Harshitha Pydimalla

Software Engineer

harshithacharis3@gmail.com • +44 7342887936 • Sheffield, England

linkedin.com/in/pydimalla-harshitha-365849193 • github.com/Harshithap37

Professional Summary

Master's graduate in Artificial Intelligence with strong foundations in software development, machine learning, and data engineering, seeking to contribute to the Atlassian Williams Racing Technology & Innovation Industrial Placement. Skilled in Python, C++, and C#, with hands-on experience in AI-driven systems, data analysis, and building scalable software tools. Passionate about applying innovation and problem-solving to support high-performance engineering, streamline enterprise systems, and deliver impactful technology that drives Formula 1 performance.

Skills

Programming Languages & Tools: Python, C++, Java, JavaScript, C#, React.js, SQL, HTML, CSS, NumPy, Pandas, Jupyter Notebook, Google Colab
Data Science & Machine Learning: Machine Learning, Data Science, Mathematics, Model Training, Model Evaluation, Feature Engineering, Natural Language Processing, Data Cleaning, Statistical Analysis
Statistical & Time Series Analysis: ARIMA, SARIMA
Data Visualization: Matplotlib, Seaborn
Software Development: Data Structures & Algorithms, Object-Oriented Programming, Git, GitHub, API/JSON Integration
Databases: MySQL, SQL, Hadoop, ELP/PLM familiarity
Other Skills: Problem-solving, Microsoft Tools, Network Performance Analysis, Scalable Machine Learning, Text Processing
Soft Skills: Problem-solving, Analytical Thinking, Communication, Team Collaboration, Adaptability, Attention to Detail, Time Management, Multitasking

Projects

Web Traffic Time Series Forecasting 02/2025 – 05/2025

- Developed time series forecasting models to predict future pageviews of Wikipedia articles across multilingual languages.
- Used SARIMA model with log-transformation and seasonal differencing to capture weekly and yearly patterns.
- Applied chronological train/validation/test split to ensure robust evaluation and prevent data leakage.

Tools: Python, Pandas, Statistical Models (SARIMA), Matplotlib, NumPy, seaborn, scikit-learn

Analysis of Physical and Mental Health Issues due to excessive screen time in youngsters 01/2023 – 04/2023

- This project analyzes the effects of excessive screen time on the physical and mental health of individuals aged 16-2 using survey data from 685 students.

- After cleaning and encoding the data, statistical tests identified significant links between screen time and health issues. A Random Forest model was developed to predict the negative impact of screen time over four hours, achieving 100% accuracy.
- Findings reveal that extended screen use is associated with headaches, sleep problems, back pain, and decreased physical activity, highlighting the need for balanced digital habits to support better health.

Tools: Survey & Data Analysis, Random Forest, CHI-Square test, Feature Correlation Analysis, Feature Engineering.

Face Recognition using LBPH Algorithm 02/2022 – 03/2022

- Developed a face recognition system using the Local Binary Patterns Histograms (LBPH) algorithm to identify and verify individuals from images.
- The project involved preprocessing face images, extracting facial features with LBPH, and training a classifier for accurate recognition.
- The model demonstrated robustness to variations in lighting and facial expressions, making it suitable for real-time applications like access control or attendance systems.

Tools: Python, Machine Learning, OpenCV, Image processing, Numpy, scikit-learn, Jupyter Notebook

AI-Powered Web Support for Formal Verification and Proof Assistants 04/2025 – 09/2025

- Developed *AxiomAI*, a web-based platform combining Large Language Models, Retrieval-Augmented Generation (RAG), and a Proof API to support learning and experimentation with proof assistants (Coq, Isabelle, Z3).
- The system enables users to ask natural language queries via a chatbot in two modes (LLM-only, LLM+RAG) and generate proof snippets, with live execution supported in Z3. Designed modular backend services (FastAPI, Flask) and a lightweight frontend with metrics logging for latency, accuracy, hallucination rate, and proof success rate.
- Evaluations showed RAG significantly improved factual grounding and reduced hallucinations compared to LLM-only baselines. The project demonstrates how AI can lower the entry barrier to formal verification while providing an extensible research and educational tool.
- **Tools:** Python, FastAPI, Flask, HuggingFace Transformers, FAISS, BM25, Hybrid retrieval, Z3, Coq, Isabelle, JavaScript, HTML/CSS, KaTeX, Prism.js, Stanage HPC, GitHub

Education

The University of Sheffield 09/2025
Masters in Artificial Intelligence

Karunya Institute of Technology and Sciences 06/2019 – 07/2023
Bachelor of Technology in Computer Science and Engineering

- Secured an overall percentage of 70.11%

Extracurricular Activities

Member of National Service Scheme 07/2019 – 04/2023

Member of Nature Club 01/2022 – 04/2023

Interests

Reading | Theological Reasearch | Singing | Travelling

Languages

English	●●●●●	Telugu	●●●●●
Hindi	●●●●○	Tamil	●●○○○