção 5.3) descreve seu modelo com o seguinte algoritmo:

1. Primeiro, coloca-se m_0 nós totalmente ligados;
2. Adiciona-se, passo a passo, um novo nó com m (sendo $m \leq m_0$) conexões que o ligam a m nós já existentes na rede, conforme a distribuição de probabilidade da atração preferencial

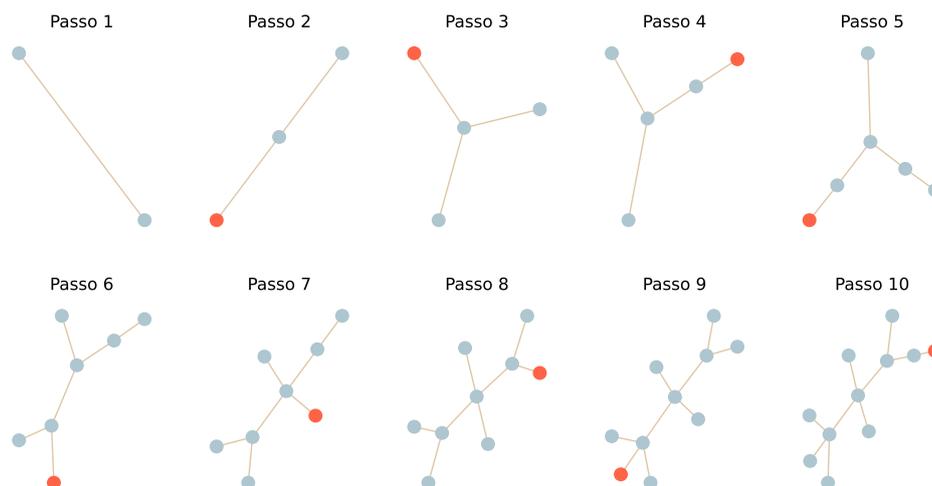
$$\Pi(k_i) = \frac{k_i}{\sum_i^n k_i}, \quad (2.27)$$

em que k_i é o *atual* grau do i -ésimo nó e $\sum_i^n k_i$ é a soma dos graus *atuais* de toda a rede.

3. Repete-se o passo 2 até que o número de nós na rede seja igual a n .

Observemos que a atração preferencial é um processo probabilístico onde um novo nó tem a possibilidade de se conectar a qualquer nó existente na rede, independentemente de ser um *hub* com muitas conexões ou um nó com poucas conexões (BARABÁSI, 2016, Seção 5.2). A Figura 2.7 demonstra graficamente como ocorre esse processo.

Figura 2.7 – Evolução do modelo de Barabási-Albert



Fonte: Adaptado de Barabási (2016, Seção 5.3).

2.3 INTERPRETAÇÃO DO MODELO DE ISING EM REDES COMPLEXAS

Após a introdução aos modelos de Ising regulares e um breve estudo sobre redes complexas, podemos agora explicar a aplicação do modelo de Ising em uma rede livre de escala. Essa adaptação permite explorar como a topologia da rede influencia as interações entre os *spins* e como essas interações podem gerar comportamentos coletivos complexos. A seguir, detalharemos a interação dos nós através das arestas e como a Hamiltoniana do sistema reflete as características dessa rede.

Cada nó da rede será associado a um *spin* σ_i , que pode assumir valores ± 1 , representando a orientação do momento magnético local. As arestas representam as interações entre os *spins*, conectando os nós e refletindo a capacidade de influência entre eles. O grau k_i de um nó descreve o número de vizinhos de um *spin*, indicando que nós com muitos vizinhos (*hubs*) terão mais interações. Em modelos físicos como o de Ising, as arestas podem ter pesos associados, representando a intensidade das interações entre *spins* vizinhos. Esses pesos são denotados como J_{ij} , representando a energia de interação entre os *spins* σ_i e σ_j . Em versão simplificada, podemos assumir que todas as interações possuem a mesma intensidade, ou seja, $J_{ij} = J$, com J sendo constante que define a energia das interações no sistema.

Na Hamiltoniana do sistema, dada pela Equação 2.18, a soma é realizada sobre todos os pares de vizinhos i e j na rede, abrangendo os pares de nós conectados por uma aresta. Assim, a interação entre os *spins* é determinada pelas arestas da rede, sendo ponderada pela intensidade da conexão J_{ij} , caso a rede contenha pesos nas arestas.

Ao implementar o modelo de Ising em uma rede livre de escala, consideramos uma rede estática, com n nós (ou partículas), cuja estrutura topológica permanece fixa ao longo da dinâmica dos *spins*. Após a construção da rede, a evolução dos *spins* segue o modelo de Ising regular, mas a topologia da rede influencia diretamente a dinâmica e o comportamento coletivo do sistema. A estrutura topológica da rede se mantém constante enquanto os *spins* evoluem com base nas interações definidas pela Hamiltoniana.

2.4 CONSIDERAÇÕES FINAIS

Neste capítulo, exploramos o modelo de Ising e suas propriedades clássicas, para então analisar sua implementação em estruturas de redes complexas. A introdução do modelo de Ising em grafos com conexões heterogêneas, como as redes livres de escala, possibilita a caracterização de comportamentos que se distanciam da grade bidimensional tradicional, que pode refletir as dinâmicas dos mercados financeiros. A teoria de redes complexas fornece base matemática robusta para compreender a heterogeneidade e a conectividade assimétrica dos mercados, onde a presença de *hubs* permite a propagação rápida de choques, fenômenos típicos de crises e bolhas. A utilização dessas redes na modelagem de sistemas financeiros amplia a compreensão dos mecanismos de propagação de informação e a formação de tendências coletivas, aspectos essenciais para o estudo do comportamento dos agentes no mercado.

3 ESTRUTURA DO MERCADO FINANCEIRO E INSTRUMENTOS DE NEGOCIAÇÃO

Antes de iniciar a modelagem de sistemas complexos, como o mercado financeiro, é fundamental compreender sua estrutura e conceitos envolvidos. O mercado financeiro é um ambiente dinâmico e multifacetado, no qual agentes interagem negociando instrumentos financeiros que representam direitos sobre fluxos de renda futuros ou índices que sintetizam o desempenho de um conjunto de ativos. Para fundamentar a modelagem quantitativa proposta, o modelo de Q-Ising será utilizado. No entanto, para sua aplicação, algumas simplificações serão necessárias. É essencial que tenhamos plena consciência das suposições que faremos ao longo do processo de modelagem.

Este capítulo está organizado em duas seções principais. Na Seção 3.1, são explorados conceitos fundamentais dos ativos financeiros, como ações, títulos de dívida, derivativos e índices, com destaque para suas funções e características. E a Seção 3.2 descreve a estrutura do mercado financeiro, abrangendo as bolsas de valores e o mercado de balcão, e discutindo como esses ambientes facilitam a negociação dos ativos.

3.1 ATIVOS FINANCEIROS

Ativos financeiros representam direitos sobre fluxos de renda futuros ou índices que sintetizam o desempenho de um conjunto de ativos. A principal característica dos ativos financeiros é que seu valor não reside em propriedades tangíveis, mas nas expectativas de rendimentos futuros. Esses instrumentos desempenham papel fundamental no mercado financeiro, facilitando a alocação de recursos e o financiamento de empresas e governos (FAMA; FRENCH, 1992).

Os principais tipos de ativos financeiros incluem:

- **Ações:** Representam frações do capital social de uma empresa, conferindo ao investidor direitos sobre lucros e, em alguns casos, sobre a gestão corporativa. Por serem instrumentos de renda variável, as ações apresentam riscos maiores, mas também podem gerar maiores retornos (FAMA; FRENCH, 1992).
- **Títulos de Dívida:** São instrumentos que representam empréstimos feitos por investidores a empresas ou governos. Os investidores recebem pagamentos de juros e o valor

principal no vencimento (ALTMAN, 1968).

- **Derivativos:** São ativos cujo valor depende de um ativo subjacente, como ações, commodities ou índices. Eles são utilizados para especulação ou *hedge*, oferecendo a possibilidade de altos lucros, mas com riscos proporcionais (HULL, 2012).
- **Índices:** Diferentemente dos ativos financeiros tradicionais, os índices não são negociáveis diretamente. Eles representam uma média ponderada de ativos subjacentes e servem como referência para o desempenho de segmentos do mercado, como o Índice Bovespa, que sintetiza o desempenho das principais ações na Bolsa de Valores brasileira (ELTON; GRUBER; BLAKE, 1996).

3.1.1 Preço de abertura, fechamento e outros parâmetros relevantes

O comportamento dos ativos financeiros é influenciado por diversos parâmetros que guiam as decisões dos investidores. Os principais parâmetros incluem os preços de abertura e fechamento, as máximas e mínimas diárias, o volume de negociação, a variação percentual e o *spread*. O preço de abertura reflete o primeiro preço de negociação de um ativo em determinado dia, enquanto o preço de fechamento serve como referência para análises de curto e longo prazo. As máximas e mínimas diárias indicam os limites dentro dos quais o ativo foi negociado. O volume de negociação fornece percepções sobre a liquidez do mercado, e o *spread* representa a diferença entre os preços de compra e venda, impactando os custos de transação.

3.2 ESTRUTURA DO MERCADO FINANCEIRO

O mercado financeiro é o ambiente no qual instrumentos financeiros são negociados. Ele é composto por diversos mercados, incluindo bolsas de valores e mercados de balcão, que oferecem diferentes níveis de regulamentação e liquidez. As bolsas de valores, por sua vez, têm horários específicos de funcionamento, com dias e períodos determinados para a negociação de ativos. No caso da B3, principal bolsa de valores do Brasil, a negociação ocorre durante os dias úteis, de segunda a sexta-feira, com exceção dos feriados. As sessões de pregão costumam começar às 10h e se encerram às 17h, sendo que, no pregão eletrônico, as negociações podem ser realizadas durante todo o dia, com horários para ajustes e ofertas. A operabilidade da bolsa nesses dias permite a formação de preços e a execução das transações.

3.2.1 Bolsas de valores

As bolsas de valores, como a B3 no Brasil e a NYSE nos Estados Unidos, são mercados organizados onde os ativos financeiros são comprados e vendidos de maneira regulamentada. Oferecem um ambiente transparente e eficiente para a negociação de ativos, permitindo a formação de preços de mercado e a captação de recursos para empresas. Neste sentido, as bolsas de valores facilitam a liquidez e a negociação.

3.2.2 Mercado de balcão

O mercado de balcão é uma alternativa às bolsas de valores, onde a negociação de ativos ocorre diretamente entre as partes, sem a intermediação de uma bolsa. Esse mercado é mais flexível e menos regulamentado, permitindo a negociação de ativos que não são listados nas bolsas, como algumas ações de pequenas empresas e instrumentos financeiros mais complexos, como derivativos. Embora o mercado de balcão ofereça maior liberdade para os participantes, ele também apresenta maior risco devido à menor transparência e regulamentação.

3.2.3 Liquidez e captação de recursos

A liquidez é um conceito básico no mercado financeiro, representando a facilidade com que um ativo pode ser comprado ou vendido sem impactar significativamente seu preço. A liquidez de um ativo depende de fatores como o volume de negociação e a profundidade do mercado. Ativos mais líquidos, como ações de grandes empresas, podem ser negociados rapidamente, enquanto ativos menos líquidos, como títulos de dívida de empresas menores, podem exigir mais tempo para serem negociados a preços justos. A capacidade de captar recursos é um aspecto fundamental das bolsas e mercados financeiros, permitindo que empresas e governos obtenham financiamento para suas operações e projetos.

3.3 CONSIDERAÇÕES FINAIS

Este capítulo fornece uma compreensão introdutória dos ativos financeiros e da estrutura do mercado financeiro, estabelecendo o contexto necessário para o desenvolvimento de modelos quantitativos aplicáveis à análise e previsão dos mercados. As séries temporais reais, que serão modeladas nos próximos capítulos, são derivadas dos preços de fechamento de diversos ativos. Nos capítulos subsequentes, aprofundaremos a modelagem do comportamento dos agentes e a implementação prática do modelo híbrido proposto, discutindo os desafios computacionais e as implicações econômicas dessas abordagens.

4 MODELAGEM BASEADA EM AGENTES E APRENDIZADO POR REFORÇO

Este capítulo adentra na modelagem baseada em agentes e sua aplicação em sistemas complexos, como o mercado financeiro. Inicialmente, na Seção 4.1, discute-se o conceito de agentes e suas características fundamentais, destacando seu papel em representar indivíduos ou entidades com comportamento autônomo e interações diversas. Em seguida, na Seção 4.2 a análise se aprofunda na metodologia de aprendizado por reforço, com a Seção 4.3 focando no *Q-learning*, algoritmo de aprendizado por reforço que permite que agentes aprendam a otimizar suas ações explorando recompensas recebidas de interações com um ambiente dinâmico e incerto. Por meio de exemplos e estudos de caso, como o problema do labirinto (OSMANKOVIĆ; KONJICIJA, 2011), ilustra-se o funcionamento e as implicações do *Q-learning* em diferentes cenários.

4.1 MODELO DE AGENTES

A modelagem baseada em agentes é uma abordagem eficaz para simular sistemas complexos, onde agentes autônomos interagem entre si e com o ambiente. Podem representar indivíduos, organizações ou entidades que tomam decisões fundamentadas em regras predefinidas ou aprendizado por reforço. O principal objetivo dessa modelagem é observar o comportamento emergente do sistema como um todo a partir das ações locais dos agentes, permitindo uma visão mais detalhada e dinâmica do funcionamento global.

Existem duas formas principais de estudar um objeto: a modelagem analítica, que simplifica o sistema em blocos conceituais para criar um modelo geral, e a modelagem sistêmica, que busca preservar as hierarquias internas do sistema (LIMA et al., 2009, p. 5279). A principal diferença entre essas abordagens está na modelagem sistêmica, que permite observar comportamentos macroscópicos emergentes, não diretamente derivados dos comportamentos individuais, ou seja, propriedades emergentes. Isso torna a abordagem sistêmica especialmente adequada para o estudo de sistemas complexos, uma vez que esses apresentam tais propriedades. Nesse contexto, a modelagem baseada em agentes se aproxima da análise sistêmica.

Organizações em Sistemas Multiagentes podem ser entendidas como entidades complexas, compostas por vasta quantidade de agentes que interagem em um ambiente estruturado com o objetivo de alcançar um propósito global (DIGNUM, 2010, p. 4). Esse propósito representa o próprio sistema ou objeto de estudo. Dessa forma, torna-se evidente que certas

propriedades emergem a partir dos papéis desempenhados por cada agente.

Há consenso entre diversos autores de que não há definição única e precisa para a modelagem baseada em agentes. Para Macal e North (2009, p. 87), nem mesmo o termo “agente” possui definição clara dentro dessa metodologia. Neste trabalho, consideraremos a modelagem de agentes como a interação entre agentes, mediada por regras. Esses agentes apresentam as seguintes características (MACAL; NORTH, 2009, p. 87):

1. Pode funcionar independentemente em seu ambiente e em suas interações com outros agentes, geralmente a partir de uma gama limitada de situações.
2. É indivíduo identificável e distinto, com um conjunto de características ou atributos, comportamentos e capacidade de tomada de decisão.
3. Possui papel que descreve como interage com outros agentes, assim como têm comportamentos.
4. Interage com seu ambiente.
5. Tem objetivos explícitos que orientam seu comportamento.
6. Pode ter a capacidade de aprender e adaptar seus comportamentos com base em suas experiências.
7. Possui atributos de recursos que indicam seu estoque atual de um ou mais recursos.

A capacidade de aprendizado dos agentes deve ser entendida como a habilidade de usar decisões passadas para influenciar futuras escolhas, dentro de certos limites. Claro, as decisões também serão impactadas pelas interações com o ambiente e com outros agentes. Os agentes serão considerados racionais, com preferências claras e a capacidade de modelar a incerteza por meio de valores esperados, buscando sempre o resultado ótimo dentro das opções disponíveis (JOSHI et al., 2013, p. 94). A modelagem baseada em agentes é interessante por sua capacidade de refletir a heterogeneidade de uma população, permitindo que cada agente tome decisões ajustadas às suas próprias circunstâncias, assumindo riscos e adaptando estratégias de investimento de maneira individualizada (WITTMANN, 2008, p. 34). No entanto, uma desvantagem dessa abordagem é a presença de suposições implícitas nos modelos, muitas vezes não claramente definidas ou discutidas (AUCHINCLOSS; GARCIA, 2015, p. 66). Por essa razão, este trabalho se esforçará para tornar todas as premissas e escolhas explícitas ao leitor.

