PRADEEP SHET

Hubballi, Karnataka India
+91-6363821470 | shetp727@gmail.com | LinkedIn | GitHub | Portfolio

Profile Summary

- Results-driven Data Scientist with hands-on experience in ML, data analysis to solve real-world business problems.
- Skilled in Python, SQL, and visualization tools like Power BI for deriving actionable insights from complex datasets.
- Proven ability to deliver data-driven solutions in domains like e-commerce, entertainment, and customer sentiment analysis.
- Proven track record in market research, requirement gathering, and both qualitative and quantitative analysis, resulting in optimized strategies and increased engagement.

Relevant Coursework

- Data Integrity
- Data Manipulation
- Machine Learning
- EDA
- Data Visualization
- Data Cleaning
- Data Modelling
- Business Impact Analysis

Experience

Trainity – Data Analyst Intern   **Jan 2024 – Mar 2024**

- Developed an optimized Power BI dashboard for efficient sales analysis and decision-making.
- Used SQL and Python to extract, preprocess, and analyses data, improving reporting efficiency.
- Identified top-performing categories, revenue trends, and payment methods to enhance sales strategies.
- Provided data-driven recommendations to optimize inventory, pricing, and operational efficiency.

Projects

Champagne Sales Forecasting | Python, Statsmodels, Matplotlib & Seaborn, Pandas & NumPy   **Feb 2025**

- Forecasted Perrin Frères Monthly Champagne Sales using time series analysis (ARIMA, SARIMA) to model trends and seasonality. Conducted time series decomposition, stationarity testing, and statistical diagnostics to optimize model accuracy. Delivered actionable insights for inventory planning, marketing, and supply chain optimization based on predictive analytics.

Restaurant Sentiment Analysis | Python, Flask, NLTK, Pandas, NumPy, Word Cloud, HTML/CSS.   **Mar 2025**

- Optimize Built a sentiment analysis model for restaurant reviews using NLP techniques and trained classifiers like Naive Bayes, Logistic Regression, and Random Forest. Preprocessed text data with NLTK, visualized trends using Word Cloud, and performed EDA with Matplotlib and Seaborn. Deployed the model using a Flask API with an interactive HTML/CSS-based web interface for real-time predictions.

Medical Insurance Cost Prediction | Machine Learning, Data Analysis, Python, Flask, HTML, CSS   **Mar 2025**

- Developed a machine learning model using Linear Regression to predict medical insurance costs based on factors like age, BMI, smoking habits, and region. Built a web interface with Flask for real-time cost predictions, and evaluated model performance using R-squared and accuracy scores. Implemented a user-friendly system for insurance providers to estimate premiums, optimizing pricing strategies and decision-making.

Movie Recommendation System | Python, Pandas, NumPy, Scikit-learn, HTML/CSS, Google Colab.   **Mar 2025**

- Developed a content-based movie recommendation system using cosine similarity to suggest movies based on genre, director, and popularity. Conducted EDA to analyze language trends, director popularity, movie status, and top-performing films. Built a Flask-based web app with an HTML/CSS frontend for interactive user recommendations..

Technical Skills

- **Analytical Tools, Languages:** Excel, Power BI, NLTK, ML, EDA, Data Modelling, Python, SQL, Google Colab, AWS
- **Skills:** Problem-Solving, Effective Communication, Decision-Making, Stakeholder Management, Analytical Thinking, Data Preparation, Statistical Analysis, Database Management, Business Intelligence.

Education

- MTech in *Computer Science and Engineering* (CSE) KLE Technological University, 2023-2025 – **CGPA: 9.11**
- B.E in *Computer Science and Engineering* (CSE) Jain College of Engineering and Technology, 2019-2023 – **CGPA: 8.9**

Achievement

- Secured **Top Performer** position in *Data Mine* during **Trajectory 2025**, the annual Techfest of Jadavpur University, demonstrating advanced analytical skills and strategic thinking.
- Secured **Top Performer** at *Trainity*, outperforming 1,000 candidates with an in-depth product rating analysis, showcasing expertise in data storytelling and visualization.