

PRANAY SONI

Aspiring Data Scientist & AI Engineer | B.Tech – Artificial Intelligence & Data Science

+91 9530370070 | spranay205@gmail.com | www.linkedin.com/in/pranaysoni301 |

<https://github.com/PranaySoni-56> | Chittorgarh, Rajasthan, India

PROFESSIONAL SUMMARY

Results-driven AI and Data Science undergraduate (CGPA: 9/10) with hands-on experience in Machine Learning, NLP, Deep Learning, and cloud-based deployments. Proven ability to build end-to-end ML pipelines, automate workflows using Docker and AWS EC2, and deliver data-driven insights through Power BI and Python. Experienced in Agile/Scrum environments and cross-functional stakeholder collaboration. Passionate about solving real-world business problems through advanced analytics and scalable AI solutions. Actively seeking a Data Science or Data Analyst internship or entry-level role.

TECHNICAL SKILLS

Programming Languages	Python, SQL (PostgreSQL)
ML / AI Frameworks	Agentic AI, LLM, Fine-tuning, RAG, Gen AI, Scikit-Learn, TensorFlow, Keras, PyTorch, XGBoost, Random Forest, BERT, Sentence-BERT
Data & Analytics	Pandas, NumPy, Matplotlib, Seaborn, EDA, Statistical Analysis, Feature Engineering
NLP & Deep Learning	NLP, TF-IDF, Sentiment Analysis, CNN, RNN, LSTM, Autoencoders, RBM
Big Data & Cloud	PySpark, AWS EC2, Docker, Cron Jobs
Visualization Tools	Power BI, Microsoft Excel
Developer Tools	Git, GitHub, Jupyter Notebook, VS Code
Methodologies	Agile, Scrum, Data-Driven Decision Making, Cross-Functional Collaboration

PROFESSIONAL EXPERIENCE

Machine Learning & Data Analytics Intern | MyShishu – Shishupuram Parenting Innovation Pvt. Ltd.

April 2025 – June 2025

- Led development of a Facebook Page Analytics Dashboard using **Facebook Graph API** to extract real-time engagement data; delivered actionable marketing insights across 10,000+ page interactions.
- Built **ML models (Random Forest, XGBoost)** to predict user engagement metrics with >85% accuracy; applied **NLP (TF-IDF, Naive Bayes)** for sentiment analysis on 5,000+ user comments.
- Automated the full data pipeline using **Docker** containers and Cron Jobs for 6-hour scheduled refresh; deployed on **AWS EC2** for scalable production performance, eliminating 100% of manual data tasks.
- Collaborated with cross-functional stakeholders in an **Agile/Scrum** environment to convert business requirements into data-driven solutions and presented findings to non-technical management.

PROJECTS

AI Recruiter Resume Screening System | github.com/PranaySoni-56

- Engineered an **NLP** pipeline using **TF-IDF** and **Sentence-BERT** embeddings with cosine similarity for semantic resume-to-job-description matching; achieved **>90% accuracy** on real-world datasets.

AI-Powered Fraud Risk Prediction System | github.com/PranaySoni-56

- Designed an **ML pipeline** for financial fraud detection using **Random Forest** and **XGBoost**; applied **SMOTE** for class imbalance handling and achieved **97.8% accuracy** with high precision-recall performance.

Image Classification using CNN | github.com/PranaySoni-56

- Implemented advanced **CNN** with **TensorFlow/Keras** for multi-class image recognition; optimized using data augmentation, dropout regularization, and hyperparameter tuning.

Google Stock Price Prediction using LSTM | github.com/PranaySoni-56

- Built an **LSTM**-based time-series forecasting model for stock price prediction using sequence modeling and dropout regularization; achieved **RMSE of 14.86** on test data.

DeepMovieRec – Movie Recommendation System | github.com/PranaySoni-56

- Implemented a deep learning recommendation system using **Stacked Autoencoders** for user-movie rating prediction; achieved **94% prediction accuracy** through unsupervised feature learning and reconstruction optimization.

EDA Case Studies (Zomato, Home Loan, IMDB, Customer Segmentation) | github.com/PranaySoni-56

- Performed comprehensive **EDA** and predictive modeling using **PCA**, **K-Means**, Hierarchical Clustering, and association rule mining (**Apriori**, **FP-Growth**) across multiple real-world datasets.

Heart Disease Prediction & House Price Prediction | github.com/PranaySoni-56

- Built classification and regression models using **Logistic Regression**, **Gradient Boosting**, and GridSearchCV; evaluated performance using R^2 , MAE, and RMSE metrics.

EDUCATION

B.Tech – Artificial Intelligence & Data Science

Sep 2023 – June 2027

Bikaner Technical University, Bikaner, Rajasthan | **CGPA: 9 / 10**

Higher Secondary Education (Science)

2023

Swami Vivekanand Govt. Model School, Begun, Rajasthan

CERTIFICATIONS

- AI Engineer Core Track: LLM Engineering, RAG, QLORA, Agents - Udemy
- Machine Learning – Udemy
- Deep Learning – Udemy
- Natural Language Processing – Udemy
- PySpark – Udemy
- SQL Bootcamp – Udemy
- Power BI – Udemy
- Python Programming – Udemy
- Microsoft Excel – Udemy
- Statistics for Data Science and Business Analysis - Udemy