

Gabriel Przytocki

Cientista de Dados | Engenheiro de Machine Learning

E-mail: gabrielhprzytocki@gmail.com

Github: <https://github.com/gprzy>

FORMAÇÃO ACADÊMICA

- **Pontifícia Universidade Católica do Paraná**
Bacharel em Ciência da Computação

Curitiba, Brasil
Fev. 2019 - Dez. 2022

EXPERIÊNCIA

- **Junto Seguros**

Curitiba, Brasil
Jan. 2022 - o momento

→ Cientista de Dados

- Criação de modelos de Machine Learning;
- Criação de *Pipelines* de Machine Learning;
- Aplicação de boas práticas de MLOPs;
- Modelagem relacional de dados (SQL Server);

Orquestração de processos com Airflow;
Computer Vision (OpenCV e Pytesseract OCR);
Pipelines de extração de dados (ETL);
Python (scikit-learn, tensorflow, numpy, pandas);

→ Engenheiro de Dados Pleno

Out. 2021 - Fev. 2022

- Processos de ETL (*Extract, Transform, Load*);
- Modelagem relacional de dados (T-SQL);
- Criação de fontes de dados (SQL Server & Tableau);

Pipelines de extração de dados (Airflow & SQL Server);
Webscraping de dados públicos (Selenium & BS5);
Consultas em SQL (T-SQL);

- **Pontifícia Universidade Católica do Paraná**
→ PIBIC Undergraduate Researcher

Curitiba, Brasil
Jul. 2021 - Jul. 2022

- Previsão de ações em Séries Temporais;
- Otimização com Algoritmos Genéticos (AG);
- Redes neurais (CNN e LSTM);

Indicadores de Análise Técnica;
Simulações de investimentos (*Backtesting*);
Modelos estatísticos (AR, ARMA & ARIMA);

→ PIBIC Undergraduate Researcher

Ago. 2020 - Ago. 2021

- Previsão de ações em Séries Temporais;
- Aplicação de modelos estatísticos (Autorregressivo);

Aplicação de Redes Neurais (CNN e LSTM);
Simulações de investimentos (*Backtesting*);

- **Mirum Agency**

→ Engenheiro de Dados

Curitiba, Brasil
Nov. 2020 - Out. 2021

- Desenvolvimento de *pipelines* de extração dados;
- Microsserviços (Google Cloud Platform);
- APIs RESTful (Python & Flask);
- *Datawarehouse* (SQL & BigQuery);

Extract, Transform & Load (ETL);
Cloud Functions & BigQuery;
Ambiente containerizado (Docker);
Web Tracking (Google Analytics, TagManager);

→ Software Developer Intern

Mar. 2020 - Nov. 2020

- Javascript, HTML5, CSS3, JQuery & Bootstrap;

Implementações de *e-mail marketing*;

- **Viasoft Korp**

→ Software Developer Intern

Curitiba, Brasil
Jul. 2019 - Dez. 2019

- Desenvolvimento *back-end* em Delphi e SQL Server;

MVVM e MVC e *framework* de testes unitários;

IDIOMAS

Português: Fluente ou Nativo

Inglês: Nível Avançado

PROJETOS

- **Avaliação do Desempenho de Técnicas de Deep Learning para Classificação do Comportamento Temporal de Ações:** Relatório concluído de iniciação científica, referente a implementação e avaliação de técnicas de redes neurais (CNN e LSTM) em comparação com o modelo estatístico Autorregressivo (AR) na previsão em Séries Temporais Financeiras. → [Link para GitHub](#)
- **Previsão do Valor Futuro de Ações a Partir do Uso de Indicadores Técnicos:** Relatório concluído de iniciação científica, referente a implementação de combinações otimizadas por algoritmos genéticos de Indicadores Técnicos na previsão de ações em Séries Temporais Financeiras, juntamente com redes CNN e LSTM. → [Link para GitHub](#)
- **Aplicação de Fatores de Decaimento em Sistemas Adaptativos de Recomendação:** Trabalho de Conclusão de Curso (TCC) em andamento, referente ao curso de Bacharelado em Ciência da Computação. Aplicações de métodos de esquecimento (*forgetting*) de informações obsoletas nos algoritmos adaptativos (*stream*) ISGD e IBPRMF.

COMPETÊNCIAS E TECNOLOGIAS

- | | |
|---|---|
| • Machine Learning: <code>scikit-learn</code> & <code>tensorflow</code> ; | Versionamento: <code>Git</code> , <code>GitHub</code> ; |
| • Estatística: Descritiva, Inferência (<code>scipy</code>); | Conteneirização: <code>Docker</code> , <code>docker-compose</code> ; |
| • Banco de Dados: <code>SQL</code> (T- <code>SQL</code>) & <code>SQL Server</code> ; | Linguagens: <code>Python</code> , <code>C++</code> , <code>R</code> , <code>Julia</code> & <code>Java</code> ; |
| • Processamento Distribuído: <code>pyspark</code> & <code>Hadoop MapReduce</code> ; | Soft Skills: <i>Problem Solving</i> , <i>Communication</i> ; |
| • Microserviços e APIs RESTful: <code>Cloud Functions</code> , <code>flask</code> ; | Cloud: <code>GCP</code> & <code>AWS</code> ; |
| • ETL: <code>Airflow</code> , <code>SQL Server</code> & <code>Cloud Functions</code> ; | Manipulação de Dados: <code>pandas</code> & <code>numpy</code> ; |
| • Visualização de Dados: <code>matplotlib</code> & <code>seaborn</code> ; | Séries Temporais: <code>statsmodels</code> & <code>pycaret</code> ; |
| • MLOPs: <code>mlflow</code> , <code>metaflow</code> ; | NLP: <code>spacy</code> & <code>nlTK</code> ; |

FORMAÇÃO COMPLEMENTAR

- *Machine Learning Specialization* - (**Stanford**, **Coursera**);
- *Deep Learning Specialization* - (**DeepLearning.AI**, **Coursera**);
- *Architecting with Google Compute Engine* (**Coursera**);
- *Elastic Google Cloud Infrastructure: Scaling and Automation* (**Coursera**);
- *Essential Google Cloud Infrastructure: Core Services* (**Coursera**);
- *Essential Google Cloud Infrastructure: Foundation* (**Coursera**);
- *Google Cloud Platform Fundamentals: Core Infrastructure* (**Coursera**);
- *Reliable Google Cloud Infrastructure: Design and Process* (**Coursera**);
- *Analyst Developer* (Certificado Profissional - **Pontifícia Universidade Católica do Paraná**);
- *Data Scientist* (Certificado Profissional - **Pontifícia Universidade Católica do Paraná**);