MRUGANK MILIND AKARTE

New York, NY 10027 (917) 900-8603
https://www.linkedin.com/in/mrugank-akarte/ mrugank.akarte@columbia.edu

EDUCATION

Columbia University New York, NY
Master of Science, Data Science Dec 2020
Relevant Courses: Machine Learning, Exploratory Data Analysis and Visualization, Probability Theory and Statistics, Statistical Inference, Algorithms.

Vishwakarma Institute of Technology Pune, India
Bachelor of Technology May 2018
Bachelor of Technology in Production Engineering, GPA: 9.44/10.00.

WORK EXPERIENCE

Ralph Lauren New York, USA
Data Scientist (Capstone Project) Sep 2020 - Present
- Analyzing historical product orders and returns data to reduce and predict propensity of product return in real time.

Nokia Bell Labs New Jersey, USA
Data Science Intern Jun 2020 - Aug 2020
- Devised an unsupervised CNN-LSTM autoencoder for anomaly detection and root cause(s) in multivariate time series.
- Model input were images of correlation matrices of different window lengths (short-medium-long term data) to capture spatial and temporal data simultaneously.
- Created model can be used to visually identify possible features responsible for anomalous behavior.
- Received Bell Labs summer intern award for outstanding innovation from president of Nokia Bell Labs.

Ellicium Solutions Pune, India
Data Science Intern Jan 2018 - May 2018
- Conducted data cleaning, manipulation and model testing for a customer retention project in Insurance domain.
- Analyzed imbalanced dataset with low churn records, large number of missing values for certain variables etc.
- Deployed logistic regression, SVM, Random Forest, Neural Networks, KNN and gradient boosting techniques to keep false positive rate within expected range required by client.
- Devised a submodule using R to capture sensory data from machines for real-time analysis.

LANGUAGE AND IT SKILLS

- Languages: Proficient in Python, SQL, R; Familiar with C++, Java, HTML.
- Packages: Tensorflow, numpy, pandas, keras, scikit-lern, dplyr, shiny, matplotlib, ggplot2, seaborn, plotly.

DATA SCIENCE PROJECTS

Columbia University, MNIST using evolution strategies Dec 2019 - Dec 2019
- Trained a NN utilizing gradient approximations and achieved 80% accuracy on MNIST dataset.

Columbia University, EDA on NATO Airstrikes during Kosovo war Dec 2019 - Dec 2019
- Created various visualizations using heat maps, time series plots, mosaic plots to understand effects of NATO airstrikes during war on people. Built D3 visualizations depicting killings and casualties.

Kaggle, Toxic Comments Classifier Mar 2018 - Mar 2018
- Developed a multi-headed model capable of detecting different types of toxicity in a sentence using RNNs. Achieved AUC-ROC of 0.978.

PAPERS

- Stochastic Flows and Geometric Optimization on the Orthogonal group, ICML 2020. Paper presents a new class of stochastic, geometrically driven optimization algorithm for optimization problem defined on orthogonal group.
- Predictive Maintenance of Air Pressure System using Boosting Trees: A Machine learning approach', ORSI, IIT Bombay 2018. Classifier with unequal class weights with a suitable threshold value achieves better performance compared to classifier using equal class weights with a threshold value of 0.5 in case of unbalanced dataset.