Pranav Gupta

Pune, Maharashtra

+917588399775 | **pranavgupta922@gmail.com** | https://www.linkedin.com/in/pranavguptaji/ |

website: https://pranav-gupta-ji.github.io

Aspiring Data Scientist, AI / ML Engineer

Highly motivated **Data Science and AI professional** with a strong academic foundation in **Physics, Chemistry, and Mathematics (B.Sc.)**. Completed a **Data Science and AI Diploma** from the reputed institute **DataMites**, gaining hands-on experience in **machine learning, deep learning, and natural language processing (NLP)**. Successfully completed an **Internship as a Data Science Intern at Rubixe AI Solutions**, where I worked on real-world datasets, built predictive and analytical models, and contributed to business-driven AI solutions. Passionate about applying data-driven techniques to solve complex problems and deliver impactful insights.

Education

Datamites Global Training Institute

Certified Data Scientist, May 2025- Dec 2025

Sant Gadge Baba Amravati University

B.Sc. Physics, Chemistry & Mathematics 2021-2024 : 71%

Work history

Data Science Consultant, Intern | Rubixe (September 2025 – January 2025)

- Collaborated with the team members to understand client requirements and business challenges. Worked on a client project where machine learning algorithms like decision tree, xgboost and random forest algorithm were applied and achieved 95% test accuracy. Played a key role in the project team and ensured timely submission of the projects.

- Constructed predictive models using machine learning algorithms, including Logistic Regression and Multiple Linear Regression, achieving an accuracy rate of 85% in attributing revenue and conversions to specific marketing touchpoints

Skills

- Python – Numpy, Pandas,Tensorflow, Pytorch ,Sklearn data visualization- tableau, power bi, SQL, Flask, Machine learning algorithms , Time series forecasting, Data analysis ,Deep learning
- Linear Regression, Logistic Regression, Clustering Classification techniques
- Natural language processing (NLP), Jupyter notebook, Git,
- Applied statistics, Statistical analysis ,Ms excel
- Effective communication, Teamwork, Adaptability, Critical thinking, Problem solving, Time management.

Certifications: Certified Data Scientist (Datamites)

Projects

ITSM System Empowered with ML

Automatically assigns ticket priority and triggers email alerts upon ticket creation. Predicts Request for Change

(RFC) failure risk to enable proactive decision-making. Forecast incidents for effective resource management.

Skills applied:

1. Data Preprocessing steps handel null values, duplicate values and categorical encoding. Feature engineering created department columns and reduced dimensions.
2. Test Multiple model and select the best Random forest model, along with a forecasting model for resources management.
3. A FastAPI endpoint was created with an automated email alert to the department with respect to the ticket.
4. Streamlit frontend created.

Home Loan Defaulter Prediction

Based on the customer's application, the system predicts whether the customer is a defaulter or not. This serves as the first-level filtering step.

Skills applied :

1. Exploratory data analysis, and then does Data Preprocessing.
2. Tested multiple models and selected the EasyEnsembleClassifier Model as the best model and also applied threshold tuning. Applied custom sklearn pipeline
3. Created Fastapi endpoint for its prediction and Streamlit used for frontend creation.

Texas Employee Salary Prediction

A machine learning project that predicts Texas state employee salaries using agency, role, and demographic data. Built with scikit-learn, it achieves **83% accuracy ($R^2 = 0.83$)** and supports data-driven payroll forecasting and compensation analysis.

Skills applied:

1. Exploratory data analysis (EDA), data preprocessing and feature engineering.
2. Tested multiple model and selected the best decision tree regression model with **r2 score of 0.832**.
3. Created Fastapi endpoint for its prediction, and Streamlit was used for frontend creation.

Rice leaf disease prediction using CNN

Applied techniques such as resizing, normalization, and data augmentation to enhance model robustness. Employed techniques such as adjusting learning rates and dropout rates to enhance the model's ability to generalize. Transfer Learning is used to optimise to reduce errors.

Skills applied:

1. Data augmentation and normalization.
2. Created a convolution neural network and its multiple variants were tested, but not get best results compared to the mobilenet model of transfer learning because very less amount of data available.
3. Created Fastapi endpoint, from which we access model for prediction.

Cell-phone Range Classifier

A machine learning project that classifies mobile phones into four price segments (low to very high) using their technical specifications. It analyzes key features like RAM, battery, display, and processor to deliver accurate, data-driven price range predictions.

Skills applied:

1. Exploratory data analysis (EDA), data preprocessing and applied feature engineering.
2. Tested multiple models and final selected model is Logistic regression with 92% accuracy.
3. Created a pipeline to streamline transformation and ml model.
4. Created Fastapi endpoint for its prediction, and Streamlit used for frontend creation