DAYI LIN Ph.D., Data Scientist

| Mobile: +1-(613)-770-8659 | Email: heylindayi@gmail.com | Site: lindayi.me
| ResearchGate: researchgate.net/profile/Dayi_Lin | Linkedin: linkedin.com/in/lindayi

SUMMARY

- Experienced in mining large-scale heterogeneous data to **uncover patterns, insights and trends** across education, software, gaming, and automotive industries
- Proven effectiveness in both industry and academia (22k+ reads on ResearchGate), collaborating internationally, publishing in top-tier journals, receiving prestigious **media coverage** including **Kotaku**, **PC Gamers**, **Gamasutra**, and national newspapers

WORK EXPERIENCE

Data Scientist at **Prodigy Game** Feb. 2019 – Present
#Statistical Modelling #Machine Learning #Spark #Python #R

- Model complex user behavior in game at both user and session level, with data of 70M+ users who generates 300M+ events per day; propose data-driven suggestions to improve game design
- Profile players based on their in-game behavior to empower personalized gaming experience
- Analyze key game features and their impact on key metrics (e.g., user retention); provide in-depth insights that drive product iteration
- Collaborate with product team and user research team to design, conduct and analyze in-game experiments for new features

Data Researcher at **BlackBerry** and **Ford Motor** Jan. 2017 – Apr. 2017
#Algorithm Design #Data Mining #Data Visualization #Kafka #Python #R

- 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 user behavioral sequence 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

Intermediate R&D Engineer at **Alibaba Group** Jul. 2014 – Sept. 2014
#Back-end Web Development

- 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

SELECTED PUBLICATIONS

(Full list available on my website)

Journal Publications (2019 Impact factors in brackets)

- Lee D, **Lin D**, Bezemer CP, Hassan AE, "Building the Perfect Game – An Empirical Study of Game Modifications, *Empirical Software Engineering* (4.457), **Accepted** Oct. 2019.
- **Lin D**, Bezemer CP, Hassan AE, "Identifying gameplay videos that exhibit bugs in computer games", *Empirical Software Engineering* (4.457), **Accepted** May. 2019.
- **Lin D**, Bezemer CP, Zou Y, Hassan AE, "An empirical study of game reviews on the Steam platform", *Empirical Software Engineering* (4.457), **Accepted** May. 2018.
- **Lin D**, Bezemer CP, Hassan AE, "An empirical study of early access games on the Steam platform", *Empirical Software Engineering* (4.457), **Accepted** Jun. 2017.
- **Lin D**, Bezemer CP, Hassan AE, "Studying the urgent updates of popular games on the Steam platform", *Empirical Software Engineering* (4.457), **Accepted** Oct. 2016.

EDUCATION

Ph.D. in Computer Science, Queen's University Sept. 2015 – Jan. 2019

#Software Analysis and Intelligence Lab (GPA 4.2/4.3) #Ultra-Large Software System Specification

- Uncovered patterns in the game stores data using **Natural Language Processing** and **Machine Learning** 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

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