ZHANG, Hongzhen

Email: zhanghongzhen1019@outlook.com   Tel: 86-187-6580-1028
My Personal Website: Hongzhen's Personal Website (hongzhengit.github.io)

Personal Statement

I am currently looking for a PhD position in disciplines of Statistics, Physics or Electrical Engineering. I got my Master's degree in engineering at Tianjin University in the School of Precision Instrument and Opto-Electronics Engineering. During the past 2 and a half years, I was working as a data analyst at Discover Financial Service, a fin-tech and card issuer company of USA. I have research or working experiences in heuristic algorithms, signal processing and machine learning.

Education Backgrounds

Tianjin University (member of 985, 211 universities) *2017.09 – 2020.01*
Master's degree in Engineering, GPA: 85/100.
Xiamen University (member of 985, 211 universities) *2013.09 – 2017.06*
Bachelor's degree in Engineering, GPA: 3.3/4.0.
Double degree in Economics, Major in Mathematical Statistics, GPA: 3.0/4.0

Publications

- **Extension of Terahertz Time Domain Spectroscopy: a Micron-level Thickness Gauging Technology**
  Hongzhen Zhang, Lili Shi, Mingxia He[*]. *Optics Communications*, 506 (2022) 127597.
- **Terahertz Thickness Measurement Based on Stochastic Optimization Algorithm**
  Hongzhen Zhang, Mingxia He[*], Lili Shi, Pengfei Wang. *Spectroscopy and Spectral Analysis*, 40(2020) 3066-3070.
- **A terahertz non-polar material detection technology based on Rouard's Method**
  with Mingxia He, Lili Shi and Pu Wang. *Invention patent*, Patent No. CN201910303091.9, Waiting for authorization.
- **A thickness measurement technology developed with terahertz spectrum**
  with Mingxia He, Lili Shi and Pu Wang. *Invention patent*, Patent No. CN201811197783.1, Authorized.
- **A multispectral imaging device for the Meibomian glands of human eyes**
  with Yanping Chen, Tianyu Zheng and Yifan Yang. *Invention patent*, Patent No. CN201610250677.X, Authorized.

Research Experiences

Research on information extraction methods for Terahertz Spectra

Tianjin University. Under the supervision of Prof. HE Mingxia. Funded by National Natural Science Foundation of China (NSFC) (Grant No.61675151).

Information extraction methods are essential to most kinds of spectrum analysis technologies. We built mathematical models to simulate the way that magnetic waves propagate in mediums and combined it with heuristic algorithms like Differential Evolution (DE) and Genetic Algorithm (GA) to extract information (thickness, refractive index, pulse positions, etc) from the spectra of terahertz waves. In the time domain,

- A multiple regression model was proposed by summing up signal pulses that propagate through single-layer mediums.
- DE or GA was leveraged to calibrate parameters for our multiple regression model which is not parametric linear, then the overall time-domain signal would be separated into pulses at different positions based on determined parameters.
- The thickness and refractive index were calculated according to Time of Flight (TOF) theory and Fresnel's Equation.

In the frequency domain,

- The transfer function for single-layer mediums could be derived by performing Fourier Transformation on our multiple regression model.
- Based on Rouard's Method, we extended this transfer function to simulate the way that magnetic waves propagate in multi-layer mediums.
- We improved DE algorithm with several self-adaptive strategies and took advantage of it to extract information from our extended transfer function.

Research on HAR models for Shanghai Composite Index

Graduation Thesis for my double degree in Economics, Xiamen University.

In this project, we built a HAR (Heterogeneity Autoregression) model with 5-min high frequency data of Shanghai Composite Index, to evaluate volatilities which are caused by traders with different expectations and risk appetites.

- To simulate behaviors of traders in different types (Short, Mid, Long), RV (Realized Volatility) with 1, 5 and 22 lags were selected as the explanation variables for our HAR model.
- Results showed that HAR model could reflect marketing volatilities in time and well illustrate the fluctuations caused by different kinds of traders in the market.

Working Experiences

Discover Financial Service (DFS, Shanghai COE)

Senior Analyst, 2020.03 – 2022.07

Volume Forecast of Home Loan Applications

This project was proposed mainly because of the significant increase in house price in USA during 2021, which contributed a lot to the momentum for home equity loan applications. More stuffs are needed to assist on these additional applications. To provide some insights:

- We selected explanation variables based on business knowledge and correlation analysis.
- The selected explanation variables were fed into a time series model as extra regressors.
- With the future application volumes predicted by our time series model, an approximated number of newly hired agents could be estimated.

NLP research on Customers' Complaints

With the continuous increase in market share that DFS products possess, lots of feedbacks and complaints are received by customer-facing agents. An automated document analysis tool is necessary to help on the analytic and classification tasks.

- We constructed a binary classifier with pre-trained BERT to recognize customer complaints from a large set of texts.
- A Random Forest classifier was leveraged to divide these complaints into a set of determined categories.
- For complaints that could not be labeled, we took advantage of a network-based community detection method to find out their natural clusters.

Data Management Platform developed with PyQt

We developed a desktop application with PyQt framework to maintain the invoice data from our third-party vendors, which includes several modules:

- The connection module for interaction between Python environment and our Data Warehouse.
- The data formatting module developed with Pandas and Xlwings for structuring the data from raw files.
- The reporting module used to track the performance of our third-party costs.

Analog Devices (ADI, Beijing Office)

Part-time algorithm intern, 2019.07 – 2020.01

Parameter Selection for Generalized Memory Polynomial (GMP) Model.

Generalized Memory Polynomial (GMP) Model is a widely-used compensation structure for linearity in Digital Pre-Distortion (DPD). To simplify the form of GMP model and reduce the circuit scale of DPD, we applied a c-Lasso algorithm mentioned in Muhammad Tabassum (2018) to get a simplified GMP model with sparse parameters.

Awards

- Innovation Award at Discover Financial Service
- Excellence Award at Discover Financial Service
- 1st prize of National Mathematical Modeling Competition (Undergraduate Group)

Programming Skills

I have solid skills in Python, MATLAB and SQL language and I am also familiar with R software. For more information about myself and my research/working experiences, please check my personal website https://hongzhengit.github.io/.