ZHOURONG LI

Phone: +1 6465750518 | E-mail: zl2977@cumc.columbia.edu | Website: https://bluedenw.github.io/

EDUCATION INFORMATION

Columbia University	**New York, USA**
Master of Science in Biostatistics	**09/2020-05/2022**

- Main Knowledge: *Probability Theory, Statistical Inference, Multi-Variable Linear Regression, ANOVA Testing, Generalized Linear Models, Mixed Effects Models, Cox Proportional Hazards Models, Interactive Graphic Representation Building, Supervised and Unsupervised Models, Survival Analysis, Relational Databases and SQL Programming, Natural Language Processing and Sentimental Analysis, Latent Variable and Structural Equation Modeling, GIS and Spatial Analysis, Principle of Epidemiology*

McGill University	**Montreal, Canada**
Double Major, Bachelor of Science in Statistics and Computer Science	**09/2015-08/2019**

- Main Knowledge: *Regression & Analysis of Variance, Probability, Statistics and Hypothesis Testing, Introduction to Stochastic Processes, Theory of Computation, Time Series Analysis, Real Analysis, Abstract Algebra, Advanced Calculus, Matrix Numerical Analysis, Applied Machine Learning, Artificial Intelligence, Algorithm Design, Computer Software Systems, Data Structures, Computer Hardware Systems, Operating Systems, Financial Theory, Micro Economics and Macro Economics Analysis and Applications*

PERSONAL SKILL

- Programing Language: Python, R, C, Java, SQL, MATLAB
- Programming Utility: R Studio, Visual Studio, Eclipse, MATLAB, Jupiter Notebook
- Microsoft Utility: Microsoft Word, Excel, Power Point Presentation, OneNote
- Adobe Utility: Adobe Photoshop CS6, Adobe Premiere Pro
- Apple Utility: Numbers, Pages, Keynotes, Final Cut Pro
- Systems: Microsoft Windows, Linux, Macintosh OS X
- Languages: Mandarin (Native), Cantonese (Native), English (Proficient), French (Basic)
- Interests: Piano (Level 8 Certify), Basketball, photography, Video Editing
- Personal Website: bluedenw.github.io
- Others: Take the CFA Level 1 Exam in August, 2021

PROFESSIONAL EXPERIENCE

Yuexiu Industrial Investment Fund Management Co., LTD.	**Guangzhou, China**
Investment Analyst Intern	**04/2020-07/2020**

- Conducted industry research, investment feasibility analysis, and wrote industry research reports
- Responsible for writing department's monthly investment reports and quarterly investment reports
- Took part in summarizing the strategic placement information of all stocks on China's Sci-Tech Innovation Board

Portfolio Management and Quantitative Strategies Investment, E Fund Management Co., Ltd.	**Guangzhou, China**
Quantitative Analyst	**10/2019-11/2019**

- Developed Python program which provided real time comparison between portfolio rate of return and benchmark via Tushare, implemented data visualization with Plotly, and verified results with Wind database
- Developed Python program which efficiently classified portfolio rate of return and benchmark rate of return by industries
- Built single-period and multi-period Brinson models for performance attribution and ranked results by Allocation Return

North American Headquarters, Synced	**Toronto, Canada**
Data Analyst	**05/2018-07/2018**

- Developed WebCrawler program in python, which automatically collected and summarized Author/Topic/Organization information from renowned machine learning meetings (i.e. CVPR, NIPS, ICML etc.)
- Developed Python program to analyze collected topics and find the most popular topic through different evaluation methods. (topic frequency, number of references, Google H5-index)

EXTRACURRICULUM ACTIVITY

***Online Course Lecturer*, Central China Normal University (Montreal, Canada)**	**07/2019-08/2019**

- Responsible for teaching the course COMP 271: Algorithms & Data Structures, introduced algorithms and data structures in computer science, i.e. AVL Trees, Red Black Trees, Greedy Algorithms, Minimum Spanning Trees, Dynamic Programming. Algorithms in Artificial Intelligence (Mini-Max, Monte-Carlo) were also being introduced
- Distributed assignments and exams to students, provided feedbacks and assessments for better understanding the course

***Technical Consultant*, Computer Task Force (Montreal, Canada)**	**01/2017-04/2017**

- Dealt with frequently asked questions from students about computer science problems concerning with system, hardware, software development and optimization
- Assembled hardware, including constructed customized computer and repaired facility of Computer Lab

