PRAJWAL K

Portfolio | Github | prajwalgbdr03@gmail.com | +91 81976 72995 | LinkedIn | Blogs | Bengaluru

PROFESSIONAL EXPERIENCE

INeuron.Ai - *Data Scientist | Intern*   Jan 2023 – Feb 2024
- Spearheaded an End-to-end **MLOps Predictive Maintenance Project**, achieving **98%+** performance across two classification models.
- Leveraged Python libraries (Pandas, Numpy, Scipy, Sklearn) for data manipulation and analysis. Implemented MLflow for seamless ML lifecycle management, Utilized Evidently AI for proactive model issue resolution, ensuring optimal performance and continuous enhancement.
- Expertly deployed models utilizing cutting-edge technologies (Docker, Streamlit, GitHub Actions) through Hugging Face spaces integration within CI/CD pipeline.
- Ensured data quality and relevance through expert cleaning, transformation, thorough Exploratory Data Analysis, and Model Monitoring.

Omdena Inc - *Data Scientist | Freelance*   Mar 2024 – Ongoing
- The project's core objective is to advance sustainable business practices through the creation of an AI-driven tool tailored for automating climate disclosure reporting. This endeavor is pivotal in facilitating businesses to effectively fulfill regulatory obligations by harnessing the power of AI and large language models (LLMs) to simplify the drafting process of reports.

PROJECTS

Consignment Pricing Prediction *(Numpy, Pandas, Scikit-Learn, XGBoost, Random Forest)*   *Github*
- Built an end-to-end machine learning pipeline, using MLOps & DVC technology, to predict the price of consignment items, and achieved a Mean Absolute Error of 3.2%.
- Implemented MLOps practices to ensure robustness, and reproducibility throughout the model development lifecycle, enhancing the project's efficiency and maintainability.
- Achieved an impressive **95%** accuracy by effectively tackling a regression problem using **Random Forest Regressor.**

Onsite Health Diagnostics – OHD *(CNN, Transfer Learning, Scikit-Learn, Tensorflow, Hugging Face)*   *Github*
- Achieved the primary project goal of integrating diverse disease prediction models into a unified website, ensuring each model maintains **>90%** accuracy metrics.
- Employed transfer machine learning techniques including Resnet50, Vgg16, Alexnet, and XgBoost to construct a robust multi-classification and regression models.
- Engineered a web application utilizing Streamlit, deploying a comprehensive Docker application via seamless integration with Hugging Face through CI/CD pipeline.

Jigsaw - Multilingual Toxic Comment Classification *(Tensorflow, HuggingFace, Transformers)*   *Github*
- Built a multilingual text classification model to predict the probability that a comment is toxic based on **4,35,775 text comments** in **7 different languages**, using the data provided by Google Jigsaw.
- Fine-tuned **BERT-Multilingual-base and XLMRoBERTa** models on the multilingual text.

SKILLS

Core Skills: Machine Learning, MLOps, Deep Learning, Computer Vision, Natural Language Processing, Time Series.

Tech Stack: Python, MySql, Tensorflow, Keras, Python (eg. scikit-learn, numpy, pandas, matplotlib), Computer vision, Natural Language Processing

Miscellaneous Technologies: A/B testing, ETL, Data science pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, Time series, Hypothesis testing, Transfer Learning, OOP, APIs, Git, Github actions, AWS, Dockers.

EDUCATION

INeuron Intelligence, Full Stack Data Science Masters   Apr 2023 – Jun 2024
Relevant Coursework: Machine Learning, Data Science Statistics, Data Analytics, Probability & Discrete Mathematics

Singhania University, Bachelor of Science in Aviation   Jun 2019 - Sep 2022
Cumulative GPA: **8.5/10**

COURSEWORK

- Machine Learning Specialization by Andrew Ng - Stanford University - **DeepLearning.AI**
- Mathematics for Machine Learning and Data Science by Luis Serrano - **DeepLearning.AI**
- Python for Data Science and Machine Learning - Jose Portilla - **Pierian Training**