Além disso, essa metodologia facilita a implementação da solução do modelo de Ising aplicado a uma rede livre de escala, permitindo a incorporação de novos comportamentos aos agentes sem alterar o modelo de Ising, como as decisões sobre o preço de compra ou venda de ações, que serão discutidas posteriormente.

4.2 INTRODUÇÃO AO APRENDIZADO POR REFORÇO

Em aprendizado de máquina, “os problemas de aprendizagem por reforço envolvem aprender o que fazer – como mapear situações para ações – de modo a maximizar um sinal numérico de recompensa” (SUTTON; BARTO, 2018, p. 2, tradução nossa). Os principais componentes de um sistema de aprendizagem por reforço incluem:

- Agente: A entidade que toma as decisões.
- Ambiente: O contexto ou mundo no qual o agente opera.
- Estado: Situação em que o agente deve realizar uma ação.
- Ações: As escolhas disponíveis para o agente.
- Recompensas: Resposta recebida após cada ação, que indica o sucesso ou fracasso da ação em relação aos objetivos.
- Memória: Armazena todas as situações em que o agente se encontrou (e suas respectivas recompensas) para compará-las com as novas situações.
- Treinamento: Processo de construção da memória.

No aprendizado por reforço “o agente tem de explorar o que já sabe para obter recompensas, mas também tem de explorar para fazer melhores seleções de ações no futuro.”¹ (SUTTON; BARTO, 2018, p. 3, tradução nossa). Este dilema é conhecido por Exploração/Exploração (ou *Exploration/Exploitation* em inglês), e constata-se que “a questão do equilíbrio entre exploração e exploração nem sequer se coloca na aprendizagem supervisionada e não supervisionada, pelo menos nas suas formas mais puras” (SUTTON; BARTO, 2018, p. 3, tradução nossa). Segundo Sutton e Barto (2018, tradução nossa), a diferença entre aprendizagem por reforço e aprendizagem supervisionada ocorre “na medida em que os pares de entrada/saída corretos nunca são apresentados, nem as ações subótimas são explicitamente corrigidas”.

O objetivo do agente é aprender uma política, que é uma função que mapeia estados para ações, para maximizar a recompensa cumulativa ao longo do tempo. Para isso, a aprendizagem por reforço utiliza diferentes algoritmos, sendo o *Q-learning* um dos mais conhecidos. Esta técnica é amplamente aplicada em diversas áreas, como jogos, robótica, controle de sistemas, e finanças.

¹“Exploit” foi traduzido como “explorar” no sentido de aproveitar ou tirar proveito de ações que já são conhecidas por trazer recompensas elevadas. Já “explore” foi traduzido como “explorar” no sentido de investigar novas ações para descobrir seu potencial.

4.3 Q-LEARNING: DEFINIÇÃO E FUNDAMENTOS

“O *Q-learning* é uma forma de aprendizagem por reforço sem modelo” (WATKINS; DAYAN, 1992, p. 55, tradução nossa), no sentido de que não precisa conhecer previamente as regras ou a dinâmica do ambiente. O termo “*Q-learning*” (ou Aprendizado *Q*, em tradução livre)² procede do conceito central do algoritmo: a função *Q*, a qual é a função que associa valor a uma ação que o agente aprende durante o processo de aprendizado por reforço. A letra “*Q*” no *Q-learning* não possui sentido definido, mas é comumente utilizada para denotar a qualidade computada pela função *Q* de determinada ação em um estado particular. A seguir, discutiremos apenas o que for relevante para o entendimento do modelo de *Q-learning* dependente do tempo, já que para o fim que nos interessa este é necessário e suficiente.

Neste sentido, a função *Q* é definida como o valor esperado da soma dos retornos futuros R_t , dados o estado atual do agente s e a ação a :

$$Q(s, a) = \mathbb{E}[R_t \mid s_t = s, a_t = a]. \quad (4.1)$$

À medida que o sistema evolui, a qualidade de uma ação pode variar; portanto, a atualização da função *Q* é descrita por (adaptado de Watkins e Dayan (1992)):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right], \quad (4.2)$$

em que $Q(s_t, a_t)$ é o valor atual da função *Q* para o estado s_t e a ação a_t ; r_{t+1} é a recompensa imediata após a ação tomada; $\max_a Q(s_{t+1}, a)$ é a estimativa do melhor valor de *Q* futuro; α é a taxa de aprendizado, que ajusta a função *Q* mais rápida ou lentamente, tornando a recompensa imediata e a diferença entre o ganho futuro e o atual mais ou menos relevantes; e γ é o fator de desconto que determina a importância dos retornos futuros em relação à recompensa imediata.

A Equação 4.2 atua ajustando o valor atual da função $Q(s_t, a_t)$ incorporando informações mais recentes ao passo da série temporal. O termo nos colchetes

$$r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)$$

pode ser entendido como erro temporal (ou TD *error* em inglês), que computa a diferença entre a recompensa esperada antes e após observar a recompensa real r_{t+1} e a estimativa da melhor recompensa futura. “Cada atualização depende do erro TD atual combinado com

²O termo *Q-learning* é bastante consolidado no contexto de aprendizado por reforço em aprendizado de máquina, portanto, seguiremos utilizando o termo original em inglês, e os termos subsequentes em tradução livre – indicando sempre o termo original em inglês.

os traços de elegibilidade de eventos passados” (SUTTON; BARTO, 2018, p. 179, tradução nossa). Se o erro temporal for positivo, isso indica que a ação a_t levou a recompensa maior do que o esperado, então o valor de $Q(s_t, a_t)$ aumenta. Se for negativo, o valor de $Q(s_t, a_t)$ diminui, indicando que a recompensa foi menor do que o esperado.

De maneira concisa, a experiência do agente é composta por uma sequência de passos e no t -ésimo passo, o agente (adaptado de Watkins e Dayan (1992)):

1. observa seu estado atual s_t ,
2. seleciona e executa uma ação a_t ,
3. observa o estado subsequente s_{t+1} ,
4. recebe uma recompensa imediata r_t e
5. ajusta os valores de $Q(s_t, a_t)$ utilizando um fator de aprendizado α , conforme Equação 4.2.

Neste sentido, o agente é treinado seguindo os procedimentos descritos acima. O processo de treinamento é faseado em etapas que chamaremos de episódios de treinamento. No primeiro episódio, o agente tem sua memória vazia, e a constrói à medida que muda de estado. Ao fim do primeiro episódio, o agente terá percepção do problema rudimentar, pois, provavelmente não teve a oportunidade de tomar todas as decisões possíveis em cada estado. Por isso, repete-se o processo para mais um episódio e as recompensas recebidas agora acumulam com as anteriores. Com número suficientemente grande de episódios, o agente terá explorado todas as ações possíveis em cada estado, resultando em uma memória bem estruturada. Isso permitirá ao agente alcançar seu objetivo de maneira eficiente e eficaz.

4.3.1 Tabela Q

A Tabela Q (ou Q -Table em inglês) é uma estrutura de dados característica do algoritmo Q -learning que age como a memória do agente. A estrutura da tabela consiste nas colunas, que representam as ações possíveis do agente, e nas linhas, que representam os estados em que este já se encontrou, conforme Tabela 4.1. Os valores dos elementos da tabela é dado pela função Q e esta é atualizada conforme a Equação 4.2. As ações possíveis são predefi-

Tabela 4.1 – Tabela Q do Q -learning

Estado/Ação	Ação 1	Ação 2	Ação 3
Estado 1	$Q(s_1, a_1)$	$Q(s_1, a_2)$	$Q(s_1, a_3)$
Estado 2	$Q(s_2, a_1)$	$Q(s_2, a_2)$	$Q(s_2, a_3)$
Estado 3	$Q(s_3, a_1)$	$Q(s_3, a_2)$	$Q(s_3, a_3)$

nidas no algoritmo. Os estados, por outro lado, são construídos durante o percurso da série temporal, conforme a definição de *Q-learning* como um algoritmo *sem modelo* apresentada por Watkins e Dayan (1992). Isso ocorre porque não há tabela Q antes da inicialização do algoritmo.

Espera-se que, em determinado momento, o agente comece a identificar os estados presentes na tabela, permitindo que o aprendizado por reforço ocorra. Caso contrário, todos os estados serão novos, e o agente nunca terá referência de recompensa.

4.4 APLICAÇÃO DO *Q-LEARNING* NO PROBLEMA DO LABIRINTO: UM ESTUDO DE CASO

Uma aplicação comum do modelo de *Q-learning* é a solução do problema do labirinto (OSMANKOVIĆ; KONJICIJA, 2011): o agente deve encontrar o caminho ótimo de um ponto inicial até um ponto objetivo, minimizando a recompensa negativa associada aos movimentos.

O modelo pode ser pensado da seguinte forma:

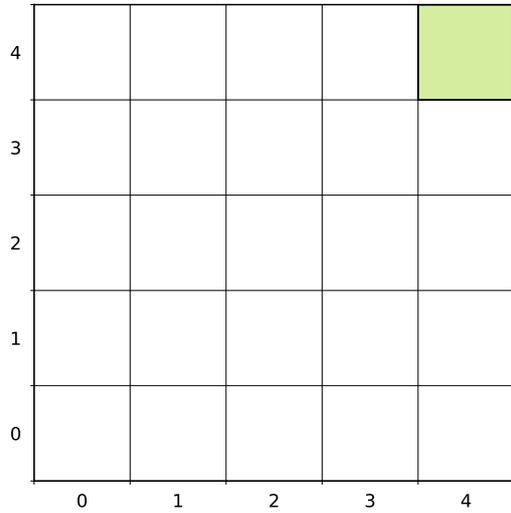
- Ambiente: labirinto com ponto inicial fixo, ponto objetivo e células com custos.
- Estado: cada célula do labirinto é representada por um estado, denotado por $s = (i, j)$, onde i e j correspondem às coordenadas das linhas e colunas, respectivamente.
- Ações: mover-se para células adjacentes (para cima, para baixo, para a esquerda, e para a direita).
- Recompensa: recompensa negativa para cada movimento, recompensa positiva ao alcançar a meta.
- Meta: chegar ao estado s_M com menor número de ações.

Na Figura 4.1a observamos uma grade com 5×5 células, de forma que o agente possa passear livremente pela grade, sempre caminhando por células adjacentes a que se encontra. O agente inicia seu movimento na célula $(0, 0)$ e pode tomar a decisão de caminhar para a célula $(1, 0)$ ou $(0, 1)$. Quando este se coloca em qualquer célula que não seja a $s_M = (4, 4)$ (destacada em verde na Figura 4.1a), recebe uma recompensa negativa fixa, e ao chegar em seu objetivo, recebe recompensa positiva e o episódio de treinamento se encerra. A Figura 4.1b mostra a distribuição de recompensas de cada célula.

No primeiro episódio do treinamento, o agente não possui nenhuma informação prévia sobre o ambiente, resultando em uma tabela Q inicial completamente vazia. Nesse contexto,

Figura 4.1 – Esquema ilustrativo do labirinto com objetivo em $s_M = (4, 4)$ e ponto de partida na célula $(0, 0)$

(a) Ambiente do labirinto

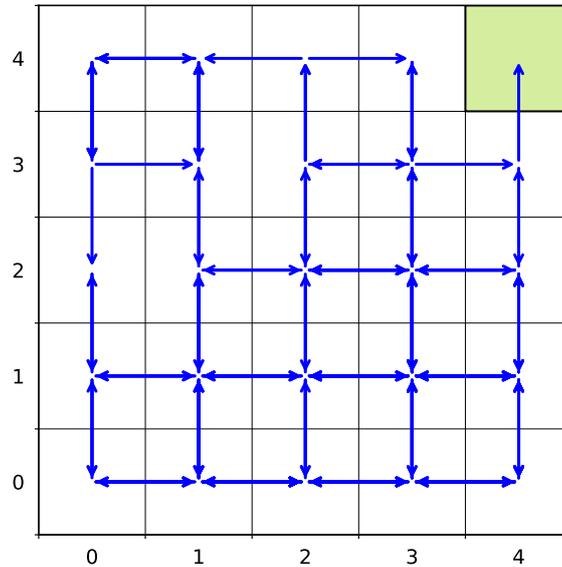


(b) Indicação das recompensas de cada célula

4	-1.0	-1.0	-1.0	-1.0	100.0
3	-1.0	-1.0	-1.0	-1.0	-1.0
2	-1.0	-1.0	-1.0	-1.0	-1.0
1	-1.0	-1.0	-1.0	-1.0	-1.0
0	-1.0	-1.0	-1.0	-1.0	-1.0
	0	1	2	3	4

Fonte: Elaboração própria.

Figura 4.2 – Trajetória percorrida por agente aleatório, saindo do ponto inicial até o objetivo

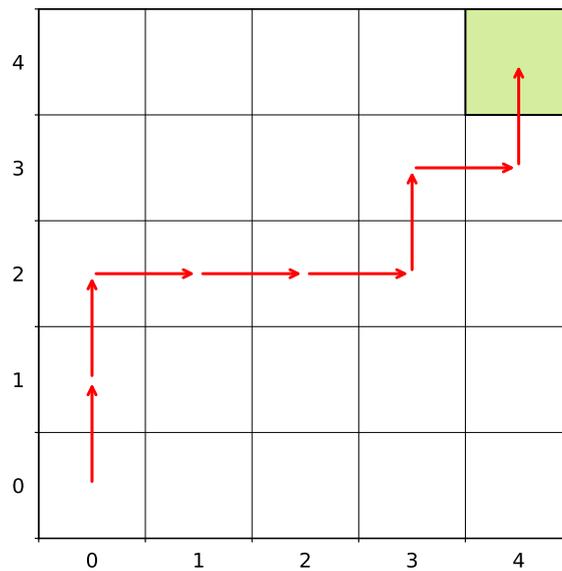


Fonte: Elaboração própria.

se eliminarmos a capacidade de aprendizado do agente (Equação 4.2), podemos considerar que ele se comporta, neste caso específico, de forma análoga a um caminhante aleatório. A Figura 4.2 ilustra o percurso de um caminhante aleatório (agente sem capacidade de aprendizado) ao explorar o labirinto. As setas indicam as transições entre células, mostrando o caminho percorrido pelo agente, que muitas vezes realiza várias voltas em torno de determinadas células antes de alcançar a meta.

Considerando agora que há atualização da tabela Q , após o treinamento, o agente consegue aprender trajetórias ótimas até a meta, estado $s_M = (4, 4)$, como demonstra à Figura 4.3. Após treinado, o agente irá percorrer sempre o mesmo trajeto, já que o dilema Exploração/Explotação somente ocorre durante o treinamento. Podemos entender que o modelo após o treinamento sofre de superajuste, no sentido de que se mudarmos o objetivo, por exemplo, o agente não saberá se comportar eficientemente. O que se pode fazer para contrapor esse desafio é retreinar o modelo para o objetivo seguinte, e é esta abordagem que usaremos no modelo específico que implementaremos junto ao modelo de Ising.

Figura 4.3 – Trajetória percorrida por agente treinado, saindo do ponto inicial até o objetivo



Fonte: Elaboração própria.

4.4.1 Enviesamento do modelo

Uma questão que surge naturalmente é: por que o agente recebe recompensa negativa em mesma intensidade em cada célula, exceto no objetivo? É natural querermos colocar recompensas menos negativas em células próximas ao objetivo, ou até mesmo positivas. Contudo, fato é que estaríamos especificando o caminho de interesse para o agente implicitamente pelas células que o beneficiassem mais. Como resultado, o agente seguiria exatamente o caminho sugerido pelas diferentes recompensas, o que pode ou não ser desejável. No caso do labirinto, o objetivo é que o agente encontre o caminho ótimo por conta própria. Se direcionarmos o caminho por meio das recompensas, a simulação perde o propósito.

