



Universidade Federal
do Rio de Janeiro

Escola Politécnica

ANÁLISE DE NOTÍCIAS DO MERCADO FINANCEIRO UTILIZANDO
PROCESSAMENTO DE LINGUAGEM NATURAL E APRENDIZADO DE
MÁQUINA PARA DECISÕES DE SWING TRADE

Lucas Gama Canto

Projeto de Graduação apresentado ao Curso de Engenharia de Controle e Automação da Escola Politécnica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Engenheiro.

Orientador: Heraldo Luís Silveira de Almeida

Rio de Janeiro
Março de 2020

ANÁLISE DE NOTÍCIAS DO MERCADO FINANCEIRO UTILIZANDO
PROCESSAMENTO DE LINGUAGEM NATURAL E APRENDIZADO DE
MÁQUINA PARA DECISÕES DE SWING TRADE

Lucas Gama Canto

PROJETO DE GRADUAÇÃO SUBMETIDO AO CORPO DOCENTE DO
CURSO DE ENGENHARIA DE CONTROLE E AUTOMAÇÃO DA ESCOLA
POLITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE ENGENHEIRO DE AUTOMAÇÃO.

Examinado por:

Prof. [TODO]Nome do Primeiro Examinador Sobrenome, D.Sc.

Prof. [TODO]Nome do Segundo Examinador Sobrenome, Ph.D.

Prof. [TODO]Nome do Terceiro Examinador Sobrenome, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
MARÇO DE 2020

Gama Canto, Lucas

Análise de Notícias do Mercado Financeiro Utilizando Processamento de Linguagem Natural e Aprendizado de Máquina Para Decisões de Swing Trade/Lucas Gama Canto. – Rio de Janeiro: UFRJ/ Escola Politécnica, 2020.

XIII, 15 p.: il.; 29, 7cm.

Orientador: Heraldo Luís Silveira de Almeida

Projeto de Graduação – UFRJ/ Escola Politécnica/ Curso de Engenharia de Controle e Automação, 2020.

Referências Bibliográficas: p. 13 – 14.

1. Aprendizado de Máquina. 2. Processamento de Linguagem Natural. 3. Mercado Financeiro. I. Silveira de Almeida, Heraldo Luís. II. Universidade Federal do Rio de Janeiro, Escola Politécnica, Curso de Engenharia de Controle e Automação. III. Título.

*Ao povo brasileiro, pela total
contribuição em minha
graduação.*

Agradecimentos

Gostaria de agradecer a todas as pessoas e situações que tornaram este momento possível. Em especial, meus pais Benedita e Manoel, pelo suporte e esforço incondicional em apoiar minha decisão de vir estudar engenharia no Rio de Janeiro, aos professores da graduação, que me fizeram evoluir no âmbito acadêmico, profissional e pessoal, em especial ao meu orientador e professor Heraldo, que não mediu esforços para me ajudar neste trabalho, e aos amigos que me apoiaram e participaram do meu processo de graduação.

Resumo do Projeto de Graduação apresentado à Escola Politécnica/ UFRJ como parte dos requisitos necessários para a obtenção do grau de Engenheiro de Automação.

ANÁLISE DE NOTÍCIAS DO MERCADO FINANCEIRO UTILIZANDO
PROCESSAMENTO DE LINGUAGEM NATURAL E APRENDIZADO DE
MÁQUINA PARA DECISÕES DE SWING TRADE

Lucas Gama Canto

Março/2020

Orientador: Heraldo Luís Silveira de Almeida

Curso: Engenharia de Controle e Automação

Com o objetivo de automatizar análises fundamentalistas de mercado, o uso de tecnologia para processamento de texto vem sendo utilizado constantemente no meio acadêmico[1] e profissional[2]. De forma a contribuir para este campo em crescimento, este trabalho discorre um estudo acerca da criação de modelos preditivos sobre a valorização ou desvalorização de ações na bolsa de valores do Brasil (B3, antiga Bovespa) a partir de notícias sobre o mercado brasileiro de forma a auxiliar decisões de Swing Trade, ou seja, compra e venda de ações dentro de uma janela de tempo maior que um dia.

Para isto, o presente projeto utiliza o framework PyText, que se baseia em conceitos de Aprendizado de Máquina, Redes Neurais e Processamento de Linguagem Natural de forma a desenvolver modelos preditivos com a tarefa de classificação textual.

