

VIGNESH GIRISH

Data Engineer

Glasgow, United Kingdom

☎ - 07760973695

✉ - vignesh.malhotra15@gmail.com

in - <https://www.linkedin.com/in/vignesh-malhotra/>

🌐 - lifeofviggy.co.uk/

EXPERIENCE

DATA ENGINEER

ADSUM TECHNOLOGIES, LONDON

SEPT 2022 – PRESENT

- Built Data Pipelines on Airflow to ensure complete data flow from source to final prediction models within 10 seconds upto 2 minutes depending on the size of the chunk of data.
- Maintaining the integrity of data with constant monitoring, testing and cleaning pipelines setup in each stage of the data flow to ensure high quality and complete data is present for the models.
- Training Text classification models to classify and categorise financial transactions with an accuracy of ~91% and group them accordingly.
- Setting up AWS S3 instances to store and retrieve data from pipelines and update data in storage, ensuring real time data updates for product dashboards within 3 seconds.
- Built a Revenue generation tool that parses accounting filings data from CompaniesHouse (gov.uk) to identify potential leads for the marketing team and help saving 40-50 hours of market research time per week.

MACHINE LEARNING ENGINEER

CAVINKARE, INDIA

JUN 2019 – AUG 2021

- Implemented a deep learning-based face-recognition attendance system to replace traditional biometrics and scaled it up to recognise up to 500 employees.
- Built and deployed a computer-vision pipeline with a consistent accuracy of around 92% on Microsoft Azure to monitor employee safety and report hazards thus reducing off-time security load.
- Developed a parsing tool to generate structured data from unstructured purchase orders and bills using OCR techniques and text similarity saving up to 20 hours of manual-hours per week.
- Worked closely with the team and upper management to build a chatbot using Microsoft LUIS to answer employee queries about company policies.
- Analysed sales and logistics data using interactive data visualization to identify seasonality and trends to predict high demand products and peak times.

SKILLS

- | | |
|---------------------------------|---|
| • Python | • Tensorflow, PyTorch, NLTK |
| • ETL, ELT | • Data Analysis (Pandas, Numpy, Matplotlib) |
| • SQL (Postgres, MySQL) | • Agile SDLC |
| • Spark, Hadoop, PySpark | • CI/CD (Git, Sentry, New Relic) |
| • Azure (DataBricks, ML Studio) | • Flask/ Django |
| • AWS (S3, Apache Airflow) | • Data Visualization (PowerBI, Tableau) |

EDUCATION

MSc DATA SCIENCE

UNIVERSITY OF GLASGOW

September 2021 – September 2022

- 2nd Place in Kaggle competition to detect and classify cancer cells

B. TECH in COMPUTER SCIENCE

SRM UNIVERSITY

July 2015 – June 2019

- Graduated First Class (77%)

EXTRA CURRICULAR

- Runner up in Code for Good Hakathon by JP Morgan Chase & Co. – 2022
- Student Mentor at MCR Pathways – 2022