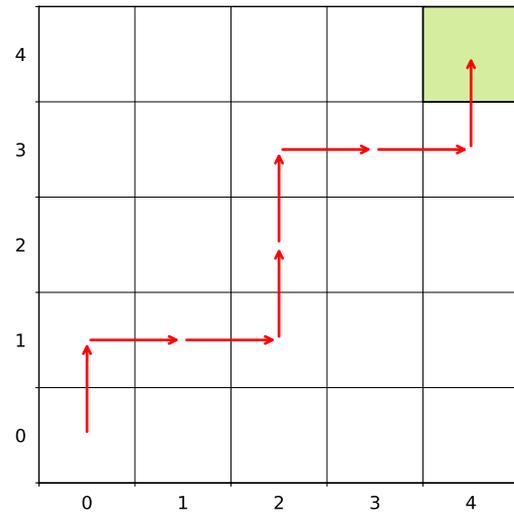
Para demonstrar esse comportamento, redefinimos as recompensas dos estados do modelo, atribuindo valor menos negativo as células diagonais (Figura 4.4a). Ao treinar o agente nessas condições, observamos que o caminho seguido agora intercepta as células diagonais (Figura 4.4b). Embora o modelo ainda possa produzir caminhos que não interceptem todas

Figura 4.4 – Representações visuais do processo do agente enviesado

(a) Indicação das recompensas de cada célula

4	-1.0	-1.0	-1.0	-1.0	100.0
3	-1.0	-1.0	-1.0	-0.5	-1.0
2	-1.0	-1.0	-0.5	-1.0	-1.0
1	-1.0	-0.5	-1.0	-1.0	-1.0
0	-0.5	-1.0	-1.0	-1.0	-1.0
	0	1	2	3	4

(b) Trajetória percorrida pelo agente treinado



Fonte: Elaboração própria.

as células diagonais, essa probabilidade diminui quanto maior for o número de episódios de treinamento.

É importante destacar que o deslocamento de uma célula diagonal para outra é um problema análogo ao do labirinto original: encontrar o caminho ótimo – em termos de recompensas acumuladas – de A para B. Não há restrições sobre como isso deve ser feito, desde que as recompensas acumuladas sejam as maiores possíveis. Assim, o agente pode seguir diferentes trajetórias entre células com recompensas menos negativas. Isso ocorre porque as bordas das diagonais ainda têm recompensas iguais, de modo que a soma final das recompensas pode ser preservada por mais de um caminho percorrido entre as diagonais. Se fosse traçado um caminho contínuo para o agente, não haveria outra opção na qual a soma das recompensas fosse maior ou igual a essa.

4.5 CONSIDERAÇÕES FINAIS

Neste capítulo, mostramos que o agente baseado em *Q-learning* atua como especialista em tomada de decisões, em específico na tarefa de encontrar o caminho ótimo entre dois pontos. No âmbito mais amplo desta pesquisa, propomos desenvolver um modelo de *Q-learning* especializado em investimentos em ativos de risco. Este especialista será utilizado para influenciar os agentes (ou partículas) do modelo de Ising analogamente ao campo externo, de modo que as grandezas físicas típicas deste modelo descrevam a série temporal de um ativo.

A escolha do *Q-learning* é justificada por sua simplicidade e consistência. Este algoritmo

de aprendizado por reforço é acessível e intuitivo, permitindo a implementação de soluções eficientes em problemas de tomada de decisão em ambientes complexos. Uma das principais vantagens do *Q-learning* é sua natureza interpretável; ao contrário de muitos modelos de Aprendizado de Máquina que operam como “caixas pretas”, o *Q-learning* oferece maior transparência nas decisões tomadas pelo agente.

Essa abordagem híbrida une a eficiência do *Q-learning*, que otimiza decisões com base em experiências passadas, com a modelagem física proporcionada pelo modelo de Ising, criando um sistema capaz de capturar tanto padrões temporais quanto as interações internas entre os agentes. A integração entre essas duas metodologias será realizada ao definir a ação do agente *Q-learning* como campo externo que influencia o modelo de Ising. Isso permitirá que o modelo de Ising seja guiado para refletir tendências de mercado, em vez de exibir comportamentos puramente estocásticos.

Entretanto, é importante ressaltar que a escolha do *Q-learning* é, em última análise, arbitrária. Outros algoritmos que sejam igualmente capazes de “interpretar” séries temporais e tomar decisões ótimas poderiam ser utilizados como alternativas viáveis, até mesmo superiores ao modelo que propomos. Modelos como SARSA, *Deep Q-Networks* (DQN) ou mesmo abordagens baseadas em aprendizado supervisionado podem ser eficazes, dependendo das características específicas do problema e dos dados disponíveis. A flexibilidade da modelagem baseada em agentes permite a inclusão de diferentes técnicas, desde que estas sejam capazes de capturar dinâmicas do sistema e fornecer soluções satisfatórias.

Os detalhes específicos do algoritmo, como definição dos estados s , escolha das recompensas r_n , e a forma exata de acoplamento entre o agente e o modelo de Ising serão abordados nos capítulos seguintes. Além disso, será investigado como essa integração impacta a convergência do modelo e sua capacidade de adaptação a diferentes condições de mercado. Em suma, a escolha do *Q-learning* é uma abordagem justificada, mas a diversidade de modelos existentes oferece opções que podem ser exploradas em futuras pesquisas, permitindo adaptação contínua às demandas de sistemas complexos e às particularidades dos dados em análise.

5 IMPLEMENTAÇÃO DO MODELO DE ISING E DE APRENDIZADO POR REFORÇO

Neste capítulo, implementamos o modelo de Ising em redes complexas e o aprendizado por reforço no mercado financeiro. Na Seção 5.1, apresentamos a formulação teórica do modelo de Ising e sua relação com redes livres de escala, destacando a influência da estrutura da rede na dinâmica do sistema. Na Seção 5.2, discutimos a implementação de estratégias de investimento baseadas em *Q-learning*, e avaliamos seu desempenho. Por fim, na Seção 5.3 discutimos simplificações necessárias para tornar viável a implementação prática do modelo de investimentos, garantindo equilíbrio entre precisão e eficiência computacional.

5.1 MODELO DE ISING EM UMA REDE LIVRE DE ESCALA

Nesta seção, será apresentada a caracterização detalhada do modelo de Ising aplicado a uma rede livre de escala, com enfoque nos aspectos físicos do sistema. A análise baseia-se nos resultados das simulações realizadas, cuja metodologia será descrita nas subseções subsequentes. Este estudo é essencial para entender o comportamento esperado do modelo na rede, fornecendo a base necessária para interpretar seu desempenho quando combinado com o modelo de *Q-learning*. Após a descrição metodológica do *Q-learning*, que será utilizado para otimizar as decisões de compra e venda de ativos de risco, avançaremos para a implementação completa do modelo, com objetivo de caracterizar as séries temporais desses ativos.

5.1.1 Procedimentos das simulações

A primeira etapa na construção do modelo de Ising em uma rede livre de escala é a modelagem da própria rede. Utilizamos o algoritmo proposto por Barabási (2016, Seção 5.2), conforme descrito na Subseção 2.2.5, para estruturar a rede. Inicialmente, definimos o número de nós n e o grau mínimo m da rede. Para implementar o modelo de Ising, basta atribuir *spin* para cada nó, inicialmente distribuídos aleatoriamente (distribuição uniforme de *spins*). Além disso, precisamos atribuir os parâmetros: energia de interação entre os nós vizinhos J (trabalharemos com interação homogênea); o campo externo, inicialmente nulo $h = 0$; e a temperatura T , variada conforme o interesse.

O algoritmo utilizado para simular o comportamento dos *spins* e estimar as propriedades termodinâmicas do sistema é o clássico algoritmo de Metropolis. Após estabelecer a configuração inicial dos *spins*, o algoritmo realiza a simulação das interações entre vizinhos. Para isso, é proposta uma alteração no estado de um *spin* em um nó específico da rede, σ_i , invertendo o seu valor. Em seguida, calcula-se a variação de energia ΔE_i associada a essa alteração, considerando as contribuições energéticas dos *spins* vizinhos. A alteração proposta é então aceita ou rejeitada com base na variação de energia calculada. Se a variação de energia ΔE_i for negativa, a nova configuração é aceita por indicar redução na energia do sistema. Se ΔE_i for positiva, a aceitação da nova configuração é feita com probabilidade que decresce exponencialmente com ΔE_i e é inversamente proporcional à temperatura (T), especificamente: $p_i = \min(1, \exp(-\Delta E_i/k_B T))$.

Tabela 5.1 – Parâmetros da Rede e do Modelo de Ising

Rede	
Parâmetro	Símbolo
Nº de nós	n
Grau mínimo	m

Ising	
Parâmetro	Símbolo
Energia de interação	J
Temperatura	T
Campo externo	h
Nº de passos MC	N_{passos}

Este processo de proposta e aceitação é realizado pelo método de Monte Carlo (MC) para cada nó da rede e repetido por N_{passos} iterações, permitindo amostragem adequada do espaço de configurações do sistema. Após um período inicial de equilíbrio $N_{\text{passos}}/2 \in \mathbb{Z}$, no qual o sistema atinge estado estatisticamente estável, as amostras coletadas são utilizadas para calcular propriedades termodinâmicas relevantes: magnetização, energia média do sistema, calor específico e suscetibilidade, conforme discutido na Seção 2.1. O algoritmo de Metropolis, portanto, facilita a análise do comportamento dos *spins* e a compreensão das transições de fase no modelo de Ising. A Tabela 5.1 resume os parâmetros relevantes para o modelo de Ising na rede livre de escala.

5.1.2 Assunção da constante de Boltzmann

Para simplificar as equações e cálculos no modelo de Ising, assumiremos que a constante de Boltzmann (k_B) é igual a 1. Essa escolha de unidade natural implica que a temperatura (T) e a energia de interação (J) são expressas em unidades adimensionais.

Ao definir $k_B = 1$, simplificamos a relação entre a energia do sistema e a temperatura para:

$$\beta = \frac{1}{T}$$

onde β é o inverso da temperatura. Com essa assunção, a constante de Boltzmann é eliminada das equações, facilitando a análise e interpretação dos resultados. Nesse contexto, tanto a energia de interação J quanto a temperatura T são tratadas como adimensionais, permitindo comparação direta e simplificada dos resultados obtidos em diferentes configurações do modelo, sem necessidade de unidades físicas explícitas.

5.1.3 Propriedades físicas do modelo de Ising na topologia complexa

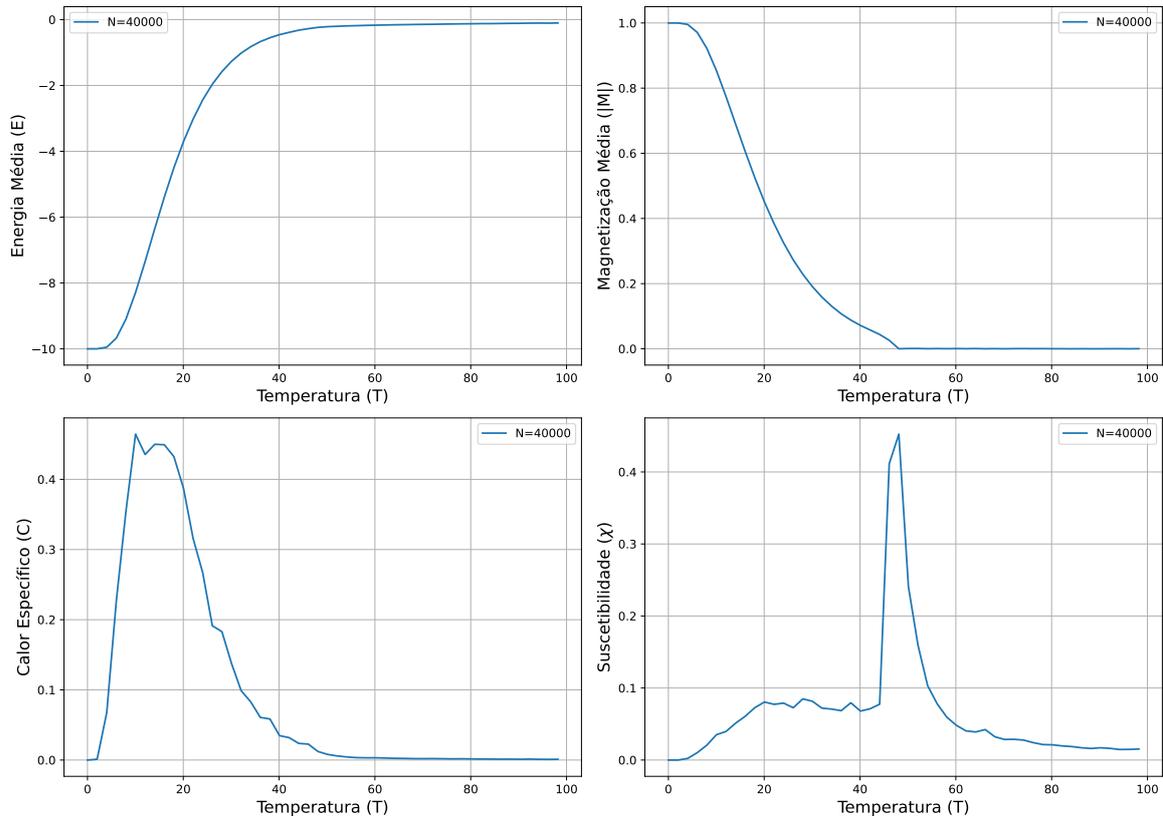
O primeiro resultado a ser analisado é a variação das quantidades termodinâmicas em função da temperatura. Este estudo é realizado utilizando a metodologia descrita, com a adição da variação da temperatura. A Figura 5.1 exibe quatro gráficos que mostram a variação do módulo da magnetização, energia média do sistema, calor específico e suscetibilidade em função da temperatura.

Para a realização da simulação, foi definida a seguinte configuração inicial:

- Número de nós (n): 40.000.
- Grau mínimo (m): 10.
- Energia de interação (J): 1, 0.
- Temperatura (T): de 0, 1 a 100, 0 com 200 pontos.
- Campo externo (h): 0, 0.
- Número de passos de Monte Carlo (N_{passos}): 2.000.

Observa-se que, embora o calor específico exiba comportamento aparentemente anômalo, a energia apresenta comportamento coerente, o que sugere que a anomalia no calor específico pode ser atribuída a erro numérico. Em contraste, a suscetibilidade revela descontinuidade significativa, com pico localizado entre $T = 40$ e $T = 60$. A magnetização média, por sua vez, corresponde a este comportamento, corroborando a observação. Portanto, este

Figura 5.1 – Variação das quantidades físicas (na ordem: Energia, Magnetização Média, Calor Específico e Suscetibilidade) com a temperatura



Fonte: Elaboração própria.

pico pode ser considerado o ponto crítico, e a temperatura associada a ele será denominada de temperatura crítica, T_c .

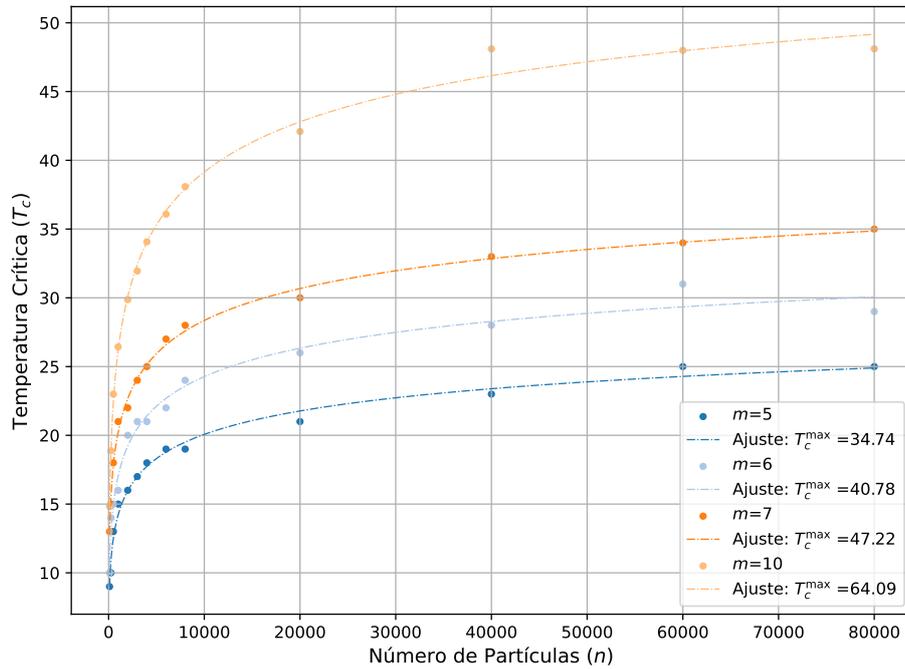
A análise do comportamento da temperatura crítica do sistema será dividida, respectivamente, em relação à variação do número de nós e número de conexões mínimas.

