SPARE-IT:

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

Image classification and Annotation using YOLO, U-NET and Mask R-CNN: Performed Instance segmentation and Annotations by fine-tuning pre-trained models like Circular-Net trained 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 using prompt engineering, and Fine-tuning techniques with Langchain on historical statistical reports, and customer queries. Increased overall customer satisfaction by 40%.

DATAMETICA

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 to find the best model that balances speed, scalability, flexibility, and parameter 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, that utilizes customer features related to Demographics, device type, plan type, billing methods, promotions, and usage patterns.

Machine Learning models: Developed supervised classification models like Logistic Regression, Random Forest and XGBoost to classify potential churn profiles to retain them with promotion policies. Enhanced the model performance through Hyper-parameter tuning and Cross validation to achieve an F1 score of 93%

CAPGEMINI

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. Performed data cleaning to handle missing values, abnormal trends, and other inconsistencies using EDA techniques. 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%

InfoSys

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%