Yangfan Liang

Email: yangfanl@andrew.cmu.edu
Phone: (+1) 206-751-3363
Address: Hamburg Hall 3008
4800 Forbes Ave, Pittsburgh, PA

Education

- **Carnegie Mellon University** — Pittsburgh, PA
  - Ph.D. Student in Information System and Management — Aug 2016 - Present
  - Secondary Master in Machine Learning — Jan 2020 - Present
- **University of Science and Technology of China** — Hefei, China
  - B.S. in Statistics — Sep 2012 - June 2016
  - Minor in Computer Science
  - Graduated with Guo Moruo Scholarship (*summa cum laude*)

Research Interest

- **Topics:** Consumer behaviour in digital markets, User-generated content, Online privacy.
- **Methodologies:** Econometrics, Machine Learning, Natural Language Processing

Selected Coursework

- **Machine Learning and Statistics**

Convex Optimization (10725)	Graduate AI (15780)
Topics in Deep Learning (10707)	Machine Learning with Large Datasets(10605)
Intro to Machine Learning (10701)	Statistical Machine Learning (10702)

- **Marketing and Economics**

Game Theory and Applications	Analytical and Structural Marketing Models
Econometrics	Microeconomics

Working Paper

- Yangfan Liang and Rahul Telang (2018), "Consumer Response to Data Breach: Evidence from Consumer Transactions." *Presented on Annual Privacy Law Scholars Conference (PLSC) 2019.*
- Yangfan Liang, Michael Smith, Daegon Cho, and Gordon Burtch (2020), "Impact of Early Digital Movie Release on Box Office: Evidence from Korean Market".
- Yangfan Liang, Ravi Koka, Viv Penninti, Sunder Kekre and Rahul Telang (2020), "Leveraging Financial News Sentiment to Generate Alpha".

Grant, Honors and Awards

- Selected to the NBER Digitization Tutorial for PhD Students — 2019
- PNC Center for Financial Services Innovation Research Grant, CMU — 2017
- Presidential Fellowship, Carnegie Mellon University — 2016
- Guo Moruo Scholarship (Highest honor for USTC undergraduate), USTC — 2015
- National Scholarship (Top 1%), Ministry of Education of China — 2014

Programming Skills

- Python, R, C, LaTeX.

Interest in AI

As a PhD student major in Information System from Carnegie Mellon University, my research interest lies in the field of consumer behaviour in digital markets and user-generated content. The recent rapid development in AI has provided us various tools to analyze unstructured data, such as text classification through BERT and video object detection through ResNet. Therefore, my interest is to apply AI in unstructured data; and then exploit their structures to get suitable features for understanding consumer behavior in digital markets.

I have taken extensive courses related to AI and have become a secondary master student in machine learning. During last summer, I utilized my learnt knowledge on a machine learning project for a fintech startup named Stocksnips. This project is to predict Stocksnips sentiment through deep learning. Stocksnips are financially oriented sentences of listed companies extracted from articles either from financial news such as WSJ or from SEC filing. Previously they manually designed bi-gram features which they believed are important for predicting sentiment based on their finance expertise, and then used SVM for binary classification through those features. However, this method relies on specific domain expertise and is not scalable to more complex features like. Instead, we applied the newly introduced NLP pretrained models - BERT and XLNet, and then finetuned those models for sentiment classification. We successfully improved this binary classification task accuracy from 87% to 96%. Currently I am working on a follow up project, which is to utilize this newly extracted sentiment of stocks as a factor for improving portfolio management performance.

This spring, I have taken part in NBER digitization tutorial and learned from Prof. Amalia Miller about digitization and fairness. Inspired by that, this summer I am working at Intel as an applied machine learning intern, focusing on topics about fairness in machine learning.