5.1.3.1 Temperatura crítica e o limite termodinâmico em função do tamanho do sistema

Agora que conseguimos obter a temperatura crítica de uma simulação, podemos variar o número de partículas n e construir o gráfico da variação de T_c . No gráfico apresentado na Figura 5.2, T_c é representada em função de n , o número de partículas na rede, para diferentes valores de m , o grau mínimo. As curvas sugerem tendência de saturação da temperatura crítica à medida que n aumenta, especialmente para valores maiores de m . Para $m = 5$, T_c se estabiliza ao nível relativamente baixo, enquanto para $m = 10$, a saturação ocorre em nível mais elevado de T_c .

O grau mínimo m é fundamental para o comportamento de T_c . À medida que m aumenta,

Figura 5.2 – Variação da temperatura crítica em função do número de partículas



Fonte: Elaboração própria.

a conectividade da rede cresce, resultando em maior robustez do sistema. Isso resulta no aumento de T_c , pois redes mais conectadas exigem maior energia térmica (temperatura) para romper as correlações entre *spins*, tornando a rede mais estável.

O ajuste sugerido é dado pela função

$$T_c(n) = T_c^{\max} \frac{n^\alpha}{n^\alpha - n_0},$$

indicando forma assintótica para T_c conforme n cresce. Para grandes valores de n , T_c tende ao valor de saturação T_c^{\max} . O parâmetro n_0 representa o ponto de inflexão, onde ocorre a transição significativa no comportamento de T_c , e α controla a rapidez dessa transição.

O gráfico mostra que a temperatura crítica se estabiliza para grandes valores de n , como esperado, pois em redes maiores a influência de novos nós na dinâmica do sistema é menos significativa. A taxa de saturação, no entanto, depende de m : redes com maior conectividade (maior m) alcançam a saturação mais rapidamente e em níveis mais altos de T_c , indicando que redes bem conectadas são mais resistentes à desordem térmica.

Essa relação reforça a ideia de que redes com maior grau mínimo apresentam transições de fase em temperaturas mais elevadas. Isso está alinhado com a física dos sistemas de *spin*, onde maior conectividade intensifica a interação entre *spins*, exigindo mais energia (temperatura) para desestabilizar o sistema. Dessa forma, espera-se que aumentar m resulte em aumento de T_c , refletindo maior estabilidade térmica da rede e deslocando a transição de

fase para temperaturas mais altas. O valor limite T_c^{\max} também deve crescer indefinidamente.

Além disso, valores maiores de m acelera a estabilização da rede, permitindo que o sistema atinja a saturação de T_c com número menor de partículas. O ponto de inflexão da curva $T_c(n)$, denotado por n_0 , se desloca para valores menores de n à medida que m aumenta. Isso sugere que a rede se aproxima rapidamente da temperatura crítica máxima, fazendo com que n/n_0 seja pequeno em comparação com redes de m menor.

Em termos físicos, conforme m cresce, a rede livre de escala se torna mais densa e interconectada, resultando em um sistema mais resistente a flutuações térmicas. Os nós (ou *spins*) em redes altamente conectadas influenciam-se mutuamente, aumentando a energia necessária para desordenar o sistema. Assim, espera-se que redes com maior conectividade exijam temperaturas mais altas para passar pela transição de fase, elevando o valor limite T_c^{\max} .

No limite termodinâmico, consideramos um sistema com número extremamente grande de partículas, onde as flutuações estatísticas se tornam desprezíveis. Nesse regime, propriedades macroscópicas, como temperatura e magnetização, tornam-se bem definidas e estáveis. A temperatura crítica é constante quando os demais parâmetros são mantidos fixos, essencial para a definição precisa das transições de fase. Em sistemas finitos, embora ainda seja possível discutir pontos críticos e transições de fase, essas quantidades podem sofrer flutuações significativas, dificultando sua caracterização. Portanto, ao conhecer o tamanho mínimo necessário, dada uma conectividade adequada, podemos trabalhar com sistemas finitos que reproduzam as características do limite termodinâmico, assegurando a estabilidade de suas propriedades.

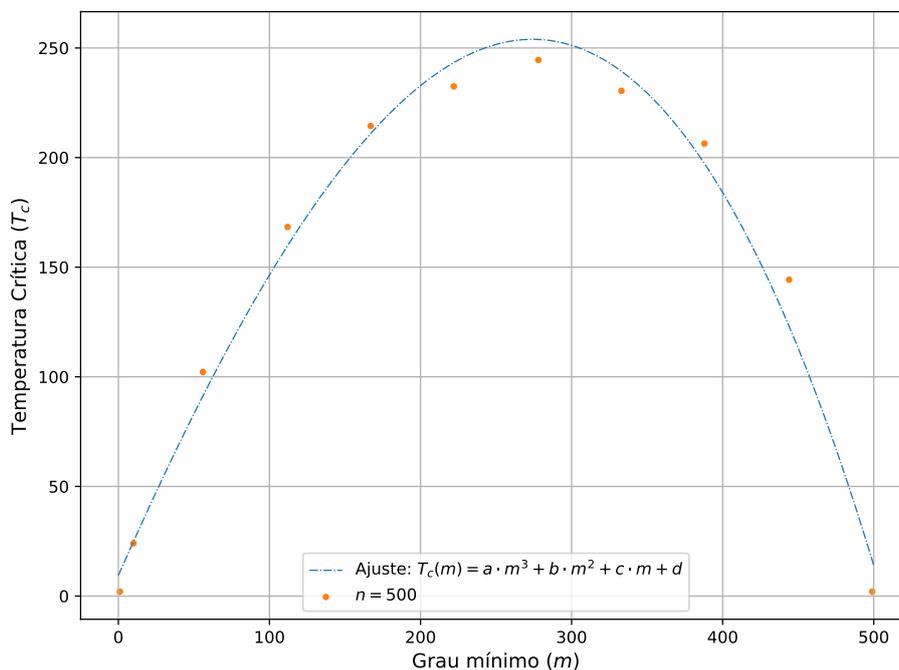
5.1.3.2 Influência do grau mínimo na convergência do limite termodinâmico

É importante entender como o grau mínimo m afeta as propriedades físicas de um sistema com n partículas. Para isso, plotamos o gráfico da temperatura crítica em função de m , mantendo $n = 500$. Isso nos ajuda a avaliar se as propriedades do limite termodinâmico são universais. A Figura 5.3 mostra que a temperatura crítica em relação ao número de conexões tem formato de parábola, ajustada por uma função polinomial de grau 3. Observamos “simetria”¹ nas redes com grau mínimo $1 + i$ e $n - i$ ($i \in (0, n)$).

É interessante notar que a temperatura crítica (T_c) do modelo de Ising em redes livres de escala com $m = 1 + i$ e $m = n - i$ é praticamente a mesma, apesar das diferenças no número de conexões das topologias. Essa semelhança pode ser explicada por dois aspectos: topológico e físico.

¹A palavra simetria está entre aspas, pois os dados foram ajustados por uma função polinomial de terceiro grau e, para $i \in (0, n)$, $T(1 + i) \neq T(n - i)$ em geral, mas os valores são consideravelmente próximos.

Figura 5.3 – Variação da temperatura crítica em função do grau mínimo – Valores dos parâmetros ajustados: $a = -2.91 \times 10^{-6}$; $b = -1.66 \times 10^{-3}$; $c = 1.56$; e $d = 9.43$



Fonte: Elaboração própria.

Sob a perspectiva topológica, a conectividade global da rede pode ser mais importante do que as conexões locais. A rede com $m = 1$ assemelha-se a uma cadeia, enquanto a com $m = n - 1$ é completamente conectada. Apesar dessas diferenças, a presença de *hubs* altamente conectados em redes livres de escala é fundamental para a propagação de informações e a sincronização das interações entre *spins*. Mesmo que a rede com $m = 1$ tenha conexões limitadas, a interação entre *spins* vizinhos é suficiente para que a transição de fase ocorra em temperatura crítica semelhante à de uma rede mais conectada.

Do ponto de vista físico, a temperatura crítica de um sistema de Ising depende do equilíbrio entre a energia de interação dos *spins* e a energia térmica. Para que a transição de fase ocorra, os *spins* devem interagir e se alinhar. Em ambos os casos, $m = 1$ e $m = n - 1$, a interação entre *spins* adjacentes é suficiente para gerar temperatura crítica semelhante. No primeiro caso, essa interação é favorecida pela presença de *hubs*, enquanto no segundo, é garantida pela conectividade total da rede.

A temperatura crítica reflete a estabilidade do estado coletivo dos *spins*. Apesar das diferenças na topologia, o sistema pode permitir correlações suficientes entre *spins* para que T_c permaneça próximo. Isso sugere que, independentemente das variações na conectividade, o sistema alcança equilíbrio termodinâmico semelhante quando a temperatura atinge T_c . O equilíbrio entre a energia de interação e a energia térmica determina o comportamento coletivo, e a temperatura crítica é o ponto em que a energia térmica é suficiente para romper a ordem imposta pelas interações.

Portanto, a proximidade de T_c em um modelo de Ising com $m = 1 + i$ e $m = n + i$ pode ser vista como uma propriedade coletiva que depende mais das interações entre os *spins* e das propriedades globais da rede do que da conectividade local. A dinâmica coletiva permite que a transição de fase ocorra sob condições termodinâmicas semelhantes, refletindo um equilíbrio entre as interações de *spins* e a energia térmica.

5.1.3.3 Diagrama de fases

O objetivo final deste trabalho é estabelecer conexão entre o modelo de Ising e a série temporal dos preços dos ativos financeiros por meio do campo externo h . Para tanto, é necessário compreender o comportamento do modelo em função desse campo. Assim, empregamos o diagrama de fases, conforme apresentado na Subseção 2.1.4 e ilustrado na Figura 2.2.

Na Figura 5.4 temos o diagrama de fases para o sistema com a seguinte configuração inicial:

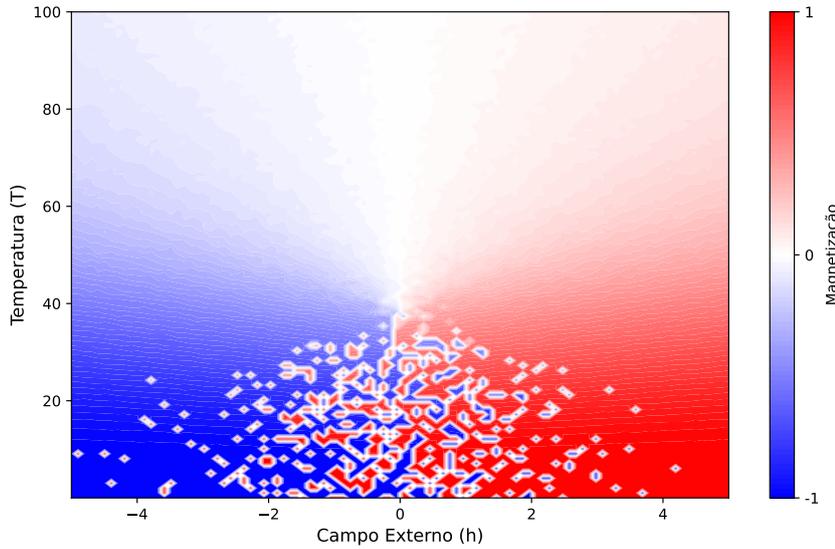
- Número de nós (n): 400.
- Grau mínimo (m): 17.
- Energia de interação (J): 1, 0.
- Temperatura (T): de 0, 1 a 100, 0 com 100 pontos.
- Campo externo (h): de $-5, 0$ a $5, 0$ com 100 pontos.
- Número de passos de Monte Carlo (N_{passos}): 2.000.

O tamanho reduzido do sistema é atribuído ao custo computacional associado ao número de simulações de Monte Carlo executadas. Apesar do sistema não estar em tamanho que possa ser considerado análogo ao limite termodinâmico, o comportamento apresentado ainda deve ser semelhante por continuidade.

Na Figura 5.4 observamos flutuações significativas que seguem padrão distinto, contrastando com a homogeneidade prevista no modelo clássico. No modelo quadrado 2D, se o número de passos de Monte Carlo for insuficiente para garantir a convergência, as flutuações tendem a ser isotrópicas ao longo do eixo $h = 0$, tornando-se mais intensas à medida que a temperatura se aproxima do ponto crítico T_c . Essas flutuações se tornam menos frequentes em temperaturas muito baixas, onde o sistema tende a atingir o estado ordenado.

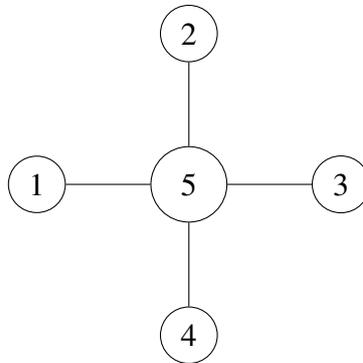
Dado que observamos precisamente o oposto, podemos interpretar que uma das causas dessas flutuações está relacionada ao acoplamento dominante do *hub*, que dificulta alcançar a configuração de menor energia do sistema (termalização). A magnetização do modelo pode

Figura 5.4 – Diagrama de fases do modelo de Ising em uma rede livre de escala



Fonte: Elaboração própria.

apresentar sinal negativo (ou positivo) mesmo quando o campo externo é, respectivamente, positivo (ou negativo), o que pode parecer inusitado. No entanto, esse comportamento pode ser demonstrado através de um modelo reduzido composto por 5 partículas, sendo uma delas um *hub*:



A inversão do *spin* de qualquer uma das partículas $i \in \{1, 2, 3, 4\}$ provoca a seguinte variação de energia:

$$\Delta E_i = 2\sigma_i(J\sigma_5 + h), \forall i \in [1, 4].$$

Dessa forma, para $0 < h < J$, a inversão do *spin* é energeticamente favorável quando $\sigma_5 = -1$ e $\sigma_i = +1$. Isso ocorre porque, sob essas condições, o termo $J\sigma_5$ prevalece sobre o campo externo h , resultando em uma variação de energia $\Delta E < 0$, que favorece a inversão do *spin* de σ_i para -1 . Este resultado não exclui a possibilidade desse comportamento ocorrer para $h > J$; ele é favorável quando $J > h$, justificando sua utilização como prova de sua ocorrência. No entanto, à medida que h se torna maior que J ($h > J > 0$), a inversão

do σ_i se torna energeticamente desfavorável. Como mostrado na Figura 5.4, a região onde a inversão é desfavorável se expande à medida que h aumenta, com a área de comportamento positivo crescendo proporcionalmente ao aumento de h .

Se o *hub* for positivo ($\sigma_5 = +1$), vemos que a inversão do σ_i para $+1$ também é favorável para $h < 0$. Portanto, a tendência observada de magnetização negativa, mesmo para o campo externo positivo (e vice-versa), pode ser explicada pela dinâmica do acoplamento dominante do *hub* com os *spins* periféricos, especialmente em temperaturas baixas, onde a probabilidade de inversão de *spin* diminui. Assim, a influência do *hub*, concomitante à interação preferencial em baixa temperatura, resulta no surgimento de magnetização contrária ao campo externo.

