Hemesh Raaja Malathi

930-333-5409 | heraaj@iu.edu | hemesh0204.github.io | Hemesh0204 | hemesh-r-m

Education

Indiana University Bloomington — Bloomington, IN
Masters in Computer Science: 3.76/4 — *August 2023 - May 2025*
Amrita Vishwa Vidyapeetham — Chennai, India
Bachelors in Computer Science: 8.63/10 — *August 2019 - May 2023*

Relevant Coursework

Data Structure and Algorithm, Machine Learning, Artificial Intelligence, Computer Vision, Database Management System, Natural Language Processing, Business Analytics, Distributed Systems, Cloud Computing

Work Experience

Data Engineering Intern — June 2024 - August 2024
Reinsurance Group of America — *Chesterfield, Missouri*
- Developed and deployed a Lambda function to automatically turn off unused EC2 instances based on 24-hour activity logs, significantly improving resource efficiency.
- Integrated the automation process with Jenkins for seamless deployment and utilized Datadog for monitoring and notifications, ensuring real-time updates were sent to a Slack channel whenever instances were turned off.
- Achieved substantial cost savings of approximately $1,000 over one month by optimizing EC2 usage and minimizing unnecessary running instances.

Data Scientist Intern — Febraury 2023 - April 2023
New Pro Data — *Madurai, India*
- Developed a robust resume parsing algorithm with NLP, achieving 95 % accuracy in extracting work experience, projects, education, skills, and key highlights, reducing manual effort by 60 %.
- Employed data cleaning and ranking pipelines for 93 % accurate job matching.
- Built and integrated an HR chatbot using Django, enhancing candidate screening and improving recruitment efficiency.

Data Analyst Intern — August 2021 - September 2021
The Sparks Foundation — *Chennai, India*
- Conducted EDA on US supermarkets using Matplotlib, Seaborn, Scipy, and Plotly, identifying operational inefficiencies and reducing inventory costs by 20 %.
- Leveraged insights from top-performing areas to make data-driven recommendations for improvement.
- Identified top-performing areas within the supermarket and analyzed the factors contributing to their success.

Projects

Algae Classification | *Python, Tensorflow* — January 2024- May 2024
- Processed microscopic algae image data from a FlowCam DB dump by the City of Bloomington, including augmentation to enhance dataset quality.
- Implemented and evaluated CNN, AlexNet, and ViT models, achieving 98% top-5 classification accuracy, enhancing algae identification reliability.
- Deployed the model in the Bloomington government office, integrated with human and statistical predictions, and added to a pipeline starting from preprocessing to result display, reducing bacteria identification time by 50%.

Mental Health Prediction | *Python, Sklearn, Tensorflow* — May 2021 - June 2021
- Built a predictive model using machine learning to assess mental health status based on diverse behavioural and demographic data, enabling early intervention.
- Applied advanced data preprocessing techniques to the Kaggle dataset, employing a range of models including regression, classification, decision trees, and boosting.
- Developed an ADA Boost model that achieved a commendable 81% accuracy rate.

Technical Skills

Languages: Java, Python, C/C++, SQL, NoSQL, HTML/CSS, JavaScript, R, Pyspark, TypeScript
Frameworks: React, Flask, Django, Hadoop, Express.js
Developer Tools: Git, Docker, Visual Studio, PyCharm, IntelliJ, Tableau, AWS
Libraries: Pandas, NumPy, Matplotlib, Sklearn, Tensorflow, Pytorch