

# Yu-Hsien Liu — Data Scientist

☎ +1 (613) 890 6260 • ✉ liu.yuhsien81@gmail.com  
🌐 yuhsienliu.github.io • in yuhsienliu • 🌐 yuhsienliu

## Summary

---

Action-oriented Data Scientist with 4 years of significant experience leading data mining projects, and applying sophisticated analytics techniques to build analytical big data solutions that enhance the financial market operations, policy analyses and research. With her strong grasp of the essence of technical issues and developed concepts, models and frameworks, Yu-Hsien Liu has also collaborated with academia and industry consultants in research projects, workshops, seminars and conferences to embrace new developments.

## Technical Skills

---

**Programming languages:** Python, R, Matlab, Java

**Cloud SaaS:** Azure, AWS

**Database:** MySQL, ElasticSearch, Hadoop HDFS

**Version control tools:** Git, SVN, Docker

**Visualization tools:** Tableau, Plotly, bokeh, matplotlib, ggplot2

**Technical writing:** L<sup>A</sup>T<sub>E</sub>X, Markdown

## Professional Experience

---

### Data Analyst

2019 – Present

*Employment and Social Development Canada*

*Ottawa – Gatineau, Canada*

- Acted as a subject matter expert in the areas of data science, big data, and incorporating new technologies with existing infrastructures
- Applied, adapted and designed R programs for statistical analysis of heterogeneous treatment effects with ML methods (RFs and NNs) at the individual, group and overall levels
- Proposed data visualization techniques to present the results of Program Evaluations, which involves a literature review, data ETL, and statistical analyses
- Delivered detailed documentations, hands-on tutorials and demo to economists on key tools and ML analytic approaches to improve efficiencies

### 📁 Project: Horizontal 3<sup>rd</sup> Cycle Evaluation of the Labour Market Development Agreements

- Efficiently designed data-driven nearest neighbour sampling methodology to build alternative comparison groups for evaluating the effectiveness of labour programs
- Competent in understanding business requirements and identifying use cases for the deployment of predictive analytics
- Produced in- and post-program outcome estimates, such as employment insurance benefits and social assistance dependence, for labour program participants

### Research Assistant – Data Science

2016 – 2019

*Bank of Canada*

*Ottawa, Canada*

- Accurately extracted, transformed, and loaded information from unstructured web-scraped textual data into automated pipelines
- Summarized recommendations on research methodologies from technical and non-technical audience
- Successfully launched data warehousing with Azure to incorporate cloud computing and storage
- Built positive working relationships with other teams (DSO and ITS) to deliver on commitments

## **📖 Project 1: Can Media and Text Analytics Provide Insights into Labour Market Conditions in China?**

- Published a novel nowcasting Labour Market Conditions Indicator (LMCI) with NLP, SVM, and machine learning techniques
- Carefully validated the usefulness of the proposed index (LMCI) by conducting rolling macroeconomic time-series, regression and sentiment analyses
- Enhanced model performances with feature engineering to evaluate accuracy of predictive analytics

## **📖 Project 2: Understanding and Predicting Chinese Monetary Policy**

- Quantitatively created a Chinese monetary policy stance index by using unstructured textual data
- Derived data from official statements from the government and commercial news articles
- Effectively communicated with specialists from other fields to drive innovation and expand departmental capabilities
- Skillfully engaged and articulated research findings to various targeted audiences

## **Publications**

---

J. Bailliu, X. Han, M. Kruger, Y.-H. Liu, and S. Thanabalasingam. Can media and text analytics provide insights into labour market conditions in China? *International Journal of Forecasting*, 35:1118–1130, 2019. <https://doi.org/10.1016/j.ijforecast.2019.03.003>.

## **Projects**

---

### **Face Recognition and Gender Classification with K-Nearest Neighbours**

*Carleton University, Ottawa ON*

**Python**

2019

### **Analysis of Trump Tweets with NLTK and LDA**

*Carleton University, Ottawa ON*

**Python**

2019

### **ARIMA and eGARCH Estimation with Shanghai Composite Index**

*Carleton University, Ottawa ON*

**MATLAB**

2018

### **Space Weather Forecast with conditional Generative Adversarial NN**

*National Research Council Canada, Ottawa ON*

**Python – Tensorflow**

2018

## **Conferences**

---

### **53<sup>rd</sup> Annual Canadian Economics Association Conference**

*Banff BC*

2019

Bank of Canada: Textual Analysis of Central Bank Communications / Banque du Canada : Analyse textuelle des communications de la banque centrale

### **52<sup>nd</sup> Annual Canadian Economics Association Conference**

*Montreal QC*

2018

Bank of Canada: Big Data and Machine Learning / Banque du Canada : Mgodonnes et apprentissage automatique

## **Education**

---

### **Master's of Data Science and Economics**

*Carleton University, Ottawa ON*

2018 - 2019

### **Joint Honours Bachelor's in Mathematics and Economics**

*University of Ottawa, Ottawa ON*

2013 - 2017