Suwesh Prasad Sah

GitHub | Hugging Face | LinkedIn | Google Scholar | suwesh081@gmail.com | +91-8926190630

Education

National Institute of Technology Rourkela, India — Dec 2020 - Jun 2024
Bachelor of Technology in Computer Science and Engineering — GPA: 7.80/10.0

St. Xavier's English High School, Bihar — Jun 2018 - Sep 2020
Intermediate in Science and Maths — Percentage: 81/100

Marigold English Boarding School, Bhaktapur — 2018
Matriculation — GPA: 3.60/4.0

Experience

Hero Housing Finance — Jun 2024 - Present
Graduate Engineer Trainee — Gurugram, Haryana, India

- Built a Retrieval-Augmented Generation (RAG) pipeline–based LLM application integrating FAQ and video knowledge bases to an intelligent chatbot and video clip generator for technical support and sales enablement in Salesforce Org.
- Software testing and implementation of test automation workflows for desktop and mobile applications.

TVS Motor Company — May 2023 - July 2023
Data Engineer, Internship — Bangalore, Karnataka, India

- Worked with the finance team, building data pipelines and frameworks on Microsoft Azure cloud platform, with Agile methodology.
- Built ETL pipelines to connect, extract and load data from Azure data lake, Delta tables, SAP HANA tables, SAP client APIs, and SQL server on Azure Data Factory.
- Created frameworks in Databricks using Python and PySpark to extract transform and load files and data from source data lake to destination.

Wobblrr IT Services — Jul 2022 - Sep 2022
Network Engineer, Internship — Delhi, India

- Simulated and developed VLAN in Cisco Packet Tracer which were later used for on-site physical implementation.

Research and Projects

Parallel Perception Network — 2024 - 2024
Research Publication (view paper) — huggingface/PPN

- Proposed and developed a novel baseline architecture for developing sophisticated models capable of true hardware-enabled parallelism, achieving neural processing speeds that mirror the high velocity of an autonomous race car.
- The model enables real-time scene segmentation and reconstruction from LiDAR perceived data and achieves a 2x speedup in model inference time compared to a sequential setup.

Multi-core Sentiment Engine — 2023 - 2023
Python, NLP — github/SentimentEngine

- Developed a rule-based text analysis algorithm based on ETL approach with multiprocessing for parallel processing of multiple text files.
- Performed web data extraction using requests library, parsing HTML content using BeautifulSoup and text pre-processing using NLTK's tokenize functions, to employ sentimental analysis by computing positive and negative scores, polarity, and more parameters.

Parking Forecasting System — 2023 - 2023
Python, ML, IoT, Time Series Prediction — github/SPS

- Implemented an IoT-based smart parking system with data collection on the cloud.
- Using sci-kit learn, two models are used to individually predict the probability of slot availability. When compared, the Multi-Layer Perceptron gives slots with more probability of availability compared to the Random Forest Classifier.

Technical Skills

Programming Languages & Frameworks: Python, C/C++, MySQL; PySpark, PyTorch, Transformers
Cloud & DevOps: Microsoft Azure, AWS, CI/CD, Automation
Operating Systems: Windows, Linux, MacOS
Other Programs: Adobe Lightroom, Adobe Photoshop
Programming Concepts: Data Structures, Algorithms, DBMS, OOPs
Languages: English, Hindi, Nepali, Maithili

Certifications

CISCO Networking Academy Introduction to Data Science Course completion. 2023

Relevant Coursework

Computer Science: Deep Learning, Machine Learning, Cloud Computing, Software Engineering, Database Management System, Computer Organization and Architecture, Data Structures and Algorithm, Distributed Systems, Object Oriented Programming, Operating Systems, Computer Networks