

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL

JEAN PIERRE JARRIER CONTI

**REDES NEURAS RECORRENTES E EXPOENTE DE LYAPUNOV
APLICADOS A SÉRIES TEMPORAIS FINANCEIRAS**

DISSERTAÇÃO

CURITIBA

2019

JEAN PIERRE JARRIER CONTI

**REDES NEURAIS RECORRENTES E EXPOENTE DE LYAPUNOV
APLICADOS A SÉRIES TEMPORAIS FINANCEIRAS**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de “Mestre em Ciências” – Área de Concentração: Engenharia da Computação.

Orientador: Heitor Silvério Lopes

CURITIBA

2019

Dados Internacionais de Catalogação na Publicação

Conti, Jean Pierre Jarrier

Redes neurais recorrentes e expoente de Lyapunov aplicados a séries temporais financeiras [recurso eletrônico] / Jean Pierre Jarrier Conti.-- 2019.

1 arquivo texto (81 f.) : PDF ; 1,63 MB.

Modo de acesso: World Wide Web

Título extraído da tela de título (visualizado em 25 set. 2019)

Texto em português com resumo em inglês

Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Curitiba, 2019

Bibliografia: f. 73-76

1. Engenharia elétrica - Dissertações. 2. Redes neurais (Computação). 3. Análise de séries temporais. 4. Aprendizado do computador. 5. Especulação. 6. Investimentos de capital. 7. Funções de Lyapunov. I. Lopes, Heitor Silvério. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. III. Título.

CDD: Ed. 22 – 621.3

Biblioteca Central da UTFPR, Câmpus Curitiba
Bibliotecário: Adriano Lopes CRB-9/1429

TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº ____

A Dissertação de Mestrado intitulada “**Redes Neurais Recorrentes e Expoente de Lyapunov aplicados à Séries Temporais Financeiras.**” defendida em sessão pública pelo(a) candidato(a) **Jean Pierre Jarrier Conti**, no dia **29 de agosto de 2019**, foi julgada para a obtenção do título de Mestre em Ciências, área de concentração **Engenharia De Computação**, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial.

BANCA EXAMINADORA:

Prof(a). Dr(a). Heitor Silvério Lopes - Presidente (UTFPR)

Prof(a). Dr(a). André Eugenio Lazzaretti - (UTFPR)

Prof(a). Dr(a). Elisangela Ferretti Manffra - (PUCPR)

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 29 de agosto de 2019.

“Um raciocínio lógico leva você de A a B. A imaginação leva você a qualquer lugar que você quiser.” — Albert Einstein

RESUMO

CONTI, Jean. REDES NEURAIAS RECORRENTES E EXPOENTE DE LYAPUNOV APLICADOS A SÉRIES TEMPORAIS FINANCEIRAS. 81 f. Dissertação – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

O estudo da formação de preço em ativos do mercado financeiro é considerado um dos tópicos mais desafiadores deste setor. A possibilidade de auferir lucro durante as oscilações intra-diárias torna a previsibilidade um dos problemas mais relevante para este segmento. Embora economistas, acadêmicos e profissionais da área tenham opiniões divergentes quanto à possibilidade de se prever preços, estudos recentes consideram existir algum grau de previsibilidade em séries financeiras. Este indício considera as séries de preços como um sistema não linear e caótico; onde neste caso, ao menos uma previsibilidade de curto prazo poderia ser obtida. Dentro deste contexto, métodos de Aprendizagem de Máquina e Aprendizagem Profunda, aliado a sistemas de apoio a tomada de decisão têm sido utilizados buscando prever a direção dos movimentos. Neste trabalho, o problema da previsão de direção do próximo minuto é abordado a partir da perspectiva de uma tarefa de classificação. Para tal, cinco ativos da Bolsa de Valores Brasileira foram escolhidos utilizando como premissa a liquidez dos ticks no período. Um método de supervisão de três classes foi aplicado e dez indicadores técnicos foram utilizados como atributos para os ativos escolhidos. Dois conjuntos de dados foram construídos onde no primeiro, o valor contínuo dos indicadores foi mantido enquanto no segundo todos os valores foram discretizados. Utilizando Máquinas de Vetores de Suporte, Floresta Aleatória, MLP e LSTM três métodos foram aplicados ao processo de classificação. O primeiro método comparou os resultados entre o conjunto contínuo e o conjunto criado através de indicadores discretizados. Mantendo apenas o classificador que apresentou o melhor resultado o segundo experimento acrescentou quatro atributos utilizando como cálculo o expoente de Lyapunov da série. O objetivo deste experimento foi investigar se os atributos poderiam contribuir para a melhora da classificação de todos os ativos. Por fim, um experimento final utilizou o cálculo do máximo expoente de Lyapunov como controle de treinamento para o classificador. O objetivo neste caso foi excluir trechos das séries onde o valor do expoente de Lyapunov indicasse presença de caos. Os resultados mostraram que embora a discretização dos indicadores tenha contribuído para a maioria dos classificadores, redes LSTM apresentaram um desempenho melhor nos conjuntos de dados contínuos na maioria dos ativos abordados. Em linhas gerais foi possível concluir que cada ativo se beneficiou de um conjunto de parâmetros específicos. Os experimentos demonstraram não haver um método que apresente um desempenho ideal para todos os ativos escolhidos no período definido.

Palavras-chave: Aprendizado Profundo; Séries Temporais Financeiras; Expoente de Lyapunov

ABSTRACT

CONTI, Jean. RECURRENT NEURAL NETWORKS AND LYAPUNOV EXPONENT APPLIED TO FINANCIAL TIME SERIES. . 81 f. Dissertação – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

The study of financial market asset pricing is considered one of the most relevant subjects in this segment. The possibility to obtain profit during the intraday oscillations turned the predictability problem relevant to this area. Although economists, academics and industry professionals have different views about the possibility of predicting prices, some recent studies consider that there is some degree of the predictability in financial series. This evidence considers the prices series as a chaotic and non-linear system where, at least, short term predictability may take place. In this way, Machine Learning and Deep Learning methods have been used together as decision support systems to forecast the movement direction of prices. In this work, the problem of predicting the next minute is addressed from the classification perspective. Five assets of the Brazilian Stock Exchange were chosen based on the ticks' liquidity in the period. A three-class supervision method was applied and ten technical indicators were used as attributes for the assets chosen. Two data sets were constructed: one using the continuous value of the indicators, and other using discretized values. Using Support Vector Machine, Random Forest, MLP and LSTM three methods were applied to the classification process. The first method compares the results between the continuous and the discretized data sets. Using the classifier that presented the best result, the second experiment added four attributes using the Lyapunov exponent calculation. The objective of this experiment was to investigate if the attributes could contribute to improve the classification of all assets. Finally, a final experiment used the calculation of the maximum Lyapunov exponent as a training control for the classifier. The objective, in this case, was to exclude portions of the series where the value of the Lyapunov exponent indicated chaos. The result showed that although indicator discretization contributed positively for most classifiers, LSTM networks showed a significant improvement in continuous data sets with superior performance in most cases. In general, it was concluded that each asset benefited from a set of specific methods. The experiments demonstrated that there is no general method that presents an optimal performance for all assets in the defined period studied.

Keywords: Deep Learning; Financial Time Series; Lyapunov Exponent

LISTA DE FIGURAS

FIGURA 1	– Exemplo da influência das notícias	18
FIGURA 2	– Exemplo de um gráfico de tendência	19
FIGURA 3	– Exemplo de uma série de retornos	25
FIGURA 4	– Exemplo do efeito borboleta	27
FIGURA 5	– Exemplo do atrator estranho de Lorenz	28
FIGURA 6	– Exemplo das trajetórias do expoente de Lyapunov	29
FIGURA 7	– Expoente de Lyapunov aplicado a um mapa logístico	30
FIGURA 8	– Exemplo de associação entre dados e classes	31
FIGURA 9	– Exemplo de uma curva ROC	33
FIGURA 10	– Exemplo da separação de um hiperplano em um classificador SVM.	34
FIGURA 11	– Unidade de memória de uma LSTM.	37
FIGURA 12	– Visão Geral do método de indicadores discretizados e contínuos.	49
FIGURA 13	– Visão Geral do método do atributo previsor de Lyapunov	51
FIGURA 14	– Deslocamento da janela para o método do atributo previsor de Lyapunov .	53
FIGURA 15	– Diagrama geral de blocos do método de discretização	54
FIGURA 16	– Deslocamento da janela para o método da janela Lyapunov	55
FIGURA 17	– Retornos minuto a minuto dos ativos de curto prazo	57
FIGURA 18	– Retornos sem <i>outliers</i> dos ativos de curto prazo	57
FIGURA 19	– Experimento 01: Variação da janela de dados na rede LSTM.	59
FIGURA 20	– Experimento 01: Resultado para todos os classificadores.	62
FIGURA 21	– Experimento 02: Diferentes janelas de Lyapunov utilizada no experimento.	64
FIGURA 22	– Experimento 02: Resultado para LSTM com métrica <i>F-measure</i>	65
FIGURA 23	– Experimento 03: Variação da janela de dados no experimento.	67
FIGURA 24	– Experimento 03: LSTM e métrica <i>F-measure</i>	68

LISTA DE TABELAS

TABELA 1	– Tabela com a visão geral dos trabalhos analisados.	43
TABELA 2	– Tabela com os ativos selecionados para análise	44
TABELA 3	– Tabela com os atributos das séries financeiras	47
TABELA 4	– Parâmetros utilizados nos experimentos.	48
TABELA 5	– Tabela com os atributos discretos das séries financeiras	50
TABELA 6	– Distribuição das classes dos ativos.	58
TABELA 7	– Experimento 01: Variação de janela LSTM.	59
TABELA 8	– Experimento 01: Resultado dos classificadores - ROC.	60
TABELA 9	– Experimento 01: Resultado dos classificadores - <i>F-measure</i>	60
TABELA 10	– Experimento 02: Diferentes janelas Lyapunov.	64
TABELA 11	– Experimento 02: Atributo previsor de Lyapunov - <i>F-measure</i>	65
TABELA 12	– Experimento 03: Janela de Lyapunov no experimento.	67
TABELA 13	– Experimento 03: Janela de Lyapunov - <i>F-measure</i>	68
TABELA 14	– Experimento 01: Resultado dos experimentos e dos testes estatísticos	77
TABELA 15	– Experimento 02: Resultado dos experimentos e dos testes estatísticos	78
TABELA 16	– Experimento 02: Resultado dos experimentos e dos testes estatísticos	79
TABELA 17	– Experimento 03: Resultado dos experimentos e dos testes estatísticos	80
TABELA 18	– Experimento 03: Resultado dos experimentos e dos testes estatísticos	81

LISTA DE SIGLAS

HME	Hipótese de Mercado Eficiente
IC	Inteligência Computacional
IA	Inteligência Artificial
AP	Aprendizado Profundo
AT	Análise Técnica
SMA	Média móvel simples
WMA	Média móvel ponderada
MOM	Indicador Momentum
IFR	Índice de Força Relativa
MACD	Convergência / Divergência das Médias Móveis
MME	Média Móvel Exponencial
A/D	Indicador de Acumulação e Distribuição
EL	Expoente de Lyapunov
MC	Matriz de confusão
F1	<i>F-measure</i>
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
RF	Random Forest
RNAs	Redes Neurais Artificiais
MLP	Multilayer Perceptron
RNNs	Redes Neurais Recorrentes
LSTM	Long Short-Term Memory
NSE	National Stock Exchange of India Limited
NB	Naive Bayes
DNN	Deep Neural Network

SUMÁRIO

1 INTRODUÇÃO	13
1.1 MOTIVAÇÃO	15
1.2 OBJETIVOS	16
1.2.1 Objetivo Geral	16
1.2.2 Objetivos Específicos	16
1.2.3 Estrutura da Dissertação	16
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1 ANÁLISE TÉCNICA	17
2.2 INDICADORES TÉCNICOS	19
2.3 SÉRIE DE RETORNO E VOLATILIDADE	24
2.4 CAOS DETERMINÍSTICO	25
2.5 EXPOENTE DE LYAPUNOV	28
2.6 CLASSIFICADORES	30
2.6.1 Métricas de desempenho	31
2.6.2 SVM	33
2.6.3 Floresta Aleatória	34
2.6.4 MLP	35
2.6.5 LSTM	36
2.7 SÉRIES FINANCEIRAS	38
2.8 TRABALHOS RELACIONADOS	39
2.8.1 Previsão da direção	40
2.8.2 Análise de Previsibilidade	41
3 METODOLOGIA	44
3.1 SÉRIES FINANCEIRAS	44
3.1.1 Pré-processamento	45
3.1.2 Aprendizado Supervisionado	46
3.1.3 Recursos utilizados e parâmetros definidos	47
3.2 MÉTODO DOS INDICADORES DISCRETIZADOS	48
3.3 MÉTODO DO ATRIBUTO PREVISOR DE LYAPUNOV	51
3.4 MÉTODO DA JANELA DE LYAPUNOV	52
4 RESULTADOS E DISCUSSÃO	56
4.1 SÉRIES FINANCEIRAS	56
4.2 EXPERIMENTO 01: INDICADORES DISCRETIZADOS E CONTÍNUOS	57
4.2.1 Experimentos Realizados	58
4.2.1.1 Determinação da Janela LSTM	58
4.2.1.2 Classificação	59
4.2.2 Discussão	60
4.3 EXPERIMENTO 02: ATRIBUTO PREVISOR DE LYAPUNOV	63
4.3.1 Experimentos Realizados	63
4.3.1.1 Determinação da janelas de Lyapunov	63
4.3.1.2 Classificação	64

4.3.2 Discussão	65
4.4 EXPERIMENTO 03: JANELA DE LYAPUNOV	66
4.4.1 Experimentos Realizados	66
4.4.1.1 Determinação de diferentes janelas de Lyapunov	66
4.4.1.2 Classificação	67
4.4.2 Discussão	68
5 CONCLUSÃO E TRABALHOS FUTUROS	70
5.1 TRABALHOS FUTUROS	71
REFERÊNCIAS	73
Apêndice A – TESTES ESTATÍSTICOS	77

1 INTRODUÇÃO

O mercado de capitais desempenha um papel fundamental no desenvolvimento econômico da sociedade moderna. Em linhas gerais, as bolsas de valores promovem o encontro de poupadores e tomadores, flexibilizando as negociações de ativos financeiros. Ao tornar isto possível, são criadas condições para que empresas tomem recursos diretamente dos poupadores, viabilizando seus projetos de investimento. Os poupadores, por outro lado, aceitam a oportunidade de participar dos projetos assumindo os riscos decorrentes das operações¹.

Ao decidirem negociar, os participantes do mercado fecham negócios causando oscilação nas cotações dos ativos financeiros ao longo do tempo. O histórico destas negociações é definido como uma coleção de informações temporais. A possibilidade de obter lucros entre os intervalos das negociações torna o problema da previsibilidade das ações um dos mais importantes assuntos deste segmento (WANG et al., 2012).

Economistas, acadêmicos e profissionais do setor divergem quanto à possibilidade de se prever preços em séries temporais financeiras. Diversas discussões em torno do tema surgiram até que Malkiel e Fama (1970) apresentaram a Hipótese de Mercado Eficiente (HME). Ela sugere que o preço atual reflete toda a informação existente e, portanto, não é possível utilizar nenhuma informação conhecida, ou passada, para prever preços. Seu estudo pode ser considerado um marco no que diz respeito à previsibilidade do mercado de ações. Por diversos anos ela foi aceita e ainda hoje causa contradições acerca do tema.

Uma segunda hipótese é chamada de *random-walk*. Ela foi apresentada por Malkiel (1973) e afirma que os valores das ações variam independentemente do histórico, ou seja, o preço de amanhã será influenciado apenas pela informação de amanhã, independente do preço de hoje. Embora estas duas hipóteses afirmem que não há como prever o preço de uma ação, no final do século 20 outros estudos passaram a refutar tais hipóteses e considerar que pode haver algum tipo de previsibilidade nos preços das ações (GROSSMAN; STIGLITZ, 1980; LO; MACKINLAY, 2002; SITTE; SITTE, 2002).

¹<https://www.investidor.gov.br/publicacao/LivrosCVM.html> - Acesso em: 01/06/2019.

Desde então, outros autores começaram a investigar o tema da previsibilidade do mercado considerando a existência do que tem sido chamado de caos determinístico. Dentro desta hipótese, os preços dos ativos se comportam como um sistema não-linear caótico e, portanto, intervalos de previsibilidade, ainda que no curto prazo, podem existir (GOMES; JUNIOR, 2010). Peters (1991) e, posteriormente, Willey (1992), constataram a presença de caos em diversas séries financeiras. Tais estudos utilizaram dimensão fractal e o expoente de Lyapunov para tal constatação. Dentro deste contexto, uma nova perspectiva sobre o tema da previsibilidade surge.

Por outro lado, embora exista algum tipo de previsibilidade, as séries financeiras envolvem uma complexidade significativa (ABU-MOSTAFA; ATIYA, 1996). Cada país adota particularidades em seu sistema financeiro e suas variações de taxa de juros, possíveis movimentações políticas, notícias locais, e outras afetam o comportamento das séries criando diferentes propriedades em cada mercado (SALES, 2017). Além disto, a este conjunto de variáveis é adicionado um alto grau de incerteza, tornando o processo dinâmico, ruidoso, não-linear e caótico por natureza (BLANK, 1991; ABU-MOSTAFA; ATIYA, 1996; GOMES; JUNIOR, 2010). Gomes e Junior (2010) recomendam cautela ao tratar o tema, uma vez que cada série financeira pode ter um comportamento específico.

Para enfrentar este desafio, a Inteligência Computacional (IC) tem sido uma grande aliada para tratar os problemas do mercado financeiro. Relatórios do setor apontam o uso de métodos de Inteligência Artificial (IA) como estratégico e fundamental para os desafios futuros neste segmento (BLOOMBERG, 2019). Na mesma direção técnicas de Aprendizado Profundo (AP) - uma sub-área de IC - tem ganho evidência em temas do mercado financeiro nos últimos anos (MINH et al., 2018; CHONG et al., 2017; KIM; WON, 2018; DIXON, 2018). A relevância de seu uso se dá ao fato de sua alta capacidade de abstração de dados não-lineares através de suas múltiplas camadas (LECUN et al., 2015).

Métodos de AP utilizam um conjunto de vetores de características como parâmetros de entrada para estabelecer previsões (LECUN et al., 2015; HAN et al., 2012). A grande maioria dos estudos relacionados a previsão do mercado financeiro utilizam indicadores técnicos - um conjunto de equações matemáticas usado para prever preços futuros (PRING, 2002) - como variáveis de entrada em modelos de AP.

Outra abordagem que tem demonstrado ser promissora no campo da previsão de séries financeiras, são métodos de AP baseados em classificação (LEUNG et al., 2000). Alguns estudos tem utilizado classificadores buscando prever não o preço, mas sim se haverá um movimento de alta ou baixa no próximo instante de tempo. A este tipo abordagem é dado

o nome de previsão de direção. Em seus estudos Kara et al. (2011) e Patel et al. (2015) demonstraram que o uso de indicadores técnicos associados a uma abordagem de previsão de direção, pode trazer ganhos significativos para o problema da previsibilidade de ações.

1.1 MOTIVAÇÃO

A maioria dos estudos relacionados à previsibilidade de preços tem como principal motivação, alimentar sistemas de apoio à tomada de decisão em tarefas de investimento. Em linhas gerais, estes sistemas podem atuar de duas maneiras: recomendando um momento favorável e, deixando ao investidor a execução da operação; alimentando robôs de investimento onde estes são capazes de executar ordens automaticamente. Enquanto na primeira abordagem ainda existe o processo de intervenção humana, a segunda opção busca retirar todas as emoções da operação tornando o processo estritamente mecânico.

Sistemas de sugestão de investimentos já vem sendo utilizados por pessoas físicas em suas aplicações financeiras. Com a recente chegada das *fintechs*², espera-se nos próximos anos ocorra um crescimento significativo em aplicações de IC à problemas do mercado financeiro. Embora muitos estudos relacionados à previsibilidade de ações já tenham sido realizados, ainda existem questões em aberto sobre o tema apresentado. O problema de encontrar métodos de previsão que funcionem de forma confiável parece ser um dos principais desafios para o segmento³. Na prática, alguns modelos apresentam um bom desempenho durante o treinamento e resultados diferentes durante a operação. Isto acarreta perdas financeiras significativas e desconfiança no modelo.

No que se refere à BM&FBovespa (Bolsa de Valores, Mercadorias e Futuros) - nome dado a Bolsa de Valores do Brasil - menos de 1% da população realiza algum tipo de investimento no mercado de ações⁴. Aversão ao risco e a falta de educação financeira podem ser os motivadores da baixa adesão de investidores da Bolsa no Brasil. Neste aspecto, séries de curto prazo podem contribuir no desempenho de sistemas de recomendação de investimento. Prazos mais curtos podem trazer sinais capazes de antever mudanças no mercado. Nesta mesma linha, operações de curto prazo tem um tempo de exposição menor, e com isto podem diminuir o risco das operações.

Atualmente, a previsibilidade de séries financeiras de curto prazo são pouco exploradas

²Termo associado à empresas do setor financeiro que usam novas tecnologias e inovação com recursos disponíveis para competir com instituições financeiras tradicionais.

³<https://www.bloomberg.com/professional/blog/machine-learning-models-often-fail-learn-quicktake-qa/>

⁴Os dados foram obtidos através do relatório de mercado disponível em: <http://www.b3.com.br> (Acesso em 01/05/2019). A população do Brasil atualmente é de aproximadamente 208 milhões de pessoas.

no mercado de ações Brasileiro e poucos estudos são encontrados nesta janela de operação. Em um estudo recente, Nascimento (2018) observou em um de seus experimentos que algumas séries de curto prazo apresentavam comportamentos diferentes em alguns dos dias da semana. Desta forma, além das contribuições de seu trabalho, Nascimento (2018) sugeriu investigar formas de identificar janelas ou momentos onde as séries poderiam ser mais previsíveis, buscando, assim, melhorar a eficácia da previsibilidade. Neste sentido, a utilização de métodos de medida de sistemas dinâmicos associados à previsão de direção de séries de curto prazo são as principais motivações deste trabalho.

1.2 OBJETIVOS

1.2.1 OBJETIVO GERAL

O objetivo deste trabalho é investigar métodos de previsão em séries de curto prazo da Bolsa de Valores do Brasil.

1.2.2 OBJETIVOS ESPECÍFICOS

- 1. Avaliar os métodos de supervisão de séries financeiras proposto na literatura a fim de supervisionar as séries definidas;
- 2. Avaliar o método proposto por Kara et al. (2011) e (PATEL et al., 2015) construindo um *Baseline* para os experimentos;
- 3. Comparar o método proposto do atributo predictor de Lyapunov aos resultados do *Baseline*;
- 4. Comparar o método proposto da janela de Lyapunov com os resultados dos experimentos anteriores.

1.2.3 ESTRUTURA DA DISSERTAÇÃO

Este trabalho está organizado da seguinte forma: no Capítulo 2 é apresentada uma revisão de literatura em conjunto com fundamentos teóricos relacionados ao trabalho proposto. No Capítulo 3 é detalhada a metodologia utilizada neste trabalho. O Capítulo 4 aborda os experimentos, resultados e discussões. Por fim, no Capítulo 5 são apresentadas as conclusões do trabalho assim como os trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este Capítulo apresenta a fundamentação teórica e a revisão bibliográfica, as quais servem de base para o desenvolvimento deste trabalho. Nas Seções iniciais serão abordados os tópicos referente ao mercado de ações, assim como, demais fundamentos relacionados à análise de previsibilidade. A partir da Seção 2.8, serão detalhados os trabalhos relacionados ao tema proposto.

