

Mrunal Sonekar

mrunalsonekar28@gmail.com 8104691694

Summary

Data scientist with around 1.4 years of experience working on real business problems. I work with Python, SQL, Pandas, and scikit-learn to clean data, spot trends, and build models that help teams make better decisions. I enjoy turning messy data into clear insights and am always looking to keep learning and improving.

Skills

- **Programming:** Python (Pandas, NumPy, scikit-learn), SQL
- **Machine Learning:** Linear and Logistic Regression, Random Forest, Hyperparameter Tuning (GridSearchCV)
- **Deep Learning:** Neural Networks using TensorFlow and Keras
- **Tools & Environments:** Jupyter Notebook, Google Colab, Git
- **Data Visualization:** Power BI, Seaborn, Matplotlib
- **Soft Skills:** Strategic decision-making, cross-functional communication, data-driven problem solving, collaborative project execution

Education

G.H. Raisoni University, Madhya Pradesh, Bachelor of Technology (B.Tech),
Computer Science and Engineering August 2019 – July 2023
• CGPA: 7.87

Experience

1.Designation: **Data Scientist**, SimplyMoveIn Technologies Pvt. Ltd., Bangalore, August 2023 – December
India 2024

- Spearheaded the implementation of a pricing forecast model using Python and scikit-learn, improving prediction accuracy by 18% using location and amenity data.
- Boosted model performance by 15% and reduced prediction errors by 22% through robust data preprocessing, including feature scaling and outlier treatment.
- Applied Linear Regression and Random Forest algorithms with hyperparameter tuning using GridSearchCV to optimize model performance.
- Created interactive data visualizations with Matplotlib to uncover key pricing patterns and inform strategic decisions.
- Improved property booking performance through an 18% RMSE reduction in pricing forecast models.

2.Data Science Intern, Rubixe AI Solution February 2024 – August
Internship Certificate 2024

- Executed end-to-end data workflows, including data cleaning, exploratory data analysis (EDA), and machine learning model development using Python (Pandas, NumPy, scikit-learn), resulting in insights that supported strategic decision-making.
- Initiated machine learning algorithms for POC projects, achieving a model accuracy of 85% and optimizing decision-making processes.
- Automated data pipelines, reducing data preparation time by 20%, and ensured high-quality datasets for analysis.
- Collaborated with cross-functional teams to translate business requirements into deployable machine learning solutions, earning recognition for technical proficiency and a solution-driven approach.

Projects

Ficzon IT Solution Client Project

- Conducted statistical analysis and data manipulation on 7,000+ records to generate actionable insights and reports.
- Engineered a lead classification model using SMOTE to address data imbalance, boosting prediction accuracy by 17% and improving lead conversion outcomes.
- Generated visual reports using Python libraries (Matplotlib, Seaborn) to highlight sales performance and regional trends for stakeholder review.

Rent Prediction

- Engineered a rental price prediction model using Linear Regression and Gradient Boosting with K-Fold cross-validation, achieving an accuracy of 85%.
- Applied advanced preprocessing techniques, including outlier handling, feature scaling, and label encoding, resulting in a 22% error reduction.
- Visualized model performance metrics (RMSE, MAE, R²) using Matplotlib and Seaborn, driving a 12% increase in property bookings.

ANN for Image Classification

- Built a deep learning model using TensorFlow and Keras to classify Fashion MNIST images, achieving 89% validation accuracy through normalized inputs.
- Tuned hyperparameters, introduced dropout layers, and optimized architecture to improve model generalization and reduce overfitting.
- Designed and fine-tuned an Artificial Neural Network (ANN) pipeline with dropout and normalization strategies to enhance generalization and reduce overfitting in image classification tasks.

Certifications

Certified Data Scientist IABAC

Data Science Certification NASSCOM