Pranshu Kumar

Boston, MA · pranshu1921@gmail.com · (617) 331-5469 · pranshu1921.github.io

Education

Northeastern University — Boston, MA
Masters of Science Analytics — Sep 2019 - May 2021

University of Petroleum and Energy Studies — Uttarakhand, India
Bachelor of Technology Computer Science — July 2015 - July 2019

Work Experience

Intel (Remote) — Dehradun, Uttarakhand, India
Summer Trainee — May 2018 - July 2018

- Successfully completed 'Intro to AI', a 8-week course covering ML , Deep Learning theory
- Successfully completed 'Deep Learning', a 12-week course covering ML algorithms, Neural Network architectures, convolutional networks, and recurrent networks
- Built a Python analytics project to predict 'FIFA 2018 World Cup's Best XI' team

Skills

Programming:	Python (pandas, numpy, scikit-learn, seaborn, nltk etc.), SQL, R
Supervised Machine Learning:	Supervised Learning (Linear/Logistic Regression, KNN, Naive Bayes, Decision Trees
Unsupervised Learning:	K-Means Clustering, DBSCAN, Recommendation Systems
Statistics:	Descriptive Statistics, Proportional Sampling, Regression Analysis, Inferential Statistics
Data Visualization:	Matplotlib, Tableau, Power BI
Deep Learning:	Natural Language Processing, Convolutional Neural Networks, Keras, Tensorflow

Projects

Personalized Cancer Diagnosis https://github.com/pranshu1921/Personalised-Cancer-Diagnosis
A machine learning case study to analyze Memorial Sloan Kettering Cancer Center (MSKCC) data for predicting the effect of genetic variations in the cancer tumors for enabling personalized medicine.

NYC Taxi Demand Prediction https://github.com/pranshu1921/Taxi-Demand-Prediction-NYC
A time-series forecasting and regression solution to find number of pickups, given location corrdinates(latitude and longitude) and time, in the surrounding regions, using data collected in Jan - Mar 2015 to predict the pickups in Jan - Mar 2016, provided by the NYC Taxi and Limousine Commission(TLC).

Facebook Friend Recommendation using Graph Mining
https://github.com/pranshu1921/Facebook-Friend-Recommenation-Graph-Mining
The project involves using data from the FacebookRecruiting challenge on Kaggle to predict missing links from a given directed social graph to recommend users.

Quora Question Pair Similarity https://github.com/pranshu1921/Quora-Question-Pair-Similarity
This is a binary classification problem for predicting whether questions pairs on Quora are duplicates or not. Used Natural Language Processing and Fuzzy Features for Advanced feature extraction. Compared Logistic Regression, Linear SVM, and XGBoost.

Amazon Fashion Discovery Engine
https://github.com/pranshu1921/Amazon-Fashion-Discovery-Engine
A content based recommendation engine for recommending apparel items or products at Amazon, using text and image data retreived from website. Suggested text based recommendations using Bag of Words (BoW), Word2Vec and TF-IDF techniques. Made image based recommendations using Convolutional Neural Network(CNN).

Amazon Fine Food Reviews https://github.com/pranshu1921/Amazon-Fine-Food-Reviews
A Natural Language Processing(NLP) based project that uses data provided on the 'Amazon Fine Food reviews' challenge posted on Kaggle, to determine the polarity of a given user review, following a score/rating of 4 or 5 considered positive, 1 or 2 negative and 3 neutral and ignored.