Suman Yadav
(318) 512-5501 | Monroe, LA | sumanydv615@gmail.com | github.com/Ydv-Suman | linkedin.com/in/suman-ydv/

Summary

Aspiring Data Scientist with strong foundations in Machine Learning, Data Analysis, and Web Development. Proven success in delivering accurate ML models and scalable web platforms. Passionate about solving real-world problems using data-driven solutions. Seeking internship/full-time opportunities to apply technical expertise and grow in a dynamic environment.

Education

University of Louisiana, Monroe, LA — Monroe, Louisiana
Bachelor of Computer Science, GPA: 3.5 — May 2027
Relevant Coursework: Data Structure and Algorithms, Internet Programming, Advanced Discrete Structures

Experience

Startups Realm Technology — Lalitpur, Nepal
Intern — May 2023 – Aug 2023

- Designed responsive user interfaces improving user engagement by 20%
- Developed adaptive user experience with HTML, CSS, and JavaScript across 10 platforms
- Maintained code integrity and facilitated collaboration using Git for 25 team members.

Extracurricular Activity

BookShare — April-2025

- Enabled 50+ daily book uploads by building a responsive book-sharing platform with AI-powered pricing suggestions for 1000+ users.
- Boosted reading engagement by 25% via a personalized recommendation engine and intuitive 'Read More' UI preview.
- Enhanced backend speed by optimizing database queries and efficiently managing 500+ book/user records.

Projects

Car Sales Price Prediction — June 2025

- Built an ML model (Random Forest) to predict car sale prices using structured vehicle data, achieving strong predictive accuracy.
- Automated the full preprocessing pipeline (missing values, encoding, scaling) with Pandas and scikit-learn.
- tuned hyperparameters via GridSearchCV and cross-validation, improving model F1-score by 18%
- Automated data cleaning and preprocessing using pandas and scikit-learn pipelines, reducing manual intervention.
- Engineered categorical features using OneHotEncoder and improved model robustness with GridSearchCV
- Delivered a reproducible ML workflow leveraging Python, boosting model maintainability and scalability.

Heart Diseases Project — June 2025

- Designed a medical ML classification system achieving 92% recall for detecting heart disease using patient clinical data such as age, blood pressure, cholesterol, and ECG results.
- Cleaned and analyzed data through EDA to enhance feature reliability; identified outliers and trends in ECG/cholesterol data.
- Benchmarked models (Logistic Regression, KNN, Random Forest); visualized results with ROC-AUC, confusion matrix, and SHAP.

Bulldozers Price Prediction — June 2025

- Developed a machine learning regression model to predict bulldozer resale prices using auction and equipment feature data, achieving 94%**+** R^2 accuracy**.**
- Engineered a regression model (Random Forest) to predict bulldozer resale prices with 94%+ R^2 accuracy on real auction data
- Conducted advanced feature engineering (e.g., machine age, sale timing, product category) boosting accuracy by 15%
- Built a robust preprocessing pipeline handling missing values, categorical encoding, and temporal transformations
- Leveraged visual analytics (Seaborn, Matplotlib) to uncover key price drivers, enhancing interpretability for stakeholders
- Tuned hyperparameters using GridSearchCV, achieving MAE within 10% of actual sale prices on validation set.

Skills

Languages: Java, Python, JavaScript, HTML, CSS
Libraries/Frameworks: React, Tailwind CSS, NumPy, Pandas, scikit-learn, Matplotlib, Seaborn
Tools: Git, VS Code, Jupyter Notebook, GitHub