



Plano de Projeto - Acadêmico

1. Participantes da pesquisa

Nome completo
Cristiane de Andrade Coutinho
Adriana
Supervisor - Geraldo Magela Severino Vasconcelos

2. Motivação e Justificativa

Tema de pesquisa: Previsão de retornos de ativos com algoritmos de deep learning

Ao emprestar recursos a instituições, cria-se um ambiente de negociação com objetivo de recebimento de retorno financeiro com base em investimentos. Dessa forma, os prêmios que são retornados dessas negociações são denominados como “valores mobiliários”. Quando se adquire um título de dívida, o investidor também assume o risco sobre esse retorno, ou seja, a possibilidade desse investimento ser desvalorizado.

Com o objetivo de maximização de resultados, criou-se toda uma estrutura tecnológica para apoiar e estabelecer a relação de tempo e causalidade nos ganhos de ativos. Para Guedes (2022), o desenvolvimento de algoritmos do setor financeiro foi inserido aos poucos no mercado de negociações e evoluiu à medida que análises mais técnicas de matemática, probabilidade e estatística foram sendo incorporadas nos chamados “robôs investidores”.

Com esse avanço tecnológico, o desenvolvimento de algoritmos com Machine Learning e Deep Learning, tornou-se cada vez mais viável e atrativo, utilizados para

seleção de ativos e otimização de carteiras de investimentos. Shimabukuro (2024), atribuiu esses avanços ao crescente aumento na velocidade de processamento computacional e também a diminuição do tamanho físico de memória. Por causa disso se tornou possível, a implementação mais barata de algoritmos de otimização e predição.

Soares (2023), destaca que algoritmos de machine learning são amplamente explorados em problemas de regressão e classificação para previsão de especificação de retornos e classificação de um determinado ativo. Em contrapartida, Shimabukuro (2024), ressalta que os fatores limitantes relacionados à aplicação de modelos de machine learning para previsão de ativos financeiros, são limitantes no que tange à grande necessidade de volume de dados, baixa sensibilidade a ruído de mercado e risco de overfitting.

Tendo isso em vista, este estudo propõem uma abordagem qualitativa a respeito da análise de diferentes tipos de arquiteturas de redes neurais, com utilização de técnicas de deep learning, a fim de capturar ruídos de mercado e generalizar o problema de forma que essas arquiteturas possam ser aplicadas a diferentes tipos de ativos financeiros.

3. Problema de pesquisa

A previsão de retorno de um determinado ativo é extremamente complexo de prever por sua natureza ruidosa e por ser difícil para algoritmos capturar suas nuances e generalizar o problema. Dessa forma, o problema de pesquisa se adequa em quais arquiteturas de deep learning podem generalizar melhor a previsão de retornos de ativos financeiros.

4. Objetivos

- **Objetivo geral:** Análise qualitativa da capacidade preditiva de retorno financeiro a partir de diferentes arquiteturas de redes neurais usando técnicas de deep learning sobre uma série temporal.
- **Objetivos específicos:**
 1. Avaliação do desempenho das arquiteturas treinadas, observando valores preditos em comparação aos valores reais.
 2. Mensurar a capacidade de generalização das arquiteturas, testando sobre diferentes bases de dados.

5. Escopo

Este trabalho visa realizar um estudo sobre diferentes arquiteturas de deep learning, com o objetivo de realizar previsões de retornos de ativos financeiros. Além disso, o estudo também vai buscar entender como essas arquiteturas podem generalizar o problema para diferentes ativos.

Será feita uma busca em portais web por base de dados que disponham dos dados em séries temporais, para que seja observado o comportamento conforme o tempo. Após esse processo de busca, será feito análises exploratórias dos dados a fim de serem encontrados padrões, anomalias, e suas principais características. Se forem encontradas anomalias, será feito um processo de limpeza, normalização e padronização dos dados.

Juntamente com essa etapa, será estudado diferentes tipos de arquiteturas de redes neurais a fim de encontrar uma solução adequada para o problema. Nesta etapa será feita uma sequência de testes com a base de dados. Dessa forma,

espera-se contribuir com novas perspectivas sobre o tema para a comunidade científica, trazendo resultados consistentes.

6. Metodologia de pesquisa

- **Tipo de pesquisa:** Exploratória.
- **Abordagem:** Qualitativa.
- **Métodos de coleta de dados:** Pesquisa bibliográfica, análise de dados públicos.
- **Técnicas de análise:** Análise exploratória de dados.

