Aftab Mallick

Bengaluru, Karnataka,India • aftabmallick007@gmail.com • 8100458986 • linkedin.com/in/aftab-mallick

github.com/Aftabmallick • aftabmallick.github.io/ml_portfolio

Experienced ML/DL Engineer skilled in training and developing advanced **Machine Learning** and **NLP** models to deliver impactful solutions for diverse projects. Proficient in **TensorFlow, Machine Learning, Deep Learning, Computer Vision,** and **NLP**.

EDUCATION

B.Tech In Computer Science and Engineering(Data Science)   **Sep 2020 - Jun 2024**
MCKV Instuitute of Engineering   **GPA:** 8.96

Coursework:

Machine Learning Techniques: Supervised Learning (Regression, Classification), Unsupervised Learning (Clustering, Dimensionality Reduction), Deep Learning (CNN, RNN).

Other Areas: Natural Language Processing (NLP), Computer Vision, Reinforcement Learning, and Time Series Analysis.

Tools & Frameworks: Apache Spark, TensorFlow, PyTorch, Keras, Scikit-learn, Jupyter Notebook, AWS SageMaker, Google Cloud AI Platform, Microsoft Azure Machine Learning.

Data Science Skills: Data Cleaning, Feature Engineering, Data Wrangling, Data Visualization, Data Mining.

SKILLS

Programming Languages: Python, Java, C ,Javascript

Python Libraries: Pytorch, Tensorflow/Keras, Scikit-learn, Numpy, Pandas, Matplotlib, Seaborn, Langchain, NLTK, Spacy

Deep Learning: Multi-layered Perceptrons/Neural Networks, Computer Vision(OpenCV, CNNs, Image Similarity, Image Classification, Object Detection, Siamese Networks, Generative Models & GANs),

Natural Language Processing: Text Representation, Language Modeling, RNNs, LSTM, NER, Attention, Transformers, BERT

Supervised Machine Learning: Classification & Regression, Decision Trees, Random Forest, GBDT, XGBoost, Time Series

Unsupervised Machine Learning: K-Means, DBSCAN, GMMs, Anomaly Detection, Recommender Systems

MLOps : MLflow, Flask, FastApi, Dockers, AWS

PROJECTS

KIDNEY DISEASE CLASSIFICATION WITH CICD ON AWS 🔗   **Apr 2024 - May 2024**

- Utilized Kaggle's kidney image dataset to train a **VGG16 CNN** model, employing **OpenCV** for image processing.
- Employed **MLflow** to track all ML experiments conducted during hyperparameter tuning and selected the best parameters .
- Implemented **DVC** for creating a machine learning pipeline, streamlining data ingestion, preprocessing and model training.
- Developed a **Flask** server for predictions and implemented **CI/CD** for continuous integration and deployment on **AWS EC2**.

RAG BASED QUESTION ANSWERING SYSTEM 🔗   **Apr 2024 - Apr 2024**

- Created a **RAG** chain using OpenAI GPT 3.5 turbo model and **langchain** framework.
- Converted PDF / Website data in text format, then utilized **OpenAI text embeddings** to convert the text into vectors.
- Uploaded the vectors to a cloud database called **Pinecone**, enabling easy access from any location.
- Retrieve answers for user questions using cosine similarity.

End to End Bengaluru House Price Prediction Model 🔗   **Mar 2024 - Mar 2024**

- Developed a **machine learning** model using **Python** to predict house prices in Bengaluru, India.
- Utilized various algorithms (**Linear Regression, Lasso Regression, Decision Tree**) through **GridSearchCV** for optimal model selection.
- Achieved an accuracy score of **89%** demonstrating model effectiveness.
- Designed and implemented the web application using **Flask** and **HTML/CSS/JavaScript** for precise price predictions.

CHATBOT FOR FOOD DELIVERY APPLICATION 🔗   **Mar 2024 - Mar 2024**

- Developed a **Dialogflow**-powered food delivery chatbot, potentially increasing user engagement by 30%.
- Implemented custom entities within **Dialogflow**, achieving 60% accuracy in food item recognition.
- Integrated **FastAPI** backend, reducing order processing time by 80%.
- Utilized **MySQL** database for efficient order data, tracking, and menu management.

WORK EXPERIENCE

Business and Data Analytics Intern   **Jul 2023 - Aug 2023**
Ybi Foundation   **Remote**

- Engineered **Artificial Neural Network (ANN)** models achieving an average accuracy of **90%** in predicting telecom customer churn, resulting in a significant reduction in customer attrition rate and substantial revenue preservation.
- Implemented **artificial neural network** solutions for banking and telecom churn forecasting, achieving a 15% increase in accuracy compared to traditional methods and enabling prompt customer retention actions.
- Analyzed financial news sentiment using **Random Forest** and **NLP** techniques, leading to a 15% increase in the accuracy of market trend predictions and providing actionable insights for investment decisions.