

# Manish Kumar Mahto

[Manishcode123@gmail.com](mailto:Manishcode123@gmail.com) | (+91) 7667056502 | Ranchi, Jharkhand, India

 [Github Profile](#)

 [Portfolio website](#)

 [Linkedin Profile](#)

## SKILLS

- ❖ **Programming Languages:** Python, Java
- ❖ **Libraries:** Numpy, Pandas, Matplotlib, Seaborn, Streamlit
- ❖ **Tools/Technologies:** SQL, Excel, Git, Github, Jupyter Notebooks
- ❖ **Data Science Skills:** Data Cleaning, EDA, Feature Engineering, Model Building, Statistical Analysis, Machine Learning, NLP, ANN, CNN, RNN, LSTM, GRU
- ❖ **Mathematics:** Statistics and Probability, Linear Algebra
- ❖ **Soft Skills:** Problem-Solving, Analytical Thinking, Collaboration, Time Management

## ACADEMIC PROJECTS

- ❖ **Movie Recommendation System** ([Github link](#))
  - **Developed a Content-Based Movie Recommendation System:** Achieved **75%** memory efficiency improvement using text vectorization and cosine similarity for personalized recommendations.
  - **Optimized and Deployed:** Created an interactive web app with **Streamlit**, showcasing movie details and recommendations with a user-friendly interface.
  - **Tools Used:** Python, Pandas, Scikit-learn, Streamlit, and Google Colab.
- ❖ **Mobile Price Predictor** ([Github link](#))
  - **Data Processing & EDA:** Cleaned and transformed a dataset of over **1,000** mobile models, engineered features, and performed exploratory data analysis with visualizations including **10+** charts and plots to uncover insights.
  - **Model Development:** Built and evaluated **5** regression models (**Linear, Ridge, Decision Tree, Random Forest, XGBoost**) with **R<sup>2</sup> scores** up to **81.57%**.
  - **Tools Used:** Python, Pandas, NumPy, Matplotlib, Seaborn, scikit-learn, XGBoost.
- ❖ **Quora Duplicate Question Pair** ([Github link](#))
  - Developed a **machine learning** model to classify duplicate question pairs from the **Quora dataset**, achieving an accuracy of **78%** using a **Random Forest Classifier**.
  - Engineered **custom features** such as common word ratios and implemented advanced **text processing techniques**, including fuzzy string matching and **text vectorization**.
  - **Tools used:** Numpy, Pandas, Matplotlib, Seaborn, Contractions, Fuzzywuzzy, nltk, Gensim, Scikit-learn, Xgboost.

## EDUCATION

- Ranchi University, Ranchi, Bsc(IT) SGPA: **8.46** | (October 2022 – June 2025)
- XII (Jharkhand Academic Council) **80.8%** | 2022
- X (Jharkhand Academic Council) **89.2%** | 2020