Bindu Latha Banisetti
banisetti.b@northeastern.edu | 1-857-867-2699| linkedIn/bindu-latha-banisetti | bindubanisettish@gmail.com

EDUCATION

MASTER OF SCIENCE IN DATA SCIENCE — Boston, MA | Dec 2023
Northeastern University

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE ENGINEERING — Vijayawada, India | Mar 2013
V R Siddhartha Engineering College

Data Science skills: Python (Numpy, Pandas), Scikit-learn, TensorFlow, Pytorch, R, GCP, AWS, Git, SQL, PySpark

WORK EXPERIENCE

DATA SCIENTIST INTERN | Spare-it — Boston, MA | Jan 2023 – Aug 2023

Instance Segmentation using Computer Vision and Report Generation using GPT models for sustainable waste management

- **Object Detection and Segmentation using YOLO, U-NET and Mask R-CNN:** Performed Instance segmentation by fine-tuning pre-trained models like Circular-Net on waste management dataset. Developed Lambda functions using AWS CDK to facilitate data pipelines.
- **LLM based summary Report generation from segmented images:** Finetuned GPT-based foundation model for abstractive summarization by employing prompt engineering, and Langchain leveraging past statistical reports, and customer queries. Increased overall customer satisfaction by 40%.

DATA SCIENTIST | Datametica — Pune, India | Aug 2018 - Dec 2021

Automatic Ticket classification using Machine Learning and Deep Learning

- **Word embeddings:** Cleaned customer texts using Stemming, Lemmatization, tag removal and other methods, and created word vectors using TF-IDF, word2vec, FastText and Glove to create features for Supervised classification.
- **Machine Learning models:** Utilized Naïve Bayes, Logistic Regression, CatBoost, and XGBoost models to classify customer tickets into 5 levels of severity and urgency. Added Human in the loop to handle exceptions and improved accuracy to 85%.
- **Deep Learning models:** Implemented RNN and LSTM models for features with more than a year of data. Applied techniques like Regularization, Early stopping, and Skip connections and achieved an accuracy of 91% AUC.
- **Dashboards using Tableau:** Designed real-time dashboards to monitor trends and anomalies and track metrics and KPIs and communicated insights and recommendations to stakeholders.

Topic modeling to classify customer complaints from various feedback channels

- **Datasets and Text Analytics:** Collected customer concerns from customer support channel, social media, and audio transcripts and conducted Text analysis using word distributions, N-gram analysis, sparsity visualization, TF-IDF and other meta data.
- **Topic Modeling using Latent Dirichlet Allocation and Latent Semantic Analysis:** Performed Topic modeling with unsupervised classification algorithms like LSA, LDA and Non-negative matrix factorization that balances speed, scalability, and sensitivity.
- **Metrics and Visualization:** Used PCA for dimensionality reduction and t-SNE to visualize 2D and 3D interpretation of complaint clusters. Attained a perplexity score of 180 and a Topic coherence score of 0.87 after multiple iterations.

Proactive Churn Prediction: Predictive Analytics for Churn Reduction

- **Data wrangling and Feature Engineering:** Created data pre-processing pipelines to perform cleaning, outlier handling, normalization, and standardization, for customer features related to demographics, device type, plan type, billing methods and usage patterns.
- **Supervised Classification models:** Implemented Logistic Regression, Random Forest and XGBoost to classify potential churn profiles. Enhanced the model performance through Hyper-parameter tuning and Cross validation to achieve an F1 score of 93%.

DATA ANALYST | Capgemini — Hyderabad, India | Oct 2016 – Jul 2018

Predictive Analytics for augmenting Customer Experience for banking product (Finacle)

- **Anomaly Detection to reduce Financial Fraud:** Developed an outlier detection model to flag fraudulent and circular transactions using transaction details, user history, user profiles and other features. Overcame the drawback of cold start problems using K-means and enhanced the model using Isolation Forest to achieve an accuracy of 92%.
- **A/B Testing to enhance consumer experience:** Conducted various A/B tests to address complaints regarding Automated customer support, Dashboard design and response, security protocols etc. to ensure proper alignment of product with expectations and ad-hoc improvements. Improved the KPIs for engagement and satisfaction score from 65 to 90%.

SOFTWARE ENGINEER | Infosys — Hyderabad, India | Mar 2014 – Sep 2016

Optimizing Data Integration for real-time Analytics and Reporting

- **ETL Pipelines for Reports and Analytics:** Developed and fine-tuned ETL script to speed up data ingestion into data warehouses to support real-time and batch data processing for dynamic dashboards.
- **SQL Query optimization:** Optimized queries using techniques such as Indexing, caching, column pruning, sharding, and locking for swift data retrieval and reducing server loads. Improved database performance by 24%.

ACADEMIC PROJECT (MASTERS)

CREATIVE WRITING COLLABORATOR USING GPT-3 — Transformer, Large Language Models

- Developed a platform where writers can collaborate with GPT-3 to brainstorm ideas, refine plots, and overcome writer's block. Enabled the model to suggest plot twists, character developments, and dialogues.
- Enhanced the creativity and productivity of writers, as evidenced by a 25% increase in word output and positive feedback.

CERTIFICATIONS & AWARDS

- **Machine Learning Specialization** and **Deep Learning Specialization certification** from **Deeplearning.ai**.
- **Google Certified Professional Data Engineer**.
- Endowed with annual **StarAward** in Datametica for successful deployments following CI/CD pipelines.