JELIN RAPHAEL AKKARA

MASTERS IN PHYSICS OF DATA

Email: jelinraphaelakkara@gmail.com Portfolio: https://jelinr.github.io/

PROJECTS / EXPERIENCE

Time Series Analysis using Transformers

- **Summary:** Collaborated with 3 peers and trained a transformer model to predict occurrence of an anomaly event from a heavily imbalanced dataset. Transformer performance (in accuracy and recall) was better and more reliable (with standard deviations of 0.01 - 0.02) compared to traditional models like CNN and DNN (with higher standard deviations).

- **Concepts/Skills used:** Time series analysis, Transformers, Principal Component Analysis, Handling imbalanced datasets using down-sampling and stratification, Communication skills, Teamwork

Fake News Classification using Multinomial Naive Bayes

- **Summary:** Built an statistical text classification model (in R) that recognizes varying degrees of fake news using the Mutlinomial Naive Bayes algorithm. Implemented algorithm in a structured and efficient manner, inspired by SQL, and managed to learn on a large dataset (of 20800 rows, each row containing 4544 words on average) in under 30 seconds.

- **Concepts/Skills used:** R Language, Statistical Natural Language Processing, Feature Selection using Mutual Information scores, Pre-processing text documents, Regex

Distributed Analysis of Big Data using Dask

- **Summary:** Collaborated with 2 peers to implement anomaly detection on a large industrial dataset (~5GB) using Dask. Developed strategy for anomaly detection and implemented algorithm in a Map-Reduce distributed approach with the help of a virtual cluster of three worker nodes.

- **Concepts/Skills used:** Big Data Analytics, Dask, Distributed Analysis, Map-Reduce paradigm, Critical Thinking

---

EDUCATION

Masters in Physics of Data — GPA (Tentative) : 28.1 / 30 — 2022 - 2024
University of Padua, Italy

B.Tech in Engineering Physics — GPA : 8.43 / 10, Graduated with Distinction — 2018 - 2022
National Institute of Technology, Calicut, India

CERTIFICATIONS AND SKILLS

- **Google Data Analytics Professional Certificate** (2022)
- **Language Proficiency in English:** IELTS Band 8, CEFR Level C1 (2022)

- **Languages:** English (native), Malayalam (native), Hindi (intermediate)
- **Programming Languages:** Python, R, SQL
- **Programming Frameworks:** Dask, PySpark, HTML, CSS, Bootstrap
- **Programming Libraries:** TensorFlow, Keras, PyTorch, Scikit-Learn, Pandas, NumPy, SciPy, Seaborn, Matplotlib