

VIGNESH GIRISH

Data Engineer

Glasgow, United Kingdom

☎ - 07760973695

✉ - vignesh.malhotra15@gmail.com

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

SUMMARY

Machine Learning Data Engineer with three years of experience in analysing large data sets and solving key business problems with data-driven solutions. Proficient in using ETL methods to extract large volumes of Data for processing and building end-to-end pipelines for production level Machine Learning models.

EXPERIENCE

DATA ENGINEER

ADSUM TECHNOLOGIES, LONDON

SEPT 2022 – PRESENT

- Building Data Pipelines on Airflow to ensure complete data flow from source to final prediction models.
- 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.
- Building Text classification models to classify and categorise financial transactions and group them accordingly.
- Setting up AWS S3 instances to store and retrieve data from pipelines and update data in storage.
- Built a Revenue generation tool that parses accounting filings data from CompaniesHouse (gov.uk) to identify potential leads for the marketing team and help with data-driven marketing.

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 for validation.
- 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
- ETL, ELT
- SQL (Postgres, MySQL)
- Spark, Hadoop, PySpark
- Azure (DataBricks, ML Studio)
- AWS (S3, Apache Airflow)
- TensorFlow, PyTorch, NLTK
- Data Analysis
- Agile SLDC
- CI/CD (GIT, New Relic)
- Flask/Django

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