Hu Chen

https://tigerrr07.github.io/tiger-website/ crishuz@foxmail.com +(86) 156-6583-8670

Education

Shandong University, Jinan, P.R.China — September 2021-Present
Master of Science in Data Science — Graduation: June 2024
Specialization: Combined Major of Business and Computer Science with Co-op
GPA: overall 88.46/100

Shandong University, Jinan, P.R.China — September 2017-June 2021
Bachelor of Science in Mathematics and Applied Mathematics
GPA: overall 88.39/100

Publications

O. AlOmeir, **E. Y. Lai**, M. Milani, and R. Pottinger. *Summarizing Provenance of Aggregation Query Results in Relational Databases.* [Short Paper]. To appear in IEEE International Conference on Data Engineering, 2021 (ICDE '21).

O. AlOmeir, **E. Y. Lai**, M. Milani, and R. Pottinger. *Pastwatch: On the Usability of Provenance Data in Relational Databases.* [Short Paper]. IEEE International Conference on Data Engineering, 2020 (ICDE '20): 1882-1885.

On-Going Work

QueryTeller: Sequence-Aware Query Recommendation Using Deep Learning.

Presentations

Developing a Data-Driven Electric Vehicle Strategy in Surrey, BC, Canada. [Co-presented]. Special Interest Group on Knowledge Discovery and Data Mining, Social Impact Session, 2020 (SIGKDD '20).

Maximizing Utilization of Electric Vehicle Charging Infrastructure in Surrey, BC Using a Data-Driven Model. [Co-presented]. UBC Multidisciplinary Undergraduate Research Conference, 2020.

UBC Computer Science Undergraduate Program Evaluation and Renewal. [Co-presented with Dr. Rachel Pottinger]. UBC Board of Governors Meeting, 2020.

Facilitating Users with SQL Query Formulation. UBC Undergraduate Three-Minute Thesis Compe- tition, 2019.

Research Experience

UBC Data Management and Mining Lab with Dr. Rachel Pottinger — March 2019-Present
Worked on *Pastwatch* as the second author and *QueryTeller* as the lead researcher and first author.

- Contributed to the project ideation by extending the smart drill-down system to aggregate queries and numerical attributes, implemented the backend system, and ran the experiments with IMDB, TPC-H, and GLEI dataset on the *Pastwatch* project.
- Inspired by real-world problems with user-database barriers seen in my past work experience and defined my research problem on query recommendation for *QueryTeller* by identifying the knowledge gaps in the existing work.
- Extracted the SQLShare and Sloan Digital Sky Survey (SDSS) dataset and empirically analyzed the sequential changes in SQL query statements posed by human users.
- Modelled our query recommendation problem as a query prediction task based on my query session analysis and presented a new approach to recommend query information (e.g., tables, attributes, functions, SQL keyword templates) by learning from the sequential knowledge ex- ploration patterns of historical users using sequence-to-sequence models.
- Implemented and adapted RNN and transformer model to SQL queries, designed and executed the experiments, analyzed the results, and wrote the *QueryTeller* paper.

UBC Data Science for Social Good Program with Dr. Raymond Ng — Summer 2019
Worked on *Developing a Data-Driven Electric Vehicle Strategy in Surrey* project.

- Partnered with the Environmental Sustainability Advisory Committee of the City of Surrey, BC to guide the development of the Surrey Electric Vehicle Transformation Strategy.
- Designed and developed a web application as a data visualizer to give the city planners a user- friendly way to interact with the data, including the spatial distribution and time trends of Surreys vehicle

- Enabled data-driven city planning by helping the city select 20 curbside charger locations for a Canadian federal funding proposal in September 2019.

Grad Course Projects

CPSC 530L AI Social Impact with Dr. Kevin Leyton-Brown Spring 2020

- Parterned with three graduate students and worked on a research project that uses deep learning techniques to improve irrigation strategies in agriculture as a collaboration with ecohydrologists in UBC Earth and Ocean Sciences.
- Defined an interdisciplinary research problem from scratch by looking into real-world issues (e.g., water crisis, the "more crop per drop" movement in agriculture) and narrowing down project scope by mapping the major challenges and stakeholder needs and soliciting experts' view.
- Extracted, explored, and processed 60GB NASA satellite data used in modelling.

COMM 635 Causal Inference in Information Systems with Dr. Arslan Aziz Spring 2020

- Used difference-in-difference and fixed effects to evaluate the impact of online platform policy changes on incentivized reviews in small electronic products, e.g., batteries and screen protectors.
- Applied NLP techniques, e.g., TFIDF, n-grams, doc2vec, for matching and sentiment analysis.
- Proved and validated with robustness check that after Amazon's ban on incentivized reviews, the number of unnatural reviews maintained while their characteristics became more similar to natural reviews, providing a proof-of-concept for evaluating platform-wide policy effects.

Industry Experience

Huawei Technologies Co., Ltd. March 2021-June 2021
Theory Lab, Intern

Zhejing Lab August 2022- February 2022
Theory Lab, Intern

Other Experience

UBC CS Undergraduate Program Renewal Project, *Admin Assistant* UBC CPSC August 2019-August 2020
304 Introduction to Relational Databases, *Teaching Assistant* Summer 2019

Honors & Rewards

2020 Natural Sciences and Engineering Research Council of Canada (NSERC) Undergraduate Student Research Award (USRA) – $4,500
2020 UBC CS Rick Sample Memorial Research Award – $2,500 2019
IVADO/Mila Deep Learning Winter School Scholarship – $500
2019 UBC Sauder School of Business Kenneth G. Young Memorial Scholarship (ranked 8/693) – $800 2018
UBC Sauder School of Business Scholarship (ranked 3/659) – $2,370
2018 UBC Trek Excellence Scholarship (top 5%) – $1,500

Community Involvement

SIGMOD 2020, *Student Volunteer* June 2020
UBC Data Science for Social Good Program, *Mentor* June 2020
UBC CS Student Society (CSSS) Coffee Chat, *Mentor* March 2020-May 2020
UBC CS Tri-Mentoring Program, *Mentor* September 2018-April 2019
Greater Vancouver Regional Science Fair, *Lab Volunteer* April 2017
Vancouver Learning Buddies Network, *Math Tutor Volunteer* January 2017-April 2017
UBC YOURS Club, *IT Team Executive* October 2015-April 2016

Skills

Programming Languages: Python, C++
Frameworks/Libraries: Pytorch, CUDA
Tools: Linux, Git, Github, VSCode