Kartik Ullal

Boston, MA 02215 · (857) 334-3460
kartikullal99@gmail.com · www.linkedin.com/in/kartikullal · Portfolio Website

EDUCATION

Master's of Science, Data Science Aug 2023
Northeastern University, Khoury College of Computer Sciences, Boston CGPA: 3.83/4

- Relevant Courses: Linear Algebra and Statistics, Data Visualization and Analytics, Machine Learning, Natural Language Processing, Large Scale Parallel Processing, DBMS, Computer Vision

EXPERIENCE

Data Engineer Coop July 2022 - Dec 2022
Wayfair LLC, Boston, Massachusetts

- Developed a self-healing dedupe functionality for their data pipeline in Google BigQuery and integrated it with Slack and Email using Python, which increased Service Level Agreement (SLA) meetup from 86% to 92%.
- Implemented a Prophet-based time series forecasting model to detect anomalies and alert stakeholders when rates deviated by over 2 standard deviations, resulting in a 30% decrease in anomaly resolution time.
- Deployed ETL procedures, data models, and Looker dashboards for FinTech and Loyalty Strategic Programs Team, satisfying data requirements of 100+ stakeholders across channels. Improved data accessibility, accelerated report creation, and enhanced decision-making.

Data Science Intern June 2020 - Dec 2020
Prayaas Corps, Jaipur, Rajasthan

- Collaborated with the organization's technology team to build a database of student information, including demographic and academic data, resulting in a 20% increase in the accuracy of student records.
- Implemented an XGBoost algorithm to predict student academic performance, resulting in an 80% accuracy rate in identifying at-risk students and a 10% reduction in required additional support.

Data Science Intern June 2019 - Dec 2019
Navlakhi Education, Mumbai, Maharashtra

- Created and optimized a PostgreSQL database to store and manage data for JEE testing software and the Mumbai Fashion Academy website, including preprocessing of data using Python to ensure data quality and consistency.
- Developed a recommendation system for the Mumbai Fashion Academy using item-item collaborative filtering, allowing students to receive personalized course recommendations based on their course selections and interests.

PROJECTS

Resume Parser and Job Similarity Scoring| GitHub | Northeastern University Jan 2022 - May 2022

- Collaborated with engineers to parse resumes from PDF to JSON using PyMuPDF and built a Named Entity Recognizer with Spacy to identify skills from the resume and job description. Achieved an 85% accuracy score on the NER model.
- Utilized Gensim and Huggingface libraries with Doc2Vec and SentenceBERT to convert skills into embeddings and calculated similarity using Cosine Similarity. Achieved a 92% similarity score between my own resume and a data science job description.

Predicting Depression, Stress, and Anxiety Scores | GitHub | Northeastern University Sept 2021 - Dec 2021

- Coordinated with engineers to build an ML model predicting depression, stress, and anxiety scores from personality traits and demographics, and visualized different parameters to find a high correlation.
- Implemented a StackCVRegressor algorithm with Lasso, Ridge, ElasticNet, Gradient Boosting, LightGBM, and XGBoost regressions to predict scores, achieving an RMSE of 10.

Detection of Covid-19 using Chest X-rays | GitHub | Hack India Crisis Hackathon March 2020 - April 2020

- Collaborated with a team of four engineers to integrate, preprocess, and resize images from five different datasets to create a complete dataset of 17,194 chest X-ray samples.
- Develop a Covid-19 detection system using Dense Convolutional Neural Networks (CheXNet) on chest X-rays, achieving an accuracy of 93.496% on testing data.

TECHNICAL KNOWLEDGE

Programming Languages: R(ggplot2, dplyr), Python(Scikit-Learn, Numpy, Pandas, PyTorch, Tableau), SQL
Skills: Machine Learning, Deep Learning, Data Analytics, Natural Language Processing, Data Visualization (Tableau, Looker, PowerBI), Kronos, Google Cloud Platform, Excel, PyTorch, BigQuery