Kunal Nehete

Data Scientist

+91- 9226716399
kunalnehete01@gmail.com

Showcased proficiency in creating and implementing Machine Learning models aimed at enhancing predictive analytics and optimizing various operational processes within the organization. Targeting opportunities in a dynamic environment to apply expertise in statistical modeling, advance data analysis, visualization and predictive modeling.

TECHNICAL SKILLS

- **Programming Languages:** Python (IDE: Jupyter notebook, VSCode), SQL
- **Data Analysis & Visualization:**
  - **Libraries:** NumPy, Pandas, Matplotlib, Seaborn, Plotly, Folium
  - **Tools:** Power BI, Tableau, Excel
- **Machine Learning:** Scikit-learn, Linear Regression, Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbours, Bagging, Boosting (AdaBoost, Gradient Boosting), XGBoost
- **Deep Learning:** Neural Networks, ANN, CNN, SGD, ADAM, Softmax, ReLU, Tensorflow, Keras
- **Text Processing:** NLTK, WordCloud, TF-IDF, Word2Vec, Bag of words
- **Time Series Analysis:** Decomposition, Durbin Watson test, Augmented Dickey Fuller test, pacf plot, acf plot, ARMA, ARIMA, SARIMA, FBProphet models
- **Statistics:** Descriptive Statistics, Inferential Statistics (including Hypothesis Testing such as T-test, Chi-square Test), ANOVA, Regression Analysis, Time Series Analysis, Multivariate Analysis, Principal Component Analysis (PCA)

PROFILE SUMMARY

- A Motivated Data Scientist with **3+** years of experience in **Retail & E-commerce**, **Transportation & Logistics** and **Construction & Engineering** domains. Recognized for assessing operational needs and developing solution models that have reduced client costs by up to **70%**, increased revenue by **52%**, and enhanced sales forecasting and customer segmentation for improved business strategies.

WORK EXPERIENCE

JDIT Business Solutions Pvt. Ltd., Pune
Since February 2022
Data Scientist

- Advanced knowledge of **ML** algorithms such as **Ensemble Models, AdaBoost, Grid Search, Random Forest, Decision Tree, SVM, Linear Regression, Ridge, Lasso, K-Fold Cross Validation,** and **KNN Imputer**.
- Proficient in developing **end-to-end Data Science pipelines**, including **Data Standardization, Feature Extraction, and Model Validation**.
- **Experienced in Feature Selection** using **Filter** Methods**, Wrapper** Methods**, Embedded** Methods, and **Feature Importance** techniques.
- Enhanced model accuracy through **Hyperparameter Tuning**.
- Skilled in **Hypothesis Testing,** Insights Generation**,** and **Root Cause Analysis**.
- Utilized **structured** and **unstructured data** to predict **retail trends**, perform **sales forecasting**, and enhance **customer segmentation** using **RFM analysis** and **Customer Lifetime Value (CLTV) modeling**.
- Extracted meaningful insights from text data, including **topics, sentiment analysis,** and **customer satisfaction** assessment.

PROJECTS

Retail Stores Sales Data Analysis and Sales Forecasting

- Investigated over **1 million sales records** from a **retail chain of 1,115 stores**, uncovering key **performance drivers**.
- **Discovered** top **10 profit-generating stores** and **sales trends** using a **rolling 30-month analysis** of **customer behavior, seasonality, promotions, and competition effects**.
- **Engineered time series forecasting models** (**ARIMA, SARIMA, Prophet, LSTM**) to **predict sales six weeks ahead**, achieving **98% accuracy** by incorporating **historical trends and external factors**.
- **Compared regression-based forecasting** (**Linear, Decision Tree, Random Forest, XGBoost**) with **time series models**, ensuring **optimal predictive performance**.

Customer Segmentation and Recommendation System for Online Retail

- Designed a customer segmentation and recommendation system for online retail customers using **RFM analysis**, **K-Means clustering**, and **PCA** (90% variance retained).
- Processed and categorized **4,234 customers over 1M+ transaction records** with **RFM segmentation**, integrating **Recency, Frequency, and Monetary (RFM) scores** to classify customers into **10 behavioural segments**.
- Constructed a **Customer Lifetime Value (CLTV) prediction model** using **BG/NBD and Gamma-Gamma models**, estimating future purchases and revenue.
- Forecasted **6-month and 12-month CLTV values**, enabling targeted marketing strategies.
- Delivered personalized top 3 **product recommendations** for **4,067 customers**, enhancing engagement and driving targeted marketing strategies.

Financial Analytics and Risk Prediction

- Evaluated a 52-feature financial data, **creating key profitability, liquidity, and risk metrics** for **company performance assessment**.

CORE COMPETENCIES

Retail

Project Management

Data Strategy

Reporting & Analytics

Data Processing and Analysis

Predictive analytics and modeling

Predictive Maintenance

Condition-Based Monitoring

EDUCATION

MTech. in Construction Management, Veermata Jijabai Technological Institute, 2018

B.E. in Civil Engineering, North Maharashtra University, 2016

RECENT CERTIFICATIONS

NPTEL certified -Introduction to Large Language Models

PERSONAL DETAILS

Date of Birth: 28th June 1995

Languages: English, Hindi, Marathi

Address: Pune, 411028

- Developed **financial forecasting, risk classification** (**99%** accuracy), and **anomaly detection** models, leveraging XGBoost, Gradient Boosting, and clustering techniques.
- Strengthened **credit scoring** & **fraud detection** using K-Means, DBSCAN, and Isolation Forest, improving risk-based **company segmentation**.

Predictive Maintenance and Condition-Based Monitoring (CBM) for Maritime Drive Systems

- **Explored 18 operational parameters** from a **ship's gas turbine (GT)** to construct **Predictive Maintenance & CBM** framework.
- **Researched advanced methodologies** by studying research papers- "A Deep Supervised Learning Approach for Condition-Based Maintenance of Naval Propulsion Systems" and "Machine Learning Approaches for Improving Condition-Based Maintenance of Naval Propulsion Plants", refining the dataset understanding, modeling approach, and feature selection.
- Built a **predictive model** with **99% accuracy**, improving performance by **2-3%** by integrating **polynomial features into PCA-generated components**.
- Resolved key challenges, including **100+ problematic values** (cleaned via regex)**,** handling **5+ multicollinear features**, and capturing **non-linear relationships** among **16+ independent** and **2 dependent variables**.

Airfare Trend Analysis & Tickets Price Prediction

- Built a predictive model to forecast air ticket prices for domestic flights in India, achieving high accuracy.
- Employed advanced data processing techniques to enhance model accuracy and performance.
- Optimized predictive performance by **comparing multiple models** and fine-tuning hyperparameters.