Abstract of Undergraduate Project presented to POLI/UFRJ as a partial fulfillment of the requirements for the degree of Engineer.

FINANCIAL MARKET NEWS ANALYSIS USING NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING FOR SWING TRADE DECISIONS

Lucas Gama Canto

March/2020

Advisor: Heraldo Luís Silveira de Almeida

Course: Automation and Control Engineering

In order to automate fundamental market analysis, the use of text processing technology has been constantly used in academic[1] and professional[2] means. To contribute to this growing field, this paper discusses a study about the creation of predictive models regarding the valuation or devaluation of shares on the Brazilian stock exchange (B3, former Bovespa) based on news about the Brazilian market in order to assist Swing Trade decisions, that is, buying and selling stocks within a time window longer than one day.

To this end, the present project uses the PyText framework, which is based on Machine Learning, Neural Networks and Natural Language Processing concepts in order to develop predictive models with the task of textual classification.

Sumário

Lista de Figuras	x
Lista de Tabelas	xi
Lista de Símbolos	xii
Lista de Abreviaturas	xiii
1 Introdução	1
1.1 Tema	1
1.2 Delimitação	1
1.3 Justificativa	2
1.4 Objetivos	2
1.5 Metodologia	2
1.6 Descrição	3
2 Fundamentação Teórica	4
2.1 Bolsa de Valores e Ações	4
2.1.1 Preços de Ações	4
2.1.2 Índice de Bolsa de Valores	6
2.2 Aprendizado de Máquina	6
2.2.1 Aprendizado Supervisionado	7
2.2.2 Aprendizado Não Supervisionado	7
2.2.3 Avaliação de Desempenho	7
2.3 Processamento de Linguagem Natural	8
2.3.1 Sintaxe	8
2.3.2 Representação Vetorial	9
2.4 Redes Neurais	9
2.4.1 Rede BilSTM Com Atenção	9
2.4.2 Rede CNN	9
2.5 PyText	9
2.5.1 Configuração	9

2.5.2	Uso	9
3	Obtenção e Pré-processamento de Dados	10
3.1	Estratégia	10
3.2	Pré-processamento	10
4	Treinamento	11
4.1	Configuração	11
4.2	Rotina de Treinamento	11
4.3	Análise	11
5	Considerações Finais	12
5.1	Conclusão	12
5.2	Trabalhos Futuros	12
	Referências Bibliográficas	13
A	Rotinas e Arquivos de Configuração	15
A.1	Rotina de Pré-processamento e Criação de Conjuntos de Dados . . .	15
A.2	Rotina de Criação de Arquivos de Configuração e Execução de Trei- namentos	15
A.3	Modelo de Arquivo de Configuração	15

Lista de Figuras

2.1	Os três níveis da HME, cada nível adiciona um tipo de informação cujo com o qual não seria possível prever um movimento de preço no mercado[3].	5
-----	---	---

Lista de Tabelas

2.1 Os 5 ativos com o maior volume e participação na B3, associados ao Ibovespa[4].	6
---	---

Lista de Símbolos

\emptyset EXEMPLO-Conjunto vazio, p. 1

Lista de Abreviaturas

iBovespa	Índice Bovespa, p.2
HME	Hipótese do Mercado Eficiente, p.4; p.5
AT	Análise Técnica, p.4; p.5
AF	Análise Fundamentalista, p.4; p.5
PNL	Processamento de Linguagem Natural, p.7

Capítulo 1

Introdução

1.1 Tema

O tema deste trabalho se resume no estudo da criação de modelos preditivos de modo que estes possam prever a valorização ou desvalorização de ações da bolsa de valores por meio do processamento de notícias do mercado brasileiro.

Deste modo, o problema a ser abordado é a identificação de quando uma notícia pode impactar positivamente ou negativamente a variação de preço de ações de forma automatizada.

1.2 Delimitação

Este trabalho se restringe ao processamento de texto em português brasileiro, tendo como foco a predição da variação de preço das ações que fazem parte da bolsa de valores do Brasil, a B3. Pela indisponibilidade de dados sobre notícias brasileiras contendo a informação do horário de lançamento da notícia, o projeto mira em predições dentro de uma janela de tempo maior que um dia, de forma a auxiliar decisões de Swing Trade, isto é, operações de compra e venda de ações numa janela de tempo maior que um dia.

Além disso, o estudo se baseia na ferramenta PyText, um framework recentemente desenvolvido pelo Facebook que providencia modelos de processamento de linguagem natural de última geração através de uma interface simples e extensível[5].

