Yuxin (Yuki) Chen

802 Academy Way, Kelowna, British Columbia, Canada V1V 0C4
(+1) 236-970-1880 | yuxin.yuki.chen@gmail.com | yukichen-yuxin.github.io | linkedin

EDUCATION

University of British Columbia, Kelowna, Canada — September 2022 – June 2023
Data Science, M.Sc. | **GPA: 91.6/100**
- Relevant Courses: Algorithms and Data Structures, Databases and Data Retrieval, Web and Cloud Computing.

Shanghai University of International Business and Economics, Shanghai, China — October 2018 – June 2022
Data Science and Big Data Technology, B.Sc. | **GPA: 88.49/100**
- Relevant Courses: C++ Programming, Python Programming, Data Mining and Machine Learning, Data Visualization, Natural Language Processing, Computational Statistics, Linear Algebra, Bayesian Statistics, Artificial Intelligence.
- Honours: First Prize Scholarship (Autumn 2020), Second Prize Scholarship (Spring 2021, Spring 2020, Autumn 2019, Spring 2019), Outstanding Student Leadership Award (2020).

University of California, Los Angeles — February 2021
Big Data and Business Applications Online Programme | **GPA: 4.3/4.3**
- A 10-week intense program consists of R Programming, Data Management, Statistical Thinking, Machine Learning, Tableau Visualization, and SPSS; with emphasis on case studies.

RESEARCH INTERESTS

Natural Language Processing | Applied Machine Learning | Information Retrieval | Dialogue Systems | Knowledge Graphs

TECHNICAL SKILLS

Programming: C++, Python, SQL, R, Golang, Lingo, Shell
Software: AIMMS, Tableau, MATLAB, SPSS, Photoshop, Director, Audition, Excel, Management Scientist
Platforms: Linux, Windows, AWS
Open-Source Frameworks: Scikit-learn, Keras, TensorFlow, Pytorch, spaCy, NLTK

PROFESSIONAL EXPERIENCE

JustPractice Technologies Inc — Vancouver, Canada, Remote
Data Scientist Capstone Intern — April 2023 – June 2023
- Performed natural language processing techniques such as tokenization, POS tagging and word embedding in SQL and Python on over 400,000 unstructured clinical notes.
- Utilized several technologies and algorithms like LDA (Topic modelling), NER (namely entity recognition), TF-IDF and Multipartite Rank to analyze the clinical notes.
- Built a pipeline to extract corresponding values of keywords and discovered their patterns to aid healthcare professionals to identify diagnoses, procedures and symptoms more efficiently.

Optimization Analytics Technology Pte Ltd., Shanghai Office — Shanghai, China
Optimization Algorithm Intern — September 2021 – March 2022
- Analyzed client needs, formulate optimization problems for network planning and resource allocation, and process data with SQL; build mathematical models using AIMMS; develop algorithms and calculate optimal results in C++/Java/Python.
- Helped companies such as Sanofi and Rio Tinto to plan production, storage and transport solutions to optimise costs.
- Exposed to large datasets and real-world challenges faced by various industries.

RESEARCH PROJECTS

Credit Card Fraud Detection — April 2023 – May 2023
- Developed a deep AutoEncoder model using TensorFlow to detect fraud credit card, which includes a multi-layer network of encoders and decoders and implemented the method of reconstructing data to find the error threshold and achieve classification of fraud cases

Dashboard app for Homicide Cases in America — April 2023
- Used Plotly and Dash both in Python and R to built an interactive dashboard app with Heroku to address the negative effects of homicides.
- The dashboard was used to provide insights into relevant questions, such as examining how homicide-related activity is distributed geographically across the US.

Statistical Analysis Based on Data Scientist Salary — March 2023 – April 2023
- Implemented an exploratory data analysis of salary and job descriptions of data scientists' jobs was conducted to

investigate the different effects of different skills, departments, and other factors on salary, and to explore the relationship between them using the Apriori algorithm.
- Predicted the salary levels using Random Forest Regression, LASSO Regression, and Kernel Regression.