Nesse contexto, o processo de termalização é dificultado pela presença de *hubs*, o que exige um número elevado de passos de Monte Carlo para que o sistema alcance um estado estatisticamente estável. É importante investigar a quantidade de passos de Monte Carlo necessária para garantir a termalização completa, uma vez que esse valor pode variar conforme o tamanho do sistema, a temperatura e a topologia da rede. Além disso, é essencial avaliar a viabilidade computacional desse número de passos. Para o estudo proposto, esse comportamento não é necessariamente um defeito, pois, no contexto específico da simulação, pode ser considerado uma característica intrínseca do modelo. Portanto, continuará sendo adotado o modelo com tais propriedades.

5.2 MODELO DE APRENDIZADO POR REFORÇO

Nesta seção, apresentamos o modelo de aprendizado por reforço especializado na análise de séries temporais ruidosas, que será utilizado para orientar a tomada de decisão no modelo de Ising. O objetivo é desenvolver um agente capaz de identificar padrões nessas séries e otimizar suas ações com base em experiências passadas. Para isso, definimos formalmente os estados, ações e recompensas do agente, garantindo que ele possa explorar e aprender de maneira eficiente. Além disso, abordamos diferentes estratégias para aproximação de estados e detalhamos o processo de treinamento e validação do modelo, destacando os principais parâmetros que influenciam seu desempenho. Também discutiremos as simplificações adotadas no modelo de investimento, uma vez que, no contexto deste trabalho, as séries temporais analisadas correspondem a ativos financeiros.

5.2.1 Estados, ações e recompensas

A primeira questão a ser resolvida é quanto à definição dos estados. É importante perceber que, para o funcionamento correto do aprendizado por reforço, o agente deve encon-

trar em sua memória, em alguma medida, o estado em que se encontra. Caso contrário, o aprendizado não ocorrerá. As quantidades possíveis de se obter de um ativo financeiro são contínuas, tornando improvável a comparação exata entre dois estados. Para lidar com essa dificuldade, é necessário aproximar os estados, e existem diversos métodos disponíveis para isso.

Uma abordagem comum em *Q-learning* é usar um método de busca, que envolve a comparação do estado atual $s' = (x', y')$ com outros estados $s_i = (x_i, y_i)$ já existentes em uma tabela ou conjunto de estados. O objetivo é encontrar um estado s_i que satisfaça uma condição específica, ou seja, o mais próximo de s' de acordo com algum critério.

Outras abordagens incluem o uso de redes neurais, como o *Deep Q-Network* (DQN) (MNIH et al., 2015; RODERICK; MACGLASHAN; TELLEX, 2017), que é eficaz na representação de funções Q complexas em ambientes de alta dimensionalidade. Além disso, árvores de decisão também podem ser aplicadas para segmentar estados e estimar a função Q , promovendo melhor interpretação da política aprendida (ROTH et al., 2019). A combinação de tabelas Q com técnicas de aproximação também é uma estratégia importante para estimar valores em estados raramente visitados. Métodos mais avançados de reforço profundo, que integram aprendizado por reforço com redes neurais, se desenvolveram para melhorar a generalização das funções Q . Exemplos relevantes incluem o *Deep Deterministic Policy Gradient* (DDPG) (LILLICRAP, 2015; TAN, 2021) e o *Asynchronous Actor-Critic Agents* (A3C) (BABAEIZADEH et al., 2016).

Optaremos por um método de busca alinhado ao principal objetivo deste trabalho: utilizar modelos que sejam robustos, mas que mantenham simplicidade de implementação e apresentem comportamento interpretável. Embora o uso de modelos mais sofisticados também seja viável, seu emprego ficará para estudos futuros, conforme os resultados e conclusões obtidos ao longo desta pesquisa.

5.2.1.1 Estados

Um estado será definido da seguinte forma:

$$s_t = \left(\frac{P_t^f}{P_{t-1}^f}, \frac{MMS_{curta}}{MMS_{longa}} \right),$$

onde P_t^f representa o preço de fechamento do ativo no dia t , e MMS é a média móvel simples, com os subscritos “curta” e “longa” referindo-se ao número de dias usados no cálculo da respectiva média móvel. A escolha desses dois parâmetros tem em vista capturar as características da série temporal no dia t . A razão entre os preços de fechamento de t e $t - 1$ reflete a tendência de curto prazo, enquanto a razão entre as médias móveis simples revela a

tendência de longo prazo².

O número de características que define um estado é arbitrário, mas deve ser equilibrado: deve-se incluir informações suficientes para distinguir os estados sem os tornar excessivamente específicos, o que dificultaria a comparação entre eles.

Para um estado ser considerado igual a um estado presente na tabela Q , utilizaremos o seguinte critério de busca: tomamos $s = (x, y)$ como o estado presente na tabela e $s' = (x', y')$ o estado que o agente se encontra que queremos comparar com s . Definimos que s' é igual a s se:

$$|x' - x| = \rho x \quad \text{e} \quad |y' - y| = \rho y,$$

onde ρ é constante que define a tolerância percentual entre os valores x e y do estado presente e os valores x' e y' do estado a ser comparado.

Se houver mais de um estado $s_i = (x_i, y_i)$ na tabela que satisfaça a condição acima, será selecionado o estado s_i que minimiza o erro acumulado, definido como:

$$E_{\text{acum}}(s') = |x' - x| + |y' - y|$$

Logo, a função que escolhe o estado com menor erro acumulado é:

$$s^* = \arg \min_i (|x' - x_i| + |y' - y_i|)$$

onde $s^* = (x^*, y^*)$ é o estado na tabela que apresenta o menor erro acumulado em relação a s' .

5.2.1.2 Ações e recompensas

A Tabela Q se dará conforme Subseção 4.3.1. Além disso, recapitulamos abaixo a Equação 4.2 responsável pela atualização dos valores Q :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right].$$

Os parâmetros α , γ e ϵ continuam com os mesmos significados neste modelo, discutidos na Seção 4.3.

Assim, definimos as ações possíveis do agente como: $a_0 = -1$ (“vender”), $a_1 = 0$ (“manter”) e $a_2 = 1$ (“comprar”). Estas três ações abrangem todas as possibilidades de participação no mercado financeiro. A recompensa para cada ação será definida como proporcional aos ganhos provocados pela ação, isto é, a recompensa $r_{t+1}(a)$, em função da ação

²Vale ressaltar que essa definição de estado é aplicável a qualquer série temporal ruidosa, independentemente de estar relacionada a ativos financeiros ou a outros domínios.

a tomada no instante t , é dada por:

$$r_{t+1}(a_t) = a_t \times (P_{t+1}^f - P_t^f), \forall a \in \{-1, 0, 1\}.$$

Com isso, se o agente realizou uma compra ($a_t = 1$) e o preço aumentou ($a_t \times (P_{t+1}^f - P_t^f) > 0$), a recompensa será positiva. Caso o agente tenha comprado e o preço tenha caído, a recompensa será negativa. O mesmo raciocínio se aplica à venda ($a = -1$): se o preço diminuir, a recompensa será positiva, e se subir, a recompensa será negativa. Quando o agente opta por não tomar nenhuma ação ($a = 0$), a recompensa será sempre nula.

5.2.2 Treinamento e validação

O processo de treino e validação aplicado ao modelo utiliza uma janela deslizante para dividir dinamicamente os dados temporais em conjuntos de treino e teste, permitindo que o agente seja avaliado continuamente ao longo do tempo. Esse método é útil em séries temporais, pois os padrões podem mudar com o tempo, e treinar o modelo em diferentes partes da série pode ajudar a capturar essas mudanças.

É importante destacar que o processo de treinamento é onde a tabela Q é construída e atualizada continuamente, dia após dia, ao longo de vários episódios. Durante a fase de teste, o agente interage com os dados apenas uma vez, podendo consultar a tabela Q construída no treinamento para tomar decisões. No entanto, a tabela Q não é mais atualizada nesta fase.

Em relação à memória do modelo (tabela Q), há duas abordagens possíveis: reiniciar a tabela periodicamente ou mantê-la, permitindo que cada novo treino a atualize com os dados mais recentes. Essas abordagens serão discutidas nas próximas seções.

Parâmetros principais:

- P_t^f : Preço de fechamento do ativo no instante t , ou seja, os dados de interesse utilizados para treinar e testar o modelo.
- ω : Tamanho da janela de treino, que define o número de instantes consecutivos usados para treinar o modelo em cada iteração.
- τ : Tamanho da janela de teste, que define o número de instantes consecutivos usados para validar o modelo após cada período de treino.
- N : Número total de pontos de dados disponíveis na série temporal.
- Episódios: Ciclos de aprendizado em que o agente explora o ambiente (conjunto de treino), toma decisões e ajusta suas estimativas de recompensas futuras.

Em cada iteração, dois conjuntos de dados são extraídos:

1. **Conjunto de treino:** No instante $t = i$, uma janela de tamanho ω é extraída da série temporal P_t^f , contendo os preços de fechamento no intervalo de tempo $[i, i + \omega)$. Este conjunto de dados, denominado $\mathcal{D}_{\text{treino}}(i)$, é usado para treinar o modelo. Formalmente, ele é definido como:

$$\mathcal{D}_{\text{treino}}(i) = \{P_t^f \mid i \leq t < i + \omega\}.$$

Ou seja, $\mathcal{D}_{\text{treino}}(i)$ contém os preços de fechamento do instante $t = i$ até $t = i + \omega - 1$.

2. **Conjunto de teste:** Imediatamente após o intervalo de treino, uma nova janela de tamanho τ é extraída para validação, contendo os preços de fechamento no intervalo $[i + \omega, i + \omega + \tau)$. Este conjunto de dados, denominado $\mathcal{D}_{\text{teste}}(i)$, é usado para avaliar o desempenho do modelo após o treino. Formalmente, ele é definido como:

$$\mathcal{D}_{\text{teste}}(i) = \{P_t^f \mid i + \omega \leq t < i + \omega + \tau\}.$$

Ou seja, $\mathcal{D}_{\text{teste}}(i)$ contém os preços de fechamento do instante $t = i + \omega$ até $t = i + \omega + \tau - 1$.

Após cada iteração de treino e teste, o índice i , que representa o ponto de partida da janela de treino, é deslocado de τ unidades, de modo que tanto a janela de treino quanto a janela de teste avancem simultaneamente. Isso garante que os dados de treino e teste sejam sempre atualizados a cada iteração. O novo valor de i é dado por:

$$i = i + \tau.$$

O processo continua enquanto houver dados suficientes para criar ambas as janelas (treino e teste). A condição de parada ocorre quando:

$$i + \omega + \tau > N.$$

Neste ponto, não é mais possível extrair a janela de treino e a de teste, e o processo termina.

Observe que, sempre que a tabela é atualizada, existem duas abordagens possíveis: em vez de simplesmente acrescentar dados aos já existentes, podemos optar por apagar seu conteúdo, adotando atualização por reescrita, e preenchê-la com dados do novo treinamento. Alternativamente, é possível manter um acúmulo contínuo de informações na tabela. A eficácia de cada método será analisada nas seções seguintes.

Ao usar uma janela deslizante, o modelo é treinado em diferentes períodos da série temporal, ajudando a capturar possíveis mudanças nos padrões ao longo do tempo. Como o conjunto de teste é sempre atualizado em cada iteração, o modelo é continuamente validado em novos dados, proporcionando avaliação mais consistente de seu desempenho. Finalmente, o

tamanho das janelas de treino (ω) e teste (τ) pode ser ajustado para atender às necessidades específicas do problema, proporcionando flexibilidade.

Em resumo, o processo de treino e validação com divisão dinâmica descrito é uma forma segura de usar dados temporais para treinar e avaliar o modelo, mantendo o controle sobre a evolução do desempenho à medida que o tempo avança.

5.3 SIMPLIFICAÇÕES NO MODELO DE *Q-LEARNING*

Para garantir a eficiência computacional e viabilidade do modelo, várias simplificações foram adotadas, que permitem preservar aspectos essenciais do comportamento coletivo dos agentes enquanto minimizam a complexidade computacional. As simplificações a seguir aplicam-se, na maioria, ao mercado financeiro; para outros domínios, podem ser necessárias outras considerações.

O modelo utiliza exclusivamente informações históricas dos próprios ativos financeiros, como o preço de fechamento diário e o volume de negociação. Esses dados servem como entradas para o algoritmo de aprendizado por reforço, limitando o conjunto de variáveis a serem analisadas. Essa simplificação permite que o modelo se concentre em dados diretamente observáveis e amplamente disponíveis, mantendo a relação com a realidade do mercado enquanto evita a complexidade associada à modelagem de aspectos como eventos externos ou variáveis endógenas mais difíceis de quantificar.

Outro ponto relevante é a ausência de custos de transação no modelo, como taxas de corretagem ou *spreads* entre os preços de compra e venda. A remoção desses custos visa simplificar as simulações computacionais, permitindo que decisões de compra e venda se baseiem apenas nas expectativas futuras de rentabilidade dos ativos, sem distorções impostas por custos externos. A implicação dessa suposição é a consideração de um mercado perfeitamente competitivo e líquido, onde transações podem ser realizadas sem custos adicionais, refletindo simplificações da realidade dos mercados financeiros.

Ainda sobre a questão da liquidez, assume-se que o mercado possui liquidez suficiente para acomodar transações unitárias sem afetar os preços dos ativos, ou seja, a compra ou venda de uma unidade de ativos por agentes não gera impacto significativo no preço de mercado. Isso implica que, no contexto deste modelo, as elasticidades de oferta e demanda são suficientemente altas para garantir que a dinâmica do mercado seja estável, mesmo com a execução de transações individuais. Portanto, a premissa de alta liquidez exclui o risco de movimentos bruscos de preços causados por ordens de grande volume ou pelo impacto de transações específicas de um único agente.

Por fim, assumimos que o agente, embora capaz de tomar decisões de compra e venda,

não exerce influência direta sobre o preço dos ativos ou sobre a dinâmica global do mercado. Ou seja, as ações de compra ou venda de um único agente não geram movimentos significativos nos preços, caracterizando mercado atomisticamente eficiente. Essa hipótese se alinha à ideia de que o agente opera em ambiente de mercado altamente competitivo, no qual as decisões individuais não são suficientes para alterar a formação dos preços, e os movimentos de mercado são predominantemente impulsionados por fatores macroeconômicos ou pela agregação das ações de todos os participantes.

5.3.1 Configuração inicial

As simulações a seguir são referentes a seguinte configuração inicial do modelo:

- N : 1180 dias de investimento.
- ω : 200 dias.
- τ : 100 dias.
- Número de Episódios: 1.000.
- MMS_{curta} : 5 dias (MMS_5).
- MMS_{longa} : 20 dias (MMS_{20}).

Os parâmetros α , γ , ϵ e ρ serão otimizados no intervalo de $[0, 1]$.

O modelo será aplicado a três instrumentos financeiros distintos: Índice Bovespa (Ibovespa, símbolo de negociação: \hat{BVSP}), Vale S.A. (VALE3.SA) e Apple Inc. (AAPL). Os dados de cada ativo foram coletados até a data mais recente à época da consulta, 23/08/2024, abrangendo o período de 1180 dias. No entanto, os primeiros 200 dias ($\omega = 200$) compõem a janela inicial de treinamento, de modo que a avaliação do modelo ocorre apenas nos 980 dias subsequentes. Vale destacar que os preços da Apple Inc. estão denominados em dólares americanos (US\$), enquanto os do Índice Bovespa e da Vale S.A. estão em reais brasileiros (R\$).

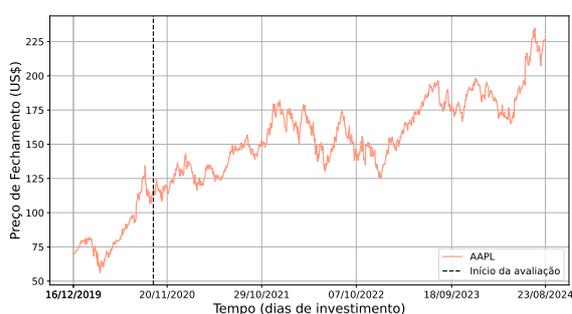
Nesse sentido, a escolha de cada ativo se deu devido suas relevâncias mercadológicas. O Índice Bovespa é o principal indicador do desempenho médio das ações negociadas na B3 (Bolsa de Valores do Brasil). Ele avalia o desempenho de uma carteira teórica que inclui as ações das empresas mais negociadas e representativas da economia brasileira, sendo classificado como índice de mercado. A Vale S.A., por sua vez, é uma das maiores mineradoras do mundo e tem sede no Brasil. As ações da Vale S.A. são títulos de renda variável, especificamente ações ordinárias, que representam a propriedade parcial da empresa sendo negociadas

na B3. Ambas foram escolhidas por serem instrumentos financeiros fundamentais no Brasil, além de possuírem extenso histórico de preços.

