

# DAYI LIN

PH.D. CANDIDATE IN COMPUTER SCIENCE

| Mobile: +1-(613)-770-8659 | Email: [dayi.lin@cs.queensu.ca](mailto:dayi.lin@cs.queensu.ca) | Site: [lindayi.me](http://lindayi.me)  
| ResearchGate: [researchgate.net/profile/Dayi\\_Lin](https://researchgate.net/profile/Dayi_Lin) | LinkedIn: [linkedin.com/in/lindayi](https://linkedin.com/in/lindayi)

## SUMMARY

- Experienced in mining large-scale heterogeneous data to **uncover patterns, insights and trends** across software, gaming, and automotive industries
- Proven effectiveness in both industry and academia (17k+ reads on ResearchGate), collaborating internationally, publishing in top-tier journals, receiving prestigious **media coverage** including **Kotaku, PC Gamers, Gamasutra**, and national newspapers
- Passionate and personally driven to keep up with new data science techniques and best practices

## WORK EXPERIENCE

**Data Researcher at BlackBerry and Ford Motor** Jan. 2017 – Apr. 2017

#Spark #Kafka #Algorithm Design #Data Mining #Data Visualization

- Designed Markov Chain-based algorithms, in combination with binary classification techniques (Random Forest) to identify orphaned or wrongly-recognized voice commands, and mistakes in user behaviors (e.g., misclicks)
- Focused on analyzing **time-series user behavioral data** from infotainment system in Ford vehicles to provide feedback to design and development teams, and in turn improve the user perceived quality of the system
- Instrumented the source code of a legacy large-scale distributed system and used **log mining** techniques to collect and stream necessary data for online analysis

**Research Assistant at Software Analysis and Intelligence Lab** Sept. 2015 – Present

#Data Scraping #Computational Data Analysis #Statistical Modelling

- Uncovered patterns in the game stores data using **Natural Language Processing** and **statistical modelling** to provide practical suggestions to game developers, thereby helping them produce higher quality games and improve user satisfaction
- Built an intelligent tool to automatically identify game bugs from gameplay videos, with a median average precision at 10 of 0.91
- Additionally, improved the interpretation and goodness-of-fit by 100% for the just-in-time cross-project defect models using context-aware mixed-effect modelling

## Intermediate R&D Engineer at Alibaba Group

Jul. 2014 – Sept. 2014

#Back-end Web Development #Operation Coordination

- Developed the back-end and administrative dashboard of Member Entertainment System, and the back-end of the membership module of Tmall, which supports multi-million pageviews per day
- Coordinated with operation staff to integrate the Member Entertainment System into several live TV shows including the Voice of China

### SELECTED PUBLICATIONS (FULL LIST AVAILABLE ON MY WEBSITE)

#### Journal Publications (Impact factors in brackets)

- **Lin D**, Bezemer CP, Zou Y, Hassan AE, "An empirical study of game reviews on the Steam platform", *Empirical Software Engineering*\* (3.275), Accepted May. 2018.
- **Lin D**, Bezemer CP, Hassan AE, "An empirical study of early access games on the Steam platform", *Empirical Software Engineering*\* (3.275), Accepted Jun. 2017.
- **Lin D**, Bezemer CP, Hassan AE, "Studying the urgent updates of popular games on the Steam platform", *Empirical Software Engineering*\* (3.275), Accepted Oct. 2016.

\* *Empirical Software Engineering* has the **highest impact factor** among *Software Engineering journals* according to *InCites*.

### EDUCATION

#### Ph.D., Queen's University (GPA 4.2 / 4.3)

2015 – Present

#Ultra-Large Scale Software Systems (ULSS) specialization #Supervisor: Prof. Ahmed E. Hassan

Directly Promoted to Ph.D. program in Computing from M.Sc. program in 2016, because of my outstanding academic performance. Gave multiple guest lectures about gaming data mining in the context of game architecture and Mining Software Repositories.

### TECHNICAL SKILLS

- **Domains:** Data Mining Machine Learning Software Engineering Game Engineering Defect Prediction
- **Skills:** Supervised & Unsupervised Learning Computational Data Analysis Predictive & Explanatory Modelling Natural Language Processing Data Engineering Data Visualization
- **Languages:** Python R Java C/C++ PHP SQL
- **Tools:** Kafka Spark GraphX D3.js \*nix Git/SVN

### SELECTED AWARDS

- Feb. 2018 Winner and Avanade Sponsored Prize, QHacks 2018
- Sept. 2017 Queen's Graduate Award, Queen's University
- Sept. 2016 Winner, Limestone City Hacks 2016
- Jun. 2015 Outstanding Graduates, XUPT