

# JACK LEITCH

+1(208) 703-9228 ◇ Boston, MA ◇ [jackmleitch@gmail.com](mailto:jackmleitch@gmail.com) ◇ [www.jackmleitch.com](http://www.jackmleitch.com) ◇ [linkedin.com/jackmleitch](https://linkedin.com/jackmleitch)

## EXPERIENCE

### Olive Ai

Machine Learning Engineer

Boston, MA

July 2022 - Current

- Designed and implemented architecture, supporting thousands of API calls daily, to automate prior authorizations reducing decision time by an average of 10 days
- Optimized ML training pipeline by decreasing time to process millions of medical documents from 20 hours to 30 minutes by implementing efficient ETL pipelines in Databricks using PySpark
- Built a highly scalable medical concept feature generation pipeline using FastAPI and Azure Kubernetes Service to be used by other core engineering teams
- Built and maintain internal-use Python packages for data processing, code generation, and feature computation to add major functionality to AI products

### IAAD-UK

Machine Learning Engineer Consultant

Remote, U.K.

Feb - Jun 2021

- Built an end-to-end license plate recognition system achieving 94% accuracy
- Developed using TensorFlow and deployed dockerized FastAPI endpoint using AWS ECS
- Reduced computing costs from MVP by 46% by cutting image processing time by 67%

### RSA Insurance Group

Data Science Intern

Horsham, U.K.

Jun - Aug 2019

- Forecasted motor insurance claims using time series methods (ARIMA) in R
- Incorporated 3rd party weather data into the model to improve short-term forecasts by 32%
- Integrated R-based forecasts into an Excel sheet that was used by 10s of actuaries with little coding experience

## PERSONAL PROJECTS

### Strava Kudos Predictor

Sep - Nov 2021

- Created a tool that predicts Kudos on Strava activities to investigate which attributes are most influential
- Developed in Python using the Strava API, Scikit-Learn, SHAP, and Optuna
- Engineered new highly predictive features using domain knowledge, e.g., type of training run

### Strava Data ELT Pipeline

May - Jun 2022

- Built a data pipeline to ingest my own Strava data using the Strava API on a weekly basis
- ELT pipeline consisted of the Strava API, MySQL, Python, S3, Redshift, and Airflow
- A data warehouse was used to build an automatically updating dashboard in Tableau and to trigger automatic re-training of my Strava Kudos Prediction model

### Essay Companion

May - Jun 2022

- Deployed an NLP-powered Google Chrome extension to summarize, paraphrase, get named entities, and find keyword synonyms from highlighted text using FastAPI, Docker, Hugging Face Optimum, and onnxruntime

## EDUCATION

Master of Science, Mathematics, Boise State University

2020 - 2022

- Division 1 NCAA track and cross-country student-athlete

GPA: 4.0

Bachelor of Science, Mathematics, The University of Edinburgh

2016 - 2020

- Multivariate Data Analysis, Statistical Computing, Probability

First-class honors (GPA: 4.0 equivalent)