1.3 Justificativa

Diante do crescente número de investidores na bolsa de valores no Brasil, nota-se uma maior preocupação da população brasileira acerca da busca por independência financeira e fontes alternativas de renda com o intuito de contribuir à economia familiar, previdência, ou mesmo utilizar este método como fonte principal de renda[6].

Ao mesmo tempo, estudos associados à inteligência artificial, aprendizado de máquina e processamento de linguagem natural continuam emergindo no meio acadêmico e auxiliando o meio profissional como nunca antes, incluindo o mercado financeiro[7].

Através destes dois fatores, o presente trabalho busca contribuir para a difusão do estudo e uso de algumas destas tecnologias sobre um assunto que gradualmente se encontra dentro do interesse da população brasileira e que colabora para uma possível instauração de uma cultura de economia e independência financeira no Brasil.

1.4 Objetivos

O objetivo geral do presente trabalho é de analisar modelos preditivos associados ao mercado financeiro que possam ser construídos a partir do framework PyText, tendo como objetivos específicos, apresentar: (1) A busca por dados de notícias e do histórico da bolsa de valores; (2) A lógica utilizada para a união destes dados de forma a construir os conjuntos de dados utilizados no treinamento dos modelos; (3) O pré-processamento dos conjuntos de dados; (4) As possíveis configurações do framework utilizado de forma a obter a melhor performance; (5) O detalhamento e a análise dos modelos finais encontrados.

1.5 Metodologia

O trabalho teve início a partir da procura por bases de dados de notícias associadas ao mercado brasileiro e escritas em português do Brasil, seguida pela obtenção do histórico das variações de preço dos ativos que compõem o iBovespa. Após isto, o histórico foi filtrado de forma a manter as informações dos 5 ativos mais significativos e das variações destes ativos que ocorreram dentro da mesma janela de tempo das notícias obtidas. Em seguida, estes dados foram unidos de forma a obter 5 conjuntos de dados para cada ativo, cada um levando em consideração uma diferente janela de tempo para indicar a valorização: de 1 a 5 dias.

Logo após, houve a etapa de pré-processamento do corpo das notícias de forma a remover possíveis ruídos e facilitar a etapa de treinamento, sem perda de contexto do conteúdo. Com os conjuntos de dados prontos, foram feitos testes no PyText com o objetivo de definir a melhor configuração possível para a natureza dos dados, e assim obter a melhor performance.

Por fim, os testes finais de cada modelo gerado foi detalhado e analisado para permitir uma conclusão e avaliação do processo como um todo.

1.6 Descrição

O capítulo 2 apresenta toda a fundamentação teórica utilizada como base para o projeto a partir de uma breve descrição de como a bolsa de valores funciona seguida de explicações sobre Aprendizado de Máquina, Processamento de Linguagem Natural, Redes Neurais e o framework Pytext.

No capítulo 3 é detalhado todo o processo executado para obtenção do conjunto de notícias e do histórico da B3, seguido do pré-processamento realizado nestes dois conjuntos e a criação dos conjuntos de dados finais utilizados para o treino, cada um associado a um ativo e uma janela de tempo específica.

Os detalhes das configurações utilizadas no PyText e o treinamento em si é especificado no capítulo 4, onde há uma discussão acerca dos parâmetros encontrados para a geração de modelos mais performáticos, além das métricas finais encontradas para cada modelo gerado.

Por fim, o capítulo 5 apresenta uma conclusão acerca dos modelos encontrados seguido por sugestões que futuramente podem ser aplicadas para a evolução do tema e uma possível melhora de desempenho dos modelos preditivos.

O código desenvolvido para o pré-processamento e geração dos conjuntos de dados e arquivos de configurações do PyText utilizados para a geração dos modelos podem ser encontrados no repositório do github referenciado em [8].

Capítulo 2

Fundamentação Teórica

2.1 Bolsa de Valores e Ações

A Bolsa de Valores é um lugar centralizado onde, além de abranger outros tipos de investimento, se negociam ações (também chamados de ativos ou papéis), isto é, parcelas do capital social de empresas de capital aberto. Atualmente a B3 (Brasil, Bolsa, Balcão) é a Bolsa de Valores oficial do Brasil que em 2017 atingiu a 5ª posição das maiores bolsas de mercados de capitais do mundo em valor de mercado, com um patrimônio de US\$ 13 bilhões[9].

