Liu Jason Tan

Website: **LiuJasonTan.com** • LinkedIn: linkedin.com/in/liujasontan

Phone: (347) 764-5660 • E-mail: liuta@umich.edu

Education

- **Master of Applied Data Science** *(August 2022)* GPA: **4.00** /4.00
  University of Michigan – School of Information, Ann Arbor, Michigan
- **Bachelor of Science in Information Systems, Cum Laude** *(May 2020)* GPA: **3.64** /4.00
  Stony Brook University, Stony Brook, New York

Work Experience

- **Stony Brook University – Senior Computer Specialist** *(October 2017-May 2020)*
  - Performed on-site troubleshooting and data backup to save precious time of professors and staff, handled essential communications for all information technology-related issues on and off-campus leading to faster response time and assisted with advanced back-end hardware and software support to provide security to all devices connected to the campus network
- **Stony Brook University - Teaching Assistant for Calculus III** *(August 2018 – December 2018)*
  - Graded hundreds of homework assignments with feedback every week leading to better grades every subsequent assignment, held office hours and responded to emails to help students with questions to provide one-on-one feedback, proctored all exams to ensure academic integrity and monitored discussion board to increase student engagement in the course

Skills

- Programmed in C, HTML, CSS, Java, **R**, **SQL** and **Python** (with libraries such as **Keras**, **TensorFlow**, **SciKit Learn**, **Altair**, **Matplotlib**, **Seaborn**, **Numpy** and **Pandas**)
- Competent in Microsoft **Access**, **Excel**, PowerPoint, Project, Word, Jira, Confluence, and IDE such as NetBeans, Eclipse and **Jupyter Notebook**
- Experienced in using data science methods such as **preprocessing**, exploratory data **analysis**, data **modeling** and data **visualization**
- Familiar with machine learning algorithms such as **neural networks**, support vector machine, **clustering**, dimensionality reduction**,** and **regression** modeling
- Performed **data mining** tasks to find patterns and similarities between item sets, vectors and sequences

Recent Projects

- **Stock Market Prediction (2019)** – Final project for Data Science course, which analyzed past stock prices to make predictions for future stock prices, using **Long Short-Term Memory (LSTM)** neural network with the **Keras** Library and **linear regression**. Spitted the data into training, testing, and validation sets, **tune the hyper-parameters** of the model and evaluated the model to **prevent overfitting**
- **Electric Vehicle Analysis (2020)** – Personal project to analyze the data of a Tesla vehicle, comparing efficiency with temperature, average speed, and driving smoothness, as well as recording battery degradation over time. Performed **data cleaning** including removing outliers and filling missing values.
- **My Voice Data Challenge (2021)** – Winner of the data challenge, which uses **Natural Language Processing** (NLP) to analyze text message sentiment regarding the Coronavirus. Wrote an automation script that provides a tool for researchers that reduced the **sentiment labeling** task from **hours to minutes**. The scalable tool performs **data cleaning**, text **encoding,** and **hierarchical clustering** using **BERT**, which creates reproducible results that will affect public policy decisions. Presented the tool and results to a **peer-reviewed conference/symposium**, consisting of top NLP or medical researchers.