YADHUKRISHNAN PANKAJAKSHAN

krish99.work@gmail.com | krishpy99.github.io | linkedin.com/in/yk2310 | Boston, MA

EDUCATION

Master of Science in Computer Science | *Northeastern University* Aug. 2023 – May 2025
Coursework: Algorithms, MLOps, Intro to Data Science, Machine Learning, NLP *GPA: 3.92/4.00*

TECHNICAL SKILLS

Languages: Python, Go, C++, Javascript, SQL, NoSQL, GraphQL
Tools & Frameworks: Flask, FastAPI, Hadoop, PySpark, Hive, Linux, Git, Tensorflow, PyTorch
Data Science: Feature Engineering, Computer Vision, NLP, Fine-tuning, Machine Learning

PROJECTS

The Oracle: A QnA bot | Python, Streamlit *Mar. 2024 - Jun. 2024*
- Fine-tuned the deepset/roberta-base-squad2 model on a subset of the MS MARCO dataset (1,000 questions).
- Developed a few-shot query classification system to categorize incoming user queries based on semantic similarity.
- Tested the model using a set of 10K questions achieving good metrics while using **Exact Match** (EM) and F1 score.

Credit Card Approval Prediction | Python *Jan. 2024 – Apr. 2024*
- Built an ML pipeline for credit card approval prediction as part of coursework for DS5110. Trained a baseline **regression** model and compared with other models like **Random Forest** and **XGBoost**.
- Performed hyperparameter tuning using cross-validation to optimize model performance on XGBoost.

InkSight: A Handwritten Text Detection App | Python, OpenCV, React Native *Jul. 2020 – Mar. 2021*
- Developed a brand new architecture integrating **CTC** (Connectionist Temporal Classification) with traditional neural network models, resulting in a **6%** reduction in **levenshtein distance** and **13%** increase in word accuracy.
- Developed a mobile application that digitizes handwritten text images using a cloud server for text prediction.

Pneumonia Detection from Xray images | Flask, OpenCV, Tensorflow *Sep. 2019 – Oct. 2019*
- Uses the "Chest X-Ray Images (Pneumonia)" dataset, for prediction modelling as part of a Kaggle contest.
- Built a convolutional neural network (CNN) using TensorFlow, with adam optimizer and binary cross-entropy loss.
- Evaluated the model's performance with **ROC** curve and **F1** score. Achieved a place in top 10 submissions.

EXPERIENCE

Teaching Assistant | Python, C, Java *Jan. 2024 - Present*
Northeastern University
- Intro to Data Science - (Fall 2024), Algorithms - (Summer 2024), Programming - (Spring 2024)

Machine Learning Engineer | FastAPI, Docker, React *Jun. 2024 - Sep. 2024*
Northeastern University
- Worked on Paws, a RAG-based language model for the university with a projected userbase of **4000+** students.
- Orchestrated lean service containers to ensure speedy deployment and reduced startup times by **30 hours**.
- Integrated **FAISS** for efficient similarity search and retrieval in the context of large-scale document databases.
- Implemented a LlamaIndex retrieval pipeline for prompts, enhancing the overall relevance of responses by **65%**.

Software Engineer | React, Spring, Jenkins *Nov. 2022 – Aug. 2023*
Barclays
- Worked with the DevOps & internal applications team and developed **10+** pipelines & features for internal usage.
- Tested ETL pipelines resulting in a **15%** increase in system reliability and reduced costs by **$500 per week**.
- Spearheaded the development of **2** REST APIs for Document Management Platform, used by **100k+** customers.
- Wrote auto-scheduled pipelines for reporting that helped generate **6** business reports and saved **$330** every week.
- Hand-rolled **Logger**, a tool-extension to a Devops web application integrating **z/OS logs** using cURL commands.

Graduate Software Engineer | COBOL, JCL, Sonar *Sep. 2021 – Oct. 2022*
Barclays
- Developed cashback marking algorithms for Barclaycard, delivering **2** mainframe projects within **2 months**.
- Contributed to migration testing team and automated data validation workflows saving **$950** weekly.
- Innovated data processing algorithms for **100K+** transactions, reducing processing to **linear time complexity**.
- Integration tested developer tools and extensions used by peers, resulting in **12%** increase in peer commit activity.
- Collaborated in writing CI/CD pipelines for new features, reducing deployment times by **15%**.