As ações são negociadas diariamente a partir das ordens de compra e venda emitidas pelas corretoras durante o pregão eletrônico, que na B3, funciona em dias úteis das 10:00 às 17:00.

2.1.1 Preços de Ações

O preço de um ativo na Bolsa de Valores pode ser determinado por diversas razões que podem se relacionar entre si, entre essas, pode-se destacar a lei da oferta e demanda, perspectivas de crescimento da empresa associada ao papel e especulação. A previsibilidade acerca de movimentações no mercado de ações normalmente pode ser baseada em Análise Técnica (AT - estudo dos movimentos do mercado baseado em métricas como preço, volume e taxa de juros[10]), Análise Fundamentalista (AF - estudo feito a partir de resultados financeiros e operacionais, indicando a saúde da empresa[11]) ou numa junção destes dois conceitos.

A validade da previsibilidade destas movimentações são questionadas por críticas com base na Hipótese do Mercado Eficiente (HME) e seus três níveis definidos em [12]:

- HME fraca: Afirma que os preços refletem totalmente a informação contida na sequência histórica dos preços. Ou seja, a AT não consegue prever os movimentos futuros pois os preços passados só podem descrever o presente.
- HME semi-forte: Afirma que os preços presentes não só refletem toda a sequência histórica de preços mas também toda informação pública sobre as organizações associadas ao ativo em questão. Neste nível de eficiência, a AF também não seria capaz de prever movimentos futuros, pois toda informação como demonstrativos de resultados ou análises orçamentárias refletiria apenas o preço presente.
- HME forte: Neste nível, é afirmado que *toda* informação conhecida sobre as organizações é totalmente refletida pelo preço presente, logo, nem mesmo aqueles com informações privilegiadas podem utilizar isto como ferramenta para prever preços futuros.



Figura 2.1: Os três níveis da HME, cada nível adiciona um tipo de informação cujo com o qual não seria possível prever um movimento de preço no mercado[3].

Não há uma resposta correta perante a validade da HME. Porém, muitos acadêmicos acreditam, pelo menos, na HME fraca[13], fazendo com que, em algumas ocasiões, seja preferível a utilização da AF, ou seja, a análise de resultados financeiros, relatórios anuais e notícias divulgadas acerca do mercado financeiro.

2.1.2 Índice de Bolsa de Valores

Com objetivo de parametrizar algumas informações intrínsecas às bolsas, estas disponibilizam diversos índices. O principal índice da B3 é o Ibovespa, que é formado pelos ativos com maior volume negociado na bolsa nos últimos meses e indica, de forma resumida, o desempenho das ações negociadas na B3. Por ser um indicador principal, muitos fundos de investimento baseados no mercado de ações estão atrelados ao Ibovespa, contribuindo para a atratividade destes ativos de maneira geral.

Tabela 2.1: Os 5 ativos com o maior volume e participação na B3, associados ao Ibovespa[4].

Código	Ação	Qtde. Teórica	Part.(%)
ITUB4	Itaú Unibanco	4.738.562.684	9,095
PETR4	Petrobras	4.520.185.835	7,038
BBDC4	Bradesco	3.873.597.664	7,028
VALE3	Vale S.A.	3.147.743.563	8,414
ABEV3	AMBEV	4.344.066.764	4,173

2.2 Aprendizado de Máquina

Pode-se definir Aprendizado de Máquina como o campo de estudo de algoritmos com o objetivo de fazer com que computadores possam agir sem serem explicitamente programados para fazer determinada tarefa. São algoritmos que analisam dados e aprendem com eles, para que assim possam fornecer uma predição de algo on mundo.

Outra definição dada por Tom M. Mitchell[14] fala que o campo de Aprendizado de Máquina busca responder a pergunta: “Como podemos construir sistemas de computadores que possam automaticamente melhorar através de experiência e quais são as leis fundamentais que governam todo este processo de aprendizado?”. Outra definição do mesmo autor[15] diz que “Um programa de computador é dito aprender com a experiência E em respeito a uma tarefa T e medida pelo desempenho P se o seu desempenho em T, medido por P, melhora com a experiência E”. Neste conceito, se fosse desejado um programa de computador que aprendesse a classificar e-mails como spam ou não, por exemplo, poderíamos fazer a seguinte associação:

- E = A experiência de ver o usuário classificar emails como spam ou não.
- T = A tarefa de classificar os emails.
- P = O número ou fração de emails corretamente classificados como spam/não spam

Geralmente, os algoritmos de Aprendizado de Máquina podem ser divididos em dois tipos: Aprendizado Supervisionado e Aprendizado Não Supervisionado.