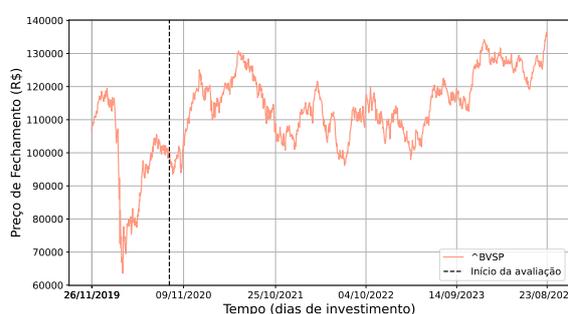
Por sua vez, a Apple Inc. foi selecionada por ser gigante global com capitalização de mercado trilionária. Assim como a Vale S.A., as ações da Apple Inc. são também títulos de renda variável, amplamente negociados por investidores institucionais e de varejo em todo o mundo. O desempenho das ações da Apple reflete frequentemente as tendências no setor de tecnologia. Os gráficos presentes na Figura 5.5 mostram os preços de fechamento dis-

Figura 5.5 – Evolução temporal do preço de fechamento

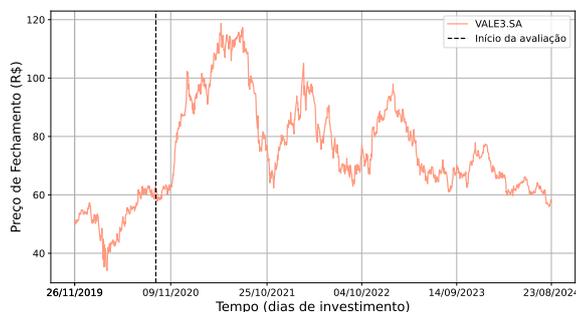
(a) Preço de fechamento da Apple Inc.



(b) Preço de fechamento do Índice Bovespa



(c) Preço de fechamento da Vale S.A.



Fonte: INVESTING (2024).

cutidos. Cada ativo exibe comportamentos globais distintos: Apple Inc. demonstra padrão predominante de crescimento; Índice Bovespa apresenta flutuação entorno de valor médio; já a Vale S.A. tem comportamento oscilatório combinado com decaimento. A princípio, esses comportamentos são genéricos e podem ou não carregar informações relevantes para modelagem. No entanto, é por meio da aplicação dos modelos que esperamos ser capazes de determinar sua relevância. Ademais, o comportamento inicial distinto de cada série temporal pode impactar o desempenho dos modelos. No entanto, o método de treinamento e validação por janela deslizante tende a limitar esse impacto ao curto e médio prazo, tornando-o insignificante na análise de longo prazo.

5.3.1.1 Análise de risco

Para analisarmos o risco de cada ativo e comparar com os resultados dos modelos utilizados nesta pesquisa, utilizaremos as seguintes métricas: Índice de Sharpe e Retorno Anualizado. Isso porque o índice de Sharpe é uma métrica usada para avaliar o retorno de ativos em relação ao risco que eles envolvem, sendo mais útil para determinar a eficiência do ativo ao gerar retorno ajustado ao risco. Já o retorno anualizado permite observar o ganho médio de ativos ao longo do tempo, o que é importante para entender a sua performance geral e compará-lo com outros ativos em horizontes temporais anuais. A Tabela 5.2 apresenta os valores das métricas calculadas, para cada ativo, no período de 1180 dias em que o modelo será treinado e validado.

Tabela 5.2 – Índice de Sharpe (\mathcal{S}) e Retorno Anualizado (\mathcal{R}) dos ativos analisados

Ativo	\mathcal{S}	\mathcal{R} (%)
AAPL	0,640	17,38
^BVSP	0,342	6,21
VALE3.SA	-0,023	-0,71

A AAPL apresenta índice de Sharpe de 0,64, valor que indica um retorno ajustado ao risco considerado satisfatório, embora não excelente ($>1,0$). Esse resultado sugere que a Apple Inc. tem conseguido compensar adequadamente o risco assumido pelos investidores, gerando retornos positivos acima da taxa livre de risco. Adicionalmente, o retorno anualizado de 17,38% reforça o desempenho da empresa, evidenciando sua capacidade de gerar ganhos consistentes ao longo do tempo. Esse desempenho posiciona a AAPL como ativo atraente para investidores que buscam equilíbrio entre risco e retorno, especialmente em contextos de mercados voláteis.

O índice ^BVSP apresenta índice de Sharpe de 0,34, valor positivo, porém inferior ao observado para a AAPL. Esse resultado indica que, embora o retorno ajustado ao risco seja razoável, ele não atinge o mesmo patamar de eficiência observado na Apple Inc. O retorno anualizado de 6,21%, embora positivo, é significativamente mais modesto que o da AAPL, refletindo desempenho mediano em termos de rentabilidade. Assim, o ^BVSP pode ser considerado opção viável para diversificação de portfólio, mas seu apelo é limitado quando comparado a ativos com retornos mais expressivos.

A VALE3.SA, por sua vez, apresenta índice de Sharpe de -0,02, valor negativo que indica que o risco associado ao ativo não foi compensado por retornos adequados. Esse resultado sugere que, no período analisado, a Vale S.A. não conseguiu gerar retornos suficientes para superar a taxa livre de risco, resultando em desempenho abaixo do esperado. Ademais, o retorno anualizado de -0,71% confirma a falta de atratividade do ativo, evidenciando tra-

jetória de perdas ao longo do tempo. Diante desses resultados, a VALE3.SA configura-se como opção pouco recomendável no curto prazo.

É importante ressaltar que, embora o índice de Sharpe e o retorno anualizado sejam métricas importantes, a tomada de decisões de investimento deve considerar também outros fatores, como o contexto macroeconômico, horizonte temporal do investimento e perspectivas setoriais. Dessa forma, a análise aqui apresentada serve como ponto de partida para avaliação dos modelos propostos nesta pesquisa.

5.3.2 Resultados das previsões: *Q-learning*

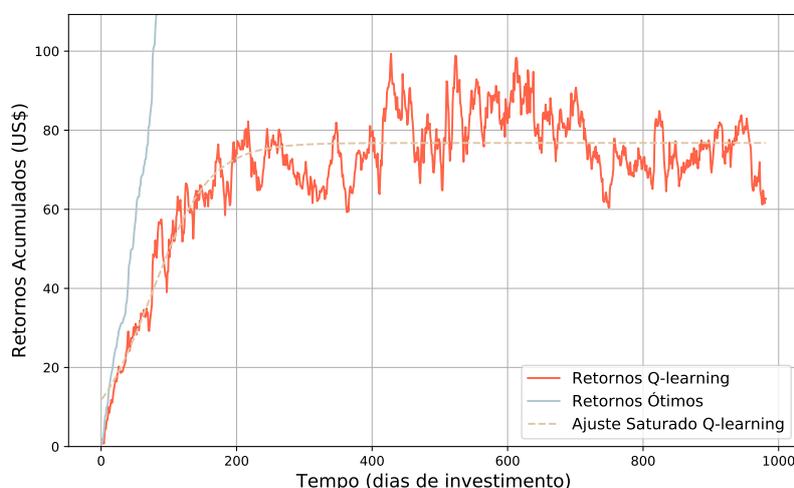
Nesta seção, apresentamos os resultados do modelo de aprendizado no qual a tabela Q é reescrita a cada janela de treinamento de ω dias e os resultados são avaliados ao longo de τ dias (modelo ARM: atualização por reescrita da memória). Esses resultados são comparados com os obtidos pelo modelo com acúmulo contínuo de memória (modelo ACM: acúmulo contínuo de memória), considerando os diferentes instrumentos financeiros analisados. Essa abordagem tem como objetivo investigar os possíveis prejuízos ou benefícios de não preservar memórias de longo prazo, mesmo considerando que o modelo já incorpora mecanismos para enfraquecer memórias antigas.

Os resultados estão representados nas figuras a seguir, onde os retornos acumulados são mostrados em função dos dias de investimento (como discutido na Seção 3.2). Assim, é possível acompanhar a evolução da quantia monetária do agente ao longo dos dias de negociação, assumindo que ele realiza apenas a compra ou venda de uma unidade do instrumento financeiro por dia. Além disso, a curva de Retornos Ótimos refere-se a um modelo idealizado que tomaria sempre as decisões que resultariam em lucro, ou seja, um modelo 100% eficiente.

Em todas as simulações realizadas (independente do ativo), o maior crescimento dos lucros ocorreu de forma linear, com taxa aproximadamente constante ao longo do tempo. Em alguns casos, observamos comportamento de saturação nos retornos acumulados – exemplo visto na Figura 5.6. Embora os retornos acumulados em cenários saturados possam superar os de comportamento linear, este último apresenta maior perspectiva de crescimento a longo prazo. Dessa forma, consideraremos como resultado ótimo aquele que apresenta comportamento linear e maior média dos retornos acumulados nos últimos 5 dias da simulação.

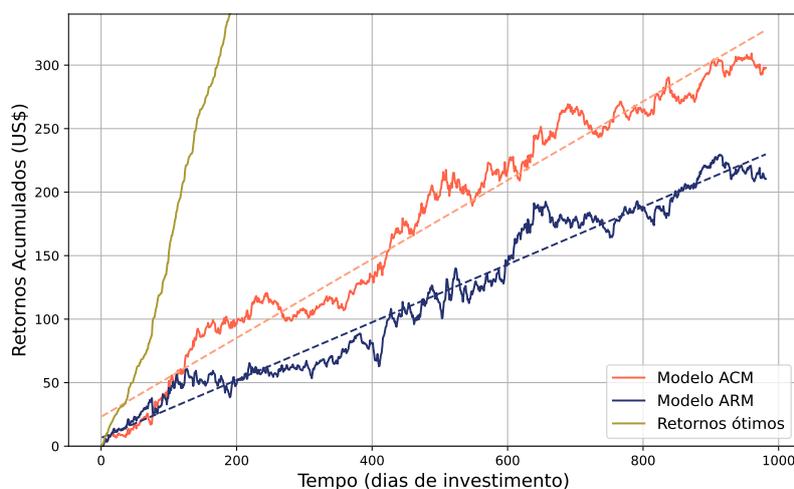
No gráfico apresentado na Figura 5.7, a análise inicial do modelo aplicado aos dados da Apple Inc. mostra que, considerando a otimização dos parâmetros discutida anteriormente, o modelo pode gerar lucros (nas simplificações previamente mencionadas). No entanto, observa-se que, em todos os casos apresentados a seguir, os lucros gerados pelo modelo de *Q-learning* são inferiores aos que seriam obtidos se o modelo fosse 100% eficiente (curva de

Figura 5.6 – Simulação do modelo de investimentos com *Q-learning* para o ativo da Apple Inc. (AAPL) com reescrita da memória – Caso saturado



Fonte: Elaboração própria.

Figura 5.7 – Simulação do modelo de investimentos com *Q-learning* para o ativo da Apple Inc. (AAPL) – Comparação dos modelos ACM e ARM



Fonte: Elaboração própria.

Retornos Ótimos).

No cenário com reescrita da memória (modelo ARM), a eficiência do modelo é de aproximadamente 16,0%³. Já no caso em que há acúmulo contínuo da memória (modelo ACM), a eficiência aumenta para 21,8%. Isso evidencia a capacidade do modelo de se autoajustar, aprimorando seu desempenho ao ter acesso a todo o histórico de informações mapeado conforme a Equação 4.2.

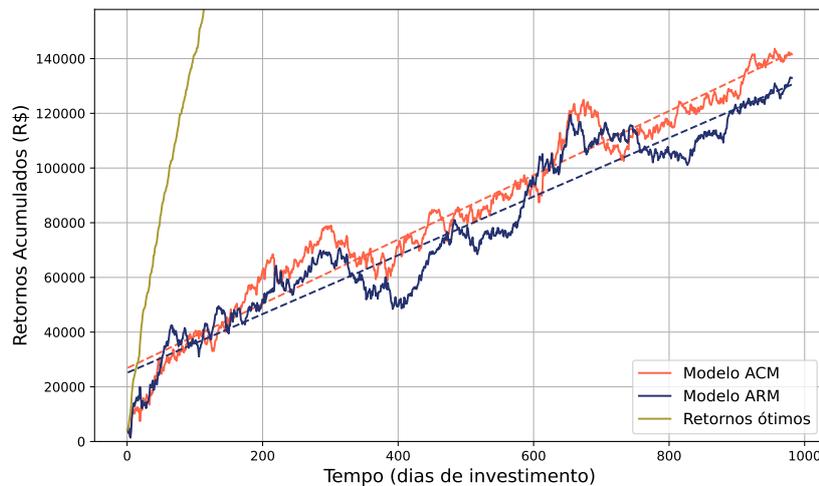
³O cálculo foi realizado com base na média da razão entre o retorno acumulado do *Q-learning* e o retorno acumulado ótimo.

Tabela 5.3 – Eficiência (η) dos modelos ARM e ACM em cada ativo

Ativo	η_{ARM} (%)	η_{ACM} (%)
AAPL	16,0	21,8
$\hat{\text{BVSP}}$	16,5	17,4
VALE3 . SA	12,6	15,3

Transpondo a análise para os ativos da Vale S.A. (Figura 5.9) e do Índice Bovespa (Figura 5.8), podemos concluir que o modelo ACM tem melhor desempenho que o ARM. Entretanto, a diferença torna-se menos significativa para estes ativos comparando com a diferença de desempenho observada na Apple Inc. A Tabela 5.3 apresenta a comparação quantitativa da eficiência de cada modelo para ativo. Em discussão futura, iremos interpretar o impacto da memória na eficiência do modelo e analisar como essa eficiência se relaciona com o risco associado a investir em cada ativo.

Figura 5.8 – Simulação do modelo de investimentos com *Q-learning* para o Índice Bovespa ($\hat{\text{BVSP}}$) – Comparação dos modelos ACM e ARM



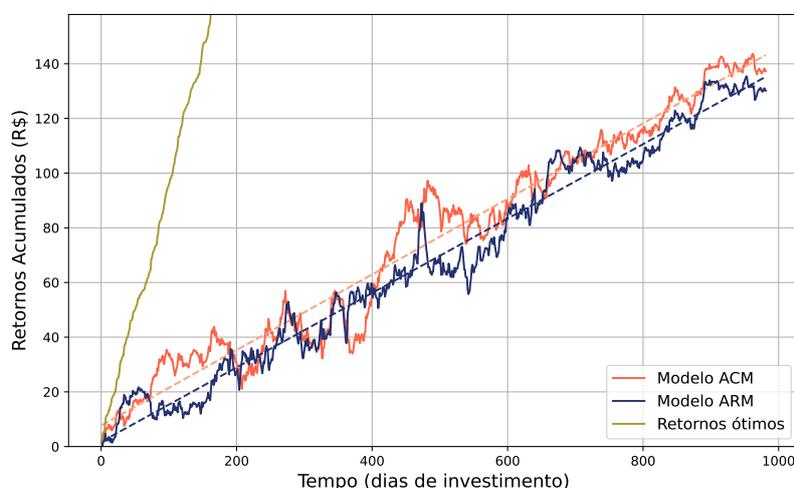
Fonte: Elaboração própria.

5.3.2.1 Análise dos parâmetros otimizados

A análise dos parâmetros será conduzida com base em suas funções descritas nas Seções 4.3 e 5.2, considerando apenas o modelo com acúmulo contínuo de memória (ACM), uma vez que apresentou o melhor desempenho. Além disso, a análise para o modelo com atualização por reescrita de memória (ARM) é completamente análoga. Os valores otimizados para cada ativo estão apresentados na Tabela 5.4.

