JEFFREY DAVID

Mobile No: +65 93731871
Email: jeffreydavid255@gmail.com
LinkedIn: https://www.linkedin.com/in/jeffrey-david-25a54a204/
GitHub: https://github.com/Jeffrey-David
Portfolio: https://jeffrey-david.github.io/
Address: Choa Chu Kang, Singapore

EXECUTIVE SUMMARY

Computer Science graduate from VIT, now pursuing M.Sc. in Information Systems at NTU, with experience at Accenture and internships in AI project development. Proficient in Data Science, Deep Learning and Big Data Engineering, I am dedicated to advancing my career in Data Science and AI while fostering innovation and problem-solving skills.

EXPERIENCE

Accenture Solutions Private Limited Bangalore, India
Big Data Engineer 11/2022 - 08/2023

- Received technical training in Apache Hadoop File System and acquired proficiency in various big data analytics tools such as Apache Hive, Spark, Cassandra, and more.
- Received recognition for outstanding performance during the Accenture Greenfield Training Program.
- Designed end-to-end data pipelines using Cloudera Distribution of Hadoop and tools like Hive and Pig.
- Increased data processing efficiency by 10% through the successful implementation of ETL processes, resulting in faster data movement from various sources.
- Implemented ETL processes to move data from various sources to the cloud-based infrastructure.
- Documented processes, procedures and best practices for Hadoop cluster management and Data Warehousing.

INTERNSHIPS

TalentKraft Pte Ltd Singapore
AI Project Development Intern 01/2024 – 03/2024

- Conducted research and data gathering on relevant data for goal setting and performance review.
- Performed feasibility studies to evaluate the viability of AI implementation.
- Designed Gen AI based solution to assist organizations spanning diverse sectors, from pharmaceuticals to manufacturing in their Goal Setting and 360 PerformanceReview Process
- Developed the website as a MERN Stack application following Agile development methodologies.
- Utilized OpenAI API to implement centralized organizational objectives tailored to department and individual KPIs.

The Sparks Foundation Vellore, India
Graduate Rotational Internship Program (Remote) 02/2021 – 03/2021

- Conducted thorough data analysis to identify patterns, trends and anomalies in diverse datasets.
- Evaluated the performance of machine learning models using appropriate metrics and fine-tuned algorithms.
- Implemented data preprocessing techniques to clean, normalize and transform raw data into suitable formats for machine learning models.

EDUCATION

Nanyang Technological University Singapore
Master of Science in Information Systems 08/2023- Present

- Achieved a Cumulative Grade Point Average (CGPA) of 4.80 /5.00.
- Completing in May 2024.
- Focusing on Data Science and Text Mining through my choice of coursework specialization.
- Courses: Text and Web Mining, Data Mining, Intrusion Detection, Research Methods for Data Analysis, Information Retrieval and Information Visualization.

Vellore Institute of Technology Vellore, India
B.Tech Computer Science and Engineering 06/2018 – 05/2022

- Graduated with a CGPA of 8.78 / 10, demonstrating consistent academic excellence and dedication to academic pursuits.
- Core Courses: Data Structures and Algorithms, Operating Systems, DBMS, Computer Architecture, Mathematics

SKILLS

- **Programming** - Python, Java, C, JavaScript
- **AI** - Machine Learning, Deep Learning, Large Language Models(LLMs), Generative AI, PyTorch
- **Data Science** - Data Engineering, Data Modelling, Web Mining, NLP
- **Data Visualization** - Business Intelligence(BI), Tableau, D3.js
- **Big Data Technologies** – Hadoop, Hive, Pig, Cassandra, Spark.
- **Software Development** - HTML, CSS, Java Servlet, JSP, SQL, Database Design, Node.js, React.js
- **Cloud Based Infrastructure** – AWS Deployment, Cloud Computing
- **Soft Skills** – Communication, Project Management, Self-Motivated, Innovation, Problem-Solving

PROJECTS

Aspect-Based Sentiment Analysis of LGBTQ Issues in Singapore

Large Language Model, Deep Learning, Natural Language Processing, Tableau 01/2024 – 04/2024

- Developed and implemented a robust methodology for sentiment analysis on LGBTQ+ discussions.
- Successfully collected diverse datasets from HardwareZone and Reddit, demonstrating proficiency in web scraping.
- Conducted thorough data cleaning procedures to ensure data integrity and quality for analysis.
- Meticulously extracted aspect terms using predefined lists and manual labelling, showcasing attention to detail.
- Achieved weighted avg accuracy of 81% in Aspect Based Sentiment Analysis after fine-tuning the Flan-T5 model.
- Demonstrated commitment to ethical data handling practices and validity considerations throughout the study.
- Provided valuable insights into attitudes towards LGBTQ+ issues in Singapore, contributing to the field of research.

Twitter Sentiment Analysis with MLP, CNN and BERT

Text Mining, Deep Learning, Natural Language Processing 08/2023 – 11/2023

- Developed a machine learning model using Multilayer Perceptron (MLP), Convolutional Neural Network(CNN) and Transformer-based BERT Model for Twitter Sentiment Analysis using PyTorch.
- Conducted data transformation and feature engineering for unstructured text data.
- Implemented data preprocessing, model training and evaluation.
- Collaborated with team members to optimize the models to achieve an accuracy of 92.57%.
- Contributed to fine-tuning BERT model for sentiment analysis on Twitter dataset.

Deep Learning Model for Analyzing Water Quality, Level and Flow of Cauvery River

Deep Learning, Environmental Analysis, Data Analysis, Predictive Modelling 11/2021 – 05/2022

- Developed 3 Deep Learning models to predict water quality, level and flow of the Cauvery River in India.
- Implemented features to predict and monitor environmental parameters that determines the water quality.
- Contributed to the understanding and assessment of water resources using advanced technology.
- Utilized the Neural Prophet model to forecast water level and water flow for the upcoming three years.
- Acknowledged for significant contribution to the Deep Learning Model for Analyzing Water Quality, Level, and Flow of Cauvery River, which is currently under consideration for publication.

Cataract Classification using CNN

Computer Vision, Deep Learning 11/2020 – 04/2021

- Developed a Cataract Classification Model using Convolutional Neural Networks (CNN) with retinal images.
- Utilized two datasets, including the Cataract dataset and Ocular Disease Recognition Dataset, for training the model.
- Preprocessed datasets to extract relevant information and remove irrelevant diseases, ensuring model accuracy.
- Implemented image processing, resizing and dataset splitting for training, testing, and validation purposes.
- Assessed model performance using metrics such as accuracy and loss to validate effectiveness.
- Employed Grid Search method for Hyper Parameter Tuning to attain the accuracy of 93.20%.

CERTIFICATIONS

- Certificate of Completion of Accenture Greenfield Training Program 01/2023 – 11/2022
- Intro to TensorFlow for Artificial Intelligence, Machine Learning and Deep Learning 03/2021 – 04/2021
- Artificial Intelligence for Beginners 04/2020 – 05/2020
- Machine Learning Using Python 04/2020 – 05/2020
- Statistical Data Analytics for Business Research 04/2020 – 05/2020
- Data Science for Engineers - NPTEL 01/2021 – 03/2021
- Programming in Java 05/2016 – 06/2016