VIGNAN VENNAMPALLY

Boston, MA 02215 | vennampally.v@northeastern.edu |8574539100 | LinkedIn: Vignan Vennampally

EDUCATION

Northeastern University, Boston, USA May 2023
Masters in Data Analytics Engineering, (GPA: 4/4)
Course Work: Data Management for Analytics | Fundamentals of Data Analytics
Indian Institute of Information Technology, Jabalpur, India July 2020
Bachelors in Electronics and Communication Engineering, (GPA: 7.6/10)
Course Work: Introduction to Machine Learning | Business Analytics with R | Probability & Statistics

TECHNICAL SKILLS

Programming: Python (Pandas, NumPy, Scikit-learn, Scipy), R, SQL, Java
Databases: MySQL, MongoDB, Neo4j, Database Design
Data Visualization: Tableau, Power BI, Seaborn, Matplotlib
Data Engineering: ETL, Data Warehousing, Hadoop, Spark, MapReduce, Hive
Technologies & Operating Systems: Microsoft Azure, AWS, Microsoft Excel, PowerPoint, Linux OS
Machine Learning & Statistics: Data Mining, Classification, Regression, Clustering Algorithms, Tree-based models, Hypothesis Testing, A/B Testing

PROFESSIONAL EXPERIENCE

Software Engineer, Data Science Application| Ericsson India Global Services, Bangalore, India Aug 2020 - Aug 2021

- Engineered the configuration of POC, Sandbox, Production Instances of Data Science Studio application in Azure and Linux Environments.
- Developed efficient Linux bash scripts and macros in Python to automate the server and log management, Memory Management tasks that boosted the server performance by 20%.
- Resolved 100+ production issues of the application that efficiently increased the bug velocity.
- Worked on 10+ Machine Learning Use case projects starting from data collection to Model deployment.

Data Analyst Intern | IIT Kanpur, India May 2019 - Nov 2019

- Conducted Extensive Data Extraction from MySQL, MongoDB Databases through efficient SQL queries.
- Built an ML model to predict the dropout rates from MOOC courses. Improved the model performance to 94.5% and increased the retention rate of customers by 25%.
- Enhanced and Redesigned the Analytical Interface of mooKIT Platform using Python framework (Plotly Dash), Power BI and Tableau to gain statistical insights of 200 course participants

PROJECTS

Time Series Analysis of human activity monitoring data

- Analyzed walking, running, climbing down, climbing up data of 15 subjects to extract time series features.
- Developed Scatter Plots for all the activities using natural visibility graph(NVG) and horizontal visibility graph (HVG) methods
- Generated Scatter Plots for Permutation Entropy and Complexity for all the human activities.

Clustering of Synthetic and Real World Data

- Implemented K-means and hierarchical clustering methods on real world-data and evaluated the cluster quality using internal and external validation methods. Generated 2D, 3D, Scatter plots for the same data.

Keyword Network and Word Frequency Analysis

- Generated and analyzed a weighted network of keywords extracted from a document using node degree and strength. Analyzed word frequencies of 10 years twitter data on a yearly basis through histograms and implemented Zipf's law to plot log-log plots of word frequencies and rank for each year. Developed bigram network graphs for each year twitter data.

University - Student - Financial Enterprise Model

- Designed a database for the University Student Bank loan ecosystem which helps students and banks in making better decisions. Modeled ER, EER, UML and Relation Data Models for the database using concepts of Normalization, Specialization, Aggregation and Categorization.
- Integrated MySQL and Neo4j database with python to draw insights using Seaborn and Matplotlib libraries

Feedback Analysis of IIITDMJ

- Collected feedback from 1000+ fest participants from 50+ different events and preprocessed the data using NLP techniques.
- Classified data using Naïve Bayes classifier as positive and negative feedback to improve the facilities.