Saurabh Vishwakarma

Chandigarh, India
+91-9695877472 sv9327545@gmail.com LinkedIn GitHub Portfolio

Projects

House Price Prediction Platform (End-to-End, Deployed) *Python, Scikit-learn, Flask, Pandas, NumPy, Render* *2025*

- Built an end-to-end regression system for house price prediction covering data preprocessing, feature engineering, model training, and inference.
- Performed exploratory data analysis, outlier detection, and error analysis to improve generalization across unseen samples.
- Deployed a production-ready web application on **Render**, exposing trained models via Flask with reproducible pipelines and documentation.

Language Modeling using GRU (From Scratch) *Python, PyTorch, GRU, Backpropagation Through Time* *2025*

- Implemented a **GRU-based language model from scratch** without using high-level PyTorch RNN abstractions.
- Manually defined update gate, reset gate, candidate hidden state, and hidden state transitions.
- Trained the model on textual data using custom training loops, PyTorch loss functions, and backpropagation through time.

Fake News Classification using Word Embeddings *Python, spaCy, Scikit-learn, Linear SVM 2024*

- Built a binary text classification system to detect fake vs real news articles using pre-trained **spaCy word embeddings**.
- Cleaned and analyzed a 22K-sample dataset; removed **8% duplicate samples** to prevent data leakage.
- Achieved **99% precision and recall** on a stratified test set using Linear SVM.
- Evaluated model performance using classification reports and confusion matrices to validate robustness.

Education

Chandigarh University *Aug 2023 – Jul 2026*
Bachelor of Computer Applications (BCA) – CGPA: 8.03 Mohali, India

IIT Madras *Aug 2023 – Jul 2027*
BS in Data Science and Applications – CGPA: 7.2 Online

Technical Skills

Programming & Libraries: Python, NumPy, Pandas, Scikit-learn, PyTorch, Tensorflow, spaCy
Machine Learning & NLP: Regression, Classification, Feature Engineering, Model Evaluation, Word Embeddings, Text Classification, Language Modeling, Sequence Modeling
Deep Learning: Neural Networks, CNNs, RNNs, GRU, Backpropagation Through Time
Deployment & Tools: Flask, FastAPI (basic), Git, GitHub, Linux, MLflow (basic)
Foundations: Probability, Statistics, Linear Algebra, Optimization

Certifications

- Foundation in Data Science – IIT Madras
- Advanced Python Programming – Coursera
- Generative AI Fundamentals – Coursera