2.1 ANÁLISE TÉCNICA

No mercado financeiro existem diversas estratégias operacionais com o mesmo objetivo: auferir lucro (WOLWACZ, 2008). Em linhas gerais, seja investidor ou especulador¹, cada participante adota as ferramentas e estratégias que melhor se adaptam às suas necessidades operacionais².

A análise técnica é uma das ferramentas mais frequentemente adotadas pelos participantes do mercado (TAYLOR; ALLEN, 1992). Ela tem como premissa que preços atuais apresentam correlação com momentos anteriores. Este tópico é frequentemente polêmico pois implica no fato do mercado não ser eficiente; conforme mencionado na Seção 1.

Para Pring (2002) a Análise Técnica (AT) pode ser definida como uma ferramenta capaz de identificar mudanças de tendências. Além disto, ela pode ser combinada com indicadores para encontrar picos e vales dentro de gráficos de preço de ações. Por outro lado, Murphy (1999) entende a AT como uma ferramenta de estudo de gráficos de ações com o propósito de realizar previsões de preços e tendências. Embora há pequenas diferenças entre as definições Pring (2002) e Murphy (1999), ambos concordam que as contribuições de Charles Henry Dow foram significativas para a base da AT.

Charles Henry Dow desenvolveu diversos princípios para compreender e analisar o

¹O especulador é um tipo de participante do mercado focado em obter lucro em operações com prazos curtos. Embora o conceito de especulador tenha uma conotação negativa, sua participação é fundamental para o mercado, pois além de tomar riscos, fornece liquidez para os demais participantes.

²<https://www.investidor.gov.br/publicacao/LivrosCVM.html> - Acesso em: 01/06/2019.

mercado. Seu principal trabalho ficou popularmente conhecido com a teoria de Dow. Charles é considerado o pai da AT pois sua teoria é a base fundamental para os principais conceitos dela (PRING, 2002; MURPHY, 1999). Em linhas gerais, há três premissas nas quais a AT é baseada. Essas premissas fazem parte dos estudos de Charles Henry Dow:

1. O mercado desconta tudo.

Esta afirmação se refere ao fato de que o analista técnico tem como premissa que tudo que afeta o preço da ação se reflete no preço da ação. Se existe um desequilíbrio entre oferta e demanda no momento, ela causará um deslocamento nos preços. Por exemplo, se a demanda está superior à oferta; os preços então devem subir, e vice-versa. As razões pelas quais este movimento ocorre se deve aos fundamentos do mercado ou da empresa sendo analisada e isto, por consequência, se reflete nos gráficos. Todavia, os gráficos não são os responsáveis pelos movimentos de alta ou baixa do mercado. O analista técnico não se preocupa com o contexto fundamentalista relacionado (MURPHY, 1999).

A Figura 1 pode servir como um exemplo relacionado a este caso. O gráfico mostra a variação de preço entre abril a novembro de 2018 da companhia Petrobras. Durante os meses de maio e junho, diversos fatores políticos e mudanças organizacionais atingiram a empresa. O conjunto dos acontecimentos, somados a fatores psicológicos e outros, contribuíram para uma queda superior a 50% dos preços.

Figura 1 - Exemplo da influência de notícias na oscilação dos preços das ações da Petrobras entre Abril e Setembro de 2018.



Fonte - MetaQuotes (2019)

2. Os preços se movem em tendências.

O conceito de tendência é um dos pontos fundamentais para a AT. Encontrar e seguir estes movimentos é algo que um analista técnico sempre busca. Para Murphy (1999) os

fundamentos desta premissa têm relação com a primeira lei de Newton, uma vez que: durante um movimento em tendência, é mais provável que a tendência permaneça do que reverta. Na Figura 2 é possível observar um exemplo com tendências de alta e baixa. Ela mostra um gráfico dos papéis do banco ITAÚ durante o ano de 2018. Após um período de queda iniciado em meados de Abril de 2018 com término em Maio, um movimento de alta se inicia a partir de Junho de 2018.

Figura 2 - Exemplo da formação de tendências nos gráficos das ações.



Fonte - MetaQuotes (2019)

3. A história se repete.

A última premissa se refere ao fato de padrões se repetirem ao longo dos anos. O analista técnico tem como premissa que o futuro é uma repetição de padrões do passado. Com alguns fundamentos na psicologia, Murphy (1999) e Pring (2002) reforçam que algumas figuras nos gráficos carregam comportamentos passados e podem indicar mercados de alta ou baixa. Ao longo dos anos analistas técnicos adquirem a habilidade de observar os comportamentos e associar estes conjunto de informações, identificando assim boas oportunidades de entradas e saídas do mercado.

2.2 INDICADORES TÉCNICOS

Conforme detalhado nos parágrafos iniciais deste Capítulo, indicadores técnicos podem ser combinados a outras ferramentas da AT para analisar o mercado. Dentro do contexto atual da AT existem inúmeros indicadores técnicos disponíveis. Neste estudo, serão utilizados um conjunto de 10 indicadores. A seleção foi realizada com base nos estudos de Kara et al. (2011) e Patel et al. (2015). Cada indicador técnico apresenta algum tipo de parâmetro de

ajuste, como o número de observações anteriores ou pesos específicos. Todos os parâmetros foram mantidos conforme reportados na literatura de Kara et al. (2011) e Patel et al. (2015).

Não existe um estudo que determine qual indicador é melhor, todavia alguns especialistas concordam que alguns deles auxiliam mais em mercados de tendência enquanto outros em mercados em ciclos ou sem tendência (TSAIH et al., 1998). Os dez indicadores escolhidos no trabalhos de Patel et al. (2015) e Kara et al. (2011) seguiram uma extensa revisão de literatura focada principalmente na conclusão de trabalhos anteriores.

Nas seções seguintes, serão apresentados trabalhos da literatura que também utilizaram indicadores técnicos como atributos de entrada para redes profundas. Em geral, métodos de previsão de direção utilizam como variáveis de entrada os indicadores apresentados nos parágrafos seguintes. Em cada parágrafo serão apresentadas às equações que definem os indicadores, assim como o princípio técnico que foi definido por Patel et al. (2015) como critério para sinalizar uma compra ou uma venda.

Dado pela Equação 1 o indicador de média móvel simples (SMA) do inglês, *Simple Moving Average* é considerado um dos indicadores mais utilizados por analistas. Considerando um período de amostras de $n = 10$ o valor de cada observação no tempo t é somado e dividido pela quantidade de amostras escolhidas. A variável C_t apresentada na fórmula deve ser interpretada como o preço de fechamento da série no tempo t . O numeral dez refere-se ao o período de dez amostras escolhidas para o cálculo do indicador.

Uma das principais características de uma SMA é que ela reflete uma versão suavizada da tendência do preço (PRING, 2002). Deve ser considerado um sinal de compra quando o preço do ativo cruzar para cima o valor da SMA. Uma sinal de venda é quando preço do ativo estiver abaixo do valor da SMA.

$$SMA_{10} = \frac{C_t + C_{t-1} + \dots + C_{t-10}}{10} \quad (1)$$

Semelhante ao indicador anterior, o indicador chamado média móvel ponderada (WMA) do inglês, *Weighted Moving Average* acrescenta um peso a cada C_n calculado. Na Equação 2 o valor de n refere-se ao período de amostras escolhidas. O valor de C_1 é o preço de fechamento no instante 1 e C_n é o preço de fechamento no instante n . A característica principal da WMA é dar um peso maior aos valores mais recentes. Por isto, um peso n vai sendo multiplicado junto com os valores de fechamento C_n da série. Seu critério técnico para compra é quando o valor de fechamento do ativo estiver acima do valor do indicador. O comportamento inverso é o critério técnico para a venda.

$$WMA = \frac{C_1 * n + C_2 * (n - 1) + \dots + C_n}{\frac{n * (n + 1)}{2}} \quad (2)$$

Segundo Murphy (1999) o indicador *Momentum* (MOM) representa uma medida da velocidade na qual os preços estão mudando em um ativo. Dado um período n , a Equação 3 estima o valor do indicador *Momentum* calculando o valor do preço de fechamento no instante t dado por C_t , dividido pelo valor do fechamento de n períodos atrás dado por C_{t-n} . O resultado final é multiplicado pelo valor 100. Deve ser considerado um sinal de compra quando o valor do indicador for positivo. Um sinal de venda deve ser considerado quando o indicador apresentar valores negativos.

$$MOM = \frac{C_t}{C_{t-n}} * 100 \quad (3)$$

Os indicadores estocásticos K% e D% são representados pelas Equações 4 e 5, respectivamente. Eles foram desenvolvidos na década de 50 e ainda são amplamente utilizados por analistas técnicos³. Seu resultado permite observar se o ativo em questão está sobrecomprado ou sobrevendido. Estas regiões de sobrecompra ou sobrevenda podem indicar uma reversão de tendência da série.

Observando a Equação 4 referente ao indicador K%, LL_{t-n} significa o menor preço no período n , enquanto HH_{t-n} o maior preço no período n . O valor utilizado como padrão para n é 14. Este mesmo valor foi utilizado neste trabalho. O valor do resultado de K% permanece sempre entre 0 e 100.

$$K\% = \frac{C_t - LL_{t-n}}{HH_{t-n} - LL_{t-n}} * 100 \quad (4)$$

A Equação 5 representa o resultado do indicador estocástico D%. Seu cálculo é uma média simples do próprio indicador K% da Equação 4. O período padrão da média é 3 e este valor foi mantido neste trabalho. Além de observar regiões de sobrecompra ou sobrevenda, os indicadores K% e D% também podem ser usados como seguidores de tendência. Neste caso, o critério técnico para compra é quando o valor corrente do indicador for superior ao seu valor anterior. O inverso vale para a venda.

$$D\% = \frac{\sum_{i=0}^{n-1} K\%_{t-i}}{n} \quad (5)$$

³<https://www.investopedia.com> - Acesso em 01/06/2019

Desenvolvido por Wilder (1978), o indicador de Índice de Força Relativa (IFR), é amplamente utilizado por analistas financeiros buscando regiões de sobrecompra ou sobrevenda. A Equação 6 representa o resultado do indicador IFR que sempre permanecerá entre o valor 0 e 100. U_p refere-se à média dos preços do período n no qual o valor da ação subiu. A variável D_p refere-se à média dos preços do período n no qual o valor da ação caiu. A variável $U_{p_{t-i}}$ representa uma subida de preço no intervalo $t - i$ e Dw_{t-i} um queda de preço no intervalo $t - i$.

O indicador IFR utiliza alguns valores definidos por padrão. O valor de n é adotado sempre como $n = 14$. Há outros dois valores também definidos para este indicador. Um deles trata-se do valor IFR a partir do qual se considera sobrecompra: $IFR > 70$ e sobrevenda $IFR < 30$. O critério técnico definido por Patel et al. (2015) para compra é quando o valor corrente estiver menor que $IFR = 30$. A venda ocorre quando o valor corrente estiver superior a $IFR = 70$.

$$IFR = 100 - \frac{100}{1 + (\sum_{i=0}^{n-1} U_{p_{t-i}}/n) / (\sum_{i=0}^{n-1} Dw_{t-i}/n)} \quad (6)$$

Convergência e Divergência das Médias Móveis (MACD) é um indicador que utiliza um conjunto de duas médias móveis do tipo exponencial buscando obter um bom resultado de tendência e momento. Deste forma é considerado um indicado simples e eficiente⁴.

A Equação 8 apresenta o cálculo para obter MACD. A operação é a diferença entre duas médias móveis exponenciais (MME) onde, conforme Equação 7, $k = 2/(1 + n)$ e $n =$ período, C_t preço de fechamento no instante t e MME_{t-1} é o valor da média móvel exponencial no instante de tempo anterior. O valor do indicador MACD é definido na na Equação 8, onde $MME(12)$ é a média móvel de 12 períodos e $MME(26)$ é a média móvel para 26 períodos. O critério técnico definido para compra é quando o valor corrente do indicador for superior ao seu valor anterior. O inverso é o critério técnico para a venda.

$$MME = C_t * k + MME_{t-1} * (1 - k) \quad (7)$$

$$MACD = MME(12) - MME(26) \quad (8)$$

Desenvolvido por um *trader* chamado *Larry Williams*, o indicador que leva o seu nome Williams %R, abreviado por %R, compara o preço corrente de um ativo, em relação ao preço

⁴<https://www.bussoladoinvestidor.com.br> - Acesso em 05/07/2019

mais alto e mais baixo dos últimos n dias. Na Equação 9 a variável H_n refere-se ao valor mais alto do preço no período enquanto L_n ao menor preço no período. De forma semelhante a alguns outros indicadores, o critério técnico para compra é quando o valor corrente do indicador for superior ao seu valor anterior. O inverso é o critério técnico para a venda.

$$R\% = \frac{H_n - C_t}{H_n - L_n} * 100 \quad (9)$$

O indicador de acumulação e distribuição (A/D) relaciona a evolução do preço do ativo com o volume de negociações. Considere um período t de negociações, onde $t = 1$ significa 1 minuto de negociações. A soma da quantidade de ações negociadas em t é considerada o volume de negociações no tempo t . O indicador A/D busca compreender o quão forte é a tendência considerando uma variável de Volume.

O indicador é definido pela Equação 10, onde C_t é o valor de fechamento no período, L_t é o menor valor no período, H_t é o valor máximo no período. O resultado é multiplicado pela variável de volume V e no fim somado ao valor anterior do indicador dado por AD_{t-1} . O critério técnico para compra no caso do indicador A/D é quando o valor corrente de A/D for superior ao seu valor anterior. A venda ocorre quando o valor corrente de A/D for inferior a seu valor anterior.

$$A/D_t = \left(\frac{(C_t - L_t) - (H_t - C_t)}{(H_t - L_t)} * V \right) + AD_{t-1} \quad (10)$$

Desenvolvido por Lambert (1983) o indicador CCI calcula o preço corrente com a média de preço de um período. Seu cálculo é apresentado na Equação 11 onde, $M_t : H_t + L_t + C_t = 3$; $SM_t = (\sum_{i=1}^n M_{t-i+1})/n$ e $D_t = (\sum_{i=1}^n |M_{t-i+1} - SM_t|)/n$. Neste indicador Patel et al. (2015) definiram que para valores de CCI entre +200 e -200, se o valor corrente do CCI for superior ao seu valor anterior o sinal é compra. Se o valor corrente do CCI for inferior ao seu valor anterior o sinal é venda. No caso do valor de CCI ultrapassar +200 o sinal deve ser de venda assim como para valores inferiores a -200 o sinal é compra.

$$CCI = \frac{M_t - SM_t}{0,015 * D_t} \quad (11)$$

2.3 SÉRIE DE RETORNO E VOLATILIDADE

A grande parte dos estudos que envolvem séries financeiras não se concentram no preço do ativo. A abordagem comumente usada é utilizar as séries de retornos. Considerando que a análise de avaliação de risco em carteiras de ativos é medida pela variação dos preços dos ativos, as séries de retornos apresentam propriedades estatísticas melhores para modelar comportamentos temporais financeiros. Além disto, por estarem livres de escala, o uso de retornos facilitam análises quando comparados com as séries de preços (MORETTIN, 2008).

Segundo Morettin (2008), há duas formas de calcular o retorno de uma série financeira. A primeira abordagem pode ser chamada de *retorno líquido simples* ou retorno linear e é dada pela Equação 12. O R_t se refere ao resultado do cálculo do retorno linear de um período t previamente determinado por minuto, dia, mês, ano, etc. ΔP_t é a diferença do preço de fechamento de um período corrente dado por P_t e do período anterior dado por P_{t-1}

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}} \quad (12)$$

Uma outra abordagem criada também por Morettin (2008) é o cálculo do *retorno composto continuamente* ou também frequentemente chamado apenas de *log-retorno*. É academicamente mais comum encontrar o termo *log-retorno* pois alguns cálculos de volatilidade e rentabilidade histórica serem mais triviais utilizando log-retorno. Sua definição é dada pela Equação 13 onde, $p_t = \log P_t$. Com exceção de retornos muito elevados, o valor do retorno linear e o log-retorno são considerados praticamente idênticos para pequenos valores e inclusive para séries de curto prazo. Este trabalho adota o cálculo de log-retorno para estimar o retorno das séries escolhidas.

$$r_t = \log \frac{P_t}{P_{t-1}} = \log(1 + R_t) = p_t - p_{t-1} \quad (13)$$

Não observável diretamente, a volatilidade também é um tipo de medida que pode ser obtida através de séries financeiras. Morettin (2008) define volatilidade como a variância condicional de uma série de retorno. Ela mede a dispersão dos retornos de uma ativo e pode ser considerada uma medida de risco.

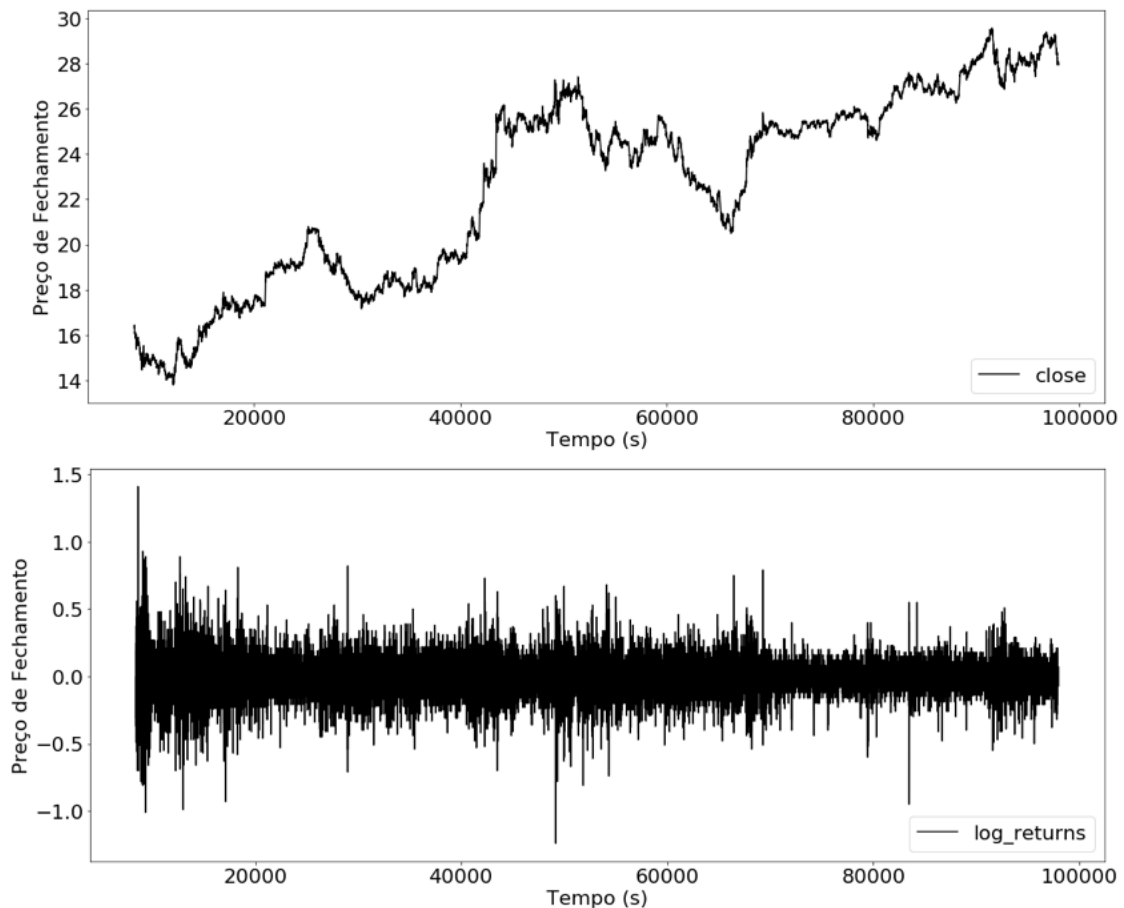
Há diversas formas de calcular a volatilidade na literatura, todavia, uma das formas mais simples é obtê-la através do cálculo do desvio padrão dos retornos. Sua definição é dada pela Equação 14, onde, n é a quantidade de amostras de retorno utilizado e, $(x - \bar{x})$ é a diferença entre cada valor da amostra em relação a média das amostras. A média das amostras dada por \bar{R} ,

é a soma de todos os retornos, calculados pela Equação 13, dividido pelo número de amostras n .

$$sR = \sqrt{\frac{\sum_i^n (R - \bar{R})^2}{n - 1}} \quad (14)$$

Na Figura 3 um exemplo de série de retornos é mostrado. Na parte superior, os valores de fechamento de cada minuto de uma série de um ativo financeiro foi plotada. Na parte de baixo da mesma imagem, seus retornos logarítmicos foram calculados.

Figura 3 - Exemplo de uma série financeira com preços de fechamento minuto a minuto e abaixo seus respectivos retornos logarítmicos.



Fonte - O autor.

2.4 CAOS DETERMINÍSTICO

Na Seção seguinte, conceitos sobre o expoente de Lyapunov serão abordados. A aplicação deste conceito está intimamente ligada à teoria do caos, e, portanto, nesta Seção será

feita uma breve revisão sobre o tema.

Henry Poincaré, no final do século XIX, lançou às bases da dinâmica caótica. Através de seu trabalho, relativo à atração gravitacional de três corpos, ele demonstrou a sensibilidade de um sistema às condições iniciais, uma das características mais marcantes do caos. Apesar das descobertas de Poincaré, o estudo do caos não teve significativa atenção nos anos seguintes. Todavia, o tema recebeu destaque novamente com as observações do meteorologista e matemático Edward Lorenz (SAVI, 2006).

Seus trabalhos relacionados à modelos atmosféricos foram os primeiros estudos do que, posteriormente, foi definido como atrator estranho na teoria do caos (LORENZ, 1963). Por volta de 1955, Lorenz juntamente com sua equipe buscava desenvolver um programa de previsão do tempo. Enquanto realizavam simulações em seu conjunto de equações, fizeram uma importante descoberta acidental. Após observarem sucessivos resultados diferentes em seus experimentos, Lorenz e sua equipe perceberam que um técnico havia truncado a terceira casa decimal de um trecho do código. Com esta pequena diferença, Lorenz constatou que uma sutil alteração nas condições iniciais do problema poderia produzir uma significativa diferença nos resultados (SILVER, 2012). Este fenômeno ficou popularmente conhecido como *butterfly effect*, ou efeito borboleta (SAYAMA, 2015).

O exemplo do efeito borboleta pode ser melhor observado na Figura 4. Para construir a Figura 4, às Equações 15, 16 e 17, conhecidas como equações de Lorenz foram utilizadas. Estas equações foram desenvolvidas por Lorenz em seus estudos referentes à dinâmica atmosférica por volta de 1960 (LORENZ, 1963). Na Equação 15, σ se chama o número de *Prandtl*; na Equação 16, ρ é o número de *Rayleigh* e por fim, na Equação 17 o β . Os valores de x , y e z influenciam diretamente no comportamento geral do sistema em função de seus valores iniciais. Para exemplificar, na Figura 4, duas séries temporais com uma pequena diferença de perturbação inicial foram sobrepostas. Os valores dos parâmetros σ , ρ e β são constantes com valores positivos e neste exemplo os seguintes valores foram definidos: $\sigma = 10$, $\rho = 30$ e $\beta = 3$. A série sobreposta está plotada em pontilhado e, para esta série, foi adicionada uma alteração de 0,00001 no valor inicial de x .