Começando pelo parâmetro de mais fácil interpretação, ϵ , responsável pelo dilema Ex-

Figura 5.9 – Simulação do modelo de investimentos com *Q-learning* para o ativo da Vale S.A. (VALE3.SA) – Comparação dos modelos ACM e ARM



Fonte: Elaboração própria.

ploração/Exploração, nota-se variação significativa entre os ativos. Para a AAPL, o valor otimizado foi próximo de 0,0, indicando que o agente que melhor otimiza os retornos acumulados prefere explorar consistentemente a tabela Q , maximizando a previsibilidade do modelo. Essa abordagem reflete a previsibilidade do mercado da AAPL, onde padrões consolidados são mais eficazes, resultando na maior eficiência observada (21,8%). Em contraste, os valores elevados de $\epsilon > 0,5$ para \hat{BVSP} e VALE3.SA apontam para a necessidade de maior exploração, provavelmente devido ao risco associado a estes ativos e às complexas dinâmicas de interação nesses mercados. Embora essa estratégia aumente a adaptabilidade, ela compromete a eficiência do *Q-learning*, que alcançou apenas 17,4% e 15,3%, respectivamente.

Tabela 5.4 – Parâmetros de aprendizado do *Q-learning* otimizados (ACM)

Ativo	α	γ	ϵ	ρ
AAPL	0,44	0,33	0,10	0,10
\hat{BVSP}	0,31	0,79	0,67	0,40
VALE3.SA	0,66	0,90	0,80	0,44

Quanto aos parâmetros α e γ , que controlam a dependência do agente à memória e o horizonte temporal das recompensas, respectivamente, observa-se um perfil específico para cada ativo. Para AAPL, o agente apresenta dependência moderada da memória ($\alpha < 0,5$) e foco em recompensas imediatas ($\gamma < 0,5$), configurando modelo mais estável e adaptado a condições de mercado previsíveis, embora menos responsivo a eventos disruptivos. Já para \hat{BVSP} , a priorização de recompensas futuras ($\gamma > 0,5$) sugere perfil mais orientado

a longo prazo, o que pode aumentar os riscos em períodos de instabilidade. Em relação à VALE3.SA, a principal diferença está no alto valor de $\alpha > 0,5$, que eleva a sensibilidade do agente a novas informações. Essa característica, embora aumente a adaptabilidade em mercados voláteis, pode ampliar os riscos em cenários mais estáveis.

O parâmetro ρ , que define a similaridade entre estados, complementa a análise dos demais parâmetros. Para a AAPL, ρ exige que um estado seja no máximo 10% diferente de outro na tabela Q para serem considerados iguais. Já para \hat{BVSP} e VALE3.SA, ρ é mais elevado, permitindo comparar estados com valores até 40% diferentes no caso da \hat{BVSP} , e 44% para a VALE3.SA. Reforçando a ineficiência do modelo de aprendizado para estes dois últimos ativos.

5.4 CONSIDERAÇÕES FINAIS

A análise inicial realizada neste capítulo evidenciou as propriedades distintas do modelo de Ising em redes livres de escala, demonstrando como a topologia da rede e os parâmetros de interação influenciam diretamente propriedades termodinâmicas do sistema. Os resultados obtidos revelaram a existência de um ponto crítico bem definido, cuja posição depende tanto do número de nós quanto do grau mínimo da rede. Além disso, destacaram a importância do acoplamento dos *hubs* na propagação das interações, especialmente sob influência do campo externo. Essas análises não apenas corroboram a capacidade do modelo de reproduzir transições de fase características dos sistemas de *spin*, mas também fornecem base para sua integração com métodos de aprendizado por reforço.

Quanto ao modelo de aprendizado por reforço, os resultados apresentados ao longo deste estudo evidenciam a influência da memória no desempenho do modelo de aprendizado por reforço na tomada de decisão em investimentos. A comparação entre os modelos ARM e ACM demonstra que a retenção da memória histórica no modelo ACM confere ganho de eficiência em relação ao modelo ARM, como evidenciado pela melhoria dos retornos acumulados em todos os ativos analisados.

A AAPL destacou-se como o ativo em que a eficiência do *Q-learning* foi mais pronunciada, atingindo 21,8% no modelo ACM. Esse comportamento pode ser atribuído à menor necessidade de exploração ($\epsilon = 0,10$) e à estabilidade do mercado (de menor risco associado, como discutido), que favorece estratégias baseadas em aprendizado de padrões históricos. Por outro lado, para os ativos \hat{BVSP} e VALE3.SA, a necessidade de maior exploração ($\epsilon > 0,5$) e a elevada dependência da memória ($\alpha > 0,5$) indicam ambientes mais dinâmicos e imprevisíveis, reduzindo a eficiência do modelo em comparação com a AAPL.

Como direções futuras, seria interessante investigar a influência de diferentes arquitetu-

ras de aprendizado por reforço, como modelos baseados em redes neurais profundas (*Deep Q-Networks* - DQN), para mitigar a limitação da discretização de estados. Além disso, explorar a influência de fatores macroeconômicos na aprendizagem do agente pode fornecer percepções sobre a relação entre o mercado financeiro e os mecanismos de decisão do modelo.

Em síntese, os resultados obtidos reforçam a importância da memória e dos parâmetros de aprendizado na adaptação do modelo ao comportamento dos diferentes ativos financeiros. A implementação de mecanismos mais sofisticados para modelagem da incerteza e do risco pode representar avanço na aplicabilidade do *Q-learning* em situações reais de investimento.

6 IMPLEMENTAÇÃO DO MODELO DE ISING NO MERCADO FINANCEIRO

Este capítulo apresenta a metodologia de acoplamento, os procedimentos de simulação e resultados obtidos, enfatizando a eficácia do modelo Q -Ising em comparação com o Q -learning convencional, bem como suas implicações para a compreensão do risco e das dinâmicas de mercado. Na Seção 6.1, detalhamos o processo de acoplamento entre o modelo de Ising e o aprendizado por reforço. A seguir, na Seção 6.2, descrevemos a metodologia utilizada nas simulações. Por fim, na Seção 6.3, apresentamos e discutimos os resultados obtidos com o modelo Q -Ising.

6.1 ACOPLAMENTO DO MODELO DE Q -LEARNING AO MODELO DE ISING

Para realizar o acoplamento entre o modelo de aprendizado por reforço e o modelo de Ising, utilizamos o campo externo h . Neste acoplamento, o agente de Q -learning age como “especialista em investimentos”, cuja opinião influencia a decisão de participação da população (partículas) no mercado. A magnetização do sistema representa a opinião média dessa população. Estamos tratando o modelo de Ising acoplado pelo campo externo ao modelo de Q -learning por “ Q -Ising”, para expressar que as ações do modelo de Ising agora são tomadas considerando o conceito de qualidade do Q -learning.

Agora, cada passo do método de Monte Carlo corresponde a um dia de investimento. Em cada dia t , o campo externo $h = h(t)$ é determinado em função da ação a_t do agente de Q -learning, definida pela expressão desenvolvida a partir da observação dos resultados, a qual apresentou o melhor comportamento:

$$h(t) := 2a_t|\bar{Q}_t| + \frac{1}{2}M_{t-1}, \quad (6.1)$$

onde M_{t-1} representa a magnetização do dia $t - 1$, atuando como fator de amortecimento para suavizar a transição entre as direções do campo externo, promovendo variações mais graduais.

A amplitude média da qualidade das ações a_t para o estado s_t é dada por:

$$\bar{Q}_t = \tanh \left(\frac{\max_a Q(s_t, a) - \min_a Q(s_t, a)}{1 + |\max_a Q(s_t, a)|} \right),$$

em que $\bar{Q}_t \in (-1, 1)$ e $a_t \in \{-1, 0, 1\}$. Assim, o sinal de $h(t)$ indica a expectativa de variação do preço no dia $t + 1$, enquanto sua intensidade é proporcional à qualidade (ou confiança) do sinal gerado pelo agente.

6.2 PROCEDIMENTOS DAS SIMULAÇÕES

Para maximizar a eficiência computacional, a implementação do modelo acoplado foi estruturada da seguinte maneira: inicialmente, realizou-se a otimização do modelo de *Q-learning*, conforme descrito na Seção 5.2, armazenando as informações de ações e as amplitudes médias diárias. Ou seja, os dados do *Q-learning* utilizados para influenciar o modelo de Ising são os mesmos que geraram os retornos acumulados mostrados na Subseção 5.3.2. Em seguida, com os dados otimizados do *Q-learning*, procedeu-se à simulação do modelo de Ising, conforme descrito na Subseção 5.1.1, ajustando os parâmetros de energia de interação J e temperatura T , visando obtermos retornos superiores aos esperados do modelo de aprendizado. Os valores dos parâmetros n e m foram definidos conforme a análise apresentada na Subseção 5.1.3.

6.2.1 Configuração inicial

Como o modelo de *Q-learning* já foi otimizado, seguindo as configurações iniciais descritas na Subseção 5.3.1, resta apenas definir os parâmetros para a criação do modelo de Ising na rede livre de escala:

- Número de nós (n): 40.000.
- Grau mínimo (m): 5.
- Energia de interação (J): varia continuamente de 0,01 a 2,0.
- Temperatura (T): varia continuamente de 0,1 a 100,0.
- Campo externo ($h(t)$): determinado pela Equação 6.1.
- Número de passos de Monte Carlo (N_{passos}): 2.000.

Os parâmetros n e m foram escolhidos com base na análise da estabilidade e do comportamento convergente do modelo de Ising na rede livre de escala.

6.3 RESULTADOS DAS PREVISÕES: *Q*-ISING

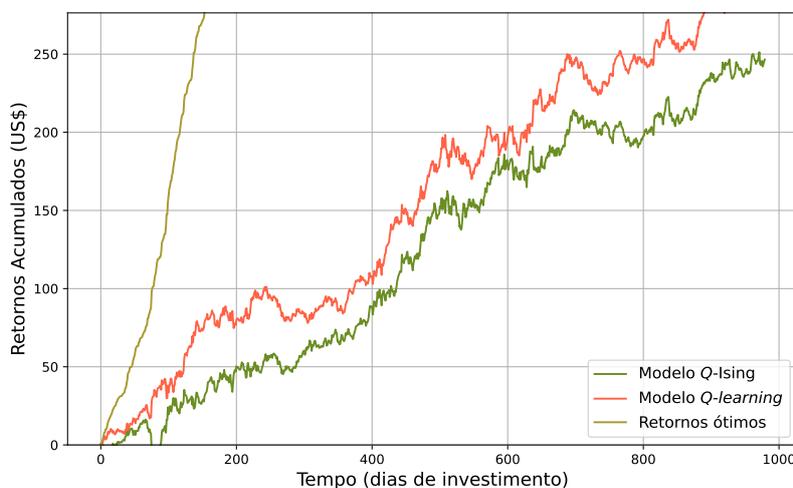
A seguir, são apresentadas Figura 6.1, Figura 6.2 e Figura 6.3, referentes ao desempenho de cada modelo em cada ativo, em que mostram os retornos acumulados em função dos dias de investimento. Cada gráfico apresenta três curvas: retornos ótimos em amarelo; retornos do modelo *Q*-Ising em verde; e retornos do modelo *Q*-learning em vermelho. Consideremos como resultado ótimo do *Q*-Ising aquele no qual os retornos acumulados tivessem menor diferença negativa – ou maior diferença positiva – com relação ao modelo *Q*-learning. Além disso, na Tabela 6.1 constam as eficiências de cada modelo, e também a eficiência relativa: $\Delta\eta = \eta_{Q\text{-Ising}} - \eta_Q$.

Tabela 6.1 – Eficiência (η) dos modelos *Q*-learning e *Q*-Ising em cada ativo em porcentagem

Ativo	η_Q (%)	$\eta_{Q\text{-Ising}}$ (%)	$\Delta\eta$ (%)
AAPL	21,8	11,8	-10,0
^BVSP	17,4	20,3	+2,9
VALE3.SA	15,3	17,8	+2,5

O primeiro resultado notável é que, apenas no ativo da Apple Inc. o modelo de *Q*-learning obteve retornos superiores ao modelo de *Q*-Ising, com eficiência relativa de -10% . Tanto para o Índice Bovespa, quanto para o da Vale S.A., o modelo de *Q*-Ising superou significativamente os ganhos do *Q*-learning, ambos com eficiência relativa maior que 2% . De qualquer forma, embora os ganhos finais do *Q*-Ising na Apple Inc. superem os do modelo *Q*-learning ARM, sua eficiência média permanece inferior.

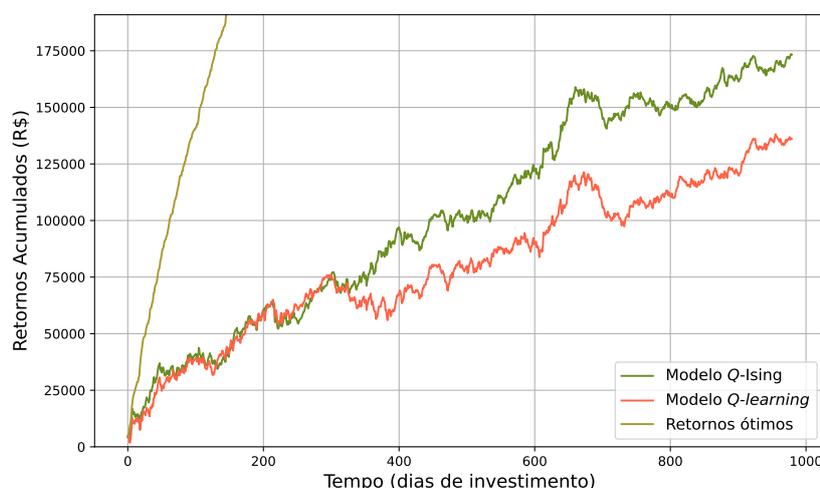
Figura 6.1 – Simulação do modelo de investimentos com *Q*-Ising para o ativo da Apple Inc. (AAPL) – Comparação com modelo *Q*-learning



Fonte: Elaboração própria.

Nos casos em que o modelo de Q -Ising superou o Q -learning em retornos acumulados, observamos que, inicialmente, as curvas de ambos os modelos se sobrepõem, divergindo em momentos posteriores. Esse ponto de divergência ocorre aproximadamente na mesma data (\hat{BVSP} em $t = 326$ dias e $VALE3.SA$ em $t = 322$ dias). No entanto, como tratamos os ativos de forma isolada, não podemos afirmar que esses fenômenos estão diretamente relacionados. Ainda assim, é possível investigar se suas causas, assumidas como independentes, coincidiram em determinado momento.

Figura 6.2 – Simulação do modelo de investimentos com Q -Ising para o ativo do Índice Bovespa (\hat{BVSP}) – Comparação com modelo Q -learning



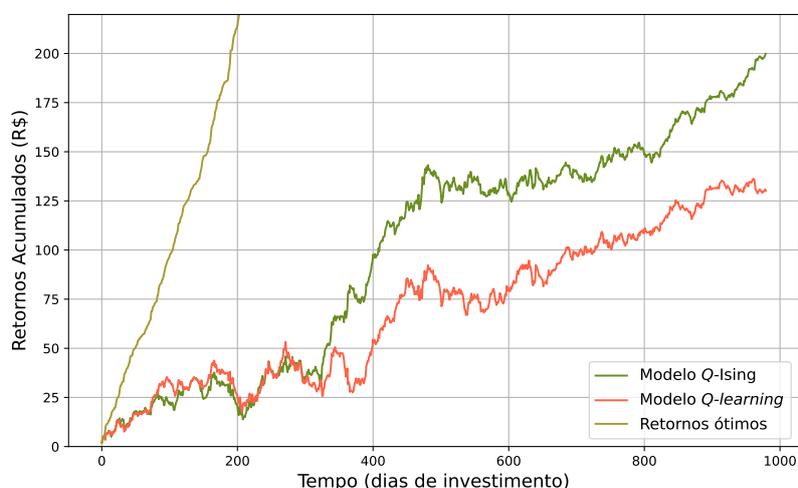
Fonte: Elaboração própria.

