

VIGNAN VENNAMPALLY

Boston, MA 02120 | vennampally.v@northeastern.edu | 8574539100 | [LinkedIn](#) | [GitHub](#)

EDUCATION

Northeastern University

Master of Science in Data Analytics Engineering, (GPA: 4/4)

Boston, USA

Expected Jun 2023

Course Work: Machine Learning for Engineering, NLP, Deep Learning & Neural Networks

Indian Institute of Information Technology

Bachelors in Electronics and Communication Engineering, (GPA: 7.6/10)

Jabalpur, India

July 2020

PROFESSIONAL EXPERIENCE

AstraZeneca

Boston, Massachusetts

Data Science Co-op

Jun 2022 - Dec 2022

- Increased KOSELUGO therapy usage by 10% through deep dive data analysis of 1TB Claims, Prescription data.
- Developed 100+ Advanced **SQL** Scripts to Identify different drivers of therapy initiation & Code business rules that accelerated potential patient identification from 50 to 300.
- Leveraged **Machine Learning** (Random Forest, XGBoost), **Deep Learning techniques** in Python to understand diagnosis combinations and identify target HCPs with 92% accuracy.
- Reduced dependency on third party service providers by 80% through developing inhouse Prescriptive and Predictive Modeling Capabilities.

Ericsson

Bangalore, India

Data Scientist

Aug 2020 - Aug 2021

- Centralized ML tasks that unify Business, AI & IT teams through engineering Data Science Studio Application in 3 different **AWS** Environments.
- Configured **Python, R, Spark** Data Processing, Machine Learning tasks in **Amazon EKS, EMR**.
- Boosted AWS EC2 Performance by 15% through developing **Linux** scripts and macros in Python that automated log management, memory management tasks.
- Operationalized 10+ **Machine Learning** projects in DSS environment saving 50+ hours of development time.

IIT Kanpur

Kanpur, India

Data Scientist

May 2019 - Nov 2019

- Partnered with Marketing, Sales team to understand business requirements and formulate hypothesis to improve Active User count by 25%.
- Extracted & Analyzed 100GB member Data from internal **MongoDB, MySQL** databases of 2 MOOC Courses.
- Developed **Supervised ML** model that predicted dropout rates of mooKIT Platform with 94.5% accuracy.
- Improvise Dynamic **Data Visualization** of 400 Member Data by redesigning analytical Interface using **Tableau**.

PROJECTS

Twitter Airline Sentiment Analysis - NLP Research Assistant

- Performed tweets extraction from 5 different airline tweets using **Twint** (Twitter Scraping Tool) to understand the consumer's voice.
- Implemented **NLP** techniques like **Stemming, Lemmatization and Vectorization** to clean & process the data.
- Identified top 10 reasons for Negative Sentiment & most common words from Negative tweets of each airline.

Face Detection in Image/Video - Deep Learning Research Assistant

- Tracked Infants breathing & sucking rates through 2 different Face Detection Methods that use concepts of integral Image, Attentional Cascade, Bounding boxes, Non-Maximum Suppression techniques.
- Reduced Number of False Positive (FP) predictions & Improved Model performance to 92% accuracy through multi-cascade Deep Learning Neural Networks.

Facebook Ads and Google Ads Data Analysis, Postman

- Implemented CRISP-DM methodology and performed **Exploratory Data Analysis (EDA)** leveraging Spends per Click metric to study the Ads campaign performance.
- Implemented **Descriptive, Inferential, and Predictive Modeling** and Analytics improved CLTV by 22 percent.

TECHNICAL SKILLS

Programming: Python (Pandas, NumPy, Scikit-learn, SciPy, Seaborn, Matplotlib, TensorFlow, Keras, PyTorch), R, SQL, C

Data Engineering & Databases: ETL, Hadoop, Hive, Apache Spark, MySQL, MongoDB

Tools & Technologies: Azure, AWS, Linux, Power BI, Docker, Kubernetes, Git, Tableau, Excel

Modeling: Machine Learning (Clustering, Classification, Recommendation), Deep Learning, NLP, Statistical Modeling,

Hypothesis Testing, A/B testing