Jay Borkar

Tel: (732)322-4048 | Email : jayborkar24@gmail.com | Greater New York City Area, United States
LinkedIn : linkedin.com/in/jayborkar Github : github.com/jayborkar

EDUCATION

Rutgers, the State University of New Jersey | Rutgers Business School Sept 2017 – Dec 2018
MS, Information Technology – Data Science GPA: 3.6 / 4

Birla Institute of Technology, India GPA: 3.5 / 4 July 2013 – May 2017
BE, Computer Science & Engineering

TECHNICAL SKILLS

Languages : Python, R, Java, SQL, C, JavaScript, HTML, CSS, PHP
Data Science : Python, R, SQL, Hadoop, MapReduce, Spark, MS Excel, Tableau, TensorFlow, Keras
RDBMS : Oracle, MySQL
OS : Windows, Linux, MacOS
Cloud. : AWS (Amazon Web Services)
Courses : Algorithms & Data Structures, Data Mining, Machine Learning, Applied Artificial Intelligence, Problem Solving with data, Data Analysis & Visualization, Probability & Statistics, Software Engineering Web Application, Database System
Libraries : NumPy, Pandas, Matplotlib, Scikit-learn, SciPy, TensorFlow, Keras, Plotly, NLTK, Surprise, MLlib, ggplot2
Certifications : **IBM Data Science** certifications, Specialization in **Data Science** by John Hopkin University (Coursera)

EXPERIENCE

Data Scientist | Home Credit - Default Risk | Capstone Project | Prof. Saed Sayad **Sept 2018 – Present**
- Designed and Implemented Data Exploration routines to do Statistical analysis and visualization in Python.
- Built Classification models in Python and TensorFlow to predict if each applicant is capable of repaying loan or not.

Data Scientist Intern | John Wiley & Sons – Hoboken, NJ, USA **June 2018 – Aug 2018**
- Automated Book Titles search task by developing a tool that uses Machine learning/NLP to perform Fuzzy String Matching.
- Automated Data cleaning by developing a tool that uses Machine learning to predict duplicate records in the database.
- Implemented solutions in Python with Dedupe, Fuzzywuzzy libraries on MySQL database having more than 450,000 Titles.

Graduate Assistant | Prof. Casimir Kulikowski | Dept. of Computer Science - Rutgers University **Sept 2017 – May 2018**
- Appointed as Teaching Assistant for CS405- Computers, AI & Society course, Taught, reviewed, graded assignments & exams.

Schlumberger Oil Field Services – Mumbai, India | Software Engineer Intern **May 2017 – July 2017**
- Upgraded documents access & storage by creating a Web application for easy access and upload of documents.
- Architected and managed the MySQL database for the web application.

Emerson Process Management – Mumbai, India | Analyst Intern **Dec 2016 – Jan 2017**
- Accelerated & Improved decision making by analyzing Emerson & IT companies Integrated Operations initiative to harness the power of Internet of Things (IoT) Technology in the Oil and Gas industry.

Oil & Natural Gas Corporation (O.N.G.C) – Mumbai, India | Software Developer Intern **May 2015 – July 2015**
- Built intelligence by creating a Web application that gives accurate/relevant clauses for the processes in Offshore group.

PROJECTS

Web based Stock Forecaster | Python, TensorFlow, MySQL, HTML, CSS, JavaScript **Jan 2018 – Apr 2018**
- Predicted future stock market prices of S&P 500 companies by developing a Web based Stock Forecaster system.
- Built Deep Learning model – LSTM, Machine Learning model - Bayesian Curve in Python for Regression, Efficiency > 95%.
- Collected raw data by Alpha Vantage API in Python, Preprocessed the large data and stored in MySQL database.

Sentiment Analysis using Deep Learning on Amazon Product reviews | Python, TensorFlow, NLP **Mar 2018 – Apr 2018**
- Classified emotional tone of product reviews into positive / negative by creating a classifier that performs Sentiment Analysis.
- Built Deep Learning model – LSTM in TensorFlow and Python, Test Accuracy achieved is greater than 87%.
- Created IDs matrix and trained the model for the whole training Amazon dataset having more than 400,000 reviews.

Large-scale Data Analytics in Recommendation Systems | Python, PySpark **Mar 2018 – Apr 2018**
- Built a recommendation system with better accuracy on Amazon large dataset having more than 100,000 ratings.
- Predicted ratings, recommended items by Matrix-factorization algorithm (SVD), Alternating Least Squares(ALS) model.
- Implemented SVD, SVD++ with Surprise, a Python Scikit and ALS in MLlib using PySpark.

Crime Prediction of NYC & Chicago | R, R-studio, Python **Apr 2018 – May 2018**
- Preprocessed raw open data (crime) of NYC & Chicago from 2001 to present having 4,000,000 rows in Python and R.
- Performed Exploratory data analysis, Correlation analysis, and Multivariate Regression in Python and R.
- Predicted number of crimes in the future by Time-Series prediction models – ARIMA, ETS in R and Python.