Praxal Patel

(516) 707-1006 | praxxals@gmail.com | linkedin.com/in/praxal

EDUCATION

New York University, Center For Data Science (GPA: 3.91/4.0) New York, NY
Master of Science in Data Science May 2021

- Relevant Coursework: Probability and Statistics, Machine Learning, Natural Language Understanding, Big Data, Deep Learning

Institute of Technology, Nirma University (GPA: 3.8/4.0) India
Bachelors of Technology in Computer Engineering May 2019

- Relevant Coursework: Deep Learning, Machine Learning, Calculus, Linear Algebra, Algorithms, Cloud Computing

WORK EXPERIENCE

Revelio Labs New York, New York
Data Scientist Jul 2021 - Present

- Employed an innovative approach to build company embeddings using Graph ML algorithms like Node2Vec.
- Working towards deploying a salary prediction model using Bayesian Additive Regression Trees.
- Deliver custom client requests by curating and structuring unstructured public profile and job postings data through the use of proprietary algorithms.

Coursera Mountain View, CA
Marketing Data Scientist Intern: ML Group Oct 2020 - Dec 2020

- Formulated a Health Dashboard to gauge continual performance of Coursera's Degree Recruitment Engine and built a system for automated hold-out sets for A/B/C tests
- Revised the Degree Forecasting model by leveraging a Bayesian approach to enhance prediction estimates of final application submitted for Coursera degrees

Wellness Space India
Data Analyst | Technical Consultant Jan 2019 - Jun 2019

- Engaged in studying Heart Rate Variability through 24-hours ECG signal data by developing SVM and LSTM models to predict a person's state of emotional arousal and analyzed key parameters associated with mental stress
- Deployed a self-monitoring application to help clients monitor stress levels and comprehend effect of various activities on heart rate; Performed anomaly detection on the ECG data for artefact correction

SKILLS & INTERESTS

- Tools & Technologies: Scikit-Learn, Tensorflow, Keras, PyTorch, Scipy, PySpark, NumPy, Pandas, Matplotlib, Plotly, GCP, AWS
- Programming Languages: Python, SQL, R, Java, C++, JavaScript, Hadoop, Spark, Hive

ACADEMIC PROJECTS

Impact of Career Development Services in Career Outcomes Sep 2020 - Dec 2020

- Quantified impact of NYU Wasserman Center for Career Development on students' actual career outcomes
- Designed a framework to determine most important career coaching activities by developing Logistic Regression (78% accuracy) and Decision Trees (87% accuracy) models having "Job Secure date" as target variable

Evaluating Summarization Tasks Using Sentence-BERT Mar 2020 - May 2020

- Devised a novel metric by employing sentence embeddings from SentenceBERT for evaluation of Summarization Tasks (Dataset- 226,711 BBC news articles)
- Validated SentenceBERT_Score correlates better to human evaluation than traditional evaluation metrics like ROUGE and BLEU

Diabetic Retinopathy Detection Using Deep Learning Jan 2018 - Dec 2018

- Identified Hard Exudates, Soft Exudates, Microaneurysm and Hemorrhages as key factors for detection; deployed CNN models to identify Retinopathy (Sensitivity-0.94) and leveraged the Ensemble Method to ascertain stage of Retinopathy

- Performed Data Augmentation and Image Segmentation leveraging Attention UNet; Conducted experiments for generalization on datasets Messidor (1200 images), DiaretDB0(130 images) and DiaretDB1(84 images)

Job Recommendation Engine Oct 2019 - Dec 2019

- Leveraged Natural Language Toolkit (NLTK 3.4) for extracting keywords about pertinent skills from job descriptions and resumes
- Developed a Recommendation Engine by computing Cosine Similarity and utilized Hit Ratio for evaluating performance of system- 46% hit rate increase over randomized recommendations for 100 recommendations

Optimal Location Prediction for Emergency Stations Jul 2018 - Jan 2019

- Formulated ML/DL Models to identify best suitable model for Travel Time Estimation; Travel Time estimated from XGBoost (RMSE: 42 seconds) was leveraged to drive K-Medoids for predicting optimal locations
- Decreased turnaround time along with reduced utility of resources; for Staten Island: average time cut down by 6 seconds utilizing 14 Fire Stations in comparison to 19 actual stations

POSITION OF RESPONSIBILITY

- Junior Data Scientist at NYU CDS responsible for classifying Neuro Typical and Atypical population under Prof. Alec Marantz
- Research Assistant at NYUAD - Understanding Indonesia's scorched earth policy using news and satellite data
- Teaching Fellow for Data Science Bootcamp at NYU STERN under Prof. Benjamin Zweig
- Junior Data Scientist under Prof. Angela Radulescu responsible for characterizing how humans visually represent objects and interpreted the difference between human representation and neural network representation