Gaurav Sah

sahgaurav165@gmail.com | 9122685588

WORK EXPERIENCE

Data Science Internship

Learn and Build | 07/2023 - 08/2023 | Jaipur , Rajasthan

Work on two major project.

House Price Prediction

- Developed and deployed a house price prediction model in Python using numpy, pandas, seaborn, and matplotlib, achieving an R-squared (R^2) score of 0.033 and a MAE of 0.368 on the test set.
- Implemented advanced data preprocessing techniques including categorical encoding and numerical scaling for enhanced model accuracy in real estate prediction.
- Applied Linear Regression effectively to forecast house prices, demonstrating strong analytical and modeling skills in competitive market environments.

IPL Data Analysis:

- Conducted EDA on over 2,000 IPL player records using Python libraries (numpy, pandas, seaborn, matplotlib) for data manipulation, visualization, and analysis.
- Developed criteria for selecting 50+ top-performing batsmen, bowlers, and all-rounders, enhancing team composition and competitiveness.
- Implemented regression analysis to predict player performance with 85% accuracy, providing actionable insights for strategic player selection and team formation**.**

PROJECTS

Image Recognition System

- Developed a plant disease detection model using TensorFlow and Convolutional Neural Networks (CNN), achieving 96% accuracy, 95% recall, and a 95% F1 score.
- Implemented image preprocessing and augmentation techniques, processing over 10,000 images to enhance model robustness and performance.
- Deployed the model on a web app using Streamlit, facilitating real-time disease detection for improved agricultural management.

Credit Card Fraud Detection

- Engineered and deployed a Random Forest classifier in Python using scikit-learn, achieving a perfect accuracy score of 1.00 and maintaining a precision of 97%.
- Implemented StandardScaler to normalize data and handle duplicates, optimizing model performance and reducing processing time by 25%.
- Validated model robustness with an F1 score of 0.84 and an ROC AUC score of 0.87, demonstrating high precision and recall in distinguishing between legitimate and fraudulent transactions.

Movie Rating Prediction

- Developed a movie rating prediction model in Python utilizing numpy, pandas, seaborn, and matplotlib for data manipulation, visualization, and analysis.
- Implemented linear regression to predict movie ratings, achieving a Root Mean Squared Error (RMSE) of 3.21, demonstrating the model's accuracy in predicting ratings on a scale of 1 to 10.
- Employed OneHotEncoder to preprocess categorical variables, enhancing model performance and achieving a coefficient of determination (R-squared) of 0.75, indicating the proportion of variance in movie ratings explained by the model.

SKILLS

TECHNICAL SKILL

C/C++ | Python | SQL | HTML

TOOLS

POWER BI | EXECL | SQL WORKBENCH

EDUCATION

Jaipur Engineering college | Jaipur , Rajasthan

Bachelor of Technology in Computer Science | 07/2025