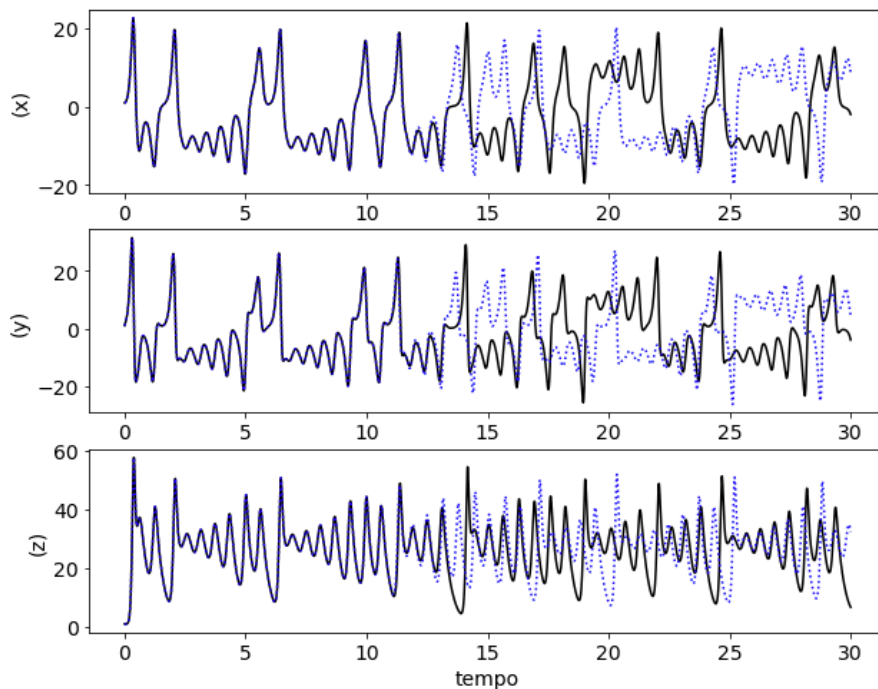
$$\frac{dx}{dt} = \sigma(y - x) \quad (15)$$

$$\frac{dy}{dt} = \rho x - y - xz \quad (16)$$

$$\frac{dz}{dt} = xy - \beta z \quad (17)$$

É possível observar que, até próximo do tempo 13, ambas permanecem coerentes e a partir de então as séries percorrem valores distintos. Este experimento reforça e demonstra a extrema sensibilidade de sistemas caóticos às condições iniciais do problema.

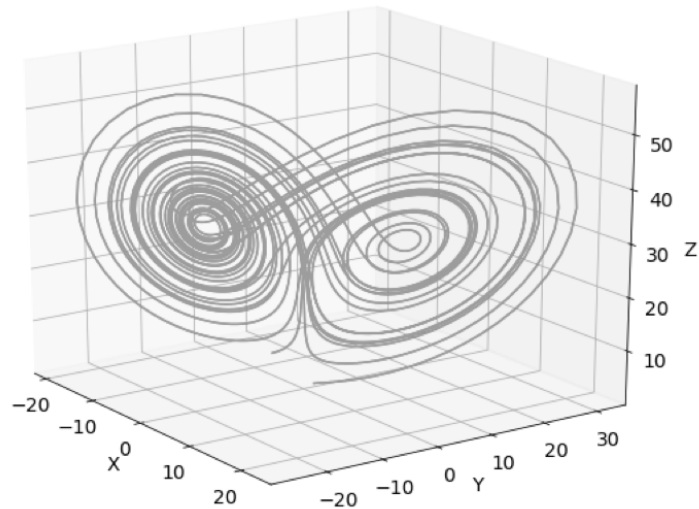
Figura 4 - Exemplo do efeito borboleta. A extrema sensibilidade de sistemas caóticos à condições iniciais. séries geradas através das equações de Lorenz.



Fonte - Adaptado de Sayama (2015).

Em função dos possíveis estados de um sistema complexo, há uma outra forma de visualizar o comportamento do sistema: o espaço de fase. O espaço de fase pode ser definido como um espaço teórico onde cada estado do sistema é mapeado para um único lugar no espaço. Um dos benefícios do espaço de fase é inspecionar visualmente para onde as trajetórias estão indo. Ao plotar, é possível observar se as trajetórias divergem ou convergem para um ponto específico. A região para onde o sistema converge é chamada de atrator. Existem vários tipos de atratores em sistemas dinâmicos. O atrator clássico demonstrado por Lorenz é chamado de atrator estranho. Em um atrator estranho o sistema alterna entre estados sem permanecer em um estado oscilatório ou fixo; é uma flutuação contínua caótica (SAYAMA, 2015).

Figura 5 - Atrator estranho ou atrator de Lorenz criado a partir das Equações 15, 16, 17 de Lorenz



Fonte - Adaptado de Sayama (2015).

2.5 EXPOENTE DE LYAPUNOV

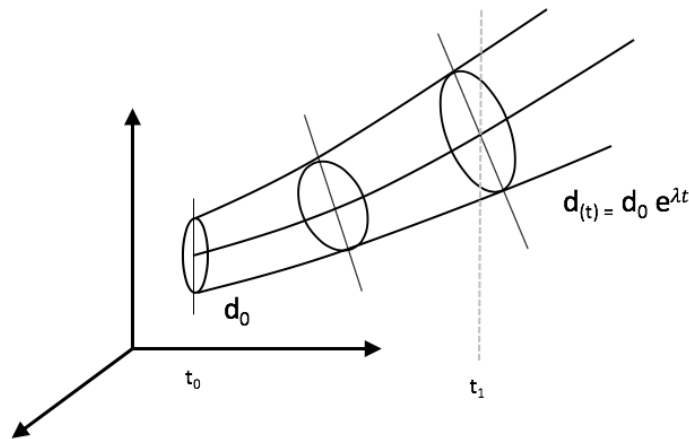
O expoente de Lyapunov (EL) tem sido considerado uma das ferramentas mais úteis para o diagnóstico de caos em sistemas dinâmicos (WOLF et al., 1985). Sayama (2015) considera o EL como uma métrica analítica para caracterizar o caos. Através do EL é possível verificar a divergência exponencial no tempo de trajetórias vizinhas. A divergência entre tais trajetórias pode ser estimada a partir de uma trajetória de referência que é definida no instante inicial de tempo. Duas trajetórias com EL positivo vão se separar rapidamente, e evoluir por trajetórias completamente diferentes durante suas iterações futuras e isto caracteriza um sistema caótico. Por outro lado, valores do EL negativos significam a ausência de caos no sistema com possíveis ciclos periódicos (SAVI, 2006). A quantidade de expoentes de Lyapunov de um sistema sempre está associada a quantidade de dimensões do sistema.

Para ilustrar melhor a abordagem do EL, considere uma esfera de diâmetro d_0 , conforme destacado na Figura 6. A partir do ponto inicial, novas trajetórias serão construídas. O objetivo é avaliar como a esfera inicial se comporta com a evolução do tempo t . A variação do diâmetro desta esfera é expressa pela Equação 18, onde e é uma base de referência. O valor de λ - definido como EL - é dado pela Equação 19. Sendo assim, caso o valor de λ seja positivo ou nulo as trajetórias serão divergentes, caracterizando o caos (SAVI, 2006).

$$d_t = d_0 * e^{\lambda.t} \quad (18)$$

$$\lambda = \lim_{t \rightarrow \infty} \frac{1}{t} * \ln \left| \frac{d_t}{d_0} \right| \quad (19)$$

Figura 6 - Exemplo de avaliação dos expoentes de Lyapunov em uma esfera que evolui ao longo do tempo.



Fonte - Adaptado de (SAVI, 2006)

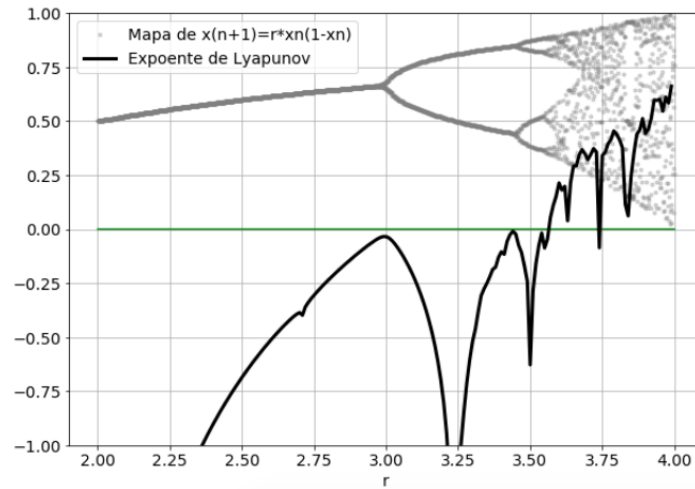
Uma forma clássica de se demonstrar o comportamento do EL sob o efeito do caos determinístico é utilizando o mapa logístico. Mapas são uma forma de representar sistemas dinâmicos no tempo discreto. O mapa logístico associa a um dado um número x_n a outro número x_{n+1} e seu comportamento é dado pela Equação 20. O resultado do EL para um mapa logístico, utilizando a Equação 20 é mostrada na Figura 7.

$$x_{n+1} = rx_n(1 - x_n) \quad (20)$$

Na Figura 7 é possível observar que, enquanto os valores de r , referente à Equação 20, permanecem menor que o valor 1, o valor do EL permanece então com valores menores que 0. Isto demonstra a ausência de um comportamento caótico no sistema naquele intervalo. Todavia, quando os valores de r atingem regiões próximas a 1,7 um comportamento caótico pode ser observado no mapa, o que é demonstrado pelos valores do EL maiores que zero.

Diversos métodos podem ser usados para o cálculo do EL. A maioria deles se diferem pela forma de analisar a trajetória principal. Este trabalho utilizará o método de Eckmann et al. (1986) para o cálculo dos expoentes. Estima-se que o uso do EL possa contribuir para os métodos de classificação, auxiliando-os à identificar regiões onde a previsibilidade da série seja maior ou menor. Mais detalhes serão apresentados no Capítulo 3.

Figura 7 - Mapa logístico calculado através da Equação 20 e seu respectivo expoente de Lyapunov. O resultado gráfico do mapa representa os atratores para cada valor do parâmetro r representado no eixo x .



Fonte - Adaptado de Sayama (2015).

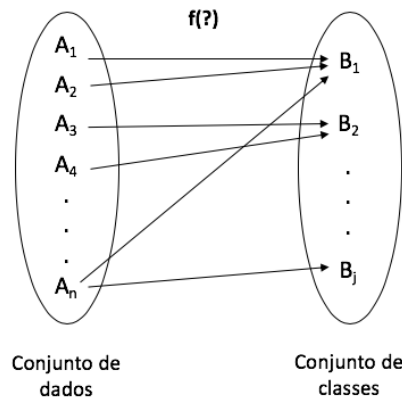
2.6 CLASSIFICADORES

A tarefa de classificação pode ser considerada uma das mais populares tarefas relacionadas ao Aprendizado de Máquina. Tal tipo de tarefa utiliza aprendizado supervisionado, que é considerado um tipo de aprendizado indutivo. Aprendizado indutivo, por sua vez, refere-se à capacidade que alguns algoritmos têm em aprender a partir de exemplos (GOLDSCHMIDT RONALDO, 2015). Os exemplos são providos através de conjuntos de dados que ainda não estão supervisionados. Realizar a supervisão de um conjunto de dados se resume a atribuir para cada saída, um valor denominado atributo-alvo ou classe (conforme mostrado na Figura 8).

A tarefa de classificação pode ser compreendida utilizando como referência a Figura 8 adaptada de Goldschmidt (2015). Ela consiste na busca de uma função que permita associar o maior número de registros A às categorias B , denominada classe. Uma vez concluída a associação, esta função pode ser aplicada a novas instâncias a fim de prever as classes nas quais estes registros se enquadrem (GOLDSCHMIDT RONALDO, 2015). Há diversos tipos de classificadores, e sua aplicação depende do tipo de problema escolhido.

Para avaliar o desempenho dos algoritmos de classificação algumas medidas de desempenho são utilizadas. O conjunto de dados, assim como, o problema a ser resolvido definem qual método medida de desempenho deve ser utilizado. Nos parágrafos seguintes, os métodos de desempenho adotados neste trabalho serão detalhados.

Figura 8 - Exemplo das associações de um conjunto de dados usado em tarefas de classificação.



Fonte - Adaptado de (GOLDSCHMIDT RONALDO, 2015)

2.6.1 MÉTRICAS DE DESEMPENHO

Um desafio na a classificação é o desbalanceamento dos conjuntos de dados de treinamento. Ou seja, considerando um problema com três classes, espera-se que cada uma delas tenham a mesma quantidade de amostras. Todavia, não é o que acontece no mundo real. Assim, métricas alternativas à acurácia⁵ devem ser consideradas para tratar conjunto de dados desbalanceados. Três abordagens foram escolhidas neste trabalho para avaliar o desempenho dos classificadores: matriz de confusão, *F-measure*⁶ e a curva ROC.

- **Matriz de confusão**

A matriz de confusão (MC), é um formato de apresentação do resultados de um processo de classificação em formato de uma tabela. Simplificando para um problema binário⁷, haverá um tabela com 2 linhas e 2 colunas. A tabela armazena a frequência de acertos do classificador por classe. Cada coluna e cada linha se referem a uma classe. A diagonal da tabela armazena os valores classificados corretamente em suas respectivas classes (HAN et al., 2012; WITTEN et al., 2016).

Os termos frequentemente usados na MC são: (TP) para resultados verdadeiros positivos; (TN) para resultados verdadeiros negativos; (FP) para resultados falsos positivos e (FN) para resultados falsos negativos.

⁵A acurácia é o tipo métrica mais utilizada como taxa de acerto do classificador. É calculado tomando o número de previsões corretas feitas e dividindo pelo número total de previsões feitas. o valor final se extrai a porcentagem (GOLDSCHMIDT RONALDO, 2015).

⁶A métrica *F-measure* também pode ser encontrado na literatura como *F1 score* ou *F-score* (HAN et al., 2012).

⁷Um problema de classificação binária se dá quando apenas duas classes fazem parte do espaço de atributos-alvo.

- **F-measure**

A métrica *F-measure* (F1) é um tipo de métrica de equilíbrio entre outras duas métricas: *recall* e *precision*. Conforme Equação 21, enquanto a métrica *recall* é definida como a razão entre verdadeiros positivos sobre falsos negativos, *precision* é a razão entre verdadeiros positivos sobre falsos positivos.

$$recall = \frac{T_p}{T_p + F_n} \quad (21)$$

$$precision = \frac{T_p}{T_p + F_p} \quad (22)$$

De forma prática, a métrica *recall* busca avaliar qual a proporção da classe X que foi corretamente classificada. Por outro lado, *precision* mostra a proporção de instâncias da classe X que foi identificada corretamente. Conforme a Equação 23, *F-measure* portanto, mostra o balanço entre *recall* e *precision* (WITTEN et al., 2016).

$$F\text{-measure} = 2 * \frac{precision * recall}{precision + recall} \quad (23)$$

- **Curva ROC e AUC**

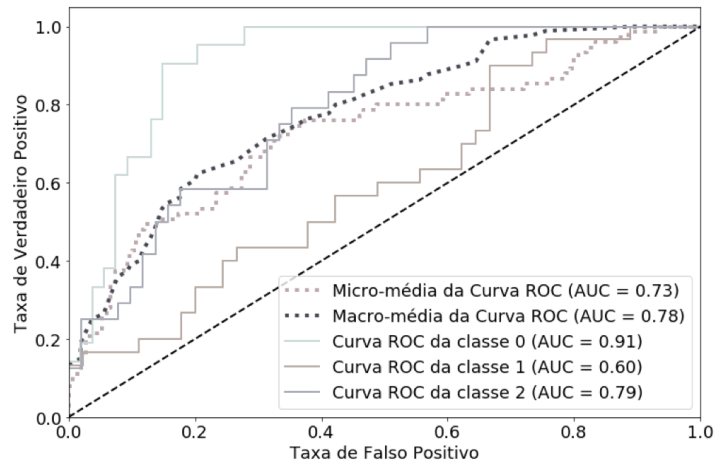
A curva ROC (ROC), do inglês *Receiver Operating Characteristic* e AUC, estão dentre as métricas mais usadas na avaliação de métodos de classificação. Curvas ROC são ferramentas visuais para comparar desempenho de classificadores. Elas utilizam os parâmetros: taxa de verdadeiro positivo (TPR), do inglês *true positive rate* e taxa de falso positivo (FPR), também do inglês *false positive rate* (WITTEN et al., 2016). TPR é dada pela Equação 24 e pode ser considerada uma medida de sensibilidade enquanto a Equação 25 representa uma medida de especificidade dado pela FPR.

Como a curva ROC traça valores para diferentes limiares de classificação, uma forma de sumarizar a ROC em um valor único é utilizar a área sobre a curva - AUC. O valor da AUC varia entre zero e um. Portanto, se todas as previsões forem erradas o AUC tem valor nulo, do contrário o valor é um. A Figura 9 exemplifica o formato de saída de uma ROC com valores de AUC inseridos na legenda.

$$TPR = \frac{T_p}{T_p + F_n} = Recall \quad (24)$$

$$FPR = \frac{F_p}{F_p + T_n} \quad (25)$$

Figura 9 - Exemplo de uma curva ROC com três diferentes classes e seus respectivos valores AUC.



Fonte - O autor.

Os parágrafos anteriores apresentaram as formas mais clássicas de avaliação de desempenho de classificadores. Tais abordagens também foram amplamente utilizadas em trabalhos relacionados que serão discutidos na Seção 2.8. Os parágrafos seguintes apresentam uma breve introdução sobre os classificadores utilizados neste trabalho.

2.6.2 SVM

O algoritmo Máquina de Vetores de Suporte ou SVM, do inglês *Support Vector Machine*, utiliza funções lineares para separar conjuntos por meio de hiperplanos⁸ (GOLDSCHMIDT RONALDO, 2015). Segundo Han et al. (2012) os primeiros estudos relacionados a SVM surgiram por volta da década de 60. Todavia, a primeira publicação foi feita por Boser et al. (1992). Embora SVM tenha sido usado inicialmente para resolver problemas de classificação, o método é amplamente usado em outras tarefas, como por exemplo problemas de regressão (YU; KIM, 2012).

O princípio de funcionamento do algoritmo SVM é construir superfícies de decisão através de hiperplanos. Desta forma, considere um problema simples de duas classes linearmente separáveis. Considerando o conjunto de dados (x_i, y_i) , no qual x_i corresponde aos atributos e y_i as classes, define-se que $y_i \in (-1, 1)$, onde a classe -1 corresponde a “vender” e a classe +1 corresponde a “comprar”. A Figura 10(a) auxilia a compreensão considerando dois atributos A1 e A2 plotados em uma visão 2D. Quando o plano é igual a dois a equação do hiperplano é dado por $w^T x + b = 0$. Na equação, o w e b são os parâmetros da reta, onde b é

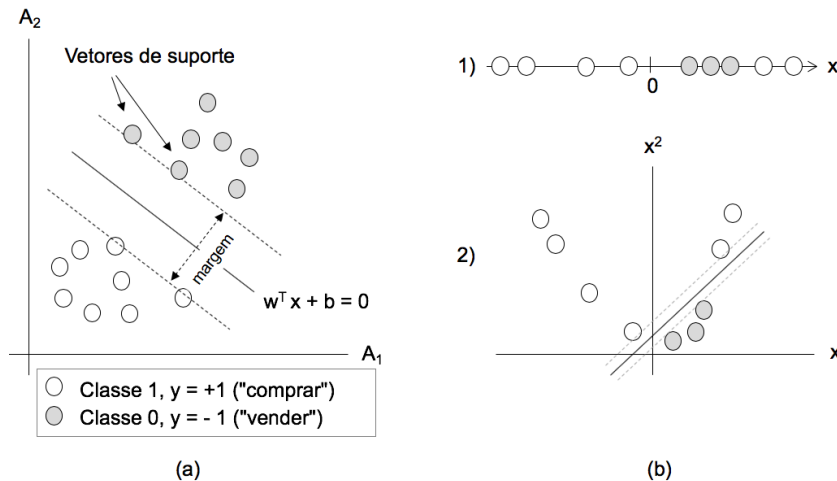
⁸Hiperplano é considerado a generalização do conceito de plano em diferentes números de dimensões. (GOLDSCHMIDT RONALDO, 2015).

um número real e w^T é o vetor transposto de w .

Ainda na Figura 10 (a) é possível observar que entre inúmeras retas seriam possíveis no espaço entre as duas classes. Desta forma, o algoritmo busca uma distância máxima entre os pontos, chamada de margem. Margem é denominada a distância entre as retas paralelas construída entre as classes. Através de um processo de otimização do algoritmo os valores de \mathbf{b} e \mathbf{w} são determinados. Tais valores representam o valor máximo para a margem. neste aspecto, a posição da reta é determinada pelos vetores de suporte, que são os pontos mais próximos do classificador linear (conforme indicado na Figura 10 (a)) (GOLDSCHMIDT RONALDO, 2015). Para obter a função de decisão, que será utilizada para classificar novos registros, técnicas para maximizar funções quadráticas são necessárias.

Para finalizar, considere um caso onde as classes (y_i) não sejam linearmente separáveis. Para tratar conjuntos de dados mais complexos ou ruidosos, SVM utiliza funções-núcleo (*kernel functions*). Um exemplo está ilustrado na Figura 10 (b), onde o eixo x sofrerá uma transformação para uma função x^2 onde assim será possível encontrar vetores de suporte para as classes. Diversos tipos de funções-núcleo existem para atender diferentes tipos de problemas de classificação.

Figura 10 - a) Exemplo da separação de um hiperplano em SVM. Adaptado de ((HAN et al., 2012)). (b) Exemplo de mapeamento de conjunto de dados x para x^2 .



Fonte - Adaptado de (GOLDSCHMIDT RONALDO, 2015).

2.6.3 FLORESTA ALEATÓRIA

Floresta Aleatória, do inglês: *Random Forest* (RF), é outro tipo de classificador usado em tarefas de aprendizado supervisionado. Semelhante ao SVM, Breiman (2001) mostrou que RF é uma ferramenta efetiva para tarefas de previsão, sejam baseadas em classificação

ou regressão. Por utilizar uma combinação de classificadores (*ensemble*⁹), esta abordagem apresenta uma melhora na acurácia de predições e maior robustez ao *overfitting*¹⁰ (BREIMAN, 2001). O termo “Floresta”, está associado ao fato da coleção de classificadores que formam a arquitetura do RF serem do tipo árvores de decisão. As árvores de decisão são tipos de estruturas de dados que armazenam informações partindo de um nó raiz, chegando até ao nó folha.

O processo de treinamento do algoritmo RF está baseado no método de *bagging*. Considere um conjunto de dados D, com d tuplas¹¹. A cada interação i, onde (i=1,2,...,k) um conjunto de treinamento D_i de d tuplas é amostrado com elementos do conjunto de dados D. Ou seja, do conjunto de dados iniciais, k amostras são construídas com instâncias e atributos escolhidos de forma aleatória. Cada conjunto D_i passa pelo processo de treinamento de uma árvore de decisão. O resultado final é, então, a média do resultado de todas as árvores (HAN et al., 2012).

2.6.4 MLP

Redes Neurais Artificiais (RNAs), são representadas como uma técnica de IA inspirada no funcionamento do cérebro por estrutura de dados computacionais (HAYKIN, 1994). As RNAs têm diferentes arquiteturas e cada uma delas são escolhidas em função do problema ou da tarefa a ser resolvida. Na maior parte das aplicações, carrega-se os dados de uma base de dados, realiza-se um pré-processamento e são apresentados para a entrada da rede. A rede processa os dados e produz uma saída de acordo com a natureza do problema (GOLDSCHMIDT RONALDO, 2015).

