Ravi Kumar

Email | +91 7355524696
LinkedIN | GitHub Portfolio Medium

Education

Indian Institute of Technology Jodhpur Graduation date: May 2023
M.Sc. in Mathematics ***CGPA: 7.77/10***

University of Lucknow Graduation date: Sep 2020
B.Sc. in Physics, Chemistry, and Mathematics ***CGPA: 8.0/10***

Skills

Languages: Python, C++, SQL, HTML, CSS, Latex

Libraries: PyTorch, TensorFlow, openai, Pandas, Numpy, Matplotlib, Scikit-Learn, etc.

Technologies & Tools: Docker, MySQL, AWS EC2, S3, Streamlit, VS Code, LaTex, Google Colaboratory, Jupyter Notebook, Flask, Git, DVC, MLFlow, DagsHub, Gemini API, GitHub Action Server (CI/CD/CD), CNN, NLP, etc.

Academic Courses: Programming Techniques, Machine Learning, Optimization, Financial Engineering, Computer Graphics, Deep Learning, Data Structures and Algorithms.

Work experience

Data Science Intern *ineuron.ai* *(Mar 2024 - present)*

- Employed advanced machine learning algorithms to forecast insurance premium prices, achieving an R^2 score of 0.88 and a Mean Absolute Error (MAE) of 2446.15 units.
- Implemented and evaluated over 9 regression models, including **XGBoost**, **CatBoost**, **Random Forest**, **Gradient Boost**, **LightGBM**, and **Linear Regression** and selected top 4 models based on performance over the training set.
- Orchestrated 10+ experiments and data pipelines with DagsHub.
- Monitored models using **MLFlow**, managed data pipelines with **DVC** on **DagsHub**, containerized application with **Docker**, and deployed them on **AWS EC2**, achieving an inference time 3 to 5 seconds. GitHub, Video, Docs

Projects

X-ray Image Classification for Pneumonia Detection *(Feb. 2024 - Mar. 2024)*

- Constructed a high-accuracy medical image classification model using advanced deep learning techniques, achieving 88.33% accuracy on the test set.
- Adopted Streamlit, **Docker**, AWS EC2, S3, **PyTorch**, CNN, resulting in a prediction latency of 3 to 5 seconds and improving scalability and deployment efficiency through CI/CD practices. GitHub

Python Database Automation Toolkit: dbautomate *(Dec. 2023 - Feb. 2024)*

- Developed the **dbautomate** Python package, streamlining data upload, storage, and deletion. Achieved a 50% reduction in data handling errors and cutting manual data management time by 50%, increasing efficiency for users.
- Utilized Python, **MySQL**, **MongoDB**, and GitHub Actions, resulting in a streamlined installation process with *pip install dbautomate==0.5.6*. GitHub, PyPI , Docs.

Movie Recommendation System *(Sep. 2023 - Oct. 2023)*

- Enhanced recommendation system accuracy by integrating 90% content-based filtering with 10% collaborative filtering, optimizing a dataset of over 10,000 IMDb and OMDb movies to increase user retention.
- Achieved a recommendation inference latency of 5-7 seconds by leveraging Flask, Azure, HTML, CSS, Bootstrap, JavaScript, and **Cosine Similarity** to enhance system efficiency and user satisfaction. Github, Demo, Blog

Stock Price Prediction *(July 2023 - Sep. 2023)*

- Implemented a data-fetching mechanism for real-time analysis, processing 10 years of historical stock price data.
- Obtained a 119.57 RMSE score on the test dataset and forecasted the next 150 days using **TensorFlow**, **LSTM**, time series analysis, machine learning, optimization, and model evaluation. GitHub.

Publication

- **Dynamics and Chaos Control of the Deformed K Map** *IIT Jodhpur* *(Jan 2023-May 2023)*
  Aishwaraya, **Kumar, R.**, Chandramouli, V.V.M.S. (2024). In: Singh, J., Anastassiou, G.A., Baleanu, D., Kumar, D. (eds) Advances in Mathematical Modelling, Applied Analysis and Computation. ICMMAAC 2023. Lecture Notes in Networks and Systems, vol 953. Springer, Cham. link.