VIPANCHI KATTHULA

vkatth2@uic.edu | +1-(312)-918-9344 |Chicago, IL

linkedin.com/in/vipanchikatthula | github.com/VipanchiKatthula

EDUCATION

University of Illinois at Chicago (UIC), Chicago, IL **August 2019 – December 2020**

Master of Science in Business Analytics with specialization in Data Analytics **GPA: 4.0**

Coursework: Data Mining, Machine Learning, Descriptive Statistics, Big Data Analytics, Business data visualization, Healthcare Analytics, Advanced database management, Deep Learning and Text Analytics.

BITS PILANI, Hyderabad, India **Graduated: May 2017**

Bachelor of Engineering in Electrical and Electronics Engineering **GPA: 3.32**

PROFESSIONAL EXPERIENCE

University of Illinois at Chicago **August 2019 – May 2020**

Graduate Teaching Assistant

- Implemented Descriptive Statistical concepts in R programming in 5 lab sessions of 70 students each
- Aided students with Bayesian statistical concepts, techniques for multivariate regression, time series, and statistical process control. Performed all assistant duties, including mentoring, teaching, researching and clerical help.

UnitedHealth Group, Hyderabad **June 2018 – June 2019**

Decision Scientist

- Worked closely with clients, understood their pain points, performed exploratory data analysis in Python and integrated machine learning models into user platforms
- Designed and deployed Tableau decision dashboards for the compliance team of a major retailer to understand possible risks associated with their practices. Helped avert potential non-compliance and regulatory risks worth more than $2.0M
- Developed Big Data workflows to predict the delivery time of the orders created which improved the operational efficiency of warehousing systems by saving 20 manhours per day
- Created real-time analytical Tableau dashboards to provide insights on performance standards for non-technical end users

Data Engineering Analyst **(ETL & SQL developer)** **June 2017 – May 2018**

- Deployed interactive Reports, Dashboards, and Visualizations for the inventory management teams to help find the decrease in demand and to forecast the demand resulting in planned operations
- Developed data pipelines and workflows to extract the data from different source systems and loaded the required entities to target databases for reporting purposes using stored procedures
- Collaborated with global businesses and developed **SQL Server data warehouse** for inventory. Created and managed database objects including tables, indexes, and procedures
- Developed jobs and performed health check of complex SQL queries for filtering and transformation of the raw data for cleansing

ACADEMIC PROJECTS

Amazon's Twitter Sentiment Analysis: *(Python, Pandas, Numpy, Spark)* **February 2020**

- Analysed sentiment from 400k tweets on amazon's handle using Spark to predict if a customer is happy or dissatisfied with delivery service

Improving EMR utilization for Chicago Clinic: *(R, ggplot2, Supervised Machine Learning Concepts)* **January 2020**

- The goal is to improve the usage of a mobile application by patients. Developed models to find key parameters influencing the patient behaviour. Applied Machine Learning concepts to get the variable importance. **Random Forest, SVM (RBF Kernel)**

.

Text similarity between documents: *(Python, scikit-learn, nltk, Natural Language Processing Concepts)* **January 2020**

- Calculated similarity between documents using TF-IDF scores to categorize a set of documents into different buckets like News, Education, etc. Performed data pre-processing and applied NLP Similarity concepts to capture the resemblance

Digit Recognition: *(Python, Keras, Neural Networks, Unsupervised Machine Learning Concepts)* **January 2020**

- Developed Convolution Neural Network (CNN) model to predict hand-drawn digits from MNIST dataset while achieving an accuracy of 97.24%

Boosting customer engagement at VMWare: *(R, ggplot2, Supervised Machine Learning Concepts)* **October 2019**

- Built decision tree models for VMWare to improve customer engagement and predict the customer retention rate and achieved a recall of 91.8 %. Machine learning algorithms used: Random Forest, XGBoost, LASSO Regression, Ridge Regression

TECHNICAL SKILLS

- **Programming and Software**: R, Python (NumPy, Pandas, Sci-Kit Learn and Matplotlib), Pytorch, Keras, sklearn, Tableau, Git, Microsoft Office Suite, SQL (SQL Server Management Studio), Alteryx
- **Databases:** Hadoop, Oracle DB**,** Microsoft SQL Server and Netezza SQL
- **Analytics:** Excellent knowledge of inferential statistics, Data Mining, Linear programming, A/B Testing
  Machine learning algorithms - K-Nearest Neighbour, Linear regression, Logistic regression, Decision and regression Trees, XGBoost, Random Forest, SVM, LASSO, Ridge Regression