Uma das arquiteturas mais simples de uma RNA é conhecida como *perceptron*. Ela consiste de duas camadas, onde o nó de entrada é representado com um conjunto de atributos e o nó de saída é resultado final do modelo. Nas arquiteturas de RNAs os nós são conhecidos como neurônios e suas conexões são feitas através de pesos ponderados (TAN et al., 2005). Em cada nó os pesos multiplicados pelas respectivas ativações e depois somados de um fator de *bias*. O valor final é aplicado a uma função de ativação. Existem diferentes tipos de função de ativação e a escolha é definida em função do problema que está sendo estudado.

Segundo Tan et al. (2005), MLP do inglês *Multilayer Perceptron* é um tipo de arquitetura de RNA que contém mais camadas intermediárias entre a camada de entrada e saída.

⁹Um método *ensemble* combina diversos classificadores, buscando assim criar vários modelos e melhorar o desempenho da classificação. Floresta aleatória é considerado um exemplo deste método (HAN et al., 2012)

¹⁰Modelo que apresenta bom resultado para um conjunto de dados treinado mas falha quando apresentado a um conjunto de dados novos.

¹¹Refere-se a uma lista com elementos em sequência.

As camadas intermediárias também são referenciadas como camada oculta e nó oculto. Todos os nós de uma camada intermediária são conectados a todos os outros nós da camada seguinte formando uma topologia totalmente conectada. Uma vez que nenhum nó é conectado a nós em camadas anteriores a arquitetura também é denominada *feedforward* (TAN et al., 2005; HAN et al., 2012).

Uma etapa importante para qualquer tipo de arquitetura de RNAs é a etapa de treinamento. Entre a saída da rede e sua saída esperada o erro de treinamento é estimado (vide a Seção 2.6 sobre atributos previsores e atributo alvo). Em função do erro, o algoritmo da rede ajusta os pesos para gerar uma saída próxima da entrada apresentada a rede. Redes MLP utilizam o algoritmo de retro-propagação do erro, ou em inglês *backpropagation*, para minimizar a função de erro entre a saída da rede e o valor real desejado (GOLDSCHMIDT RONALDO, 2015).

Utilizando o método de gradiente decrescente, o algoritmo de retro-propagação opera em duas fases. Conforme descrito por Tan et al. (2005) na primeira etapa (fase direta) os valores dos pesos da interação anterior são usados para calcular o valor da saída de cada neurônio. Os novos valores são agora repassados para a camada seguinte. Na segunda etapa (fase de retorno) a fórmula é usada no sentido inverso. O algoritmo permite desta forma utilizar os erros da camada $k+1$ para estimar o erro da camada k .

2.6.5 LSTM

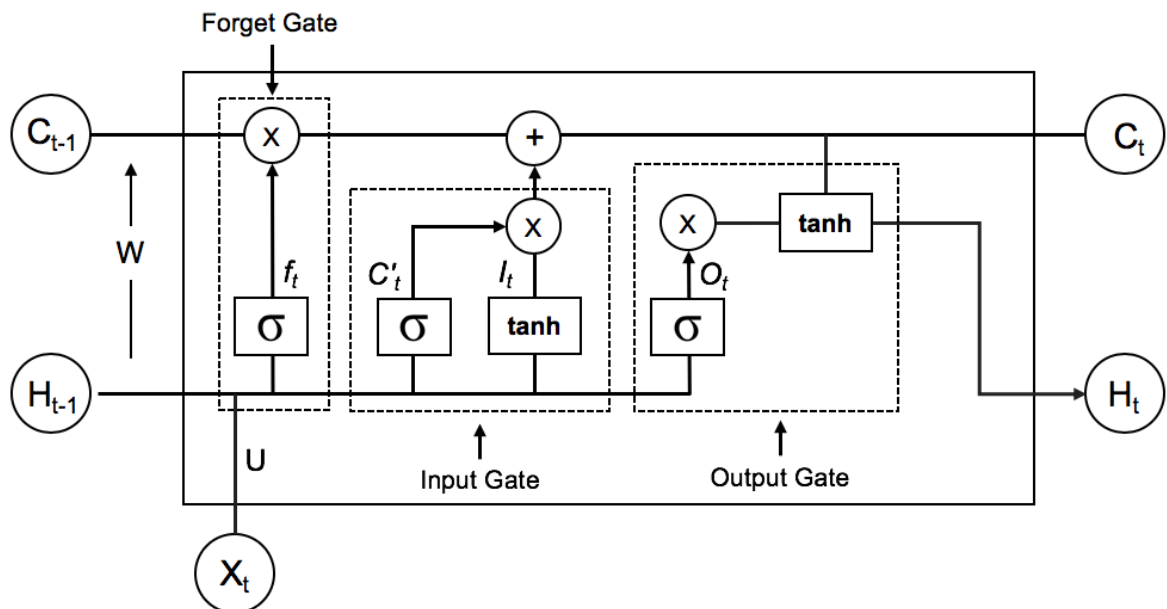
Arquiteturas de RNAs do tipo *feedforward* (como apresentado na Seção 2.6.4) não apresentam robustez para tratar problemas que envolvem dados sequenciais ou alguma relação temporal entre as entradas. Por não existir nenhuma memória associada a dados anteriores em sua arquitetura, o método falha em prever sequência de palavras, reconhecer falas ou ainda realizar traduções. Por outro lado, a arquitetura de Redes Neurais Recorrentes (RNNs) usa suas conexões com as camadas anteriores e, a partir disto, consegue armazenar representações dos eventos de entrada em forma de ativações (HOCHREITER; SCHMIDHUBER, 1997).

Redes LSTM, do inglês *Long Short-Term Memory*, foram propostas por Hochreiter e Schmidhuber (1997) e são uma das variações de arquitetura de RNNs. Sua construção busca capturar dependências temporais adicionando à arquitetura o que é conhecido como unidade de memória. Hochreiter e Schmidhuber (1997) demonstra em seu trabalho as diversas limitações das arquitetura clássicas propostas em trabalhos anteriores. Um dos principais comportamentos limitantes; conhecido como *vanishing gradient*; faz com que eventos de um passado próximo possam ser sobrescritos ou perdidos por novas informações. Arquiteturas clássicas de RNN

podem associar correlações com aproximadamente dez amostras anteriores, o que em alguns tipos de problemas não traz ganho significativo.

Para lidar com o problema do *vanishing gradient*, um conjunto de portas dentro da unidade de memória é responsável por determinar quais informações devem ser esquecidas, armazenadas ou lembradas para a próxima entrada de unidade computacional (HOCHREITER; SCHMIDHUBER, 1997). Uma unidade de memória da arquitetura de uma rede LSTM está apresentada na Figura 11. Nesta Figura são mostrados às três portas: *input gate*, *forget gate*, *output gate* e sua respectiva posição na arquitetura.

Figura 11 - Unidade de memória de uma LSTM.



Fonte - Adaptado de Colah (2019).

Cada uma das três portas tem uma função específica dentro da rede. A porta *forget gate* determina o que a rede deve “esquecer”. Esta etapa é definida pela Equação 26, onde σ refere-se à função *sigmoid*; W_f o peso do neurônio referente ao *forget gate*; H_{t-1} o valor de saída do bloco anterior, referente ao tempo anterior; b_f é o *biases* da porta e por fim, C_{t-1} é a memória do estado anterior.

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}C_{t-1} + b_f) \quad (26)$$

A outra porta mostrada na Figura 11 é a porta *input gate*, que é responsável por informar qual a nova informação deve ser salva na rede. Nesta etapa o vetor também é criado

através da função \tanh . Duas equações fazem parte da arquitetura da porta *input gate*. A etapa c'_t é representada pela Equação 27, onde σ refere-se à função *sigmoid* e \tanh a função tangente hiperbólica. W_c é o peso do neurônio referente ao *input gate* e b_i é o *biases*. Por fim C_{t-1} é a memória do estado anterior.

$$c'_t = f_t c_{t-1} + i_t \tanh(W_{xc} x_t + W_{hc} h_{t-1} + W_{cf} + b_f) \quad (27)$$

A etapa i_t da porta *input gate* é representada pela Equação 28, onde, σ refere-se à função *sigmoid*; W_i o peso do neurônio referente ao *input gate*; H_{t-1} o valor de saída do bloco anterior, referente ao tempo anterior; b_i é o *biases* da porta e por fim, C_{t-1} é a memória do estado anterior. O resultado das duas etapas da porta *input gate* são multiplicados e depois somados com a saída do *forget gate*.

$$i_t = \sigma(W_{xi} x_t + W_{hi} h_{t-1} + W_{ci} C_{t-1} + b_i) \quad (28)$$

A última porta mostrada na Figura 11 é a porta *output gate*. Esta é a porta de saída onde a rede decide qual será a sua saída. A porta *output gate* é representada pela Equação 29. Semelhante as etapas anteriores, σ refere-se à função *sigmoid*; W_o o peso do neurônio referente ao *output gate*; H_{t-1} o valor de saída do bloco anterior, referente ao tempo anterior; C_{t-1} é a memória do estado anterior; finalmente b_o é o *biases* da porta.

$$o_t = \sigma(W_{xo} x_o + W_{ho} h_{t-1} + W_{co} C_{t-1} + b_o) \quad (29)$$

2.7 SÉRIES FINANCEIRAS

A grande parte dos métodos desenvolvidos para previsão de séries financeiras utilizam janelas de negociação de dias, semanas, meses ou anos. Os dados destas janelas de negociação são comumente utilizadas pois são mais fáceis de ser encontrados de forma gratuita nas Bolsas Mundiais¹². Este trabalho traz contribuições ao explorar séries com intervalos de minutos, consideradas séries de curto prazo. A obtenção destas séries é mais complexa pois depende dos dados chamados de *tick-by-tick* ou também conhecido como dados de alta frequência¹³.

Embora não exista uma definição muito clara para dados de alta frequência, autores

¹²<http://www.b3.com.br> - Acesso em 01/06/2019.

¹³Dados de alta frequência também são conhecidos como *High Frequency Data* e utilizados em robôs HFTs *High Frequency Trading* (KEARNS; NEVMYVAKA, 2013)

concordam que sua principal propriedade é a grande granularidade de seus dados (KEARNS; NEVEMYVAKA, 2013). Dados de alta frequência armazenam todos os negócios realizados em um determinado período sob forma de eventos. Cada evento é chamado de *tick*. Cada *tick* é considerado uma negociação do mercado e seu horário de execução é armazenado em microssegundos. Utilizando as informações de alta frequência é possível construir as séries de ativos com diversas granularidades diferentes.

Neste trabalho as séries foram obtidas através de um FTP disponibilizado pela BOVESPA. O FTP estava disponível gratuitamente até final de 2018. Embora ainda acessível, o serviço está descontinuado ¹⁴. No total, dez ativos com a maior liquidez de *ticks* no período foram previamente selecionados para este estudo.

2.8 TRABALHOS RELACIONADOS

Há uma vasta quantidade de artigos científicos que investigam o uso de técnicas de inteligência computacional aplicada à previsão de séries temporais financeiras. A maior parte deles tem como principal objetivo alimentar sistemas de apoio à tomada de decisão. Conforme apresentado no Capítulo 1, a possibilidade de obter lucros durante os deslocamentos dos preços torna o problema da previsibilidade um dos mais importantes assuntos deste segmento (ABU-MOSTAFA; ATIYA, 1996). A maior parte dos trabalhos analisados utiliza prazos, mercados e ativos específicos. E cada um deles busca encontrar boas oportunidades de ganho e assim sugerir melhores oportunidades de entrada para investidores.

Cada mercado, ativo ou período é abordado de uma forma distinta. Os fatores que contribuem para esta diversidade de estratégias podem ser: (i) a disponibilidade dos dados em alguns países não são públicas e envolvem custos para aquisição de bases, tornando assim difícil o acesso às mesmas janelas temporais e a validação de alguns métodos¹⁵; (ii) o próprio interesse do grupo de pesquisa em avaliar algum período específico seja: a bolsa de seu país, algum ativo ou um prazo de interesse; (iii) por fim, segundo Cavalcante et al. (apud VANSTONE; FINNIE, 2009), alguns estudos não são compartilhados com a comunidade científica pois ao obterem bons retornos, seus desenvolvedores decidem manter seu capital intelectual confidencial. De forma geral, estes aspectos dificultam obter um método de comparação entre os estudos avaliados.

Nas próximas Seções são apresentados os artigos científicos analisados relacionados

¹⁴<ftp://ftp.bmf.com.br/marketdata/> - Acesso em 01/07/2019

¹⁵A bolsa de valores do Brasil é um exemplo deste caso. Dados de microestrutura de mercado ou prazos muito curtos dependem de compensação financeira para serem obtidos.

ao objetivo desta pesquisa. As bases utilizadas nas consultas para esta revisão de literatura foram as bibliotecas digitais do google Scholar, IEEEExplore e ScienceDirect. Os trabalhos foram divididos em dois grupos. Na Seção 2.8.1 o problema da previsibilidade é abordado apresentado os trabalhos que utilizam métodos de classificação como estratégia de solução. Na Seção 2.8.2 são apresentados trabalhos referentes ao mercado de ações do Brasil. Por fim uma a Tabela 1 consolida os estudos analisados em ambas as Seções.

2.8.1 PREVISÃO DA DIREÇÃO

A abordagem mais clássica em trabalhos de previsão de séries financeiras é estimar o preço futuro de uma série. Estudos recentes tem demonstrado que pode-se obter melhores resultados estimando a direção do movimento ao invés do preço (CHEN et al., 2003).

Em seu estudo, Leung et al. (2000) demonstrou que diferentes classificadores atingiram um desempenho superior aos métodos tradicionais de previsão de preço. O método comparou a quantidade de vezes que seu modelo de previsão acertou a direção dos índices S&P 500¹⁶, FTSE 100¹⁷ e Nikkei 225¹⁸ em um horizonte mensal. Leung também observou que o desempenho dos sistemas de estratégia foi melhor quando utilizado métodos de classificação como tomada de decisão.

Considerando $D = F(S1_{t-1}, S2_{t-1})$ onde, D é uma variável categórica que representa a direção do movimento, S1 o índice S&P 500 e S2 o Yen Japonês, Huang et al. (2005) utilizou SVM para prever a direção dos movimentos semanais do índice Nikkei 255. O estudo adotou uma supervisão binária onde valores de fechamento semanal superiores à semana anterior eram classificados com “+1” e o inverso como “-1”. Em um estudo mais recente, Kara et al. (2011) também comparou o desempenho de SVM com o de RNAs para a Bolsa de Istambul. Sua abordagem utilizou uma estratégia de previsão de direção e um conjunto de 10 indicadores técnicos.

Com uma estratégia de supervisão semelhante à Huang et al. (2005), Saad et al. (1998) também utilizou uma abordagem de classificação binária em 10 ações americanas. Saad associou à supervisão das classes um intervalo de confiança. Desta forma, retornos superiores a 2% eram classificados com “+1” e valores inferiores a 2% eram considerados “-1”.

Ballings et al. (2015) forneceu um estudo comparativo avaliando o desempenho dos

¹⁶Índice da Bolsa de Valores Americana composto por quinhentos ativos listados nas Bolsas de NYSE ou NASDAQ.

¹⁷Índice que representa um conjunto de 100 ações representadas na Bolsa de Valores de Londres.

¹⁸Índice referente as 255 maiores empresas da Bolsa de Valores de Tóquio.

principais métodos *ensemble*, incluindo RF, para previsões de direção um ano a frente. Entre os métodos avaliados RF foi o que obteve o melhor resultado para o problema. A métrica utilizada para avaliação foi a AUC. Os autores reforçam o poder dos algoritmos *ensemble* e principalmente RF.

Patel et al. (2015) avaliou o desempenho da direção do movimento das ações da Bolsa do mercado Indiano, a NSE do inglês *National Stock Exchange of India Limited*. O estudo comparou quatro diferentes classificadores, dentre eles: RNN, SVM, RF e NB do inglês *Naive Bayes*. Patel utilizou um conjunto de 10 indicadores de análise técnica de uma abordagem discreta, ou seja, os valores eram apresentados à rede como “+” ou “-1”. As métricas utilizadas para avaliar o desempenho dos classificadores foram acurácia e F1. A abordagem de Patel et al. (2015) referente à discretização dos indicadores será adotada neste trabalho.

Li e Liao (2017) conduziram um estudo comparando diversos classificadores, dentre eles: SVM, MLP e LSTM. Diferente dos outros trabalhos apresentados, seu método de supervisão utilizou três classes, onde uma delas considerava um estado neutro. Utilizando uma janela de curto prazo, Dixon et al. (2017) usou indicadores técnicos e AP em 43 papéis da CME¹⁹. De forma semelhante a Li e Liao (2017), Dixon et al. (2017) supervisionou o conjunto de dados com três classes. A saída das três classes foram definidas como -1,0,1 onde, o valor de -1 foi usado para sinalizar um movimento de “alta”, +1 foi usado para sinalizar um movimento de “baixa” e 0 foi considerado uma região de incerteza. O critério entre as classes 0 e +1 e 0 e -1 foi definido buscando realizar um balanceamento entre as classes. As contribuições de Li e Dixon relacionadas à supervisão das classes serão adotadas neste trabalho.

Fischer e Krauss (2018) compararam o desempenho da classificação de redes LSTM com RF e Deep Neural Networks (DNN). O objetivo foi compreender se a capacidade de aprendizado sequencial das redes LSTM poderia beneficiar a predição de direção das ações do índice da Bolsa americana. Fischer acompanhou o processo de supervisão observado em outros trabalhos (HUANG et al., 2005; PATEL et al., 2015) e construiu uma classificação binária com os dados diários das ações. Os resultados foram apresentados utilizando acurácia e valores estatísticos de um sistema de estratégia construído. A rede LSTM demonstrou robustez e melhores resultados para extrair informações sequenciais das séries financeiras escolhidas.

2.8.2 ANÁLISE DE PREVISIBILIDADE

A aplicação de medidas dinâmicas relacionadas ao caos são utilizadas em diversos trabalhos sobre séries financeiras. O grau de interesse parte da hipótese de que tais séries possam

¹⁹CME - *Chicago Mercantile Exchange* - é a Bolsa de mercadorias com sede em Chicago.

carregar algum comportamento caótico em alguma fração temporal. Evidenciar janelas onde não há algum indício de caos poderia abrir uma possibilidade de previsão, ao menos de curto prazo (GOMES; JUNIOR, 2010). A grande parte da literatura investiga as séries como um todo, buscando identificar se existe presença de caos na série completa. Outros trabalhos abordam o uso do EL como método de controle ou dentro de janelas temporais.

Rafuzaman (2014) propôs um sistema composto pelos valores da séries financeira e parâmetros de controle do caos. Além dos valores de abertura, máxima, mínima e fechamento, a utilização da variância, *Shannon entropy* e EL foram considerados como variáveis de entrada do sistema. Um parâmetro de defasagem ou atraso na série com valor igual a um foi escolhido em seu experimento. O sistema proposto foi concebido para prever um horizonte de 15 amostras a frente.

Em um estudo puramente matemático, Tsionas e Michaelides (2017) aplicaram uma técnica baseada no conjunto da inferência Bayesiana e o EL. Retornos diários de séries de seis países diferentes foram avaliadas em uma janela de 11 anos. Os autores observaram a presença de caos antes e após a crise nos seis países estudados. O Brasil não fazia parte do estudo proposto.

Saad et al. (1998), previamente citado na Seção 2.8.1, também abordaram o tema da análise de previsibilidade em seu trabalho. Em uma etapa preeliminar, os autores conduziram três experimentos utilizando diagrama de fase, dimensão de correlação e EL. Os testes utilizando EL foram considerados os mais eficientes para se determinar o caos.

Tabela 1 - Visão geral dos trabalhos analisados.

Citação Autor	Bolsa	Tempo	Entrada	Classificador	AP	Supervisão	Métrica
Saad et al. (1998)	NYSE	Diário	Outros	PNN	Sim	2 classes	Outros
Leung et al. (2000)	NYSE / FTSE	Mensal	Outros	PNN	Não	2 classes	Outros
Chen et al. (2003)	TSE	Mensal	Macroeconomia	PNN	Não	2 classes	Outros
Huang et al. (2005)	NIKKEY	Semanal	Outros	SVM	Não	2 classes	Outros
Kara et al. (2011)	Istambul	Diário	Indicadores	SVM e RNs	Não	2 classes	Acurácia/F1
Rafuzaman (2014)		Diário	Outros	-	Não	-	Outros
Patel et al. (2015)	NSE	Diário	Indicadores	RNN, SVM, RF, NB	Não	2 classes	Acurácia/F1
Ballings et al. (2015)	Europa	Anual	Outros	-	Não	2 classes	AUC
Dixon et al. (2017)	CME	Intradiário	Indicadores	DNN	Não	3 classes	Outros
Li e Liao (2017)	SSE	Diário	Indicadores	MLP, RNN, SVM, NB, LSTM	Não	3 classes	Acurácia
Tsionas e Michaelides (2017)	Diversas	Anual	Outros	-	Sim	-	Outros
Fischer e Krauss (2018)	NYSE	Diário	Retorno	RF, DNN, LSTM	Não	2 classes	Acurácia

Fonte - O autor.

3 METODOLOGIA

Este capítulo apresenta a metodologia adotada neste trabalho onde, três diferentes métodos são avaliados. Na sequência, a Seção 2.7 detalha os processos utilizados na construção dos conjuntos de dados dos experimentos. Na Seção 3.2 o método de discretização de indicadores proposto por Patel et al. (2015) é apresentado em uma abordagem comparativa ao método tradicional de valores contínuos. Na sequência, a Seção 3.3 apresenta o segundo método, onde atributos construídos com o expoente de Lyapunov são inseridos ao conjunto de dados do método anterior. Finalmente, a Seção 3.4 aborda outra estratégia utilizando o expoente de Lyapunov como método de controle de janelas caóticas.

3.1 SÉRIES FINANCEIRAS

Dentro do contexto apresentado, este trabalho utiliza os dados de alta frequência do Índice BOVESPA para construir séries com frequência de um minuto. A Tabela 2 apresenta os ativos escolhidos para uma análise inicial assim como sua classificação setorial.

Tabela 2 - Tabela com os ativos selecionados para análise

Código	Nome de Pregão	Atividade Principal	Classificação Setorial
PETR4	PETROBRAS	Petróleo, Gás E Energia	Exploração
VALE3	VALE	Mineração	Materiais Básicos
ITUB4	ITAUUNIBANCO	Atividade Bancária	Financeiro
BBDC4	BRADESCO	Operações Bancárias	Financeiro
ABEV3	AMBEV S/A	Distribuição Bebidas.	Consumo não Cíclico
ITSA4	ITAUSA	Gestão	Financeiro
BBAS3	BRASIL	Atividade Bancária	Serviços Financeiros
B3SA3	B3	Administrar Mercados	Serviços Financeiros
KROT3	KROTON	Educação	Consumo Cíclico
PETR3	PETROBRAS	Petróleo, Gás e Energia	Exploração

Fonte - B3 S.A. - Brasil, Bolsa, Balcão

3.1.1 PRÉ-PROCESSAMENTO

Uma vez definidos os ativos e os indicadores técnicos que serão utilizados nas análises iniciais, uma etapa de pré-processamento foi aplicada aos dados. Nesta etapa, condições de contorno e remoção de *outliers*¹ foram aplicadas. Uma etapa final detalhará a normalização dos atributos da série.

Neste trabalho foram utilizados dados históricos com início em 14 de maio de 2018 e término em 17 de abril de 2019. Embora o ideal seja realizar os experimentos em diferentes ciclos de macro tendência (ELTON et al., 2009), as restrições relacionadas à obtenção de dados minuto a minuto impediu esta análise. Todavia, séries de curto prazo apresentam uma volatilidade superior a dados de prazo maior. Desta forma, em um dia de negociação diferentes ciclos surgem garantindo, assim, esta robustez.

