VIGNAN VENNAMPALLY

vennampally.v@northeastern.edu |+18574539100 | LinkedIn | GitHub

EDUCATION

Northeastern University Boston, USA
Master of Science in Data Analytics Engineering, (GPA: 4.0/4.0) Expected June 2023
Course Work: Machine Learning, Natural Language Processing, Deep Learning & Neural Networks, AWS Cloud

Indian Institute of Information Technology Jabalpur, India
Bachelors in Electronics and Communication Engineering, (GPA: 7.6/10.0) Jul 2020

PROFESSIONAL EXPERIENCE

Northeastern University Boston, Massachusetts
NLP Research Assistant Jan 2022 - Mar 2022

- Implemented **Question Answering System** on SQuAD2.0 **Text** data using **RoBERTa**, **ALBERT** Models that achieved F1 Score of 85%, Exact Match score of 77%. Conducted Comparative Study of BiDAF+Attention, **BERT**, **T5** Models.

Deep Learning Research Assistant Apr 2022 - Jun 2022

- Implemented **Face Detection mechanism** in **Image/Video** data using MTCNN, **OpenCV** libraries reducing False Positive predictions by 20% compared to Viola-Jones technique. Explored State-of-the-art algorithms like **YOLO**.

AstraZeneca Boston, Massachusetts
Data Scientist Jun 2022 - Dec 2022

- Leveraged **Machine Learning**, **Deep Learning** techniques using **Python (TensorFlow, Keras)** to identify 100 High Potential Customers of Koselugo therapy increasing Product Sales by 10%.
- Built **Predictive Modeling** techniques that achieved early diagnosis of 900 potential patients reducing Time to Treatment Initiation by 40%.
- Analyzed 1TB IQVIA LAAD Data in **Snowflake** using **SQL** Scripts to answer 50+ business questions and identify 18 Key Performance Indicators of Product Usage.
- Utilized model explainability techniques like **SHAP** and **LIME** that increased transparency & interpretability of ML models to **Sales, Marketing** Team reducing errors by 20%.

Ericsson Bangalore, India
Data Engineer Aug 2020 - Aug 2021

- Centralized ML & AI tasks by deploying a highly scalable **Dataiku** DSS application across 3 **AWS** Environments using **Amazon EKS** achieving 30% increase in Data Processing time.
- Leveraging **Amazon ECR** for model storage operationalized Development & Deployment of 2 Machine Learning models in DSS via **Docker** containers saving 100+ hours monthly.
- Improved **AWS EC2** performance by 15% through **Linux** Scripting that automated memory, log management tasks reducing manual intervention by 40%.

R3 Media Labs Kanpur, India
Data Scientist May 2019 - Jan 2020

- Developed XGBoost Model for **Customer Churn** that increased customer retention of mooKIT Platform by 15%.
- Analyzed 100GB of **Structured, Unstructured** Click-Stream data to identify User engagement strategies that increased **Active User** count of the platform by 10%.
- Improved Time-to-insight by 20% through redesigning dynamic visualization of 400 member data using **Tableau**.
- Conducted **A/B testing** to evaluate new changes to the platform increasing User engagement by 12%.

PROJECTS & RESEARCH EXPERIENCE

Movie Recommendation System - Recommendation System

- Developed a collaborative filtering movie **recommendation** system using IMDB data, achieving 80% accuracy in predicting user ratings and a MAP score of 0.6.

Weather Forecasting - Time Series Analysis

- Analyzed **Time series** air-quality data from the Beijing Municipal Environmental Monitoring Center and implemented forecasting models (ARIMA, LSTM, GRU) to predict 10 weather parameters with an accuracy of 93%

TECHNICAL SKILLS

Programming: Python (Pandas, NumPy, Scikit-learn, Scipy, OpenCV, TensorFlow, Keras, PyTorch), R, SQL, C
Data Engineering & Databases: ETL, Hadoop, Hive, Apache Spark, MySQL, MongoDB
Tools & Technologies: Azure, AWS, Linux, MATLAB, Power BI, Docker, Kubernetes, Git, Tableau, Excel, OpenCV
Modeling: Machine Learning (Clustering, Classification, Recommendation), Deep Learning, NLP, Statistical Modeling