Tong Su

(437) 684-3329 | tong.su@mail.utoronto.ca | https://sue-tong.github.io/

EDUCATION

University of Toronto 09/2020 - Present
BSc. in Computer Science Specialist and Statistics Major Toronto, ON
Major GPA: 3.9/4.0 CGPA: 3.8/4.0
Relevant Courses: Natural Language Computing (98), Neural Networks and Deep Learning (96), Algorithm Design & Analysis (91), Operating Systems (85)

PROFESSIONAL EXPERIENCE

Full Stack Developer | Angular, Django, .NET, Azure 05/2023 - Present
Northbridge Toronto, Canada

- Collaborate with the Digital Solution Team to successfully develop a robust insurance website, catering to a customer base of over 10,000 users, encompassing both **Angular** frontend and **Django** backend components
- Utilize **.NET** and **RestAPI** to integrate and process millions of customer data, facilitating smooth data flow from the frontend to the Duckcreek wave policy system
- Implemented **YAML Pipelines** in Azure to seamlessly manage the development-to-release process, including **Resource Groups**, **App Services**, **Docker** container, **Key Vault**, and **PostgreSQL** database on **Azure Cloud**, and successfully migrated the application to the new landing zone.

Data Analyst Assistant | R, PowerBI, OBIEE, Tableau, Pandas 05/2022 - 09/2023
Ontario Public Service Toronto, Canada

- Visualized and Analyzed 10,000 posted transaction records using **Power BI** and dashboards, producing insightful reports that were presented to the Treasury Board, aiding in data-driven decision-making
- Processed and cleaned millions of datasets utilizing **Pandas** and **R** data analysis techniques to assess internal performance for the Internal Audit Branch, resulting in improved data accuracy and facilitating more effective audits

RESEARCH EXPERIENCE

Fine-tuning Pre-trained Models for Low-Resource Neural Machine Translation 01/2023 - Present
Supervisor: Annie Lee, Department of Computer Science, University of Toronto

- Developed advanced strategies to enhance low-resource translation by harnessing the power of pre-trained denoising models, pivoting techniques, and transfer learning methodologies
- Conducted an in-depth analysis of **mBART** model performance, optimizing hyperparameters (adapters, patience, dataset size) and achieving superior results compared to the baseline, emphasizing the approach's effectiveness

Analyzing Instagram Post in Sexual Assault Organizations: Image Analysis and Statistical Approaches 09/2022 - Present
Supervisor: Jia Xue, Faulty of Information, University of Toronto

- Utilized **Azure**-based image analysis to dissect a dataset of **200,000+** Instagram images, revealing insights into how sexual assault organizations leverage social media for promotion
- Performed thorough **linear regression** analyses in **R** to uncover patterns in Instagram engagement metrics, leading to actionable recommendations for improved outreach and increased social media impact

Impact of Compression Ratio on Convolutional Neural Network (CNN) Accuracy in Medical Imaging 09/2021 - 04/2022
Supervisor: Pascal Tyrrell, Department of Statistical Science, University of Toronto Link to final paper

- Explored the influence of image size reduction (compression and downsampling) on **CNN** accuracy for medical imaging with **PyTorch**, incorporating confidence estimation to evaluate preprocessing effects on algorithm training
- Developed and optimized image compression and downsampling models using the **Pillow** library, improving CNN performance in medical imaging through efficient preprocessing of medical images

PUBLICATIONS

- Lee, E. A., Danukarjanto, L., Huang, S., Sharmin, S., Hung, S. Y., Su, T. **Exploring Student Motivation in Integration of Soft Skills Training within Three Levels of Computer Science Programs, SIGCSE TS 2024**
- Ranathunga, S. D., Nayak, S., Thillainathan, S., Huang, Su, T., Peng, X., & Lee, A. **Exploiting Domain-Specific Parallel Data on Multilingual Language Models for Low-Resource Language Translation (submitted), EACL 2024**
- Xue, J., Shi, H., Su, T., Zhang, Q. (in preparation) **Leveraging Instagram for Public Engagement: A Study on Visual Communication Strategies of Canadian Sexual Assault Support Organization**
- Xue, J.*, Zhu, S.*, Shi, H., Su, T., Wang, Y., Zhang, Q. (in preparation) **Social Network Analysis of Canadian Sexual Assault Support Organizations on Instagram: Bonding, Bridging, and Outreach Dynamics**

