

Liu Jason Tan

(347) 764-5660 | liu.jason.tan@columbia.edu | <https://www.linkedin.com/in/liujasontan>
<https://github.com/ljtan> | <https://liujasontan.com>

EDUCATION

Columbia University

Master of Science in Computer Science

Track: Vision, Graphics, Interaction, and Robotics

New York, NY
Expected May 2027

University of Michigan - Ann Arbor

Master of Applied Data Science, GPA: 4.00 /4.00

Relevant Courses: Data Mining, Supervised Learning, Unsupervised Learning, Cloud Computing

Ann Arbor, MI
Aug 2022

Stony Brook University

Bachelor of Science, GPA: 3.64 /4.00

Relevant Courses: Object-Oriented Programming, Data Structures, Database Design

Stony Brook, NY
May 2020

WORK EXPERIENCE

Morgan Stanley

Associate, Operational Risk Capital Analytics (Full Time)

- Reconstructed a critical operational risk capital model from scratch, leveraging **Python** to replace a legacy implementation with 10k+ lines of clean, efficient code.
- Optimized **natural language processing** and **machine learning** models to enhance operational risk incident quality assurance, reducing manual review workload by over 50% and increasing data accuracy for risk event analysis.
- Served as a **subject matter expert** in capital reporting by collaborating with global teams to achieve strategic goals and deliver robust results that exceeded stakeholders' and regulators' expectations.
- Led and coached** 2 team members and 5 consultants by setting clear guidelines, managing timelines, and running daily standups, enabling the team to deliver mission-critical capital **analytics solutions** ahead of regulatory deadlines.

New York, NY
Aug 2022 - Present

Poisera

Data Analytics Intern

- Applied **web scraping** and **API** integration pipelines to collect and analyze large-scale public datasets, directly informing product development decisions, leading to significant user interface improvements.
- Conducted user interviews and qualitative research to **identify critical insights**, driving key management decisions and boosting customer satisfaction.

Remote
Jun 2021 - Aug 2021

ACADEMIC PROJECTS

- MyVoice Data Challenge (First Place)- Engineered an **NLP pipeline** to analyze sentiment in text message surveys regarding COVID-19. Automated data cleaning, text encoding, and hierarchical **clustering** employing BERT to improve the efficiency of research to generate deeper insights.
- S&P 500 Stock Performance Forecasting - Achieved 62% precision with a **random forest classifier**, a substantial improvement over the 20% precision of a dummy classifier. Categorized stocks into top, middle, and bottom tiers using key equity metrics such as price-to-earnings ratio, dividend yield, and volatility.
- Social Media Monitoring - Developed a full-stack application for real-time sentiment and topic monitoring of company discussions. Utilized **supervised and unsupervised learning** techniques, including **BERT** for emotion classification (e.g., surprise, anger, disgust) and **non-negative matrix factorization** for topic **clustering** (e.g., account issues, ordering issues, service issues), to gain actionable insights from social media interactions.

SKILLS

Programming languages: Python, Java, MATLAB, R, SQL

Python Packages: NumPy, Pandas, Keras, TensorFlow, SciKit Learn, SciPy, Matplotlib, PySpark, NLTK, OpenCV

Tools: Jupyter Notebook, GitHub, Bitbucket, Jira, API, Agile Workflows, Co-Pilot

Machine Learning: Regression, deep neural networks, classification, clustering, dimensionality reduction