+91 8525880083
Aurangabad, Maharashtra
maria.ayush@gmail.com

Ayush Maria

DOB: 28/04/1998
github.com/AyushMaria
linkedin.com/in/ayushmaria

Education

Computational and Software Techniques MSc (Computational Intelligence for Data Analytics) **09/2021 — 09-2022**
Cranfield University *Bedford, United Kingdom*
Bachelor of Technology in Computer Science **07/2016 — 07/2020**
Vellore Institute of Technology *Tamil Nadu, India*

Technical Skills

Languages: Python (Pandas, Numpy, Matplotlib, Seaborn, Sci-Kit Learn, PyTorch, Keras, Flask, NLTK), Firebase, AWS, PowerBI
Core Skills: Supervised Learning, Exploratory Data Analysis, Feature Engineering, Boosting Algorithms, BERT, GloVe

Work Experience

Data Scientist)| Instilmotion Labs | Hyderabad, Telangana, India **04/2023 — present**
- Incorporating Object Avoidance/Detection in drones using Deep Learning frameworks like YOLOv3
- Implementing data analysis on drone data for market research using Python and PowerBI
- Designing software for Automatic Tray Retrieval System (ATRS) deployed at Rajiv Gandhi International Airport, Hyderabad using ReactJS, Streamlit, OpenCV and Docker
- Automating drones using AI based algorithms like A*, D* to navigate waypoint missions

Skills Acquired: Python, Software Development, Firebase, Deep Learning, React, Docker, Streamlit

Software Developer - Data Analysis(Freelance Role) | Tactalyse | Groningen, Netherlands (Remote) **08/2022 — 04/2023**
- Quantified football player performance by implementing data analysis on football related data by using Pandas, Matplotlib
- Successfully explained metrics through graph visualisations for football players by generating personalised reports using libraries Tkinkter, Fpdf
- Deployed the PDF generating software online to scale the software using Flask, Openpyxl, Werkzeug and GIT
- Implemented project related KPIs by tracking the work for the Software Engineers using Google Sheets

Skills Acquired: Project Management, Software Development, CI/CD, GIT

Data Handling Intern | Mahyco | Jalna, Maharashtra, India **06/2018 — 07/2018**
- Designed ETL pipelines for Data Warehousing using SQL
- Performed data wrangling to transform raw dataset to relevant values using Pandas Data Frames
- Classified different types of crops by rating them on a scale of 1-10 measured by features like health and type using XGBoost Models with 95% accuracy
- Explained results to team members to highlight important trends in data using visualisation through Plotly

Skills Acquired: Data Visualization, Machine Learning Classification, Data Narration, Critical Thinking

Projects

Natural Language Processing (NLP) with Disaster Tweets *GitHub Link* **07/2022 — 08-2022**
- Examined and transformed a raw dataset containing tweets from Twitter to understand indications of disaster along with Stemming, Stop words removal, TF-IDF rescaling for feature engineering using Pandas, Seaborn and NLTK
- Inspected and compared predictions made by Long Short Term Memory (LSTM) Neural Network (82% accuracy) and BERT Transformer (80% accuracy)

Analysing Expected Goals (xG) in Football *GitHub Link* **05/2022 — 08/2022**
- Extracted football event data from Statsbomb open data repository using Beautiful Soup and Python Requests, Tranformed raw JSON dataset into Pandas Dataframes by using mathematical measures to wrangle the data, Loaded the Data to Google Drive
- Compared chances created by footballers measured by Probability based metrics resulting in player profiling and improvement of transfer market investment using Seaborn for Data Visualization and Logistic Regression, XGBoost, CatBoost and LightGBM Classifiers based on Log Loss and F1 score for Machine Learning

Vortex Core Detection *GitHub Link* **03/2022 — 05/2022**
- Executed Data Wrangling, Exploratory Data Analysis and Feature Engineering using Seaborn and Sci-Kit Learn to successfully engineer a Hard Voting Classifier based on Random Forest, XGBoost and AdaBoost Models
- Inspected trends in Vortex Data and predicted vortex core locations validated using Confusion Matrices and AUC with 97.15% accuracy

COVID19 - Time Series Analysis *GitHub Link* **11/2021 — 01/2022**
- Performed Data Wrangling to Transform Data using Pandas and Spark
- Studied the performance of naive K-means Clustering Algorithm with regards to Spark ML-lib to identify difference in computational performance, Designated states in the United States with similar trends of COVID cases using K-means clustering with K=4

Interests and Activities

Activities: Football, Cooking, Basketball, Dancer (B-boying, Hip-Hop), Music, Gaming, Anime, Science Fiction
Languages: English, Hindi, Elementary Spanish