Realizando análise preliminar, comparemos o dia em que esse evento ocorre com o comportamento do preço nessa data (Figura 5.5). No caso do Índice Bovespa, em $t' = 326$ dias, o preço de fechamento se encontra na transição entre dois picos. Já para a Vale S.A., em $t'' = 322$ dias, o preço está retomando o crescimento logo após a primeira queda significativa. Ainda que sem validação rigorosa, é natural conjecturar que o melhor desempenho do modelo de Q -Ising esteja relacionado à queda brusca das ações. Como um evento dessa magnitude não ocorre na mesma escala para o ativo da Apple Inc., o controle de volatilidade da opinião do agente Q -learning – promovido pelo Q -Ising – torna-se desnecessário, o que explicaria sua menor eficácia nesse caso.

6.3.1 Análise dos parâmetros otimizados

Segue-se a análise dos resultados do Q -Ising, considerando os parâmetros T e J do modelo de Ising, otimizados com base nos retornos acumulados. Os valores para cada ativo estão apresentados na Tabela 6.2.

Figura 6.3 – Simulação do modelo de investimentos com Q -Ising para o ativo da Vale S.A. (VALE3.SA) – Comparação com modelo Q -learning



Fonte: Elaboração própria.

Para baixas temperaturas, a energia térmica é insuficiente para desorganizar o alinhamento dos *spins*, resultando em estados ordenados, no qual a magnetização média assume valor elevado, próximo ao máximo possível. Esse comportamento foi observado para o ativo AAPL, com $T = 10,55$, indicando que os agentes tendem a compartilhar opinião comum (*spins* alinhados). Por outro lado, para \hat{BVSP} e VALE3.SA, a temperatura otimizada foi significativamente mais alta, permitindo que a energia térmica se tornasse dominante. Nesse regime, os *spins* ficam desordenados, refletindo cenário no qual os agentes apresentam opiniões divergentes sobre a compra ou venda dos ativos.

O parâmetro J representa a tendência dos agentes a alinhar suas opiniões com os demais. Valores mais altos de J indicam que a influência mútua entre os agentes é mais forte, tornando o alinhamento das decisões mais relevante. No caso de AAPL ($J = 1,15$), a interação entre agentes é relativamente moderada, permitindo alinhamento, mas sem dominância absoluta da dinâmica coletiva. Já para \hat{BVSP} ($J = 1,57$) e VALE3.SA ($J = 2,00^1$), a interação entre os agentes é mais intensa, o que sugere que, mesmo em temperaturas altas, há pressão significativa para a formação de padrões de comportamento coletivo, ainda que o estado final seja mais desordenado devido à influência térmica. Os valores mais elevados de J podem ser mais relevantes no contexto dos *hubs*, promovendo concordâncias de opiniões localizadas, enquanto globalmente divergem.

¹Este é o valor máximo permitido pela otimização, o que pode indicar que o modelo teria desempenho ainda melhor para $J_{\max} > 2,00$. No entanto, essa limitação não compromete a análise que se segue.

6.3.1.1 Interpretação conjunta dos parâmetros

A relação entre T e J determina o comportamento global do sistema:

- **AAPL**: Baixa temperatura e interação moderada resultam em estados relativamente ordenados, com agentes tendendo a opiniões mais homogêneas.
- **^BVSP**: Temperatura elevada e interação mais intensa levam a regimes dominados pela desordem térmica, mas ainda com alguma correlação entre agentes.
- **VALE3 . SA**: Alta temperatura e forte interação indicam sistemas onde a influência dos vizinhos é significativa, mas as flutuações térmicas impedem convergência clara das opiniões.

Essa análise sugere que, para diferentes ativos, a dinâmica dos agentes oscila entre regimes mais ordenados (forte consenso) e mais desordenados (opiniões divergentes), dependendo do equilíbrio entre temperatura e interação. Notavelmente, mercados caracterizados por maior dispersão de opiniões refletem melhor a natureza do mercado de ações, favorecendo o desempenho do modelo em termos de retornos acumulados.

Tabela 6.2 – Parâmetros do modelo de Ising otimizados

Ativo	T	J
AAPL	10,55	1,15
^BVSP	55,76	1,57
VALE3 . SA	71,85	2,00

Um aspecto interessante é que o modelo de Q -Ising apresentou seus melhores desempenhos justamente em mercados com maior risco. Esse risco pode ser interpretado tanto pelas próprias quedas bruscas dos preços quanto pelos valores dos parâmetros otimizados do modelo, sugerindo que a adaptação a dinâmicas mais caóticas é característica do Q -Ising.

6.4 CONSIDERAÇÕES FINAIS

O modelo de Q -Ising demonstrou capacidade de avaliar risco de diferentes ativos, principalmente em termos das quantidades físicas otimizadas T e J . Ao menos, em comparação com a análise feita de cada ativo utilizando-se do Índice Sharp e do Retorno Anualizado. Um aspecto surpreendente foi que o modelo apresentou melhor desempenho justamente nos ativos com maior risco associado, enquanto o Q -learning teve melhor desempenho no ativo com maior retorno ajustado ao risco. Isso reforça a importância do papel do modelo na

distribuição da decisão de compra e venda entre n agentes, capturando dinâmicas coletivas que um único agente isolado não considera. No mais, evidencia a relevância do acúmulo contínuo da memória, especialmente em séries temporais mais estáveis.

Além disso, a interpretação dos parâmetros físicos do modelo sugere paralelo interessante com características do mercado financeiro. A temperatura (T) parece estar associada à especulação e à dispersão de opiniões entre os agentes, enquanto a interação (J) pode representar a coesão entre investidores e a força das tendências de mercado. Essa analogia abre espaço para uma investigação mais profunda do potencial do modelo na análise de risco e formação de preços.

No entanto, algumas limitações devem ser consideradas. O valor máximo imposto a J pode ter restringido o desempenho do modelo, sugerindo que simulações com maior espaço de busca para esse parâmetro poderiam levar a resultados ainda mais precisos. No mais, a estrutura da rede utilizada é estática, enquanto mercados reais são dinâmicos e adaptativos, bem como o parâmetro J poderia ser dinâmico e não homogêneo. Investigações futuras poderiam explorar redes temporais ou mecanismos de reforço que permitam que a estrutura da interação entre agentes evolua ao longo do tempo.

Por fim, uma possível linha de pesquisa futura envolve a incorporação de variáveis externas, como fluxo de notícias financeiras e indicadores macroeconômicos, para testar a robustez do modelo em cenários mais realistas. Aliás, diferentes escalas temporais poderiam ser estudadas para avaliar a aplicabilidade do Q -Ising tanto no curto quanto no longo prazo.

7 CONCLUSÕES

Os resultados obtidos neste estudo evidenciam o potencial da abordagem híbrida entre o modelo de Ising e o aprendizado por reforço na modelagem do comportamento dos mercados financeiros. A aplicação do *Q-learning* e sua integração com o modelo de Ising permitiram análise aprofundada das interações entre agentes financeiros e suas decisões de investimento, destacando a influência da memória e das interações locais na previsão do comportamento dos ativos.

Os testes realizados com diferentes ativos financeiros mostraram que a eficiência do modelo varia conforme as características do mercado. Para ativos mais previsíveis, como AAPL, o modelo demonstrou desempenho superior, sugerindo que ambientes mais estáveis favorecem a aplicação de técnicas baseadas em aprendizado de padrões históricos. Já para ativos como \hat{BVSP} e VALE3.SA, onde há maior risco associado ao investimento, bem como menor retorno, e necessidade de exploração, o modelo apresentou queda no desempenho. Isso reforça a hipótese de que mercados mais dinâmicos exigem técnicas que combinem aprendizado por reforço com métodos sofisticados de previsão e gestão de risco.

A interpretação dos parâmetros físicos do modelo de Ising sugere relação interessante com variáveis econômicas. A temperatura pode estar associada à dispersão de opiniões no mercado, enquanto a interação entre vizinhos representa a coesão entre investidores. Essas conclusões indicam que a modelagem física pode oferecer novas perspectivas sobre fenômenos financeiros reais, possibilitando avanços na análise de risco e na tomada de decisões estratégicas.

Apesar dos avanços apresentados, algumas limitações do modelo devem ser consideradas. A estrutura da rede utilizada é estática, enquanto os mercados reais são dinâmicos e adaptativos. Além disso, a discretização dos estados no *Q-learning* pode limitar a capacidade de generalização do modelo. Como trabalhos futuros, propõe-se a investigação de abordagens baseadas em redes neurais profundas, como *Deep Q-Networks* (DQN), e a inclusão de variáveis macroeconômicas no processo de decisão do agente.

Além disso, embora a validação do modelo tenha sido realizada em janelas de treino-teste, ela não considerou a sincronização com a publicação dos dados financeiros. Caso essa abordagem fosse adotada, a performance real do modelo poderia ser inferior à apresentada nesta pesquisa, uma vez que a avaliação ocorreria diariamente. Ainda nesse sentido, a separação da otimização entre o *Q-learning* e, posteriormente, o modelo de Ising pode ter impactado negativamente o desempenho. Esse problema poderia ser mitigado por abordagens nas quais a otimização ocorra de forma contínua, à medida que o modelo percorre os

dados financeiros em tempo real.

Em suma, este estudo demonstrou a viabilidade do uso de modelos físicos na análise de mercados financeiros reais, abrindo caminho para pesquisas que explorem ainda mais a interseção entre física estatística, aprendizado de máquina e análise de séries temporais ruidosas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world-wide web. **nature**, Nature Publishing Group UK London, v. 401, n. 6749, p. 130–131, 1999. 11, 16
- ALTMAN, E. I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. **The Journal of Finance**, JSTOR, v. 23, n. 4, p. 589–609, 1968. 20
- AUCHINCLOSS, A. H.; GARCIA, L. M. T. Brief introductory guide to agent-based modeling and an illustration from urban health research. **Cadernos de saude publica**, SciELO Public Health, v. 31, p. 65–78, 2015. 23
- BABAEIZADEH, M. et al. Reinforcement learning through asynchronous advantage actor-critic on a gpu. **arXiv preprint arXiv:1611.06256**, 2016. 42
- BARABÁSI, A.-L. **Network Science**. [S.l.]: Cambridge University Press, 2016. 12, 16, 17, 32
- BAXTER, R. J. **Exactly solved models in statistical mechanics**. [S.l.]: Elsevier, 2016. 5
- BONDY, J. A.; MURTY, U. S. R. **Graph theory**. [S.l.]: Springer Publishing Company, Incorporated, 2008. 12
- BURY, T. Statistical pairwise interaction model of stock market. **The European Physical Journal B**, Springer, v. 86, p. 1–7, 2013. 2
- CIVIDINO, D.; WESTPHAL, R.; SORNETTE, D. Multiasset financial bubbles in an agent-based model with noise traders' herding described by an n-vector ising model. **Physical Review Research**, APS, v. 5, n. 1, p. 013009, 2023. 1
- DIGNUM, V. **Handbook of research on multi-agent systems: semantics and dynamics of organizational models**. 1. ed. [S.l.]: Information Science Publishing, 2010. ISBN 9781605662572; 1605662577; 1605662569; 9781605662565. 22
- DOROGOVTSEV, S. Lectures on complex networks. 2010. 11
- ELTON, E. J.; GRUBER, M. J.; BLAKE, C. R. The persistence of risk-adjusted mutual fund performance. **Journal of Business**, JSTOR, p. 133–157, 1996. 20
- FAMA, E. F.; FRENCH, K. R. The cross-section of expected stock returns. **The Journal of Finance**, Wiley Online Library, v. 47, n. 2, p. 427–465, 1992. 19
- GUIMARAES, W.; LIMA, L. Self-organizing three-dimensional ising model of financial markets. **Physical Review E**, American Physical Society, v. 103, n. 6, p. 062130, 2021. 1, 5
- HOFSTAD, R. V. D. Random graphs and complex networks. **Lecture notes**, 2013. 14
- HUANG, R.; SHETH, A. Transition prediction in the ising-model. **PLOS ONE**, Public Library of Science, v. 16, n. 9, p. e0259177, 2021. 1, 5
- HULL, J. **Options, Futures, and Other Derivatives: Global Edition (Custom for QMUL)**. [S.l.]: Pearson, 2012. 20

INVESTING. **Dados financeiros**. 2024. <<https://br.investing.com>>. Acesso em: 24 de ago. de 2024. 48

ISING, E. Beitrag zur theorie des ferromagnetismus. **Zeitschrift für Physik**, Springer, v. 31, n. 1, p. 253–258, 1925. 4

JOSHI, A. et al. A multi-language clir with classification by rational agents. In: **2013 International Conference on Information Systems and Computer Networks**. [S.l.: s.n.], 2013. p. 94–97. 23

KRAWIECKI, A. Microscopic spin model for the stock market with attractor bubbling on scale-free networks. **Journal of Economic Interaction and Coordination**, Springer, v. 4, p. 213–220, 2009. 1, 2

LILLICRAP, T. Continuous control with deep reinforcement learning. **arXiv preprint arXiv:1509.02971**, 2015. 42

LIMA, T. F. M. de et al. Modelagem de sistemas baseada em agentes: Alguns conceitos e ferramentas. **Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Brasil**, p. 25–30, 2009. 22

LOPES, J. L. **Ciência e liberdade: escritos sobre ciência e educação no Brasil**. [S.l.]: Editora Ufrj, 1998. i

MACAL, C. M.; NORTH, M. J. Agent-based modeling and simulation. In: **Proceedings of the 2009 Winter Simulation Conference (WSC)**. [S.l.: s.n.], 2009. p. 86–98. 23

MNIH, V. et al. Human-level control through deep reinforcement learning. **nature**, Nature Publishing Group UK London, v. 518, n. 7540, p. 529–533, 2015. 42

OSMANKOVIĆ, D.; KONJICIJA, S. Implementation of qlearning algorithm for solving maze problem. In: IEEE. **2011 proceedings of the 34th international convention MIPRO**. [S.l.], 2011. p. 1619–1622. 22, 27

RODERICK, M.; MACGLASHAN, J.; TELLEX, S. Implementing the deep q-network. **arXiv preprint arXiv:1711.07478**, 2017. 42

ROTH, A. M. et al. Conservative q-improvement: Reinforcement learning for an interpretable decision-tree policy. **arXiv preprint arXiv:1907.01180**, 2019. 42

SALINAS, S. R. **Introdução à física estatística**. [S.l.]: Edusp, 1999. 10

SINGH, S. P. The ising model: Brief introduction and its application. In: **Solid State Physics-Metastable, Spintronics Materials and Mechanics of Deformable Bodies-Recent Progress**. [S.l.]: IntechOpen, 2020. 4

SORNETTE, D. Self-fulfilling ising model of financial markets. **arXiv preprint arXiv:physics/0503230**, 2005. Disponível em: <<https://arxiv.labs.arxiv.org/html/physics/0503230>>. 1, 5

SORNETTE, D. Physics and financial economics (1776–2014): puzzles, ising and agent-based models. **Reports on progress in physics**, IOP Publishing, v. 77, n. 6, p. 062001, 2014. 4

SORNETTE, D. Financial price dynamics and phase transitions in the stock markets. **The European Physical Journal B**, Springer, v. 84, n. 2, p. 195–202, 2021. 1, 5

STROGATZ, S. H. Exploring complex networks. **nature**, Nature Publishing Group UK London, v. 410, n. 6825, p. 268–276, 2001. 11

SUTTON, R. S.; BARTO, A. G. **Reinforcement learning: An introduction**. [S.l.]: MIT press, 2018. 24, 26

TAKAISHI, T. Large-scale simulation of multi-asset ising financial markets. In: IOP PUBLISHING. **Journal of Physics: Conference Series**. [S.l.], 2017. v. 820, n. 1, p. 012016. 1

TAN, H. Reinforcement learning with deep deterministic policy gradient. In: IEEE. **2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA)**. [S.l.], 2021. p. 82–85. 42

TORQUATO, S. Toward an ising model of cancer and beyond. **Physical biology**, IOP Publishing, v. 8, n. 1, p. 015017, 2011. 4

WATKINS, C. J.; DAYAN, P. Q-learning. **Machine learning**, Springer, v. 8, p. 279–292, 1992. 25, 26, 27

WITTMANN, T. **Agent-based models of energy investment decisions**. [S.l.]: Springer Science & Business Media, 2008. 23