A série completa foi segmentada em três diferentes conjuntos: treinamento, validação e teste. Da série completa, 20% das amostras finais foram utilizadas para teste. Dos 80% restante, 70% foi reservado para treinamento e 30% para validação. Os valores escolhidos seguiram como referência os valores adotados na literatura avaliada, conforme Tabela 1.

Em todas as séries da Tabela 2 também foram aplicados ajustes com horário das negociações. Segundo Vicente et al. (2012), em um relatório técnico do Banco Central, há uma relação significativa entre a volatilidade de um ativo e seu horário de negociação. Nos primeiros minutos de negociação, assim como no fechamento do mercado há uma volatilidade superior à apresentada durante a janela de almoço. Desta forma, assim como no trabalho de Chong et al. (2017), as primeiras e as últimas amostras de cada dia foram suprimidas das séries utilizadas. Considerando estas premissas, todas as séries deste trabalho utilizaram uma janela de horário com início às 10:30 e fim às 16:30.

Em tarefas de processamento de dados, o tipo de unidade de medida usada em um atributo pode afetar o resultado do experimento. Transformações como, *gramas* para *quilogramas* aumentará o intervalo para este atributo e, assim, as redes poderão atribuir um peso maior para este atributo (HAN et al., 2012). Outra questão semelhante são atributos diferentes com valores e intervalos diferentes na mesma série. Embora o caso agora não seja relativo à sua unidade de medida, o fato de haver grande diferença na amplitude do valor pode beneficiar alguns atributos.

Em algumas séries, notícias locais referentes à troca de direção em algumas empresas,

¹*Outliers* são consideradas amostras que apresentam uma diferença de valor significativa das demais observações do grupo.

assim como recentes eventos com barragens no Brasil influenciariam significativamente alguns intervalos dos conjuntos de dados das séries escolhidas. Estes valores apresentaram uma oscilação superior à 10% entre alguns minutos. Estes intervalos foram considerados *outliers* e extraídos dos conjuntos. Desta forma, após a série estar com todos seus atributos inseridos foi aplicada a normalização dos dados conforme a Equação 30.

$$Norm = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (30)$$

3.1.2 APRENDIZADO SUPERVISIONADO

Na Seção 2.6 foram apresentados conceitos sobre classificadores assim como tópicos referentes à tarefa de classificação. Conforme Goldschmidt Ronaldo (2015), a supervisão de dados se resume em atribuir um ou mais valores para cada instância em um conjunto de dados. Sendo assim, dependendo da abordagem de supervisão escolhida, diferentes classes podem ser criadas para um dado problema. Esta tarefa acaba por criar um problema de desbalanceamento de classes, uma vez que a estratégia de supervisão poderá beneficiar um conjunto de registros. Neste caso, o classificador será melhor treinado para aquele conjunto onde existem mais amostras.

Neste trabalho o modelo de supervisão adotado seguiu como referência os trabalhos de Dixon et al. (2017) e Li e Liao (2017). Desta forma, três diferentes classes foram criadas. Suas respectivas saídas são: $\{-1, 0, 1\}$ onde, o valor de $\{-1\}$ refere-se a queda no próximo minuto; $\{+1\}$ à alta no próximo minuto e $\{0\}$ à uma região de incerteza.

A classe referente à região de incerteza $\{0\}$ foi criada buscando obter um balanceamento entre as classes e privilegiar movimentos de alta ou baixa com maior relevância. Alguns estudos utilizaram apenas duas classes em seus métodos (HUANG et al., 2005; PATEL et al., 2015). Todavia, desta forma, o classificador será treinado para prever movimentos de alta ou baixa de qualquer amplitude. Este cenário não é desejável principalmente para séries de curto prazo, pois alguns movimentos podem não justificar uma operação.

Para criar as classes foi utilizada a variável *retorno*. Ela foi calculada através da Equação 13. Uma vez que a classe $\{0\}$ não significa um alta ou uma baixa, um foi estimado um valor para o atributo *retorno*. Este valor estabeleceu um critério de transição entre as classes. Utilizando estatística descritiva, o desvio padrão dos *retornos* minuto a minuto foi utilizado para analisar qual o melhor valor de transição da região de incerteza. As séries apresentavam desvio padrão médio de 0,05% entre cada minuto. Desta forma, o cálculo foi realizado para

cada série individualmente e os valores permaneceram entre 0,03% e 0,05%. Assim, valores acima da região de incerteza foram considerados pertencentes a classe $\{+1\}$. Valores inferiores a região de incerteza foram considerados pertencentes a classe $\{-1\}$.

Todavia, o critério da região de incerteza não garantiu que as classes fossem perfeitamente balanceadas. Desta forma, foi realizada uma nova análise estimando o quanto cada ativo estava balanceado. O critério final da escolha dos ativos deste trabalho foram definidos a partir do critério do balanceamento das classes. Uma vez que o objetivo era compreender o comportamento dos métodos em séries de curto prazo, o desbalanceamento das classes poderia de alguma forma influenciar ou beneficiar alguns tipos de classificadores. Desta forma, foi definido um máximo de 3%. Sendo assim, os ativos os quais o balanceamento das classes estivessem com valores superior a 3% foram desconsiderados deste estudo. Por fim, a Tabela 3 consolida os atributos escolhidos para cada série de dados. Com exceção do atributo *DATETIME* e *CLASSE*, a normalização foi aplicada a todos os demais atributos.

Tabela 3 - Tabela com os atributos das séries financeiras

Atributo	Descrição
DATETIME	Hora e minuto de cada amostra.
SMA	Média móvel dos últimos 10 minutos.
WMA	Média ponderada últimos 10 minutos.
MOM	Momentum dos últimos 10 minutos.
K%	Estocástico K%, 5 e 3 minutos.
D%	Estocástico D%, 5 e 3 minutos.
IFR	Índice de força relativa últimos 14 minutos.
MACD	MACD, média 12 e 26 minutos.
R%	Indicador %R dos últimos 14 minutos.
A/D	Acumulação e distribuição 3 e 10 minutos.
CCI	Indicador CCI dos últimos 14 minutos.
VOLUME	Volume do minuto corrente.
FECHAMENTO	Valor de fechamento do minuto.
CLASSE	Classes $\{-1,0,1\}$

Fonte - O autor.

3.1.3 RECURSOS UTILIZADOS E PARÂMETROS DEFINIDOS

Todos os experimentos foram realizados utilizando servidores com processador Intel Core i7 3.30Ghz e GPU Nvidia, modelo Titan XP. O ambiente foi isolado através da ferramenta *virtualenv* que rodou utilizando sistema operacional Ubuntu 14.04 LTS. Todos os códigos

escritos utilizaram a linguagem de programação Python.

Para as tarefas de aprendizado de máquina, os *frameworks* utilizados foram: Keras², Sklearn³ e Pandas⁴. A criação dos indicadores técnicos foi realizada utilizando a biblioteca TA-Lib⁵. Finalmente, para os métodos de Lyapunov foi utilizada a biblioteca *nolds*⁶, uma biblioteca pronta com vários métodos de sistemas dinâmicos.

Todos os experimentos utilizaram os parâmetros apresentados na Tabela 4 para a configuração dos classificadores. Com exceção do classificador LSTM, os parâmetros utilizados no modelos SVM, RF e MLP foram mantidos dentro do padrão do *framework Sklearn* utilizado. Uma vez que o objetivo era avaliar o desempenho dos métodos, e a influência do acréscimo de atributos, nenhum parâmetro foi ajustado para obter melhor desempenho dos classificadores.

Tabela 4 - Parâmetros utilizados nos experimentos.

Parâmetros	Valores
SVM	<i>kernel</i> =linear
RF	estimadores=100
MLP	camadas(10, 10, 10)
LSTM Camadas Ocultas	32
LSTM <i>Batch Size</i>	6500
LSTM Épocas	50
LSTM Janela avaliadas	(60 / 120 / 240 / 480)
LSTM <i>loss</i>	<i>categorical_crossentropy</i>
LSTM Otimizador	<i>adam</i>
LSTM Ativação	tanh / softmax
LSTM <i>Dropout</i>	0,25

Fonte - O autor.

3.2 MÉTODO DOS INDICADORES DISCRETIZADOS

Conforme discutido no Capítulo 2, estudos tem demonstrado ser promissor estimar a direção do movimento para tratar o problema da previsão de séries temporais financeiras. Dentre estes estudos encontram-se os trabalhos realizados por Kara et al. (2011) e Patel et al. (2015). Em seus trabalhos, a previsão de direção foi abordada como um problema de classificação. Na

²<https://keras.io/>

³<https://scikit-learn.org/>

⁴<https://pandas.pydata.org/>

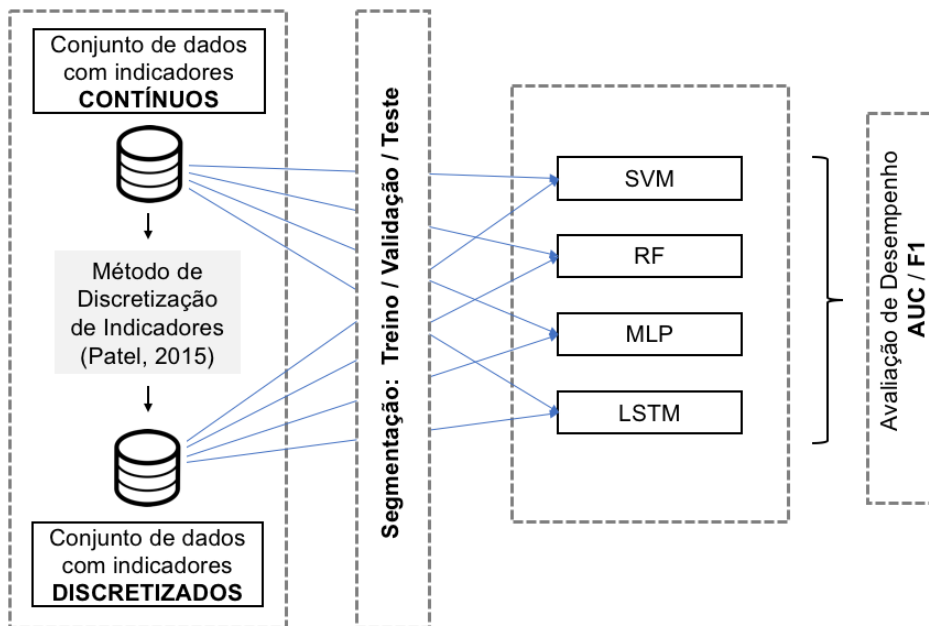
⁵<https://github.com/mrjbq7/ta-lib>

⁶<https://pypi.org/project/nolds/>

Seção 2.8.1, foi discutido os principais tipos de classificadores adotados na maior parte dos estudos recentes.

Nesta primeira etapa foram aplicados métodos já conhecidos na literatura em séries de curto prazo definidas na Seção 2.7. O principal objetivo foi observar o desempenho destes métodos e criar um *baseline* a fim de comparar os resultados com os experimentos seguintes. Uma visão geral do método pode ser observada na Figura 12.

Figura 12 - Visão Geral do método de indicadores discretizados e contínuos.



Fonte - O autor.

O conjunto de dados final da Seção 2.7 foi definido como conjunto de dados com indicadores contínuos e todos os seus atributos podem ser observados na Tabela 3. Utilizando o critério técnico de cada indicador apresentado na Seção 2.2, um novo conjunto de dados chamado conjunto de dados com indicadores discretos⁷ foi construído utilizando o método proposto por Patel et al. (2015). O novo conjunto de dados manteve os atributos do conjunto anterior, exceto os atributos referente aos dez indicadores. Estes atributos foram substituídos por dez novos atributos discretos.

Para discretizar os indicadores foram utilizados os critérios técnicos detalhados na Seção 2.2. Valores de {0} e {1} foram usados para preencher os novos atributos. Para os indicadores SMA e WMA, quando o valor do preço corrente ultrapassava o valor do indicador

⁷O termo discreto ou discretizado será utilizado neste trabalho para referir-se ao conjunto de dados onde os atributos foram discretizados. A discretização neste caso, se deve ao fato do conjunto de dados não ser preenchido com valores contínuos e sim com valores "0" ou "1".

para cima o novo atributo armazenava o valor {1} e o inverso {0}. Para valores positivos do indicador MOM seu novo atributo armazenou o valor {1} enquanto em valores negativos do indicador o valor armazenado foi {0}.

Os indicadores K%, D%, MACD, R%, A/D e CCI entre 200 e -200, tem um comportamento semelhante. Quando o valor corrente destes indicadores eram superiores ao seu valor anterior os atributos foram preenchidos como {1} e o contrário como {0}. Valores do indicador IFR superior a 70 foi atribuído {0} enquanto valores menores que 30 {1}. Por fim, para CCI maior que 200 seu novo atributo armazenou o valor {0} enquanto em valores inferiores a -200 o valor armazenado foi {1}. A Tabela 5 consolida os atributos discretos para cada série de dados.

Tabela 5 - Tabela com os atributos discretos das séries financeiras

Atributo	Descrição ou Critério Técnico
DATETIME	Hora e minuto de cada amostra.
SINAL SMA	$SMA > V_{cur} = 1; SMA < V_{cur} = 0$
SINAL WMA	$WMA > V_{cur} = 1; WMA < V_{cur} = 0$
SINAL MOM	$MOM > MOM_{t-1} = 1; 0$
SINAL K%	$K\% > K\%_{t-1} = 1; 0$
SINAL D%	$D\% > D\%_{t-1} = 1; 0$
SINAL IFR	$IFR > 70 = 0; IFR < 30 = 1$
SINAL MACD	$MACD > MACD_{t-1} = 1; 0$
SINAL R%	$R\% > R\%_{t-1} = 1; 0$
SINAL A/D	$A/D > A/D_{t-1} = 1; 0$
SINAL CCI	$CCI > 200 = 0; CCI < 200 = 1; CCI > CCI_{t-1} = 1; 0$
VOLUME	Volume do minuto corrente.
CLASSE	Classes {-1,0,1}

Fonte - O autor.

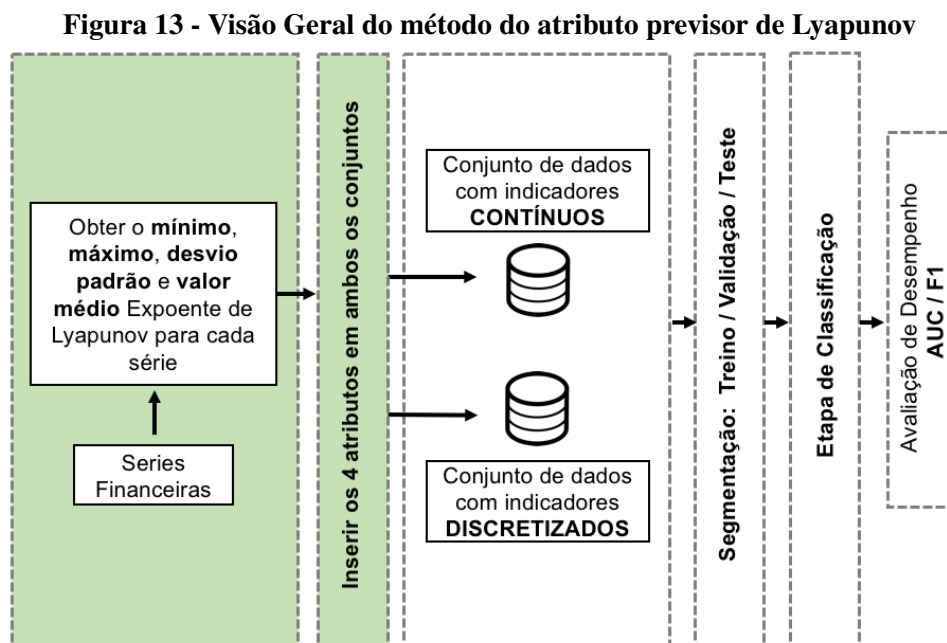
Uma vez consolidada a etapa de construção do conjunto de dados discreto, foi avaliado o desempenho dos classificadores SVM, RF, MLP e LSTM. Para isto, os conjuntos de dados foram separados em treinamento, validação e teste. A segmentação seguiu o mesmo processo detalhado na Seção 3.1.1. Embora todos os parâmetros dos classificadores tenham sido mantidos em seus valores padrão, quatro diferentes janelas foram avaliadas para LSTM {60, 120, 240, 480}. Neste experimento prévio o melhor resultado da combinação dos parâmetros e janelas serão apresentados no Capítulo seguinte.

Os resultados da classificação foram avaliados com as métricas área sobre a curva ROC - (AUC) e F1 (vide Seção 2.6.1). As mesmas métricas foram aplicadas em todos os

classificadores e ambos os conjuntos de dados. Os conjuntos escolhidos estavam com as classes balanceadas e, portanto, uma métrica mais simples como a acurácia poderia ter sido usada. AUC e F1 foram utilizadas pois são utilizados na maioria dos estudos analisados, desta forma, pode ser mais fácil a comparação. Neste mesmo aspecto trabalhos futuros podem se beneficiar de ter estes valores mais acessíveis.

3.3 MÉTODO DO ATRIBUTO PREVISOR DE LYAPUNOV

Esta Seção descreve o método chamado de atributo predictor de Lyapunov. Ele foi inspirado no trabalho de Übeyli (2010) e o algoritmo proposto por Eckmann et al. (1986) foi utilizado em seu cálculo. A Figura 13 mostra uma visão geral dos principais elementos que compõem o método. Nesta abordagem foram utilizados os conjuntos de dados da Seção anterior. O destaque desta etapa está no acréscimo de quatro atributos em ambos os conjuntos. Isto foi realizado em uma etapa inicial, conforme destacado na Figura 13.



Fonte - O autor.

Conforme descrito na Seção 2.5, o expoente de Lyapunov é considerado um tipo de métrica analítica para caracterizar o caos. Através do EL é possível verificar a divergência exponencial no tempo de trajetórias vizinhas em uma série temporal. O método de cálculo do EL utilizado foi o proposto por Eckmann et al. (1986). Visto que não é o o escopo deste trabalho implementar o algoritmo de Eckmann et al. (1986), foi utilizada a biblioteca Nolds (2019) para o processamento do cálculo do EL pelo método de Eckmann et al. (1986). A função desenvolvida

para estimar os atributos é mostrada no Algoritmo 1.

Algorithm 1: Função de construção dos atributos de *Lyapunov*

Result: Vetor com quatro atributos de Lyapunov

Parâmetros: w, i, st, c ;

Série = Carrega o valor de fechamento da série;

while *final da série não for atingido* **do**

for $i < (\text{comprimento}(\text{Serie}) - w)$ **do**

 Seleciona o conjunto entre i até janela;

 Calcula vetor o Lyapunov do conjunto selecionado;

 Extrai o menor valor de Lyapunov do vetor ;

 Extrai o maior valor de Lyapunov do vetor ;

 Extrai a média dos valores de Lyapunov do vetor ;

 Extrai o desvio padrão dos valores de Lyapunov do vetor ;

$i = i + 1$;

end

end

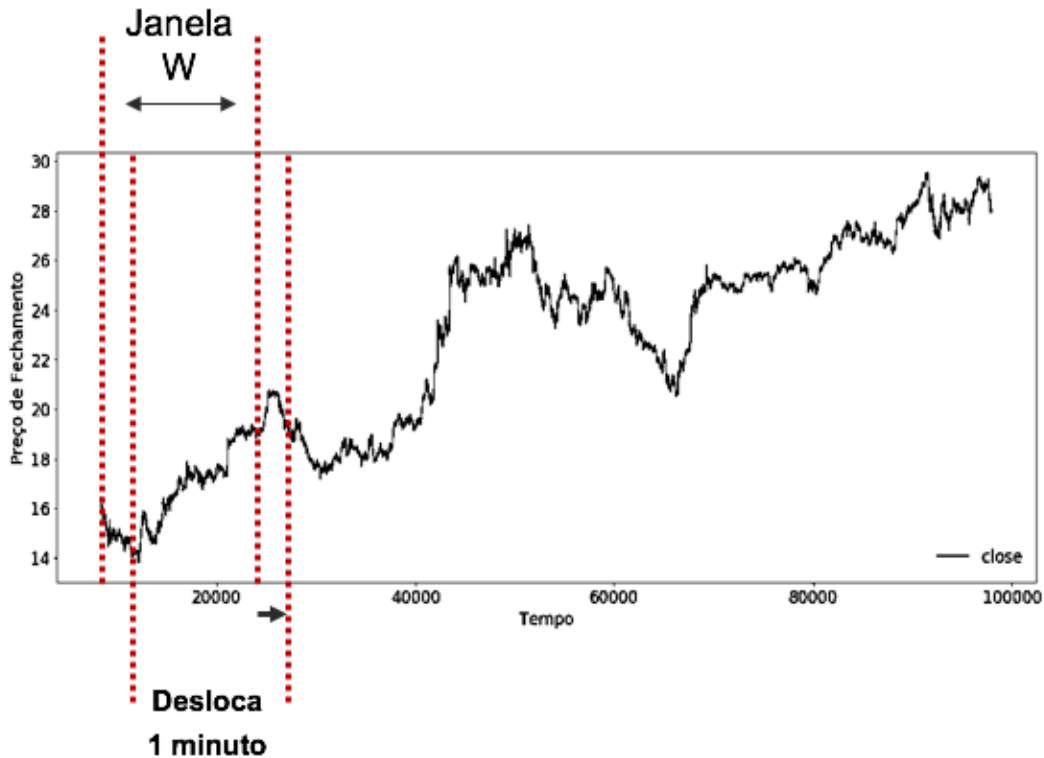
Para cada uma das séries todo o conjunto de dados é percorrido deslizando uma janela de tamanho w minuto a minuto. A Figura 14 auxilia a compreensão deste entendimento. Um conjunto de quatro opções de janelas foram avaliadas, dentre elas: $\{60, 120\}$ utilizando o atributo fechamento do minuto e $\{60, 120\}$ utilizando o atributo SMA (conforme destacado na Tabela 3).

O método foi avaliado utilizando apenas o classificador que obteve o melhor desempenho da Seção 3.2. A segmentação do conjunto de dados seguiu o mesmo processo detalhado na Seção 3.1.1 e o resultado das diferentes janelas foram avaliados no conjunto de dados contínuo e no conjunto de dados discreto. Para avaliar o resultado da classificação foi utilizando a métrica F1 e AUC. Os melhores resultados foram comparados com o método anterior proposto na Seção 3.2.

3.4 MÉTODO DA JANELA DE LYAPUNOV

Nesta etapa foi proposta outra abordagem usando o expoente de Lyapunov (EL). Este método atua como um filtro inicial nas séries de dados. Para isto, o valor do maior EL foi utilizado para identificar regiões com comportamento caótico nas séries e, caso sejam encontradas, removê-las do conjunto. Uma visão geral dos principais elementos que compõem o método é mostrada na Figura 15 onde o destaque desta etapa está na fase inicial do diagrama.

Figura 14 - Exemplo visual do deslocamento da janela para o método do atributo predictor de Lyapunov. A imagem auxilia a compreensão do Algoritmo 1 onde, a cada minuto quatro novos atributos são calculados.



