GAURAV PATIL

MACHINE LEARNING | DATA SCIENCE | PYTHON DEVELOPER

+91 9420714180 gppatil.work@gmail.com Linkedin Github Portfolio

Machine Learning, Deep Learning and Data Science enthusiast skilled in building end-to-end ML applications using Python, Scikit-learn, and Power BI. Experienced in data preprocessing, feature engineering, model development, and interactive deployment. Passionate about applying analytics to drive data-based decisions and improve business outcomes.

TECHNICAL SKILLS

- **Languages:** Python, SQL
- **Data Analysis:** Pandas, NumPy
- **Machine Learning:** Scikit-learn, XGBoost, CatBoost, DecisionTree, GradientBoosting, Feature Engineering
- **Deep Learning:** TensorFlow, Keras, Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN)
- **Visualization:** Power BI, DAX, Matplotlib, Seaborn
- **Deployment and Tools:** Flask, Streamlit, Docker, Git, GitHub

PROJECTS

STUDENT PERFORMANCE PREDICTION SYSTEM

CODE

Python | Scikit-learn | Flask | ETL

Developed an end-to-end machine learning application to predict **student academic performance** using demographic and academic features. The system incorporates a modular **ML pipeline** for **data preprocessing, model training, evaluation, and deployment.**

- **Comparative Model Evaluation:** Evaluated eight regression algorithms and achieved an **R2 score of 0.89** using **Linear Regression**, based on the project's reported results.
- **Modular ML Pipeline:** Implemented separate components for **data ingestion, data transformation, and model training** to organize the machine learning workflow.
- **Data Processing & Model Management:** Developed a structured workflow for preprocessing input data, training models, and storing trained model artifacts for inference.
- **Application Deployment & Error Handling:** Built a **Flask** application for predictions and implemented **logging** and custom exception handling to support debugging and error management.

BANK CUSTOMER CHURN PREDICTION

CODE

Python | TensorFlow | Keras | Scikit-learn | Streamlit | ANN

Built an **artificial neural network**-based machine learning application to **predict customer churn** using banking customer data. The project combines **data preprocessing, deep learning classification, and interactive visualization** to explore customer churn patterns.

- **Neural Network Development:** Designed an **Artificial Neural Network (ANN)** using **dense layers, ReLU activation** for hidden layers, and a **sigmoid output** layer for binary classification.
- **Data Preprocessing:** Applied **feature encoding** and **scaling** using **Pandas, NumPy,** and **Scikit-learn** to prepare customer data for model training.
- **Model Training & Optimization:** Configured the training process using the **Adam optimizer** and **binary cross-entropy loss** to classify customers according to their likelihood of churn.
- **Interactive Application Development:** Deployed a **Streamlit** application supporting CSV uploads, churn probability predictions, and visual exploration of customer churn patterns.

CAR PRICE PREDICTION

CODE

Python | Scikit-learn | Feature engineering | Hyperparameter tuning

Developed a machine learning **regression** model to predict car prices using vehicle-related attributes. The project focuses on **data preprocessing, comparative model evaluation, feature engineering**, and **hyperparameter optimization.**

- **Data Preprocessing & Feature Engineering:** Applied **label encoding, one-hot encoding, feature scaling,** and **feature engineering** to prepare input data for regression models.
- **Regression Model Comparison:** Implemented and compared **Random Forest, Ridge Regression, Lasso Regression, and K-Nearest Neighbors (KNN)** to evaluate predictive performance.
- **Hyperparameter Optimization:** Used **RandomizedSearchCV** to tune the Random Forest model and identify a suitable set of **hyperparameters.**
- **Model Performance Evaluation:** Achieved a reported **R2 score = 0.87** with the optimized **Random Forest model**, demonstrating the model's ability to explain variation in car prices on the evaluated data.

SPORTS MARKET SALES ANALYSIS

CODE

Power BI | DAX

Developed an interactive **Power BI dashboard** to analyze **sports market sales data** and identify revenue trends, profitability, and customer purchasing patterns. The dashboard provides **KPI visualizations** and sales breakdowns to support **data-driven business analysis.**

- **Sales Data Analysis:** Analyzed data covering **17,000+** customers to identify revenue, profit, and customer purchasing trends.
- **KPI Dashboard Development:** Created an interactive **Power BI dashboard** tracking **$24.9M** in revenue and **$10M** in profit through key performance indicators.
- **Sales Trend Analysis:** Visualized weekly **sales performance** to **monitor trends** and understand changes in sales over time.
- **Category-Wise Performance Analysis:** Compared sales performance across categories and generated **business insights** using **interactive visualizations.**

EXPERIENCE

PRODIGY INFOTECH PVT. LTD. *Jan 2025 – Feb 2025*

Position: Web Development Intern (remote)

EDUSKILLS *Oct 2024 – Dec 2024*

Position: Web Full Stack Development Intern (remote)

EDUCATION

Bachelor of Engineering - Information Technology *Pune, MH*

PES Modern College of Engineering *2022-2026*

Savitribai Phule Pune University

Overall CGPA: 8.1 | **4th YEAR SGPA: 8.97**

CERTIFICATIONS

- Introduction to Data Science
- Microsoft Power BI Desktop for Business Intelligence
- SQL for Data Analytics
- Power BI Workshop