HONORS AND AWARDS

Dean's List Scholar, University of Toronto 2020, 2021, 2022, 2023
Deloitte Digital Difference Camp Competition, 6th Position 2022
Tableau Data Analysis competition "Data Stories of Singapore" 3rd Position 2018
Singapore and Asia School Mathematics Competition Gold 2016-2017

TEACHING EXPERIENCE

STA130: An Introduction to Statistical Reasoning and Data Science 01/2023 - 12/2023
Instructor: Joshua S. Speagle (Winter 2023), Scott Schwartz (Fall 2023) University of Toronto

- Collaborated closely with professors to design and implement dynamic statistical materials for lectures and tutorials, resulting in a **20% increase** in student engagement and significantly improved learning outcomes
- Innovatively designed and led **in-person weekly tutorials and office hours**, providing comprehensive support for 500+ students grappling with complex statistical and data science concepts

Machine Learning Software Fundamentals Certificate 11/2023 - Present
Course Provider: Data Science Institute University of Toronto

- Supported three courses for **over 100 data science learners**, including "Introduction to Unix Shell, Git, and Github", "Introduction to Python", and "Building Research Software."
- Demonstrated leadership by organizing and leading **virtual support sessions** for a large student cohort, ensuring comprehensive assistance and fostering improved understanding of complex statistical and data science concepts.

PROJECTS

Enhancing Bayesian CNN Models for Galaxy Image Classification with Saliency Mapping link to final project
Team Leader (4-Members Team)

- Incorporate the **saliency mapping** methods to provide interpretability and enhance the understanding of the **Bayesian CNN** model's decision-making process
- Applied Bayesian CNN model to a dataset from the Galaxy Zoo survey to investigate the galaxy images that have extreme **cross-entropy losses** and calculate the similarity scores to find their most similar pairs.

Water Filter Application Development | Kotlin link to final project
Team Leader (7-Members Team)

- Spearheaded the creation of a dynamic **Kotlin**-based application focused on promoting water conservation awareness and featuring an innovative water filter prototype.
- Functioned seamlessly as both the **product owner** and **backend developer**, overseeing project success by designing and implementing **ViewModels** for diverse application activities.

Analysis of the Flow of Visitors to Ontario Public Libraries | R, Linear Regression link to final project
Individual project

- Conducted a comprehensive analysis of visitor flow patterns in Ontario public libraries utilizing **R**, resulting in actionable insights that guided library administrators in optimizing resource allocation and services
- Demonstrated expertise in data wrangling, exploratory data analysis, and **linear regression** techniques while investigating the relationship between visitor counts and other factors for Ontario public libraries

LEADERSHIP AND VOLUNTEER SERVICE

UTCSSA 09/2020 - Present
Marketing Dircter

- **Communicated** with various stakeholders, including the school, students, and sponsors, to ensure effective **collaboration** and promote the school and its events.
- Assisted in organizing the orientation for over **2000** new students and their parents in **5 cities** over a **3-month** period, contributing to a smooth and successful onboarding process.

Heartware Tuition Programme 05/2018 - 05/2019
Tutor Leader

- Acted as the **tutor leader** and managed a team of 10 peer tutors in one primary school
- **Communicated** with primary school students, teachers, and other tutors, practiced being patient with others

SKILLS

- **Languages**: Python, JavaScript, TypeScript, HTML5, CSS3, Kotlin, Java, C#, C, SQL, R, SAS, yaml
- **Frameworks**: Tensorflow, Scikit-Learn, PyTorch, Numpy, Pandas, Pillow, Django, React, Angular, Node.Js
- **Software and Tools:** Tableau, PowerBI, OBIEE, Excel, PowerPoint, Git, Visual Studio Code, Figma, Azure