Tung Pham

✉ tungpham9802@gmail.com   in linkedin.com/in/tung-pham-icl9802   ⌗ github.com/ttp20   ☏ +84 365 265 433

EDUCATION

Imperial College London, *London, United Kingdom*
BSc in Mathematics with Statistics   *Oct. 2020 – Jun. 2023*
Grade: **First Class Honours (78.39%) - British Honours Degree System for Undergraduates**
Relevant Coursework: Computational Linear Algebra, Applied Probability, Optimisation, Time Series Analysis, Introduction to Statistical Learning, Scientific Computation, Methods for Data Science, Survival Models, Statistical Modelling 1, Network Science, Principles of Programming

WORK EXPERIENCE

VinBigData, *Hanoi, Vietnam*
AI Engineer Intern - Vingroup's AI Engineer Training Program   *Jun. 2023 – Present.*
- Selected out of more than 500 applicants to participate in Vingroup's highly-competitive AI Engineer Training Program.
- Finished Phase 1: Fundamental Knowledge Training. Courses taken: Probability and Statistics, Linear Algebra, Machine Learning, Deep Learning, Computer Vision, Natural Language Processing and AI Ethics
- Phase 2: Professional Internship will commence in December

Reckon Digital, *London, United Kingdom*
Software Engineer Intern   *Jun. 2022 – Sep. 2022*
- Learnt to use the company's main tech stacks and frameworks such as Django, Django REST Framework and web development in general
- Resolved tickets from one of the company's ongoing projects by coding, testing and debugging the back-end of some websites, thus ensuring the completed website met client's requirements

RESEARCH EXPERIENCE (STATISTICS PROJECTS)

Estimating parameters with and without stratification: A theoretical comparison
Guide: Dr. Heather Battey, Imperial College London   *May. 2023 – Jul. 2023*
- Compared two estimators for a specific quantity: one using ordinary least squares and the other being a common estimator used in medical statistics
- Performed data simulation, log-likelihood derivation, estimator computation, and variance analysis, along with exploring their asymptotic distribution.
- Result: Assuming that the target variable is normally distributed, the estimator proposed by medical statistics has higher asymptotic efficiency

Modelling Santander Cycle Hires using Self-Exciting and Mutually-Exciting Point Processes
Guide: Dr. Francesco Sanna Passino, Imperial College London   *May. 2022 – Jun. 2022*
- Applied the theory of Homogeneous, Non-homogeneous Poison, Self-Exciting and Mutually-Exciting Hawkes point processes and proposed five statistical models to model the patterns of bike-sharing departure and arrival times across 840 bike-sharing stations in London
- Evaluated the goodness of fit of each model using various statistical hypothesis tests such as the p-value or Kolmogorov-Smirnov tests
- Result: The model incorporating both self and mutually exciting behavior best represented the patterns of cycle hires

DATA SCIENCE PROJECTS

Credit Card Fraud Detection
Machine Learning training course - Vingroup's AI Engineer Training Program
- Preprocessed and performed various feature engineering methods such as correlation analysis, SMOTE, Box-Cox transformation on a large dataset of over 400 columns to improve model performance
- Achieved out of sample of over 95% ROC-AUC score and over 80% fraud recall using various fine-tuned Gradient Boosting models such as XGBoost, LightGBM, CatBoost.

Drivers' Drowsiness Detection
Computer Vision training course - Vingroup's AI Engineer Training Program
- Used MTCNN, FaceMesh from Mediapipe, Eye Aspect Ratio (EAR) to extract and vectorize important facial features from drivers' videos
- Implemented various deep learning models such as Inception Resnet combined with LSTM, or EAR with CNN to predict drivers' drowsiness with fair accuracy and recall scores (better than human judgement of 50%)

SKILLS & OTHERS

Technical Skills: Multivariable Calculus, Statistical Theory, Statistical Modelling, Stochastic Processes, Survival Analysis, Numerical Linear Algebra, Optimisation, Machine Learning
Computing: Python, R, SQL, JavaScript, MATLAB, Julia
Coursera Courses: Mathematics for Machine Learning and Data Science Specialization, Machine Learning Specialization, Deep Learning Specialization, Generative Adversarial Network Specialization
Languages: Vietnamese (native), English (IELTS 8.5 - fluent)