

Harshvardhan Singh

Data Science Intern

Bihar, India

📞 +91 8294374503

✉ hv7275384@gmail.com

Summary

Highly motivated and results-driven Data Science student from IIT Madras with a strong foundation in Machine Learning (ML), EDA, and Python development. Proven ability to build end-to-end data pipelines, including Retrieval-Augmented Generation (RAG) systems and predictive models, seeking a challenging Data Science Internship to apply and expand technical expertise.

Education

Ongoing **BS in Data Science**, *Indian Institute of Technology (IIT) Madras*, Chennai, India
Currently building core competencies in advanced statistics, machine learning, and algorithm design.

Technical Skills

Programming	Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn), R
Machine Learning	Regression, Feature Engineering, Data Preprocessing, Model Optimization, EDA, RAG
Databases & Tools	JSON Structuring, Version Control (Git)
Soft Skills	Problem-Solving, Critical Thinking, Project Management, Teamwork
Languages	Hindi (Native), English (Fluent)

Projects

2025 **RAG System for Audio Transcription**, *Personal Project*

- Engineered a Retrieval-Augmented Generation (RAG) pipeline designed to transcribe audio files and output the results in a clear, structured JSON format.

- Implemented robust preprocessing techniques, a data chunking pipeline, and optimized retrieval logic to ensure high-quality and contextually relevant transcriptions.
- Demonstrated expertise in handling unstructured data and transforming it into usable, programmatic output.

2024 **House Price Prediction Model**, *Academic Project*
Accuracy: 83%

- Developed and fine-tuned a regression-based Machine Learning model to accurately predict housing prices.
- Conducted extensive Exploratory Data Analysis (EDA) to identify key features and relationships influencing price volatility.
- Performed advanced Feature Engineering and implemented rigorous model optimization techniques, resulting in a verifiable 83% prediction accuracy.

2024 **100 Days of Code**, *Personal Initiative*

- Successfully completed a 100-day dedicated coding challenge focused on deepening proficiency in Python programming.
- Applied concepts across various disciplines, including ML workflows, data manipulation with Pandas/NumPy, and advanced data visualization.