***Secretary*, McGill Student Chinese Brush Arts Club (Montreal, Canada)**	**01/2016-05/2017**

- Responsible for scheduling time and place of weekly internal meeting, planning club activities and events, making strategies for contacting and negotiating with sponsors
- Published advertisements and posted passages in social media platforms to update news and events of the club

ACADEMIC RESEARCH PROJECT

Individual Researcher, **Survival Analysis of Nasopharyngeal Carcinoma (directed by Dr. Caixia Li)   11/2019-12/2019**
- Built survival models for 654 nasopharyngeal carcinoma patients, studied the effects of the prognostic factors, including age, sex, pleural invasion and tumor size
- Found that age and tumor size had a non-linear effect on the risk of recurrence, then employed a logarithmic Gaussian process (GP) for modelling smooth non-linear effects

Individual Researcher, **State Space Model and Kalman Filters (directed by Prof. David Stephens)   01/2019-05/2019**
- Introduced principal knowledge, mathematical representations and deductions of State Space Models and Kalman Filters
- Discussed possible improvements of numerical stability of posterior variance matrix in Kalman Filters Algorithm--used sequentially updating the singular value decomposition (SVD) during the evaluation of the variance matrix
- Built Kalman Filters and State Space models in R using Monthly US domestic Enplanements data (1985-1996) and UK gas consumption data (1960Q1 to 1986Q4)
- Wrote 36 pages final report includes project details and R-codes

Team Leader, **Improved Baselines of Published Paper (directed by Prof William L. Hamilton)   03/2019-04/2019**
The goal of this project was to reproduce and improve the baselines used and referenced in a paper and evaluate whether the benefits of Paragraph Vectors are truly as good as the authors claim them to be
- Reproduced the best baseline (i.e. Naive-Bayes-SVM) referenced in the published paper: *Le & al. Distributed Representations of Sentence and Documents*, the reproduced baseline achieved an accuracy of 92.096% on the test set. This is +0.876% above the baseline reported in the article (91.22%)
- Implemented new baseline using word2vec technique, with careful preprocessing setting and optimized parameters, obtained an accuracy of 94.234%, outperformed Paragraph Vectors in the article (92.58%) by +1.654%

Team Leader, **Modified MNIST Project (directed by William L. Hamilton)   03/2019-04/2019**
This project attempts to perform an image analysis prediction challenge on a modified MNIST dataset. The goal is to train the best Deep Learning Model to recognize the number that occupies the most space in an image
- Built and tested multiple Deep Learning Models using different combinations of convolutional layers, tested different data preprocessing techniques including Image Binarization, Segmentation and Rotation
- Submitted the final Deep Learning Model with seven Convoluted-2D Layers and two Hidden Layers, attained validation accuracy of 97.30% and validation loss of 0.1560. Ranked 14 out of 120 groups among the whole class.

Team Leader, **IMDB Sentiment Analysis (directed by William L. Hamilton)   02/2019-03/2019**
- Built machine learning models using Logistic Regression, Support Vector Machines and Stochastic Gradient Descent as classifier respectively. And use N-gram and TF-IDF weighting in feature extraction pipelines
- Submitted the final Machine Learning Model using Logistic Regression, Tri-gram feature selection and remove 'stop-words', attained prediction accuracy with 91.088%, ranked 8 out of 120 groups among the whole class in Kaggle

Team Leader, **Predicting the Popularity of Reddit Comments (directed by William L. Hamilton)   01/2019-02/2019**
This project used a dataset of 12000 instances which was split into training, validation and test sets. Available features included text comments, the number of replies comments received, an indication of whether each comment was the root of a thread and an indicator of controversiality
- Built linear regression model via both Gradient Decent Algorithm and Closed Form Approach to predict the popularity of Reddit comments
- Submitted final model using Closed Form Approach with fine tuning hyper-parameters, achieved MSE of 1.25357 on the test set (very good result). Attained final grade 100% in this project

Individual Researcher, **Artificial Intelligence Final Project (directed by Prof. Jakie Cheung)   01/2018-05/2018**
- Wrote AI search algorithm in Java to resolve Tablut Chess--an ancient chess game. Implemented Minimax Algorithm, Minimax with Alpha-beta Prunning Algorithm and Monte-Carlo Algorithm
- Designed two algorithms for two adversarial chess players, namely Muscovite (Black Chess) and Swedes (White Chess).
- Utilized AI search algorithm to beat both Random AI and Greedy AI with win rate of 100%. Wrote a final report consists of Technical Approach, Pros/Cons of the Chosen Approach and Future Improvement