Omkar Firame

+919404350615 | [E-Mail](#) | [Linkedin](#) | [GitHub](#)

SUMMARY:

Data Scientist with strong Machine Learning foundation and 3+ years of experience in developing robust ML models using predictive data modeling, analyze data to deliver insights and implement action-oriented solutions to complex business problems.

TECHNICAL SKILLS:

Python, R, Machine Learning, Deep Learning, PyTorch, Iguazio, KubeFlow Pipeline, Azure Databricks, Dask, Web Scraping (Selenium), Git,

EXPERIENCE:

Data Scientist - Wynum Automation Services Pvt. Ltd | April 2023 - Present
- Managed a project implementing an information retrieval system for extracting relevant press releases from online sources.
- Applied advanced data preprocessing techniques to cleanse and structure content, enhancing categorization and analytical capabilities.
- Developed custom web scrapers using Python (Beautiful Soup, Requests) for precise and targeted content extraction. Initiated an independent project to automate DOCX file generation from text data, showcasing proactive and innovative problem-solving.

Data Scientist - Tata Consultancy Services | Feb 2021 - March 2023
- Spearheaded end-to-end data initiatives involving data extraction, transformation, and loading from Azure Data Lake for streamlined processes in Data Preparation, Cleaning, and Model Building.
- Engineered and implemented robust pipelines, utilizing KubeFlow, a custom Manual pipeline, and the mlrun MLOps orchestration framework on the Iguazio platform, ensuring seamless automation on a scheduled basis.
- Orchestrated data transformations to meet API requirements, systematically schematizing datasets for enhanced compatibility and integration. Applied Dask distributed computing for efficient handling of large-scale datasets, optimizing data processing and analysis workflows.

Intern - Thoughtworks Technologies India Pvt. Ltd | Jan 2020 - Sept - 2020
- Pioneered the application of machine learning in the medical domain, focusing on predicting drug properties based on chemical composition.
- Automated raw data processing with Python scripts and key libraries (chemoPy, PyBioMed, Modred), streamlining data preparation for predictive modeling.
- Demonstrated expertise in drug toxicity prediction, securing 6th position in the Tox21 challenge (2014-15) with an outstanding average AUC score of 0.8347.

PROJECTS:

Text Summarization - [[Link](#)]
- Developed a Text Summarization personal project, leveraging Streamlit for an intuitive user interface that enhances user experience and interaction.
- Implemented a modularized structure, integrating the T5 small model for efficient and effective text summarization, showcasing a systematic and organized approach to project development.

EDUCATION -

M.TECH - Mathematical Modeling And Simulation - Centre for Modelling and Simulation, SPPU, Pune | 2020

B.E - Mechanical Engineering - Bharti Vidyapeeth's College of Engineering, Kolhapur | | 2016