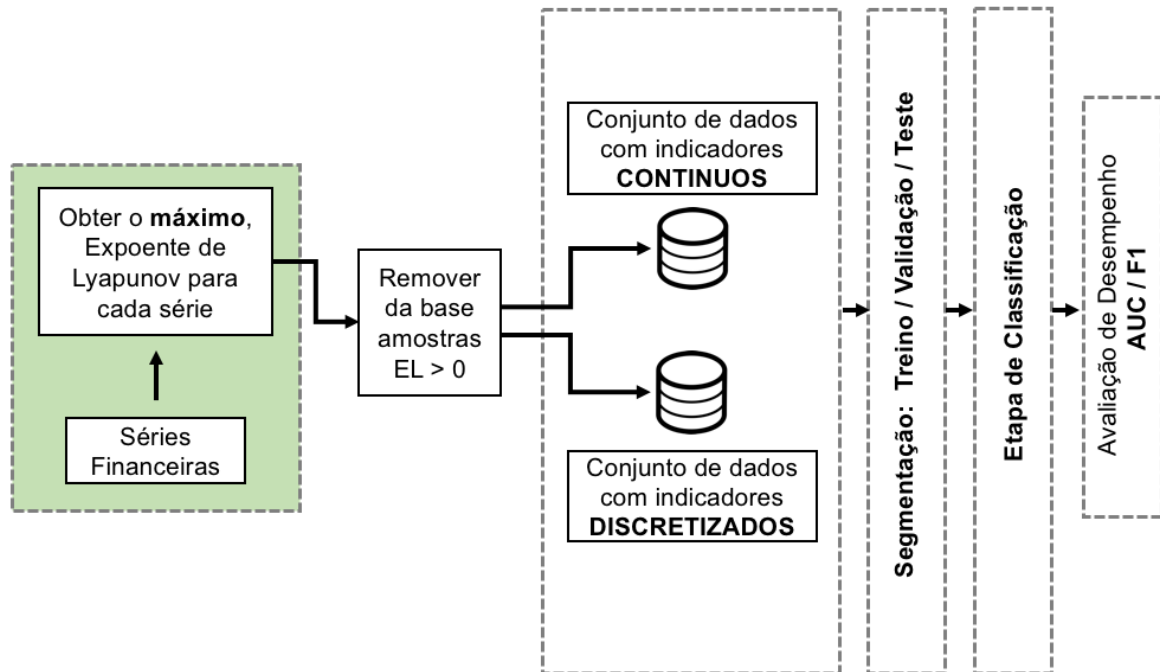
Fonte - O autor.

O restante do diagrama segue uma abordagem similar ao método da Seção 3.3.

O método utilizado para o cálculo do maior EL foi proposto por Eckmann et al. (1986). O seu cálculo retorna um vetor com o valor maior, mínimo e médio do EL. Uma vez que o interesse da busca era por regiões de maior perturbação caótica, apenas os maiores valores de Lyapunov eram interessante para o método. Desta forma, apenas o valor do maior expoente de Lyapunov foi utilizado.

A documentação da biblioteca Nolds (2019) assim como o trabalho de Eckmann et al. (1986) propõem não utilizar um conjunto de dados com poucas amostras. O trabalho de Übeyli (2010) utilizou um conjunto de 240 amostras em seus experimentos. Embora o problema de Übeyli (2010) não envolva séries financeiras sua abordagem foi tomada como referência neste trabalho. Desta forma, quatro diferentes tamanhos de janela foram usadas para estimar o maior EL das séries. Dado que os conjuntos de dados utilizaram valores minuto a minuto, experimentos preliminares utilizaram janelas de 120, 240, 360 e 480 amostras. O algoritmo 2 mostra a rotina construída para estimar o maior expoente de Lyapunov e a Figura 16 auxilia o entendimento de como as janelas foram segmentadas pelo algoritmo 2.

Figura 15 - Diagrama geral de blocos do método de discretização



Fonte - O autor.

Algorithm 2: Função do atributo com o maior expoente Lyapunov

Result: Vetor com o maior expoente de Lyapunov

Parâmetros: w, i, st, c ;

Série = Carrega o valor de fechamento da série;

while final da série não for atingido **do**

for $i < (\text{comprimento}(\text{Serie}) - w)$ **do**

 Seleciona o conjunto entre i até janela;

 Calcula vetor o Lyapunov do conjunto selecionado;

 Extrai o maior valor de Lyapunov do vetor ;

$i = i + w$;

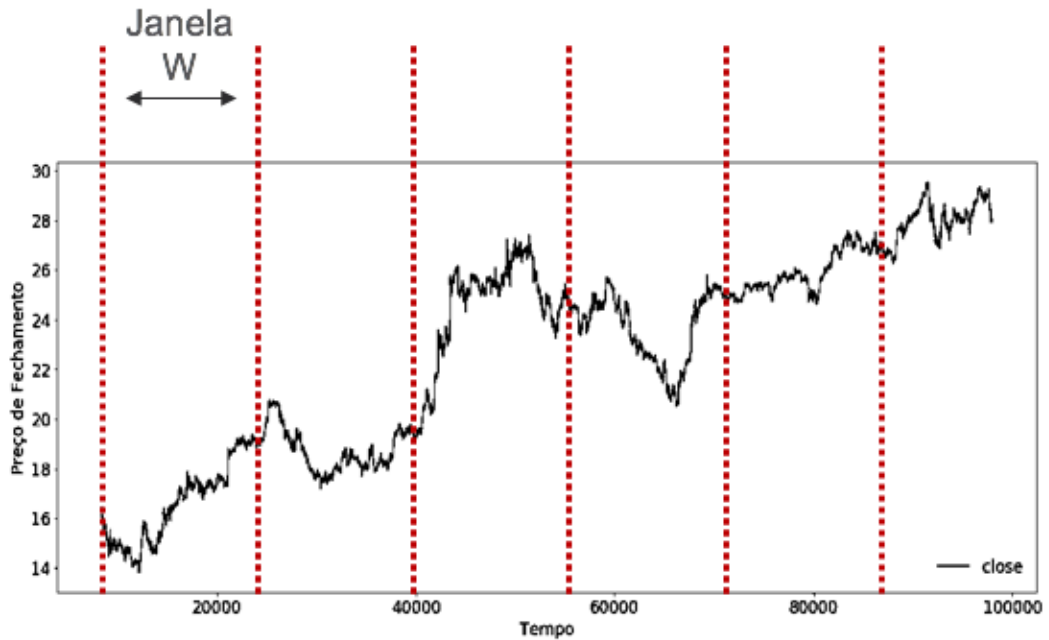
end

end

Após a execução do Algoritmo 2, todas as instâncias com valores de Lyapunov superiores a zero foram removidas das séries. Este experimento utilizou os conjuntos de dados com os atributos de Lyapunov criados para o experimento da Seção 3.3. A segmentação dos conjuntos de dados seguiu o processo detalhado na Seção 3.1.1. O resultado do processo de classificação foi realizado utilizando a métrica F1 e AUC. Os melhores resultados foram

comparados com o método proposto na Seção 3.2.

Figura 16 - Exemplo visual do deslocamento da janela para o método da janela Lyapunov. A imagem auxilia a compreensão do Algoritmo 2 onde, uma janela de tamanho w é usada e em cada janela o máximo expoente é calculado.



Fonte - O autor.

4 RESULTADOS E DISCUSSÃO

Este Capítulo apresenta os resultados dos experimentos realizados com os métodos propostos no Capítulo 3. A Seção 4.1 apresenta os resultados obtidos do pré-processamento e da supervisão de dados realizados nas séries de dados financeiras. A Seção 4.2 apresenta os resultados obtidos através dos experimentos realizados com indicadores técnicos contínuos e discretizados. Na Seção 4.3 os resultados da inclusão de atributos de Lyapunov no conjunto de dados são apresentados. Por fim, na Seção 4.4 os resultados relacionados ao experimento com a janela de Lyapunov são apresentados.

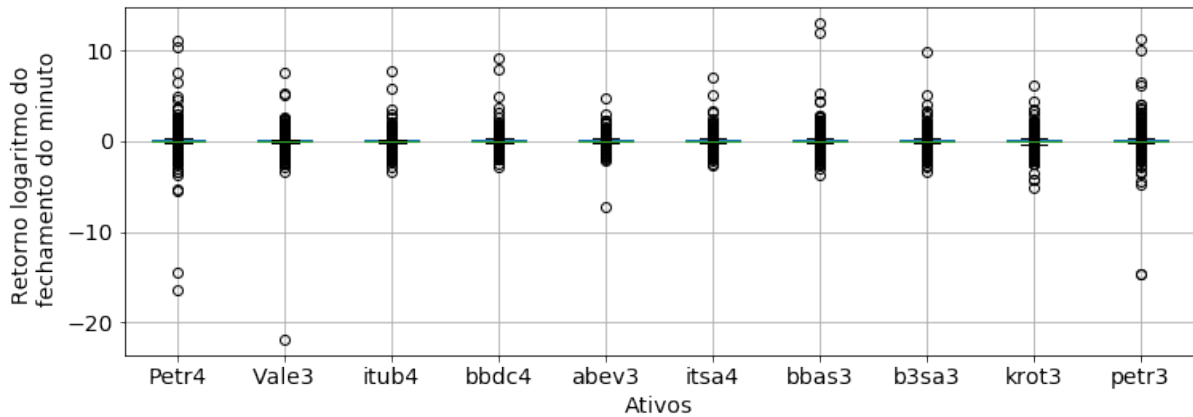
4.1 SERIES FINANCEIRAS

Esta Seção apresenta os resultados do pré-processamento e da supervisão de dados realizado nas séries financeiras e detalhado na Seção 3.1.1 e Seção 3.1.2, respectivamente. Inicialmente, dez séries da BM&FBOVESPA foram selecionadas para uma análise inicial. Na Figura 17 é mostrado um gráfico bloxplot com o resultado inicial dos retornos logaritmos. Os valores apresentados foram calculados minuto a minuto através da Equação 13 para todos os ativos selecionados.

Conforme detalhado na Seção 3.1.1, em todas as séries foram aplicados ajustes nos horários das negociações. As instâncias das séries que foram influenciadas por notícias locais criando *outliers* foram suprimidas da série. O resultado final dos retornos logaritmos sem os *outliers* para todas as séries é mostrado na Figura 18.

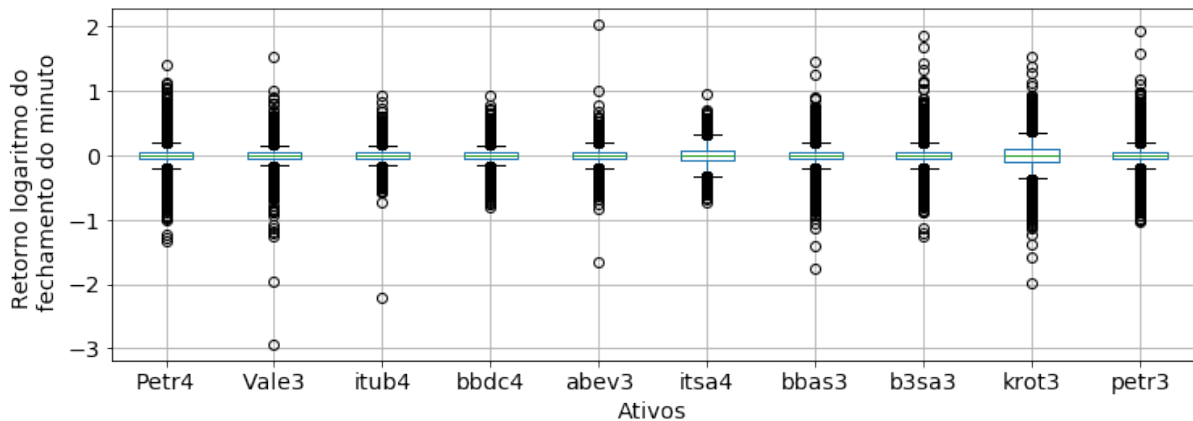
Conforme discutido na Seção 3.1.2, três diferentes classes foram criadas. Suas respectivas saídas são: -1,0,1 onde, o valor de -1 refere-se a queda no próximo minuto; +1 à alta no próximo minuto e 0 à uma região de incerteza. A região de incerteza foi criada buscando um balanceamento das classes. A fim de evitar que o desbalanceamento das classes pudesse afetar os experimentos, o desvio padrão das classes foi calculado e o resultado é mostrado na Tabela 6. Foi estabelecido um critério de desvio padrão de 3%. Desta forma, ativos com o valor do desvio padrão superior a 3% foram desconsiderados deste estudo. Na Tabela 6 os ativos

Figura 17 - Retornos minuto a minuto dos dez ativos de curto prazo da BM&FBOVESPA.



Fonte - O autor.

Figura 18 - Retornos minuto a minuto dos dez ativos de curto prazo da BM&FBOVESPA sem outliers.



Fonte - O autor.

escolhidos foram sinalizados através de um asterisco na coluna ATIVO.

4.2 EXPERIMENTO 01: INDICADORES DISCRETIZADOS E CONTÍNUOS

Esta Seção apresenta os resultados da comparação do desempenho dos métodos propostos por Kara et al. (2011) e Patel et al. (2015) detalhados na Seção 3.2. O principal objetivo deste experimento é criar um *baseline* com métodos da literatura a fim de comparar seus resultados com os experimentos seguintes. Dois conjuntos de dados com indicadores técnicos contínuos e discretizados foram criados. Com estes conjuntos, quatro diferentes classificadores foram avaliados através das métricas *F-measure* e ROC.

Tabela 6 - Tabela com a distribuição das classes.

ATIVO	CLASSE (-1)	CLASSE (0)	CLASSE (+1)	DESVIO PADRÃO
PETR4	37%	25%	38%	4%
VALE3	30%	39%	30%	5%
ITUB4	29%	41%	29%	7%
*BBDC4	32%	36%	32%	2%
*ABEV3	32%	37%	31%	3%
ITSA4	28%	44%	28%	9%
BBAS3	31%	39%	31%	5%
*B3SA3	34%	32%	34%	1%
*KROT3	32%	37%	32%	3%
*PETR3	34%	32%	34%	1%

Fonte - O autor. (*Ativos escolhidos para este trabalho.)

Dentro deste contexto, o experimento levanta a hipótese H1, onde: indicadores discretizados têm um desempenho superior de classificação quando comparados à indicadores contínuos para os cinco papéis de curto prazo da BMF&BOVESPA.

4.2.1 EXPERIMENTOS REALIZADOS

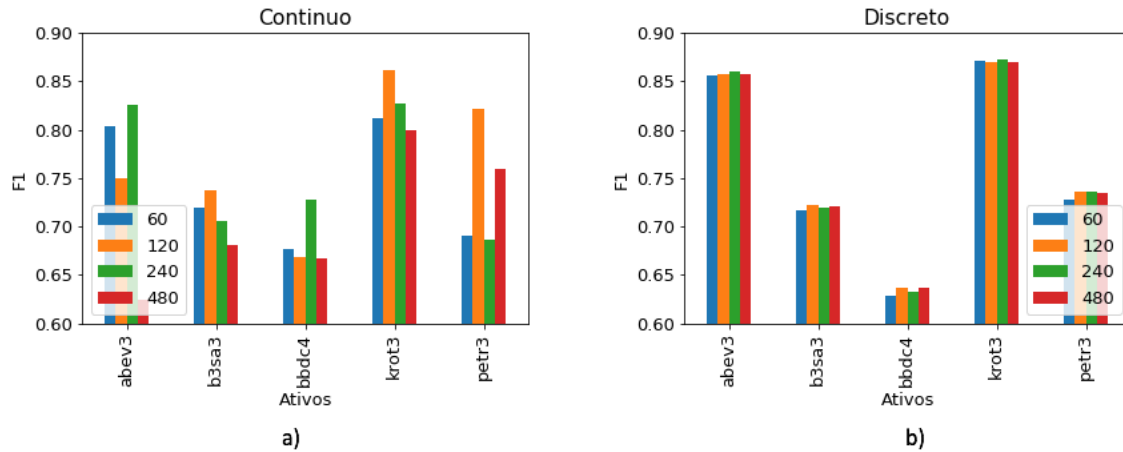
4.2.1.1 DETERMINAÇÃO DA JANELA LSTM

Para avaliar a hipótese H1, um experimento preliminar buscou compreender o impacto do tamanho das janelas nas redes LSTM. Conforme detalhado na Seção 2.6.5, redes LSTM buscam armazenar representações dos eventos de entrada em forma de ativações (HOCHREITER; SCHMIDHUBER, 1997). Sua configuração inicial depende de ajustes iniciais pois, caso isto não seja feito a rede pode não convergir para um aprendizado mínimo satisfatório.

Um experimento com quatro diferentes tamanhos de janela foi realizado e seu resultado é mostrado na Figura 19. A imagem a) na Figura 19 mostra o resultado das quatro janelas para o conjunto de dados discreto enquanto a imagem b) mostra o resultado das janelas referente ao conjunto contínuo.

Os valores da Figura 19 podem ser observados na Tabela 7 onde, os nomes das colunas terminados pela letra *C* referem-se ao conjunto de dados contínuo enquanto os nomes terminados com a letra *D* representam o conjunto discreto. Como alguns ativos tiveram um desempenho melhor com tamanhos de janelas diferentes, a Tabela 7 consolida as janelas definidas por cada ativo. O cálculo da diferença entre o menor e o maior resultado para cada

Figura 19 - Resultado da variação de diferentes janelas de dados na rede LSTM. A imagem a) mostra o resultado para o conjunto contínuo. A imagem b) apresenta os gráficos para o conjunto discreto.



Fonte - O autor.

um dos conjuntos de dados é apresentado na coluna com o nome *DIFF*. Os valores em negrito mostram o parâmetro escolhido para cada ativo.

Tabela 7 - Resultados obtidos na variação da janela LSTM para o conjunto contínuo e discreto com a métrica *F-measure*. As colunas em negrito foram escolhidas para os experimentos seguintes.

ATIVO	DIFF	60C	120C	240C	480C	DIFF	60D	120D	240D	480D
ABEV3	32%	0,804	0,750	0,826	0,624	0%	0,856	0,857	0,860	0,857
B3SA3	8%	0,720	0,737	0,705	0,681	1%	0,717	0,722	0,720	0,721
BBDC4	9%	0,676	0,669	0,728	0,667	1%	0,628	0,636	0,632	0,637
KROT3	8%	0,812	0,861	0,827	0,800	0%	0,871	0,870	0,872	0,870
PETR3	20%	0,691	0,822	0,686	0,759	1%	0,728	0,736	0,736	0,734

4.2.1.2 CLASSIFICAÇÃO

Os resultados consolidados da classificação do primeiro experimento são mostrados na Tabela 8 e 9. Em ambas as Tabelas, os resultados dos quatro classificadores são apresentados em colunas que referem-se ao conjunto de dados discreto e contínuo. Para os classificadores SVM, RF e MLP o resultado apresentado refere-se a melhor classificação do conjunto de validação. Para o classificador LSTM, o resultado apresentado refere-se o valor médio da classificação de doze rodadas no conjunto de validação. Testes estatísticos e os respectivos resultados foram incluídos no Apêndice A.

As colunas da Tabela 8 e 9 com os resultados do conjunto de dados contínuo, apresentam a letra *C* após o nome do respectivo classificador. Da mesma forma, a letra *D* representa o resultado para o conjunto de dados discreto.

Tabela 8 - Resultados obtidos do modelo contínuo e discreto com métrica ROC.

ATIVO	SVM_C	SVM_D	RF_C	RF_D	MLP_C	MLP_D	LSTM_C	LSTM_D
ABEV3	0,500	0,839	0,660	0,836	0,579	0,854	0,850	0,890
B3SA3	0,557	0,787	0,641	0,780	0,605	0,799	0,810	0,800
BBDC4	0,561	0,770	0,648	0,764	0,624	0,777	0,790	0,780
KROT3	0,621	0,848	0,682	0,795	0,639	0,845	0,880	0,900
PETR3	0,501	0,786	0,635	0,783	0,515	0,798	0,830	0,820

Tabela 9 - Resultados obtidos do modelo contínuo e discreto com métrica F-measure.

ATIVO	SVM_C	SVM_D	RF_C	RF_D	MLP_C	MLP_D	LSTM_C	LSTM_D
ABEV3	0,155	0,786	0,528	0,782	0,346	0,798	0,810	0,860
B3SA3	0,280	0,673	0,437	0,671	0,415	0,690	0,690	0,720
BBDC4	0,220	0,556	0,378	0,599	0,360	0,573	0,600	0,630
KROT3	0,409	0,784	0,583	0,726	0,411	0,775	0,830	0,870
PETR3	0,177	0,682	0,447	0,692	0,210	0,700	0,076	0,730

A Tabela 8 mostra os resultados com a métrica ROC, enquanto a Tabela 9 utiliza a métrica *F-measure*. Como ambas as métricas são utilizadas na maior parte da literatura, optou-se em manter seus resultados em tabelas diferentes. Esta separação buscou facilitar a comparação destes estudos com estudos futuros. Todavia, nos próximos experimentos os resultados serão apresentados apenas utilizando a métrica *F-measure*. Um gráfico consolidado contendo todos os classificadores e as duas métricas podem ser observados na Figura 20.

4.2.2 DISCUSSÃO

O experimento preliminar, referente a janela LSTM, demonstrou pouca influência dos resultados para o conjunto de dados discreto. Neste caso, a mesma janela foi adotada para todos os conjuntos. O contrário ocorreu com o conjunto de dados contínuo. Para estes, foram observados desempenhos diferentes para cada ativo e, desta forma, as janelas foram escolhidas em função do melhor desempenho.

Conforme reportado nos estudos de Patel et al. (2015), o método de discretizar os indicadores apresentou um desempenho superior quando comparado ao contínuo para os

classificadores SVM, RF e MLP. Todavia, redes LSTM apresentaram resultados superior para o ambos os métodos. Em relação ao conjunto de dados contínuo, o classificador LSTM apresentou um desempenho superior em todos os ativos quando comparado com os demais classificadores. Para comparar se houve diferença significativa no classificador LSTM entre os resultados dos conjuntos, contínuo e discreto, o teste estatístico de Wilcoxon¹ foi utilizado com um intervalo de confiança de 95%. Os testes foram aplicados nas amostras com a métrica *F1-measure*. Os testes falharam em rejeitar a hipótese H0 para três dos cinco ativos demonstrando diferenças entre os conjuntos apenas em dois casos. Informações adicionais sobre os testes podem ser analisadas no Apêndice A.

