Yeung-Yeung LEUNG, Ph.D.

mail@leungyeungyeung.de | leungyeungyeung | http://leungyeungyeung.de | linkedin.com/in/yeungyeungleung/

SKILLS

Programming

- Python
- R
- Matlab
- Unix
- SQL

Data processing

- data augmentation, noise filtering, artfact and outlier removal, PCA, tSNE, UMAP

Statistics

- permutation, bootstrap, mixed models

Image analysis

- noise filtering, alignment, morphological operations, segmentation, 3D rigid transform

Time serious analysis

- fourier transform, ARIMA, VAR, LSTM, GRU

NLP

- RegEx, Web scrapping, BART, Bert, TFIDF

Machine learning

- Bayesian optimization, grid/ random search, hierarchical, k-means, XGBoost,LSTM, GRU, stacked 1D-CNN, SHAP, VGG-16, VAE

LANGUAGES

Cantonese	-	Native
Mandarin	-	Professional
English	-	Professional
German	-	Intermediate

CERTIFICATES

Grade 8 Piano
Driving License German Class B

EXPERIENCE

Data Scientist - Omics data - Machine learning application, GSK *2023 - present*

- Develop ML pipeline for target nomination, causal network inference and in silico perturbation
- Validation, benchmarking and optimization of statistical and machine learning models

Research Associate, Imperial College London *2021 - 2023*

- Annotation of patient disease profile via natural language processing of clinical documents
  - Optical character recognition, RegEx pattern matching, Name Entity Recognition, pre-trained Bert text embedding
- Genome wide association study of neurodegenerative diseases
  - Detect potential causal genetic variants in a sample that comprised of 2.6 millions Single Nucleotide Polymorphisms (Linear Mixed Effect Model, LDSC, MAGMA, gene set enrichment)
  - Evaluate the causal relationship of hit genes using eQTL data (Bayes Factor colocalization), single cell RNAseq (MAGMA cell type specificity), and tissue transcriptomics (WGCNA)
  - Designed a browser application for users to query genotypes information (R shiny)

Principal Investigator, The Alan Turing Institute Data Science Study Group *2020-2021*

- Modelling Amyloid Beta Plaque Formation in Alzheimer's Disease using VGG-16 and VAE
  - Facilitate the project delivery of a team of 10 participants
  - Spatial transcriptomics and pathological image feature extraction
  - Image segmentation, pre-trained VGG-16 embedding, VAEs training and optimization

AI society NLP group project, University College London *2020-2021*

- Explore machine learning methods for crypto currency price prediction.
  - Web scrapping, social media sentiment analysis, TFIDF, DistilBART text summarization

Director, London Postdoc Network *2020 - Present*

- Website establishment, organization of medium-scale (300+audiences) virtual events

Research Associate, University College London *2018 - 2020*

- Identify the dynamic features in brain oscillation that influence motor-behavioral outcome
  - Time frequency spectral analysis, linear mixed effects models, permutation statistics
  - Generate recurrent neural networks (LSTM, GRU) and Bayesian optimized XGBoost models
  - Using SHAP approach to evaluate the key determinants that contribute model's decision
- Generated a data processing pipeline for aligning neurophysiological and behavioral data
  - Multi-sources analogue data conversion, alignment, filtering and noise removal
  - Neuronal spike sorting: filtering, SVD whitening, PCA, superparamagnetic clustering
- 3D registration of MRI brain image and stereotaxic coordinates for recording location estimation
- Performed brain implants and electromyography surgeries as the second surgeon

EDUCATION

Computational Neuroscience (PhD), University of Heidelberg, Germany *2010 - 2016*

- Grade: "magna cum laude"
- Investigation of a novel molecular pathway in synaptic vesicle release
- Validation of a mathematical model of 3D calcium dynamics in synaptic boutons
  - Designed a high-throughput pipeline for quantitative time serious image analysis
  - image alignment, noise filtering, artifact removal, automatic ROI detection

Physiology (MPhil), The Chinese University of Hong Kong *2008 - 2010*

- Multi-electrode field potential measurement of memory formation in rat brain slices
- Teaching of 40-60 medical students at practical laboratory session

Research Intern, Toronto Western Hospital, University of Toronto *2006 - 2006*

- Genetic association study: PCR optimization and analysis of sequencing results

University exchange student, University of Groningen, The Netherlands *2006 - 2006*

Molecular Biotechnology (B.S.), The Chinese University of Hong Kong *2004 - 2007*

SCHOLARSHIPS AND AWARDS

- **Academic Merit**, Shaw College, The Chinese University of Hong Kong *2006*
- **Swire University Scholarship**, John Swire & Sons (Hong Kong) Limited *2004 - 2007*