Rodrigo Aragao

433 N 32nd Street, Philadelphia, PA 19104 | (267)-366-5923
rodrigo01.aragao@gmail.com | www.linkedin.com/in/rb-aragao | https://rod31.github.io

EDUCATION

Drexel University | Philadelphia, PA *September 2021 – June 2025*
Bachelor of Science in Computer Engineering / Minor in Neuroscience

SKILLS

Programming: Python, MATLAB, C, C++, SQL
Programming Frameworks: PyTorch, TensorFlow, HuggingFace, PySpark, Pandas, Matplotlib, Tkinter, Streamlit
Machine Learning & NLP: Transformer-based models (BERT), SVM, Sentiment Analysis, Feature Engineering, OpenAI GPT API, FAISS
Software: Microsoft Office, Power BI, SQL Management Studio, Spyder IDE, VS Code, Azure Databricks, Azure Synapse

EXPERIENCE

GS1 US **Ewing, NJ**
Data Analytics Engineer Co-Op *April 2024 – September 2024*
• Developed PySpark notebooks and SQL views in Azure Databricks for barcode prefix analysis used across global trade datasets.
• Built Power BI dashboards and conducted SQL-based QA tests on large datasets to validate data pipelines in Azure Synapse.
• Created and automated data curation scripts within Azure environments to improve accessibility and consistency for stakeholders.
• Supported design and documentation of data pipelines and ELT workflows, streamlining product data accessibility for analytics.

Drexel University, Cognitive Neuroengineering and Wellbeing Laboratory **Philadelphia, PA**
Undergraduate Researcher *September 2023 – September 2024*
• Extracted and engineered features from EEG recordings using MATLAB (EEGLAB), transforming raw brainwave data into structured epoch-level power metrics for downstream analysis.
• Developed and evaluated Support Vector Machine (SVM) model to classify fast vs. slow reaction times, incorporating k-fold cross-validation, performance metrics, and decision boundary visualizations, achieving 61% prediction accuracy.
• Automated the full ML pipeline, from signal processing to model training and interpretation, supporting research on the role of theta and alpha frequency waves in cognitive attention in closed-loop neurocognitive tasks.

Fulton Bank, IT Chief Data Office **Lancaster, PA**
Data Science & Visualization Co-Op *April 2023 – September 2023*
• Created Python scripts to perform data quality profiling (completeness, uniqueness, and consistency) in Lending & Risk datasets.
• Built and enhanced Power BI reports, including Delinquency and Risk Rating dashboards for key business insights.
• Contributed to data governance by documenting and automating validation checks on historical datasets.

Drexel University College of Medicine, Neural Circuits Engineering Laboratory **Philadelphia, PA**
Undergraduate Researcher *July 2022 – September 2022*
• Assisted in optimizing computational models of ion channels in *Drosophila* neurons using Python and Machine Learning.

PROJECTS

Optogenetics Device for Engineering Neural Circuits – Senior Design Project *September 2024 – June 2025*
• Engineered a modular, Arduino-based LED control system with 96-well stimulation, enabling configurable intensity, color, pulsing, and timing for high-throughput optogenetics experiments in *Drosophila* neural circuits.
• Developed a dynamic Python GUI hosted on Raspberry Pi for real-time protocol control and experiment execution via serial communication, enhancing accessibility for neurobiology research.
• Integrated temperature monitoring with automated fan control and CSV-based metadata logging for experiment reproducibility.
• Enabled saving/loading of LED protocols and well assignments using structured JSON, supporting repeatable and scalable testing.

Succession TV Series Natural Language Processing (NLP) & Chatbot – Machine Learning Project *August 2024 – Present*
• Extracted and preprocessed full dialogue scripts for all Succession seasons using PyMuPDF and custom regex heuristics to isolate each character's lines.
• Performed comprehensive exploratory data analysis, including lexical diversity analysis, topic modeling, sentiment analysis, and word cloud visualizations, to gain insights into character speech patterns.
• Developed a sarcasm detection model using fine-tuned BERT to classify character's dialogues; conducted full ML pipeline.
• Built a FAISS similarity index on SentenceTransformer embeddings to retrieve contextually relevant quotes as prompts.
• Fine-tuned an OpenAI LLM (GPT-4o-mini) with targeted prompt engineering to emulate each character's unique style and tone.
• Implemented and launched a Streamlit web app featuring character selection, chat-history persistence, and multiple reply modes, showcasing end-to-end AI prototyping and user-facing deployment.

IMDB Movie Review Sentiment Classifier – Machine Learning Project *June 2024 – July 2024*
• Evaluated multiple models (Logistic Regression, SVM, Random Forest, and BERT) on IMDB reviews to classify movie reviews as positive or negative with optimized accuracy; applied cross-validation and grid search, achieving over 90% accuracy on unseen data.