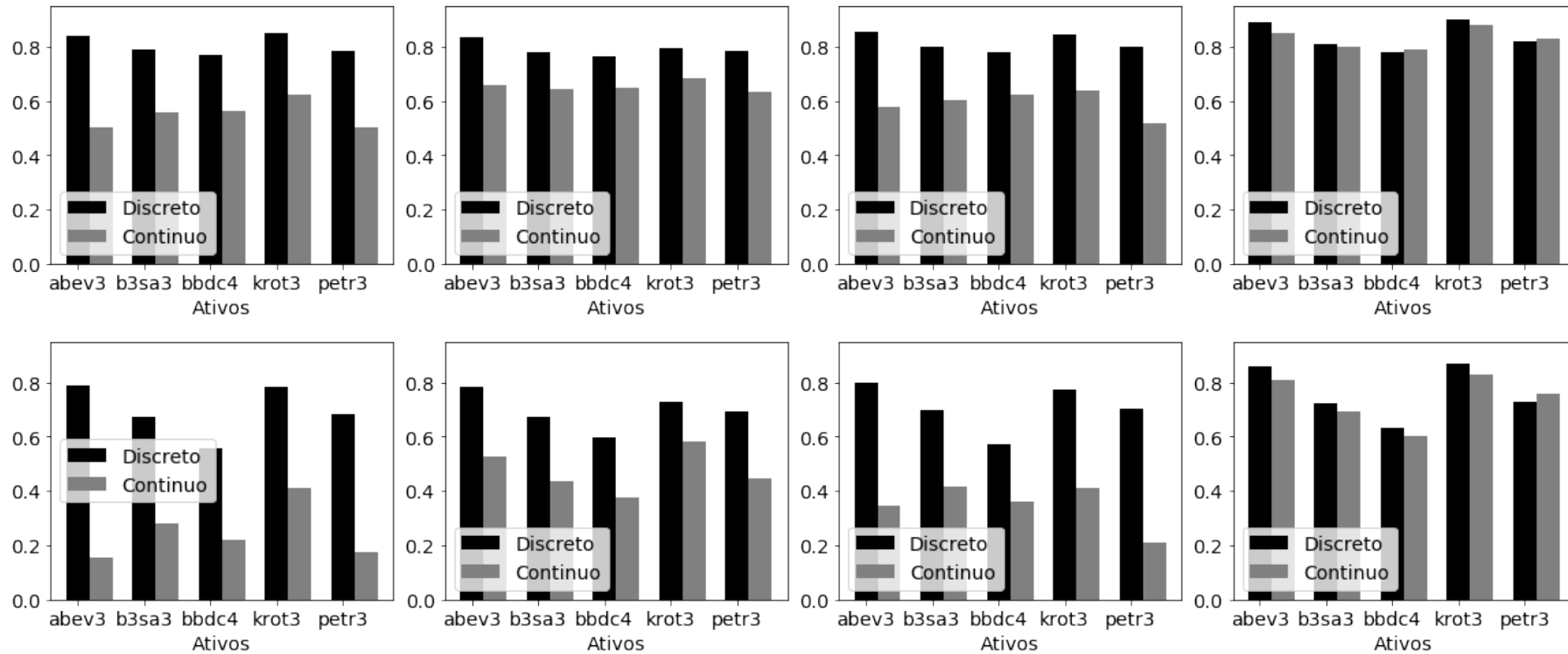
Durante seus experimentos Patel et al. (2015) comparou seu método com classificadores MLP, RF e NB² e nenhuma abordagem de AP utilizando LSTM foi utilizada em seu experimento. Embora observando a Tabela 9 é possível verificar que os conjuntos discretos se beneficiam dos resultados médios, isto é refutado nos testes estatísticos através da Tabela 14. Os resultados dos experimentos refutaram a Hipótese H1. O método proposto por Patel de discretizar indicadores não se apresentou melhor para todos os ativos quando foi utilizado redes LSTM. Isto pode estar relacionado ao fato da capacidade de redes LSTM capturarem padrões temporais de sinais. A Figura 19 ajuda a compreender isto uma vez que, alterando os valores das janelas do conjunto de dados discretos não se observa alterações. O contrário no entanto ocorre no conjunto contínuo se beneficiando dessa técnica.

Os resultados desta Seção serão utilizado como *baseline* para os experimentos da Seção seguinte. Desta forma, apenas o classificador LSTM será utilizado. Os dois conjuntos de dados também serão utilizados nos experimentos seguintes uma vez que o classificador LSTM apresentou um desempenho superior em ambos os conjuntos.

¹O teste de Wilcoxon é um teste de hipóteses não paramétrico utilizado quando se deseja comparar duas amostras relacionadas. O objetivo é avaliar se os postos médios populacionais diferem.

²Classificador *Naive Bayes*

Figura 20 - Resultado do experimento 01 com métricas ROC e F-measure para todos os classificadores no conjunto de dados discreto e contínuo. Para o experimento com a rede LSTM, a janela de 60 amostras foi utilizada no conjunto de dados discreto. Para o conjunto de dados contínuo as janelas foram escolhidas conforme a Tabela 7. Os resultados do classificador LSTM são a média de doze rodadas de experimentos (mais detalhes no Apêndice A).



Fonte - O Autor.

4.3 EXPERIMENTO 02: ATRIBUTO PREVISOR DE LYAPUNOV

Esta Seção apresenta os resultados do método do atributo previsor de Lyapunov detalhado na Seção 3.3. Os conjuntos de dados durante este experimento seguem o descrito na Seção 4.1 onde foram adicionados quatro atributos previsores propostos e descritos na Seção 3.3. O principal objetivo deste experimento foi avaliar o desempenho do classificador LSTM nos conjuntos de dados contínuos e discreto através do acréscimo dos quatro atributos previsores. Dentro deste contexto, este experimento levanta a hipótese H2, onde: o acréscimo de quatro atributos previsores, construídos usando o coeficiente de Lyapunov, melhoram o desempenho final dos cinco papéis de curto prazo em ambos os conjuntos de dados.

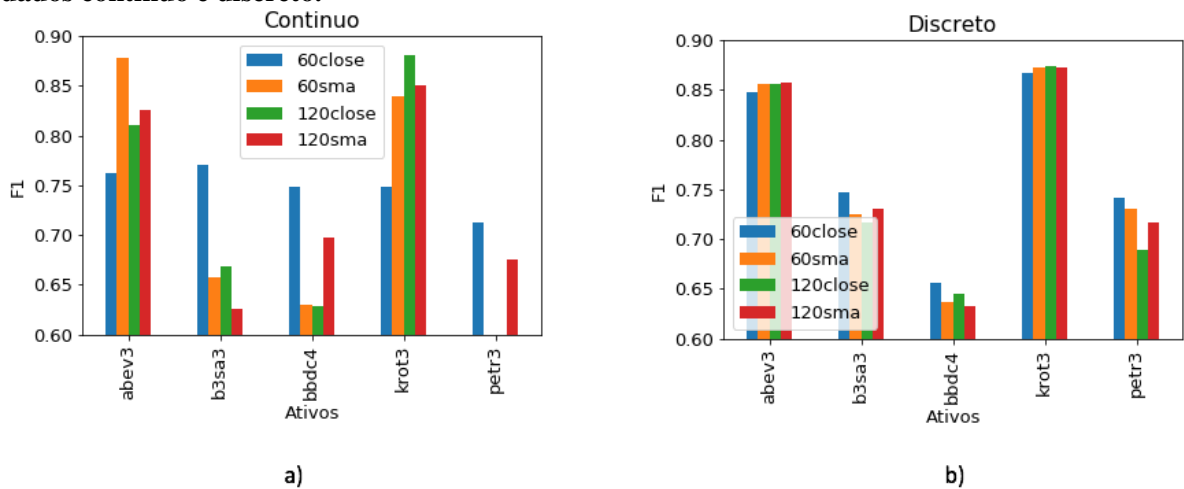
4.3.1 EXPERIMENTOS REALIZADOS

4.3.1.1 DETERMINAÇÃO DA JANELAS DE LYAPUNOV

Conforme detalhado na Seção 3.3, os atributos utilizados no método previsor de Lyapunov utilizam janelas deslizantes. Um experimento preliminar foi realizado para compreender o comportamento de diferentes janelas na classificação. Dois tamanhos diferente de janela, (60,120), com tipos de entradas (preço de fechamento, média móvel) foram avaliados. O valor de 60 refere-se ao período de 60 minutos enquanto 120, ao período de 120 minutos. Duas séries foram fornecidas para o cálculo dos atributos. Uma delas utilizou o preço de fechamento e outra a média móvel do preço de fechamento. Os resultado para ambos os conjuntos de dados podem ser observados na Figura 21. Enquanto a imagem a) mostra o resultado para o conjunto de dados contínuo a imagem b) apresenta o resultado para o conjunto de dados discreto.

Na legenda da Figura 21, a descrição *60close* refere-se à variável com o preço de fechamento de 60 minutos. A descrição *120close* à variável com o preço de fechamento de 120 minutos. A descrição seguinte, *SMA*, refere-se à média móvel para ambos os períodos. O resultado numérico cada diferente tipo de janela é mostrado na Tabela 10. Em cada coluna, os valor terminados com a letra *C* referem-se ao conjunto de dados contínuo enquanto *D* ao discreto. Da mesma forma, o acréscimo da letra *S* refere-se ao atributo *SMA* da média móvel. O cálculo da diferença entre o menor e o maior resultado para cada janela em cada conjunto de dados e apresentado na coluna com o nome *DIFF*. Os valores em negrito mostram o resultado de maior valor. Estes resultados em negrito foram os escolhidos para cada ativo.

Figura 21 - Ativos e diferentes janelas de Lyapunov utilizadas no experimento com o conjunto de dados contínuo e discreto.



Fonte - O autor.

Tabela 10 - Resultados da variação da janela deslizante de Lyapunov para o conjunto contínuo e discreto com a métrica F-measure.

ATIVO	DIFFC	60C	60SC	120C	120SC	DIFFD	60D	120D	120SD	120SD
ABEV3	15%	0,762	0,878	0,811	0,825	1%	0,848	0,856	0,856	0,858
B3SA3	23%	0,770	0,657	0,668	0,625	4%	0,747	0,725	0,717	0,730
BBDC4	18%	0,749	0,630	0,628	0,697	4%	0,656	0,636	0,645	0,632
KROT3	19%	0,749	0,839	0,881	0,851	1%	0,867	0,872	0,874	0,872
PETR3	30%	0,712	0,548	0,597	0,675	8%	0,741	0,730	0,689	0,717

4.3.1.2 CLASSIFICAÇÃO

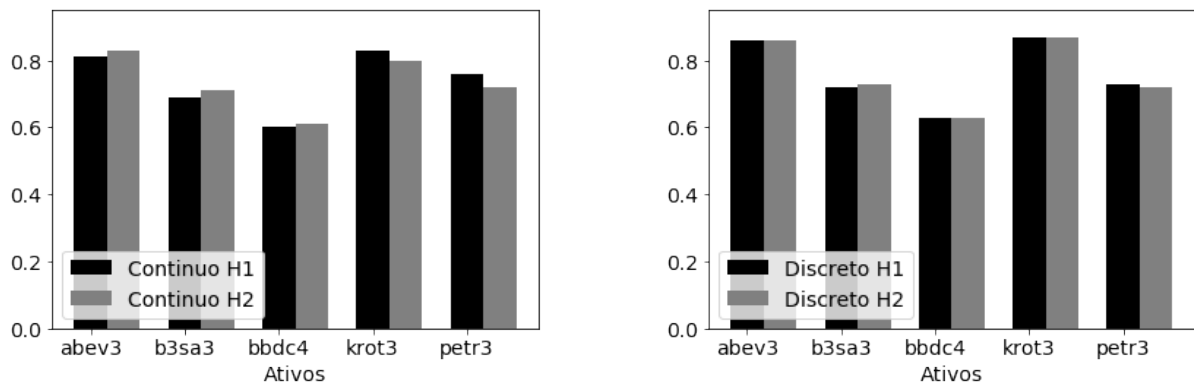
A Tabela 11 apresenta os resultados obtidos da classificação utilizando o modelo do atributo predictor de Lyapunov. As colunas com a descrição H2 já consideram a escolha da melhor janela estimada na Seção 4.3.1.1. Em cada coluna da Tabela 11, os valores terminados com a letra *C* referem-se ao conjunto de dados contínuo enquanto *D* ao discreto. As tabelas comparam a hipótese H1 proposta no experimento 01 com os resultados da hipótese H2 proposta neste experimento.

De forma gráfica, a Figura 22 mostra os valores consolidado na Tabela 11. A imagem a) mostra a comparação entre os resultados de H1 e H2 para o conjunto de dados contínuo. Na imagem b) são mostrados os resultados para o conjunto de dados discreto.

Tabela 11 - Resultados obtidos com modelo atributo predictor de Lyapunov na métrica *F-measure*.

ATIVO	LSTM H1C	LSTM H2C	LSTM H1D	LSTM H2D
ABEV3	0,81	0,83	0,86	0,86
B3SA3	0,69	0,71	0,72	0,73
BBDC4	0,60	0,61	0,63	0,63
KROT3	0,83	0,80	0,87	0,87
PETR3	0,76	0,72	0,73	0,72

Figura 22 - Resultado do experimento 02 com classificador LSTM e métrica *F-measure*. A imagem a) apresenta os resultados com conjunto de dados contínuos enquanto a imagem b) mostra o resultado com o conjunto de dados discretos. Os resultados para ambos os conjuntos são a média de doze rodadas de experimentos (mais detalhes no Apêndice A).



Fonte - O autor.

4.3.2 DISCUSSÃO

Inicialmente foi proposto um experimento para avaliar o impacto de diferentes tamanhos das janelas nos conjuntos de dados discreto e contínuo. A dispersão das porcentagens da coluna *DIFFD*, apresentadas na Tabela 10, mostra que o conjunto discreto apresentou uma estabilidade maior quando comparado ao contínuo. Todavia, em um contexto geral, o conjunto contínuo teve um impacto maior em seus resultados. Os valores finais são reforçados visualmente pela imagem da Figura 21. Assim, para o conjunto de dados contínuo o melhor resultado foi utilizado. Como o conjunto de dados discreto não apresentou ganho significativo a janela com o desempenho melhor para a maioria dos ativos foi escolhida.

Conforme H2, a expectativa do experimento era que o acréscimo dos atributos pudesse contribuir para a melhora de todos os ativos nos conjuntos discretos e contínuos. Observando os valores apresentados na Tabela 11 verifica-se uma discreta melhora para alguns papéis

no conjunto contínuo e uma estabilidade no conjunto discreto. De forma semelhante ao experimento da Seção 4.2 o teste de Wilcoxon foi utilizado para comparar se houve diferença significativa entre a hipótese H1 e H2. Observando os resultados da Tabela 15 e 16 no Apêndice A, os testes falharam em rejeitar a hipótese H0 para quatro dos cinco ativos demonstrando diferenças em apenas um dos casos para o conjunto contínuo e em dois casos para o discreto.

Em linhas gerais o desempenho do método proposto não apresentou ganhos significativos. Dessa forma, a hipótese H2 foi refutada para o experimento do atributo previsor de Lyapunov. Observou-se que enquanto alguns papéis apresentaram uma discreta melhora com o método, em ambos os conjuntos de dados, outros não obtiveram ganho de desempenho. Um comportamento observado durante os experimentos foi que para alguns papéis, por exemplo ABEV3, o método apresentou ganhos no conjunto contínuo, mas apresentou estabilidade no conjunto discreto.

4.4 EXPERIMENTO 03: JANELA DE LYAPUNOV

Esta Seção apresenta os resultados do método da Janela de Lyapunov detalhado na Seção 3.4. Os conjuntos de dados utilizados durante estes experimentos seguem o descrito na Seção 4.1. O principal objetivo deste experimento foi avaliar o desempenho dos papéis de curto prazo retirando do conjunto de dados valores onde o coeficiente de Lyapunov é superior a zero. Dentro deste contexto, este experimento levanta a hipótese H3, onde: remover trechos das séries onde os valores do expoente de Lyapunov sejam superiores a zero beneficiam o desempenho para os cinco papéis de curto prazo.

4.4.1 EXPERIMENTOS REALIZADOS

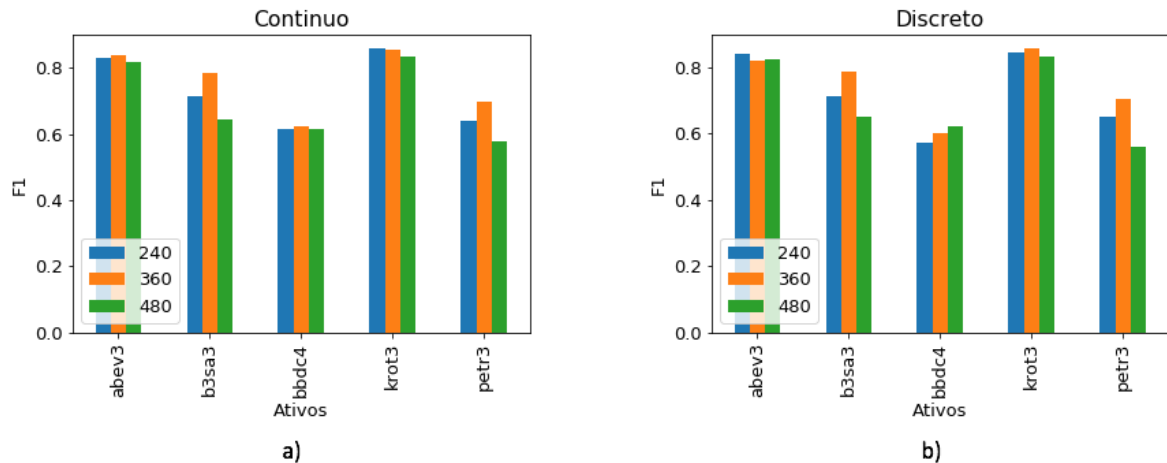
Para avaliar a hipótese H3, foram utilizados os resultados da Seção 4.3. Desta forma, os atributos previsores de Lyapunov que tiveram melhor desempenho foram mantidos para cada um dos ativos (vide resultados em negrito na Tabela 11). Ambos os conjuntos de dados também foram utilizados durante os experimentos.

4.4.1.1 DETERMINAÇÃO DE DIFERENTES JANELAS DE LYAPUNOV

Um experimento preliminar avaliou o impacto de diferentes tamanhos de janelas na classificação. Três diferentes valores foram escolhidos para avaliar o impacto. Os valores foram definidos baseado em múltiplos de 60 minutos. Desta forma, foi considerado janelas de 240, 360 e 480 minutos. O resultado para ambos os conjuntos de dados são mostrados na Figura

23. Enquanto a imagem a) mostra os resultados para o conjunto de dados contínuo a imagem b) apresenta os resultados para o conjunto de dados discreto. A Tabela 12 consolida as janelas escolhidas para cada um dos conjunto de dados em cada ativo.

Figura 23 - Resultado da variação de diferentes janelas de dados no experimento. A imagem a) mostra os resultados para o conjunto discreto. A imagem b) apresenta os resultados para o conjunto contínuo.



Fonte - O autor.

Tabela 12 - Ativo e Janela de Lyapunov utilizada no conjunto de dados contínuos e discreto.

ATIVO	CONJUNTO CONTÍNUO	CONJUNTO DISCRETO
abev3	360 amostras	240 amostras
b3sa3	360 amostras	360 amostras
bbdc4	360 amostras	480 amostras
krot3	360 amostras	360 amostras
petr3	360 amostras	360 amostras

4.4.1.2 CLASSIFICAÇÃO

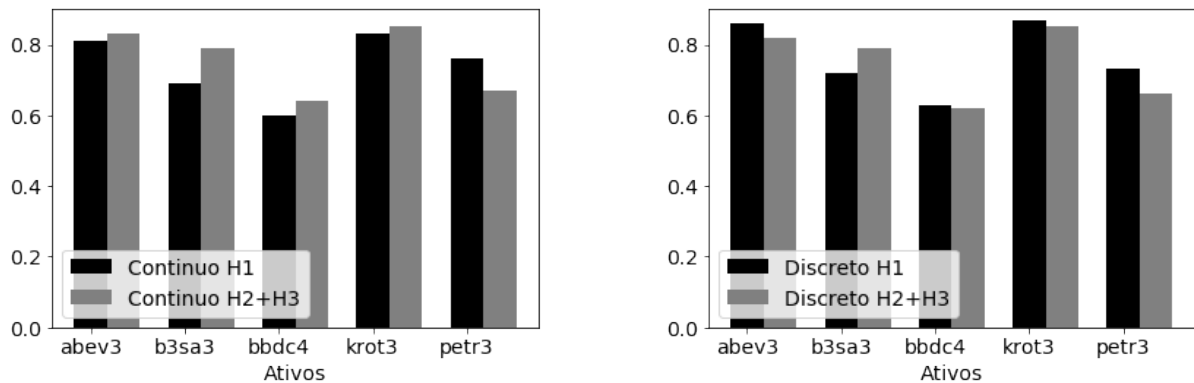
A Tabela 13 apresenta os resultados da classificação obtidos utilizando o modelo proposto da janela de Lyapunov. As colunas da Tabela terminadas com a letra *C* referem-se ao conjunto de dados contínuo enquanto aquelas terminadas com a letra *D* referem-se ao discreto. Da mesma forma, o acréscimo da descrição *H2+H3* refere-se ao fato de que os experimentos da hipótese *H3* mantêm os resultados do experimento anterior *H2*. Desta forma, a Tabela 13 compara os resultados de ambos os métodos contra a hipótese *H1*.

Tabela 13 - Resultados obtidos com modelo da janela de Lyapunov com métrica *F-measure*.

ATIVO	LSTM H1C	LSTM H2H3C	LSTM H1D	LSTM H2H3D
ABEV3	0,81	0,83	0,86	0,82
B3SA3	0,69	0,79	0,72	0,79
BBDC4	0,60	0,64	0,63	0,62
KROT3	0,83	0,85	0,87	0,85
PETR3	0,76	0,67	0,73	0,66

A Figura 24 mostra os valores da Tabela 13 consolidados em duas imagens. A imagem a) mostra uma comparação entre H1 e H2+3 para o conjunto de dados contínuo e a imagem b) mostra o resultado para o conjunto discreto.

Figura 24 - Resultado do experimento 3 com classificador LSTM e métrica *F-measure*. A imagem a) apresenta os resultados com conjunto de dados contínuos enquanto a imagem b) mostra o resultado com o conjunto de dados discretos.



Fonte - O autor.

4.4.2 DISCUSSÃO

Inicialmente um experimento preliminar foi proposto. O experimento buscou avaliar o impacto de diferentes tamanhos de janelas lyapunov na classificação. Observando o resultado mostrado na Figura 23 não foi possível observar grande dispersão nos valores como anteriormente observado no experimento 02 (vide Figura 21). Com exceção do ativo B3SA3, que apresentou ganhos em janelas distintas, os demais tiveram resultado praticamente estável. Desta forma, para o conjunto discreto três diferentes valores de janela foram escolhidos. Para o conjunto contínuo apenas a janela com o valor de 360 amostras foi utilizada. Os resultados da janela de Lyapunov por ativo podem ser observados na Tabela 12.

Conforme H3, a expectativa do experimento era que o método das janelas de Lyapunov pudessem contribuir para a melhora de todos os ativos nos conjuntos discretos e contínuos. Observando os valores apresentados na Tabela 13 é possível identificar que o conjunto contínuo foi o que se beneficiou mais do modelo proposto. A maior parte dos ativos apresentaram ganho de desempenho.

Novamente o teste de Wilcoxon foi utilizado para comparar se houve diferença significativa entre a hipótese H1 e H2+3. Observando os resultados da Tabela 17 e 18 no Apêndice A, os testes rejeitaram a hipótese H0 para quatro dos cinco ativos demonstrando diferenças em quatro séries. Das quatro séries, três delas se beneficiaram do método. Todavia, para o conjunto de dados discreto apenas um papel se beneficiou do método.

Dessa forma, a H3 foi refutada para o experimento janelas de Lyapunov. Observou-se que enquanto alguns papéis apresentaram melhora com o método no conjunto contínuo outros apresentaram desempenho negativo no conjunto discreto. Embora o método apresentou ganhos para alguns ativos o mesmo apresentou perdas em alguns papéis. Um comportamento semelhante ao observado no experimento anterior novamente apareceu neste experimento. Para alguns ativos, por exemplo BBDC4, enquanto observou-se ganhos no conjunto contínuo o mesmo apresentou perdas com o conjunto discreto.

5 CONCLUSÃO E TRABALHOS FUTUROS

Mercados financeiros cumprem um papel fundamental no desenvolvimento da sociedade moderna. Os movimentos dos ativos nas Bolsas de Valores tornam o estudo da previsibilidade de preços um dos mais interessantes neste segmento. Nos últimos anos o avanço das ferramentas de inteligência computacional tem contribuído para o problema da previsibilidade. Em geral, este assunto abrange não só o interesse de especuladores mas também de grandes gestores de recursos que utilizam sistemas de apoio à tomada de decisão em tarefas de investimento. Portanto, prever o próximo movimento do mercado é uma tarefa de grande relevância.

Neste trabalho, o problema da previsão de direção foi tratado como uma tarefa de classificação. Series históricas com dados de curto prazo foram utilizadas e quatro diferentes classificadores foram avaliados. Três experimentos foram conduzidos utilizando dados de cinco ativos de curto prazo da BM&FBovespa.

