Parth Naik

2661 E 7th Street Apartment D, Bloomington, IN 47408

naikpa@iu.edu | LinkedIn: www.linkedin.com/in/parth-naik-10937a151 | GitHub: https://github.com/parthnaik13

EDUCATION

Master of Science in Data Science **May 2020**

Indiana University, Bloomington, IN **GPA 3.89/4.0**

Relevant courses: Applied Machine Learning (Logistic Regression, Decision Trees, SVMs, Naïve Bayes, Neural Networks), Statistics (Statistical Inference using R), Elements of AI (Search, HMMs, Ensemble Learning), Deep Learning (Auto Encoders, CNNs, RNNs, VAEs, GANs, Reinforcement Learning)

Bachelor's in Computer Engineering **August 2014 - May 2018**

Savitribai Phule Pune University, Maharashtra, India **First Class with Distinction**

WORK EXPERIENCE

Kabbage, Atlanta, GA **June 2019 - August 2019**

Data Science Intern

- Worked with the data science and strategy teams to model users' weekly cashflow to provide tailored suggestions.
- Filtered out target users based on the continuity and frequency of their transactions using Spark.
- Developed a deep learning model for quantile regression.
    - Model predicts 9 quantiles for the next 4 weeks given previous year history.
    - Implemented a hybrid CNN-LSTM model architecture.
    - Achieved a bias of 1.9% (a bias of 5% was decided upon to be acceptable).
- Collaborated with the strategy team to develop and test strategies for the downstream use of the forecasts.

Intel Indexer, San Francisco, CA **February 2018 - May 2018**

Machine Learning Intern

- Developed machine learning Python notebooks aimed at providing users with crucial information regarding tickers of various stock markets such as London Stock Exchange, Shanghai Exchange, National Stock Exchange of India.
    - These notebooks are currently being presented to potential customers and investors.
- Provided next day predicted prices, z threshold, Bollinger bands and wealth generated for test predictions to users for each ticker via these notebooks.
- Used neural networks (Multi-Layer Perceptrons) for making next day predictions for tickers.
- Calculated technical stock features (standard deviation, moving averages, on-balance volume etc.) for analytics.
- Coordinated with the development team members daily to tailor the analytics to customer needs.

PROJECTS

Connect4 game-playing AI using Reinforcement Learning **December 2019 - Current**

- Developed a Connect4 game environment using Python and OOP concepts.
- Working to develop a Deep Reinforcement Learning AI using epsilon-greedy approach and policy gradients.
- Training starts as 2 RL agents compete to win. The final objective is an unbeatable Connect4 AI.

Santander Product Recommendation Challenge **September 2018 - November 2018**

- Designing a recommendation model for bank customers given historical data was the challenge.
- Reduced the memory footprint of data by 88% by choosing optimal internal representations for the data.
- Applied Synthetic Minority Oversampling Technique (SMOTE) to avoid biasing the models due to class imbalance.
- Used and tuned XGBoost classifier for making predictions.
    - The XGBoost models not only had 90%+ accuracy on the cross-validation sets but also had very low log loss indicating high confidence in the predictions.
- Predicted for each customer the probabilities that the customer would purchase the offered 24 products.
- Suggested the top 7 products to be marketed to the customer based on these probabilities.

TECHNICAL SKILLS

- **Languages**
  Proficient in: Python, C++, SQL
  Familiar with: R, JAVA, JavaScript (JS), HTML, Flask, Spark
- **Databases:** MySQL (SQL), MongoDB (NoSQL)
- **Operating Systems:** Windows, Linux, MacOS
- **Python Data Science Libraries:** Numpy, Pandas, Matplotib, Scikit-Learn, TensorFlow, Keras