SURBHIT KUMAR

+91-9038946209

surbhit3812@gmail.com ◇ https://linkedin.com/in/surbhit-kumar ◇ http://surbhit01.github.io

WORK EXPERIENCE

MAVENIR SYSTEMS, Bengaluru - Member of Technical Staff-2 Aug '24 - Present
ML Engineering Team

- Developed **regression models** using **Random Forest** and **XGBoost algorithm** to predict CPU and Memory utilization of system
- Achieved **R2 score of 0.89** on the testing dataset by performing **hyperparamter tuning** on the models
- Conducted **drift analysis** using statistical methods to detect data shifts and maintain model robustness
- **Handled skewed feature distributions** using multiple data transformation techniques to improve the model
- Worked on pipelines for data aggregation, preprocessing and model training to streamline the modelling process
- Worked on the deployment of machine learning models and services using **docker** and **kubernetes**
- Built **RESTful APIs** for several services of the application with robust error handling using **python**

WIPRO TECHNOLOGIES, Bengaluru - Senior Project Engineer (AI/ML) Sept '20 - Aug '24
Lab45 ai360 Platform Development

- Implemented models like **LSTM** and **Random Forest** for predicting device performance with 78% accuracy
- Worked on the integration of multiple **Large Language Models (LLM)** in the platform for multiple use cases
- Developed proof-of-concepts to implement **text summarization** methods tailored for different use cases
- Led the development of **content analysis** and question-answering features from documents by leveraging **LangChain**
- Integrated **Pinecone DB** for efficient storage of **document vector embeddings** thus enhancing **information retrieval**
- Performed in-depth **statistical analysis** of 20+ variables to understand their impact on the device functioning

PROJECTS

Movie Recommendation System (Project Link)

- Built a **collaborative filtering movie recommendation system** based on the MovieLens-100k dataset.
- Implemented **K-Nearest Neighbors** to identify similar users and generate personalized movie recommendation.
- Created a feature matrix based approach that ranked recommendations based on normalized distance scores.

Product Review Analysis (Project Link)

- Implemented a **phrase sentiment analysis** model to predict the corresponding rating given a product review
- **Finetuned BERT** by training on 15000+ reviews with rating as the target variable (1-5) achieving **82% accuracy**

English to Spanish Machine Translation (Project Link)

- Developed a **machine translation** model utilizing state-of-the-art techniques in Natural Language Processing
- Trained 10000 pairs of english spanish parallel sentences to build an **encoder decoder LSTM** using **tensorflow**
- Implemented **attention mechanism** and **beam search** to enhance the model's translation performance

SKILLS

Areas of interest	Data Science, Machine Learning, NLP, Deep Learning, Computer Vision
Tools/Framework	SKlearn, Tensorflow, Pandas, Numpy, LangChain, NLTK, Spacy, Docker
Languages	Python, SQL, C#

EDUCATION

Manipal Institute of Technology, Manipal Bachelor of Technology, Instrumentation and Control *2016 - 2020*