No primeiro experimento, métodos já conhecidos na literatura foram aplicados nas séries com o objetivo de criar um *baseline* para os experimentos seguintes. A partir da série dos cinco ativos, dois conjuntos de dados foram construídos utilizando indicadores técnicos contínuos e discretos. Dos resultados foi possível concluir que: (i) o conjunto de dados discreto teve um desempenho melhor apenas para os classificadores SVM, RF e MLP; (ii) em comparação com os outros classificadores, LSTM apresentou um desempenho superior em ambos os conjuntos de dados. Quando comparado entre os conjuntos discretos e contínuo, a LSTM apresentou um desempenho superior em dois ativos do conjunto discreto quando comparado ao contínuo.

No segundo experimento foi calculado o expoente de Lyapunov do valor de fechamento das cinco séries de curto prazo. O resultado foi utilizado para construir quatro atributos. Os atributos foram adicionados aos conjuntos construídos no experimento anterior. Apenas o desempenho do classificador LSTM foi avaliado uma vez que apresentou um desempenho superior nos resultados do *baseline*. A partir dos resultados foi possível concluir

que o método contribuiu com o desempenho de um papel em cada conjunto de dados.

No terceiro experimento o máximo expoente de Lyapunov foi utilizado como um filtro inicial nas séries de dados. O valor do maior EL foi utilizado para identificar regiões com perturbações caóticas nas séries e, uma vez encontradas, removê-las do conjunto. O resultado foi comparado ao *baseline* proposto e o conjunto contínuo se beneficiou com a melhora de desempenho em quatro dos cinco papéis analisado. No conjunto discreto, o método contribuiu com o desempenho de apenas um papel. Os demais apresentaram perda de desempenho.

Através dos experimentos, é possível afirmar que os classificadores LSTM apresentaram um desempenho superior quando comparado aos outros classificadores. O método proposto por Patel et al. (2015) de discretização de indicadores não apresentou melhora em todos os casos. Todavia, o método demonstrou maior robustez e menor perda de desempenho durante experimentos preliminares em tamanhos diferentes de janelas Lyapunov. Isto sugere que a estratégia proposta por Patel et al. (2015) atue também como um filtro eliminando alguns ruídos em dados.

O acréscimo dos atributos de Lyapunov se mostrou promissor e, embora tenha contribuído para o aumento de desempenho da classificação em poucas séries, seu uso em conjunto com a janela de Lyapunov trouxe ganhos para algumas séries no conjunto contínuo. Todavia, ao remover trechos das séries, através do método da janela de Lyapunov, o resultado não demonstrou benefício para a classificação no conjunto discreto e em uma série do conjunto contínuo.

A partir dos ativos e prazos abordados neste trabalho, conclui-se que, não há um método que seja vencedor em todos os casos. Cada papel conta uma história diferente através de suas características. Em cada janela de estudo, as características podem variar e comportamentos distintos podem surgir. Esta conclusão vai de encontro com a prática de alguns operadores do mercado financeiro, que afirmam que as características do mercado mudam frequentemente e se especializar em determinados mercados e ativos pode ser a melhor escolha.

5.1 TRABALHOS FUTUROS

Neste trabalho, o coeficiente de Lyapunov foi utilizado com o objetivo de auxiliar classificadores na forma de um atributo ou identificando regiões onde possíveis comportamentos caóticos estivessem acontecendo. Neste sentido, outras ferramentas de estimativa de caos poderiam ser inseridas em conjunto com o método do atributo predictor de Lyapunov proposto. Diagrama de fase, correlação de dimensão e *sample entropy* são algumas das sugestões. De

forma geral, tais ferramentas buscam identificar o grau de previsibilidade em séries temporais auxiliando à identificar características caóticas e randômicas.

Também foi observado que algumas séries apresentam desempenho melhor em alguns métodos. Neste sentido, um estudo de clusterização poderia tentar identificar se alguns métodos apresentam comportamento semelhante para grupos de ativos com características semelhantes. Um ponto de partida seria iniciar abordando alguns setores específicos extraindo características destas séries e buscando observar se as mesmas apresentam alguma semelhança. Neste sentido, uma classificação inicial das séries pode contribuir para determinar quais métodos podem ser melhor aplicados em cada estudo.

Embora obter conjuntos de dados de curto prazo seja um desafio; aplicar os métodos propostos neste trabalho em uma quantidade maior de ativos podem trazer contribuições relevantes para os estudos de previsibilidade. Em uma abordagem semelhante ao proposto no paragrafo anterior, um estudo de clusterização de mercados globais poderia buscar auxiliar a identificar se existe influência na característica do mercado em função do método escolhido.

REFERÊNCIAS

- ABU-MOSTAFA, Y. S.; ATIYA, A. F. Introduction to financial forecasting. **Applied Intelligence**, v. 6, n. 3, p. 205–213, 1996.
- BALLINGS, M. et al. Evaluating multiple classifiers for stock price direction prediction. **Expert Systems with Applications**, v. 42, n. 20, p. 7046–7056, 2015.
- BLANK, S. C. Chaos in futures markets: a nonlinear dynamical analysis. **The Journal of Futures Markets**, v. 11, n. 6, p. 711, 1991.
- BLOOMBERG. **Tecnologia Decodificada**. 2019. <https://www.bloomberg.com/>. (Acesso em 03/10/2019).
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: ACM. **Proceedings of the fifth annual workshop on Computational learning theory**. [S.l.], 1992. p. 144–152.
- BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.
- CAVALCANTE, R. C. et al. Computational intelligence and financial markets: A survey and future directions. **Expert Systems with Applications**, v. 55, p. 194–211, 2016.
- CHEN, A.-S.; LEUNG, M. T.; DAOUK, H. Application of neural networks to an emerging financial market: forecasting and trading the taiwan stock index. **Computers & Operations Research**, v. 30, n. 6, p. 901–923, 2003.
- CHONG, E.; HAN, C.; PARK, F. C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. **Expert Systems with Applications**, v. 83, p. 187–205, 2017.
- COLAH. **Colah Understanding LSTMs**. 2019. <https://colah.github.io/posts/2015-08-Understanding-LSTMs>. (Acesso em 01/06/2019).
- DIXON, M. Sequence classification of the limit order book using recurrent neural networks. **Journal of Computational Science**, v. 24, p. 277–286, 2018.
- DIXON, M.; KLABJAN, D.; BANG, J. H. Classification-based financial markets prediction using deep neural networks. **Algorithmic Finance**, IOS Press, v. 6, n. 3, p. 1–11, 2017.
- ECKMANN, J.-P. et al. Liapunov exponents from time series. **Physical Review A, APS**, v. 34, n. 6, p. 4971, 1986.
- ELTON, E. J. et al. **Modern Portfolio Theory and Investment Analysis**. New Jersey, US: John Wiley & Sons, 2009.
- FISCHER, T.; KRAUSS, C. Deep learning with long short-term memory networks for financial market predictions. **European Journal of Operational Research**, v. 270, n. 2, p. 654–669, 2018.

- GOLDSCHMIDT RONALDO, P. E. B. E. **Data mining: conceitos, técnicas, algoritmos, orientações e aplicações**. RJ: Editora Elsevier, 2015.
- GOMES, E. F. G.; JUNIOR, K. M. Mercados financeiros: Sistema caótico ou aleatório? **Revista de Economia Mackenzie**, v. 8, n. 1, p. 103–123, 2010.
- GROSSMAN, S. J.; STIGLITZ, J. E. On the impossibility of informationally efficient markets. **The American Economic Review**, v. 70, n. 3, p. 393–408, 1980.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining: Concepts and Techniques**. MA,US: Morgan Kaufmann Publishers, 2012.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. NJ, US: Prentice Hall PTR, 1994.
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997.
- HUANG, W.; NAKAMORI, Y.; WANG, S.-Y. Forecasting stock market movement direction with support vector machine. **Computers & Operations Research**, v. 32, n. 10, p. 2513–2522, 2005.
- KARA, Y.; BOYACIOGLU, M. A.; BAYKAN, Ö. K. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. **Expert Systems with Applications**, v. 38, n. 5, p. 5311–5319, 2011.
- KEARNS, M.; NEVMYVAKA, Y. Machine learning for market microstructure and high frequency trading. **High Frequency Trading: New Realities for Traders, Markets, and Regulators**, 2013.
- KIM, H. Y.; WON, C. H. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. **Expert Systems with Applications**, v. 103, p. 25–37, 2018.
- LAMBERT, D. R. Commodity channel index: Tool for trading cyclic trends. **Technical Analysis of Stocks & Commodities**, v. 1, p. 47, 1983.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436, 2015.
- LEUNG, M. T.; DAOUK, H.; CHEN, A.-S. Forecasting stock indices: a comparison of classification and level estimation models. **International Journal of Forecasting**, v. 16, n. 2, p. 173–190, 2000.
- LI, W.; LIAO, J. A comparative study on trend forecasting approach for stock price time series. In: **Proceedings of the 11th IEEE International Conference on Anti-Counterfeiting, Security, and Identification (ASID)**. Xiamen,CN: IEEE Press, 2017. p. 74–78.
- LO, A. W.; MACKINLAY, A. C. **A Non-Random Walk Down Wall Street**. New Jersey, US: Princeton University Press, 2002.
- LORENZ, E. N. Deterministic nonperiodic flow. **Journal of the atmospheric sciences**, v. 20, n. 2, p. 130–141, 1963.

- MALKIEL, B. G. **A Random Walk Down Wall Street**. New York, US: W. W. Norton Company, Inc, 1973.
- MALKIEL, B. G.; FAMA, E. F. Efficient capital markets: A review of theory and empirical work. **The Journal of Finance**, v. 25, n. 2, p. 383–417, 1970.
- METAQUOTES. **Metatrader5**. 2019. <https://www.metatrader5.com/en>. (Acesso em 01/06/2019).
- MINH, D. L. et al. Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network. **IEEE Access**, v. 6, p. 55392–55404, 2018.
- MORETTIN, P. A. **Econometria Financeira. Um Curso em Séries Temporais Financeiras**. São Paulo, SP: Editora Blucher, 2008.
- MURPHY, J. J. **Technical Analysis of the Financial Markets. A Comprehensive Guide to Trading Methods and Applications**. Londres, UK: Penguin, 1999.
- NASCIMENTO, R. C. **Avaliação de Oportunidades de Investimento no Mercado Futuro Brasileiro na Escala de Dezenas de Segundos**. Dissertação (Mestrado) — Universidade Federal do Espírito Santo, 2018.
- NOLDS. **Nolds 0.5.2 documentation**. 2019. <https://nolds.readthedocs.io/en/latest/>. (Acesso em 08/06/2019).
- PATEL, J. et al. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. **Expert Systems with Applications**, v. 42, n. 1, p. 259–268, 2015.
- PETERS, E. E. A chaotic attractor for the s&p 500. **Financial Analysts Journal**, v. 47, n. 2, p. 55–62, 1991.
- PRING, M. J. **Technical Analysis Explained: The Successful Investor's Guide to Spotting Investment Trends and Turning Points**. New York, US: McGraw-Hill Professional, 2002.
- RAFIUZAMAN, M. Forecasting chaotic stock market data using time series data mining. **International Journal of Computer Applications**, v. 101, n. 10, p. 27–34, 2014.
- SAAD, E. W.; PROKHOROV, D. V.; WUNSCH, D. C. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. **IEEE Transactions on Neural Networks**, v. 9, n. 6, p. 1456–1470, 1998.
- SALES, L. O. A. **Testando a Hipótese de Passeio Aleatório no Mercado de Ações Brasileiro**. Dissertação (Mestrado) — FGV Escola de Economia de São Paulo, 2017.
- SAVI, M. A. **Dinâmica Não Linear e Caos**. [S.l.]: Editora E-papers, 2006.
- SAYAMA, H. **Introduction to the Modeling and Analysis of Complex Systems**. New York, US: Open SUNY Textbooks, 2015.
- SILVER, N. **The Signal and the Noise. The Art and Science of Prediction**. Londres, UK: Penguin, 2012.

- SITTE, R.; SITTE, J. Neural networks approach to the random walk dilemma of financial time series. **Applied Intelligence**, v. 16, n. 3, p. 163–171, 2002.
- TAN, P.-N.; STEINBACH, M.; KUMAR, V. **Introduction to Data Mining**. MA, US: Addison-Wesley, 2005.
- TAYLOR, M. P.; ALLEN, H. The use of technical analysis in the foreign exchange market. **Journal of International Money and Finance**, v. 11, n. 3, p. 304–314, 1992.
- TSAIH, R.; HSU, Y.; LAI, C. C. Forecasting s&p 500 stock index futures with a hybrid ai system. **Decision Support Systems**, v. 23, n. 2, p. 161–174, 1998.
- TSIONAS, M. G.; MICHAELIDES, P. G. Neglected chaos in international stock markets: Bayesian analysis of the joint return–volatility dynamical system. **Physica A: Statistical Mechanics and its Applications**, v. 482, p. 95–107, 2017.
- ÜBEYLI, E. D. Recurrent neural networks employing lyapunov exponents for analysis of ecg signals. **Expert systems with applications**, v. 37, n. 2, p. 1192–1199, 2010.
- VANSTONE, B.; FINNIE, G. An empirical methodology for developing stockmarket trading systems using artificial neural networks. **Expert Systems with Applications**, v. 36, n. 3, p. 6668–6680, 2009.
- VICENTE, J. V. M. et al. **Avaliando a Volatilidade Diária dos Ativos: a hora da negociação importa?** [S.l.], 2012.
- WANG, B.; HUANG, H.; WANG, X. A novel text mining approach to financial time series forecasting. **Neurocomputing**, v. 83, p. 136–145, 2012.
- WILDER, J. W. **New Concepts in Technical Trading Systems**. Califórnia, US: [s.n.], 1978.
- WILLEY, T. Testing for nonlinear dependence in daily stock indices. **Journal of Economics and Business**, v. 44, n. 1, p. 63–76, 1992.
- WITTEN, I. H. et al. **Data Mining. Practical Machine Learning Tools and Techniques**. MA,US: Morgan Kaufmann, 2016.
- WOLF, A. et al. Determining lyapunov exponents from a time series. **Physica D: Nonlinear Phenomena**, v. 16, n. 3, p. 285–317, 1985.
- WOLWACZ, A. **Táticas Operacionais de Posição em Ações**. São Paulo, SP: Nelogica, 2008.
- YU, H.; KIM, S. Svm tutorial - classification, regression and ranking. In: _____. **Handbook of Natural Computing**. Baden,DE: Springer-Verlag Berlin Heidelberg, 2012. v. 1.

APÊNDICE A - TESTES ESTATÍSTICOS

Tabela 14 - Experimento 01: Resultado dos experimentos e dos testes estatísticos de Wilcoxon com LSTM.

Experimento	ABEV3		B3SA3		BBDC4		KROT3		PETR3	
	DC	DD	DC	DD	DC	DD	DC	DD	DC	DD
1	0,71	0,86	0,73	0,71	0,60	0,63	0,81	0,87	0,77	0,73
2	0,81	0,86	0,63	0,72	0,56	0,63	0,80	0,87	0,68	0,74
3	0,82	0,86	0,69	0,72	0,41	0,63	0,82	0,87	0,78	0,73
4	0,84	0,86	0,71	0,72	0,63	0,63	0,84	0,87	0,70	0,73
5	0,82	0,86	0,75	0,72	0,59	0,64	0,83	0,87	0,80	0,73
6	0,81	0,86	0,68	0,72	0,58	0,63	0,81	0,87	0,84	0,74
7	0,80	0,86	0,69	0,72	0,67	0,63	0,81	0,87	0,79	0,73
8	0,80	0,86	0,62	0,72	0,62	0,63	0,87	0,87	0,80	0,73
9	0,85	0,86	0,74	0,72	0,68	0,64	0,83	0,87	0,79	0,74
10	0,80	0,86	0,72	0,73	0,72	0,63	0,87	0,87	0,75	0,74
11	0,82	0,86	0,59	0,72	0,60	0,63	0,86	0,87	0,77	0,73
12	0,80	0,86	0,70	0,72	0,49	0,63	0,84	0,87	0,68	0,74
Média	0,81	0,86	0,69	0,72	0,60	0,63	0,83	0,87	0,76	0,73
Wilcoxon	p=0.002	diferente	p=0.058	mesma	p=0.1816	mesma	p=0.004	diferente	p=0.083	mesma

*DC=Conjunto de dados contínuo. *DD=Conjunto de dados discreto.

Tabela 15 - Experimento 02: Resultado dos experimentos e dos testes estatísticos de Wilcoxon com LSTM para os ativos ABEV3 e B3SA3.

Experimento	ABEV3				B3SA3			
	DC	DCL	DD	DDL	DC	DCL	DD	DDL
1	0,71	0,72	0,86	0,86	0,73	0,71	0,71	0,73
2	0,81	0,83	0,86	0,85	0,63	0,73	0,72	0,73
3	0,82	0,86	0,86	0,86	0,69	0,62	0,72	0,73
4	0,84	0,87	0,86	0,86	0,71	0,72	0,72	0,73
5	0,82	0,85	0,86	0,85	0,75	0,72	0,72	0,74
6	0,81	0,87	0,86	0,86	0,68	0,74	0,72	0,74
7	0,80	0,83	0,86	0,86	0,69	0,68	0,72	0,74
8	0,80	0,83	0,86	0,86	0,62	0,72	0,72	0,74
9	0,85	0,81	0,86	0,86	0,74	0,72	0,72	0,74
10	0,80	0,80	0,86	0,86	0,72	0,66	0,73	0,73
11	0,82	0,86	0,86	0,85	0,59	0,74	0,72	0,73
12	0,80	0,81	0,86	0,85	0,70	0,70	0,72	0,73
Média	0,81	0,83	0,86	0,86	0,69	0,71	0,72	0,73
Wilcoxon	p=0.025	diferente	p=1	mesma	p=0.624	mesma	p=0.002	diferente

*DCL=Conjunto de dados contínuo com atributo previsor de Lyapunov. *DDL=Conjunto de dados discretos com atributo previsor de Lyapunov.

Tabela 16 - Experimento 02: Resultado dos experimentos e dos testes estatísticos de Wilcoxon com LSTM para os ativos BBDC4, KROT3 e PETR3.

Experimento	BBDC4				KROT3				PETR3			
	DC	DCL	DD	DDL	DC	DCL	DD	DDL	DC	DCL	DD	DDL
1	0,60	0,60	0,63	0,63	0,81	0,77	0,87	0,87	0,77	0,76	0,73	0,72
2	0,56	0,68	0,63	0,64	0,80	0,86	0,87	0,87	0,68	0,53	0,74	0,73
3	0,41	0,59	0,63	0,64	0,82	0,68	0,87	0,87	0,78	0,79	0,73	0,72
4	0,63	0,48	0,63	0,63	0,84	0,84	0,87	0,87	0,70	0,79	0,73	0,72
5	0,59	0,65	0,64	0,63	0,83	0,88	0,87	0,87	0,80	0,76	0,73	0,72
6	0,58	0,62	0,63	0,63	0,81	0,83	0,87	0,87	0,84	0,82	0,74	0,72
7	0,67	0,60	0,63	0,63	0,81	0,81	0,87	0,87	0,79	0,65	0,73	0,73
8	0,62	0,59	0,63	0,63	0,87	0,82	0,87	0,87	0,80	0,66	0,73	0,72
9	0,68	0,58	0,64	0,63	0,83	0,87	0,87	0,87	0,79	0,76	0,74	0,72
10	0,72	0,64	0,63	0,63	0,87	0,74	0,87	0,87	0,75	0,66	0,74	0,72
11	0,60	0,63	0,63	0,63	0,86	0,80	0,87	0,87	0,77	0,82	0,73	0,72
12	0,49	0,62	0,63	0,63	0,84	0,73	0,87	0,87	0,68	0,68	0,74	0,72
Média	0,60	0,61	0,63	0,63	0,83	0,80	0,87	0,87	0,76	0,72	0,73	0,72
Wilcoxon	p=0.75	mesma	p=1	mesma	p=0.184	mesma	p=1	mesma	p=0.119	mesma	p=0.002	diferente

*DCL=Conjunto de dados contínuo com atributo previsor de Lyapunov.*DDL=Conjunto de dados discretos com atributo previsor de Lyapunov.

Tabela 17 - Experimento 03: Resultado dos experimentos e dos testes estatísticos de Wilcoxon com LSTM para os ativos ABEV3 e B3SA3.

Experimento	ABEV3				B3SA3			
	DC	DCJL	DD	DDJL	DC	DCJL	DD	DDJL
1	0,71	0,83	0,86	0,83	0,73	0,78	0,71	0,78
2	0,81	0,83	0,85	0,82	0,63	0,79	0,72	0,78
3	0,82	0,84	0,86	0,82	0,69	0,79	0,72	0,79
4	0,84	0,83	0,86	0,83	0,71	0,79	0,72	0,79
5	0,82	0,83	0,85	0,82	0,75	0,78	0,72	0,78
6	0,81	0,84	0,86	0,81	0,68	0,78	0,72	0,79
7	0,80	0,83	0,86	0,83	0,69	0,78	0,72	0,79
8	0,80	0,83	0,86	0,82	0,62	0,79	0,72	0,79
9	0,85	0,84	0,86	0,83	0,74	0,79	0,72	0,78
10	0,80	0,83	0,86	0,83	0,72	0,78	0,73	0,79
11	0,82	0,84	0,85	0,81	0,59	0,79	0,72	0,78
12	0,80	0,83	0,85	0,82	0,70	0,78	0,72	0,79
Média	0,81	0,83	0,86	0,82	0,69	0,79	0,72	0,79
Wilcoxon	p=0.005	diferente	p=0.001	diferente	p=0.002	diferente	p=0.001	diferente

*DCJL=Conjunto de dados contínuo com janela Lyapunov. *DDJL=Conjunto de dados discretos com janela Lyapunov.

Tabela 18 - Experimento 03: Resultado dos experimentos e dos testes estatísticos de Wilcoxon com LSTM para os ativos BBDC4, KROT3 e PETR3.

Experimento	BBDC4				KROT3				PETR3			
	DC	DCJL	DD	DDJL	DC	DCJL	DD	DDJL	DC	DCJL	DD	DDJL
1	0,60	0,64	0,63	0,62	0,81	0,85	0,87	0,85	0,77	0,66	0,73	0,66
2	0,56	0,61	0,63	0,61	0,80	0,85	0,87	0,85	0,68	0,66	0,74	0,67
3	0,41	0,63	0,63	0,62	0,82	0,84	0,87	0,86	0,78	0,68	0,73	0,67
4	0,63	0,64	0,63	0,63	0,84	0,85	0,87	0,85	0,70	0,64	0,73	0,65
5	0,59	0,65	0,64	0,62	0,83	0,85	0,87	0,85	0,80	0,65	0,73	0,69
6	0,58	0,65	0,63	0,63	0,81	0,85	0,87	0,85	0,84	0,67	0,74	0,63
7	0,67	0,66	0,63	0,61	0,81	0,85	0,87	0,84	0,79	0,66	0,73	0,67
8	0,62	0,64	0,63	0,63	0,87	0,86	0,87	0,85	0,80	0,67	0,73	0,68
9	0,68	0,64	0,64	0,64	0,83	0,86	0,87	0,86	0,79	0,68	0,74	0,63
10	0,72	0,63	0,63	0,61	0,87	0,85	0,87	0,86	0,75	0,67	0,74	0,68
11	0,60	0,64	0,63	0,62	0,86	0,85	0,87	0,85	0,77	0,69	0,73	0,65
12	0,49	0,64	0,63	0,63	0,84	0,85	0,87	0,85	0,68	0,68	0,74	0,69
Média	0,60	0,64	0,63	0,62	0,83	0,85	0,87	0,85	0,76	0,67	0,73	0,66
Wilcoxon	p=0.076	mesma	p=0.015	diferente	p=0.026	diferente	p=0.001	diferente	p=0.003	diferente	p=0.002	diferente

*DCJL=Conjunto de dados contínuo com janela Lyapunov. *DDJL=Conjunto de dados discretos com janela Lyapunov.