

HAODONG LING

☎ (510)590-0176 ✉ haodong.ling@berkeley.edu 🌐 milanmarks.github.io

Education

Department of Statistics, University of California, Berkeley

Aug. 2023 – current

M.A. in Statistics

School of Data Science, Fudan University

Sep. 2018 – Jun. 2023

B.S. in Data Science

GPA: 3.59, rank: 21/79

- *Math*: Linear Algebra, Probability Theory, Stochastic Process, Statistics, Regression Analysis, Statistical Learning, Time Series Analysis, Optimization
- *Programming*: Data Structure, Computational Statistics, Numerical Algorithm, Distributed System, Artificial Intelligence, Computer Vision

Publication

1. **Haodong Ling**, Haolun Shi, Nan Yuan, Yuan Ji and Xiaolei Lin (2023). qTPI: A quasi-Toxicity Probability Interval design for Phase I trials with multiple-grade toxicities. *Statistical Methods in Medical Research*, doi: 10.1177/09622802231176034

Research Experience

Development of a qTPI Design for Phase I Clinical Trial

Mar. 2022 – Oct. 2022

Advisor: Prof. Xiaolei Lin, School of Data Science, Fudan University

- Developed a qTPI design for Phase I clinical trial. Designed a novel measure for multiple levels and types of toxicity and incorporated it into a Bayesian framework of the interval based designs.
- Compared operational characteristics between the qTPI design and CRM, BLRM, BOIN, mTPI2 designs. Demonstrated the qTPI design achieved better safety, accuracy and reliability.
- Elaborated the numerical simulation and the hypothetical trial sections of the manuscript. The manuscript has been submitted to Statistical Methods in Medical Research.

Detection of cancer from cell-free DNA sequences

May 2021 – Dec. 2021

Advisor: Dr. Jie Fu, Mila - Quebec AI Institute

- Cooperator: Dr. Yi Zhang from Euler Genomics, Beijing and Dr. Min Lin from Sealab, Singapore.
- Studied topics related to deep learning, especially transformer models and deep reinforcement learning theories.
- Used a pre-trained DNABERT model, added adapter layers, and fine-tuned the model to detect cancers from cell-free DNA sequences.

Stock Market Prediction Based on Sentiment Analysis of reddit Comment

May 2021 – July 2021

Advisor: Prof. Anand BHOJAN, School of Computing, National University of Singapore

- Cooperator: Jason Ong from University College London.
- Attended the summer workshop in the School of Computing, National University of Singapore. Finished the "AI/ML for Financial Services" course with grade A.
- Used NLP models like Vader and FinBERT to extract sentiment features from reddit comments. Used machine learning models, such as KNN, Random Forest and Long Short-term Memory (LSTM) model to predict stock price. Poster is displayed on website: 🌐 sws.comp.nus.edu.sg

Selected Course Projects

NASA Airfoil Self-noise Prediction | *R, Bayesian model, Markov Chain Monte Carlo method*

Spring 2022

- Used Bayesian architecture and constructed hierarchical model for prediction of NASA airfoil self-noise.
- Established several models and conducted model selection. Estimated the model using Markov Chain Monte Carlo (MCMC) method. Evaluated the model by diagnosing the Markov chain's convergence and calculating the mean square error.

Music Recommendation System on KKBOX | *Spark*

Spring 2022

- Constructed a distributed music recommendation system on from KKBox Music Data.
- Realized user-based, item-based collaborative filtering algorithms and weighted alternating least square algorithm based on Pyspark. The ALS model achieved highest accuracy of 0.38.
- Analyzed how different parallelism setting affect the performance and discussed the best setting for different data and cluster size.

Computer Vision Course Project | *Pytorch, deep learning, transformer*

Fall 2021

- mid-term project: Implemented the Segmentation Transformer (SETR) model using Pytorch on cityscapes dataset for semantic segmentation task, which achieved 60% pixel-wise accuracy.
- final project: Proposed a Hybrid Vision-Transformer model architecture, which merge convolution layers for better feature extraction in the original Vision Transformer (ViT) model. The model achieved 71% accuracy on CIFAR-10 dataset.

Gomoku AI Project | *python*

Fall 2021

- Implemented α - β pruning, Monte-Carlo tree search, and multiple reinforcement learning algorithms like SARSA, q learning, REINFORCE, and self-teaching ADP for the Gomoku game.
- Scored over 1500 points on the rating board and beat several classic Gomoku algorithms and most of my classmates' algorithms.

Technical Skills

Programming: R, Python

English:

- TOEFL (104, Reading 29 + Listening 27 + Speaking 22 + Writing 26)
- GRE (323, verbal 153 + Quantitative 170 + Writing 3.5)

Honors / Awards

scholarship 2019-2021: Excellent student scholarship

Contemporary Undergraduate Mathematical Contest in Modeling (CUMCM) 2020: Third Price in Shanghai

The Interdisciplinary Contest in Modeling (ICM) 2021: Honorable Mention