Ravi Thej Neeli

(857) 210-7432 | ravithej080394@gmail.com |https://www.linkedin.com/in/ravi-thej-1bb11569/
GitHub portfolio: https://github.com/RaviThej0803

EDUCATION

Northeastern University, Boston, MA
Candidate for a Master of Science in Data Analytics ***Dec 2019***
Relevant courses: Predictive Modelling, Big-Data Systems and Intelligence Analytics, Applications of Artificial Intelligence, Data Visualization and Tools, Database Management Systems, Web Technologies

SASTRA University, Chennai, TN, India
Bachelor of Engineering in Computer Science and Engineering ***May 2015***

TECHNICAL KNOWLEDGE

Data Science related: Statistical Analysis, Machine Learning, Deep learning, Data Visualization, Data Mining, Data Pipeline
Languages: Python, R, Java, JavaScript, HTML
Frameworks / Platform's: NLTK, Flask, PySpark, PyCharm, Docker, Tableau, Power BI, Dash, MapReduce, Hadoop
Database: MySQL, SQL Server 2012, Oracle, PostgreSQL, Hive SQL
Cloud Platform: AWS Athena, AWS Sage Maker

PROFESSIONAL DEVELOPMENT

Stanford University, *Machine Learning (Coursera MOOC by Andrew Ng*) ***Apr 2019 - Present***
Udemy, *Machine Learning A-Z* ***Mar 2016***

WORK EXPERIENCE

Graduate Teaching and Research Assistantship – **Northeastern university, Boston, MA** ***Mar 2019 - Dec 2019***

- Guided graduate students by making them understand various unsupervised and Supervised learning Algorithms, **hyperparameter tuning** etc. in python
- Review weekly assessments and holding office hours for answering student queries

Data Engineer – **WIPRO Limited, Bangalore, India** ***May 2015 - Mar 2018***

- Performed exploratory data analysis by taking data in chunks and grouping the variables, thus reduced error percentage by 3.45%; utilized **data augmentation** to prevent overfitting.
- Trained classification models such as **SVM and logistic regression** etc., and was able to achieve 92% accuracy on unseen data, this feature led to increase in mobile downloads by 70% in 2018
- Implemented ETL process to extract data from OLTP sources into the data mart utilizing SSIS packages with **checksum** and timestamp as **incremental load technique**
- Worked with Pyspark notebooks in Databricks which includes writing complex queries, filtering, processing and reducing the **data frame and RDD**.
- Verified many ETL **Data Completeness, data transformation** scripts in order to assure the data quality from source to target.
- Developed **automated scripts** for some of the In-house mobile application mobile application features, to reduce manual intervention by 60 – 70% of the time.
- Involved on Test Case Design for the new features as well as updated the old **Accepted Test Protocols(ATP)**

ACADEMIC AND PERSONAL PROJECTS

PUMA – Location Based Analysis ***Sept 2019 – Dec 2019***

- Developed a **dash** application which forecasts their sales over a period using **prophet** in python
- Used different classifiers such as **Random Forest**, etc. to predict the transaction volume category for a state in a year
- Technologies Used: Python, Dash, Tableau, prophet

VIACOM – Ads Market Targeting | Experiential Learning ***Mar 2019 - May 2019***

- VIACOM sponsored project, developed interactive dashboards and **predicted estimated CPM rates** using different regressors in **KNIME** and Python
- Worked on **impressions conversion funnel** on Viacom's social demographic targeting data for better advertising

Mining Web Pages using NLP | Northeastern University, Boston ***Jan 2019 - Apr 2019***

- The objective is to summarize the text data present on the webpages, data from the webpages was scraped using pyquery package, cleansed and extracted the useful words from the **corpus** using NLTK package
- Used word embedding algorithms such as **Glove and word2vec** on the raw data to vectorize the data
- Preprocessed the text using **spacy** libraries