7. Cronograma de Trabalho

- **SPRINT 1** 04/08 - 15/08

Artefato 1: Formulário de Submissão de projeto

Resposta ao formulário, contendo informações relevantes do projeto. Deve descrever com clareza o problema de pesquisa, seus objetivos e hipóteses juntamente com um escopo macro do que será desenvolvido durante cada sprint.

Artefato 2: Documento de plano do projeto

O plano do projeto deverá conter os passos necessários para condução durante o módulo. Especificando o cronograma de trabalho em sprint, detalhando cada artefato. Uma breve introdução contendo a justificativa do trabalho. O plano de pesquisa também deverá conter a descrição dos problemas de pesquisa, escopo, metodologia e resultados esperados.

- **SPRINT 2** 18/08 - 29/08

Artefato 1: Introdução ao tema da pesquisa com suas devidas revisões de literatura

- Escrita da introdução, destacando pontos importantes sobre sua relevância, revisão bibliográfica, cenário atual da problemática e contribuições científicas a serem desenvolvidas.

Artefato 2: Justificativa do tema escolhido

- Escrita da justificativa do projeto, ressaltando as motivações que levaram a escolha do tema.

Artefato 3: Objetivos definidos da pesquisa

- Definição e escrita dos objetivos gerais e específicos da pesquisa.

Artefato 4: Hipóteses da pesquisa

- Definição e escrita das hipóteses da pesquisa.

Artefato 5: Problemática da pesquisa

- Delimitação do problema de pesquisa, levando em consideração as arquiteturas de redes neurais escolhidas.

Artefato 6: Descrição dos dados utilizados

- Descrição da fonte utilizada para a coleta de dados
- Descrição das variáveis do dataset
- Descrição do tipo de cada variável

Artefato 7: Definição das arquiteturas de rede neurais a serem utilizadas

- Descrição das arquiteturas utilizadas
- Justificativa para a utilização das mesmas

- **SPRINT 3** 01/09 - 12/09

Artefato 1: Análise exploratória dos dados

- Quantidade de exemplos de cada variável
- Verificação de dados nulos ou vazios
- Estatística descritiva média para variáveis numéricas
- Estatística descritiva moda para variáveis categóricas
- Verificação de outliers
- Distribuição de cada variável
- Matriz de correlação

Artefato 2: Limpeza de dados

- Remover ou imputar valores ausentes
- Eliminar duplicatas
- Tratamento de outliers
- Normalização / padronização
- Encoding de variáveis categóricas

Artefato 3: Materiais e métodos que serão utilizados utilizados na pesquisa

- Descrição dos métodos utilizados para a limpeza dos dados
- Descrição da análise exploratória de dados

Artefato 4: Implementação das arquiteturas utilizadas

- Implementação das arquiteturas utilizadas
- Avaliação dos treinamentos e resultados obtidos

- **SPRINT 4** 15/09 - 26/09

Artefato 1: Implementação das arquiteturas utilizadas

- Refinamento das arquiteturas utilizadas

- Reavaliação das arquiteturas

Artefato 2: Materiais e métodos

- Descrição das arquiteturas implementadas.
- Descrição dos processos metodológicos utilizados.

Artefato 3: Revisão de pares

- Submeter o projeto a algum professor diferente para avaliação do escopo e desenvolvimento.

● **SPRINT 5 - 29/09 - 10/10**

Artefato 1: Refinamento das arquiteturas utilizadas

- Após os momentos de avaliação, submeter as arquiteturas a um processo de refinamento.

Artefato 2: Conclusão do estudo

- Reunião dos resultados obtidos para a escrita da conclusão do estudo.

Artefato 3: Apresentação final do estudo

- Elaboração de um slide de apresentação em inglês.

8. Resultados esperados

É esperado que esse estudo aponte novos entendimentos sobre predição de retornos usando deep learning assim como novas abordagens e melhores resultados da literatura atual.

9. Observações gerais sobre a pesquisa a ser desenvolvida.

10. Referências bibliográficas

- [1] GUEDES, Anderson Cerqueira et al. **Computação evolutiva para otimização de carteiras de estratégias de negociação no mercado financeiro.** 2022.
- [2] SOARES, Taís Rigor. **Previsão de Retornos de Ativos no Mercado Financeiro Através de Métodos de Otimização.** 2023. Dissertação de Mestrado.
- [3] SHIMABUKURO, Camilo Ilzo. **Deep Learning aplicado à predição de retornos do Ibovespa: uma análise do desempenho da rede neural LSTM utilizando log-retornos e diferenciação fracionária.** 2024. Tese de Doutorado. Universidade de São Paulo.