2.2.1 Aprendizado Supervisionado

Neste tipo, o algoritmo é inicialmente servido por uma série de dados rotulados cujo resultado já é conhecido. A ideia é que o algoritmo aprenda a criar uma estratégia para chegar ao resultado baseando-se nesses dados de modelo inicial. O aprendizado supervisionado pode ser dividido em problemas de regressão ou classificação.

Problema de Regressão

Neste tipo de problema, os dados de entrada (parâmetros) são mapeados em uma função contínua. Por exemplo, um algoritmo cujo objetivo fosse prever o preço dos imóveis na cidade do Rio de Janeiro baseando-se em dados como área útil, bairro, número de vagas na garagem, etc é considerado como um problema de regressão, pois o resultado final será um número contínuo, neste caso, o preço dos imóveis.

Problema de Classificação

Neste caso, os parâmetros são mapeados de forma a classificar os dados em categorias distintas. Por exemplo, um algoritmo utilizado para prever se um tumor é benigno ou maligno a partir de dados como o tamanho, rugosidade do tumor e idade do paciente é considerado um problema de classificação, pois o resultado final será a categoria na qual o tumor pertence.

TODO: falar sobre balanceamento

2.2.2 Aprendizado Não Supervisionado

Neste tipo de aprendizado, não existe um conjunto inicial de dados e resultados, ou seja, nos permite abordar problemas onde temos pouca ou nenhuma ideia do que nossos resultados devem aparentar. Um exemplo, seria um algoritmo onde, utilizando uma coleção de 1000 artigos publicados por uma universidade, fizesse um agrupamento desses artigos, baseando-se em diferentes variáveis como frequência de palavras semelhantes, número de páginas, etc.

2.2.3 Avaliação de Desempenho

TODO: falar sobre as métricas utilizadas no pytext

2.3 Processamento de Linguagem Natural

Processamento de Linguagem Natural (PNL) se resume ao campo de estudo das tecnologias utilizadas para ajudar computadores a entenderem a linguagem natural dos humanos, é também considerado uma subárea da Inteligência Artificial. Pode ser usado em diversas aplicações[16], como por exemplo:

- Aplicativos de tradução de idioma, como o Google Translator
- Processamento de palavras, que empregam PNL para a correção gramática
- Resposta de Voz iterativa em call centers, de forma a responder adequadamente conforme a requisição de usuários
- Assistentes pessoais, como OK Google, Siri, Cortana e Alexa

Geralmente, o PNL abrange um pré-processamento de texto antes deste ser transformado em uma forma inteligível por computadores, de forma a remover ruídos ou facilitar o processamento, e isto pode ocorrer de diversas formas.

2.3.1 Sintaxe

A sintaxe se refere à forma de como as palavras se organizam em uma sentença para que se obtenha sentido gramatical. Estas são algumas técnicas de sintaxe que podem ser utilizadas no pré-processamento de texto:

- **Stemização:** É a transformação de palavras flexionadas para sua forma radical. Por exemplo, as palavras “estudos”, “estudar” e “estudando” se transformariam apenas em “estud”, mas a palavra “tiver” se transformaria em “tiv” e “tenho” se transformaria em “tenh”.
- **Lematização:** Semelhante à Stemização, porém, a palavra é resumida para seu lema, fazendo com que se alcance um nível maior de abstração. Neste caso, tanto a palavra “tiver” como “tenho” se transformaria no lema “ter”.
- **Remoção de *stopwords*:** A remoção de “palavras de parada”, ou seja, palavras como “a”, “de”, “o”, “da”, “que”, “e”, “do” é útil pois, na maioria das vezes, não são informações importantes para construção do modelo.

Além destas, outras técnicas mais simples são utilizadas, como a transformação de caracteres maiúsculos para minúsculos.