Task-oriented Question Answering for Traditional Chinese Medicine (TCM) May 2021 – May 2022
Advisor: Dr. Liangliang Liu
- Constructed a TCM knowledge graph based on *Compendium of Materia Medica* using Python and Neo4j.
- Designed and implemented an algorithm for joint-model question answering and information retrieval (an NLP dialogue system) for TCM diagnosis and prescription recommendation.
- Part of a national project on TCM knowledge mapping, co-working with the Chinese Academy of Sciences and the Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences.

Chinese Text Retelling Algorithm Based on Multi-level Semantic Units March 2021 - October 2021
Advisor: Prof. Chengcheng Hao
- Proposed a text retelling method for automated semantic augmentation aimed towards dialogue systems.
- Developed an algorithm to replace extracted Chinese words with synonyms determined by context features, sentence templates, and candidate sets, using a bidirectional LSTM model based on the attention mechanism.
- Journal article in press.

Tag-based Hybrid Bayesian Personalised Ranking for Literature Recommendation June 2020 - October 2020
- Developed a recommender system using hybrid Bayesian personalized ranking and weight coefficient algorithms, based on user interactions extracted from Wanfang Data with Python.
- Awarded Honourable Mention in the 2nd "Huiyuan Sharing" National Open Data Innovation Contest.

An Analytical Model for Epidemic Information Dissemination in Social Networks May 2020
- Developed a multi-group epidemic framework based on SIS and SIR epidemic models using MATLAB.
- Proposed a PSIH model extracting and analyzing real social media texts to obtain the regularity of information dissemination, helping platforms detect misinformation and fight fake news.
- Awarded First Prize in the 9th SUIBE Mathematical Contest in Modelling.

Community Health Pass App April 2020
- Built an app with App Inventor for communities and residences to manage health profiles and entry records of residents, by using a web API to generate and read QR codes; improving community health during the pandemic.
- Research published on *Computer Knowledge and Technology.*

Digital Government Strategies and Models for Foreign Trade Policy Optimization July 2019 - December 2019
- Led a team of 4 to analyze foreign trade performance impacted by digital government strategies, based on data collected through visiting over 20 enterprises; proposed models for policy optimization.

PUBLICATIONS

- Xu, Y., & Chen, Y. (2022). Analysis of Chinese Text Retelling Algorithm Based on Semantic Unit. *Application of Integrated Circuit*, 39(02), 60–62. doi:10.19339/j.issn.1674-2583.2022.02.021. (Chinese)
- Xu, Y., Chen, Y., & Gu, Z. (2020). Digital Technologies and Cultural Innovation of Rosewood Furniture. *Public Communication of Science & Technology*, 12(22), 93–95. https://doi.org/10.16607/j.cnki.1674-6708.2020.22.032 (Chinese)
- Chen, Y. (2020). Design and Development of Community Health Pass System with App Inventor. *Computer Knowledge and Technology*, 16, 79–81. https://doi.org/10.14004/j.cnki.ckt.2020.1976 (Chinese)
- Chen, Y. (2017). JSON Format Transmission of Interactive Data between Arduino Client and Server. *Electronics World*, 24, 132–133. https://doi.org/10.19353/j.cnki.dzsj.2017.24.066 (Chinese)

LEADERSHIP EXPERIENCE

SUIBE Water Drops Volunteering October 2018 – June 2022
Lead Volunteer
- Volunteer twice a month (4 hours per session) at the Shanghai Sunshine School for Special Needs, teaching games and sports to children with autism and developmental disorders; spread awareness and promote acceptance.
- Assist in coordinating volunteers, planning activities, fundraising, and managing the volunteer database.

SUIBE Math Club June 2019 - September 2020
President & Peer Advisor
- Planned and coordinated weekly math workshops; tutored Calculus and Linear Algebra to SUIBE students.
- Recruited and trained peer tutors, matching them with over 200 tutees upon request.

Academic Affairs, Students' Union of the School of Statistics and Information, SUIBE June 2019 - September 2020
Vice-President
- Successfully hosted the annual SUIBE Computer Fair, coordinating with invited speakers to lead seminars on AI.
- Organized coding bootcamps; held 16 coaching sessions on programming for both IT and non-IT students.
- Designed and updated e-newsletters; managed social media analytics on WeChat, tripling readership within a year.