2.3.2 Representação Vetorial

Após o pré-processamento, diversas técnicas podem ser utilizadas para a transformação do texto em números.. TODO

Bag of Words

TODO

Vetores Baseados em Redes Neurais

TODO

2.4 Redes Neurais

TODO

2.4.1 Rede BiLSTM Com Atenção

TODO

2.4.2 Rede CNN

TODO

2.5 PyText

TODO

2.5.1 Configuração

TODO

2.5.2 Uso

TODO

Capítulo 3

Obtenção e Pré-processamento de Dados

TODO

3.1 Estratégia

TODO

3.2 Pré-processamento

TODO

Capítulo 4

Treinamento

TODO

4.1 Configuração

TODO

4.2 Rotina de Treinamento

TODO

4.3 Análise

TODO

Capítulo 5

Considerações Finais

TODO

5.1 Conclusão

TODO

5.2 Trabalhos Futuros

TODO

Referências Bibliográficas

- [1] LIU, Z., ZHU, H., CHONG, T. Y. “An NLP-PCA Based Trading Strategy On Chinese Stock Market”, *Advances in Social Science and Education and Humanities Research*, v. 334, n. 2, pp. 80–89, jul. 2019.
- [2] SEDLAK, M. “How Natural Language Processing is transforming the financial industry”. <https://www.ibm.com/blogs/watson/2016/06/natural-language-processing-transforming-financial-industry-2/>, 2016. Acessado em Dezembro/2019.
- [3] SMITH, D. J. “Efficient-Market Hypothesis (EMH) and Random-Walk Theory”. <https://stockmarketsupertrader.com/theory/efficient-market-hypothesis-emh-and-random-walk-theory/>, 2018. Acessado em Dezembro/2019.
- [4] “Índice Bovespa (Ibovespa)”. http://www.bmfbovespa.com.br/pt_br/produtos/indices/indices-amplos/indice-ibovespa-ibovespa-composicao-da-carteira.htm. Acessado em Dezembro/2019.
- [5] ALY, A., LAKHOTIA, K., ZHAO, S., et al. “PYTEXT: A SEAMLESS PATH FROM NLP RESEARCH TO PRODUCTION”, 2018.
- [6] DO PAVINI, A. “Cresce número de pessoas físicas como profissionais na Bolsa”. <https://exame.abril.com.br/seu-dinheiro/cresce-numero-de-pessoas-fisicas-como-profissionais-na-bolsa/>, 2019. Acessado em Dezembro/2019.
- [7] BACHINSKIY, A. “The Growing Impact of AI in Financial Services: Six Examples”. <https://towardsdatascience.com/the-growing-impact-of-ai-in-financial-services-six-examples-da386c0301b2>, 2019. Acessado em Dezembro/2019.
- [8] CANTO, L. G. “Stock Market Predictor”. <https://github.com/lgcanto/stock-market-predictor/>, 2019. Acessado em Dezembro/2019.

- [9] MOREIRA, M. “Fusão entre BM&FBovespa e Cetip cria a B3, 5ª maior bolsa de valores do mundo”. <http://agenciabrasil.ebc.com.br/economia/noticia/2017-03/fusao-entre-bmfbovespa-e-cetip-cria-b3-5a-maior-bolsa-de-valores-do-mundo>, 2017. Acessado em Dezembro/2019.
- [10] WAWRZENIAK, D. “O Que É Análise Técnica?” <https://www.bussoladoinvestidor.com.br/o-que-e-analise-tecnica/>, 2018. Acessado em Dezembro/2019.
- [11] WAWRZENIAK, D. “O Que É Análise Fundamentalista?” <https://www.bussoladoinvestidor.com.br/o-que-e-analise-fundamentalista/>, 2018. Acessado em Dezembro/2019.
- [12] MALKIEL, B. G. “Efficient Market Hypothesis”, *Finance*, pp. 127–134, 1989.
- [13] “Análise Técnica x Análise Fundamentalista”. <https://www.tororadar.com.br/investimento/analise-tecnica/analise-tecnica-x-fundamentalista>. Acessado em Dezembro/2019.
- [14] JORDAN, M. I., MITCHELL, T. M. “Machine learning: Trends, perspectives, and prospects”, *Science*, v. 349, pp. 80–89, jul. 2015.
- [15] MITCHELL, T. M. *Machine Learning*. McGraw Hill, 1997.
- [16] GARBADE, M. J. “A Simple Introduction to Natural Language Processing”. <https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32>, 2018. Acessado em Dezembro/2019.

Apêndice A

Rotinas e Arquivos de Configuração

- A.1 Rotina de Pré-processamento e Criação de Conjuntos de Dados
- A.2 Rotina de Criação de Arquivos de Configuração e Execução de Treinamentos
- A.3 Modelo de Arquivo de Configuração