Jeffrey Tsang

jet253@cornell.edu | (718) 340-8646 | jeffreytsang.github.io | linkedin.com/in/jeffrey-tsang

EDUCATION

Cornell University, Graduate School | Ithaca, NY   Dec 2019
Master of Professional Studies in Information Science | conc. in Data Science (GPA: 4.0/4.0)
Cornell University, College of Agriculture and Life Sciences | Ithaca, NY   May 2019
Bachelor of Science in Information Science | conc. in Data Science and Interactive Tech (GPA: 3.79/4.0)
Honors: *magna cum laude*
Activities: Co-President of Information Science Student Association
Teaching Assistant: Intro to Data Science • Data-Driven Web Applications
Relevant Coursework: Data Science • Linear Algebra • Probability/Statistics • Software Engineering • Web Dev • Data-Driven Web Apps • Machine Learning • Statistical Learning • Python • Object-Oriented Programming • Language & Information (NLP)

Skills

Data Science: SQL (MySQL) • pandas • NumPy • Matplotlib • Keras • sklearn • tidyverse • Jupyter Notebook • Tableau • Git
Programming: Python • R • Java • HTML • CSS • JavaScript • D3.js • Vue.js • PHP • WordPress

PROFESSIONAL EXPERIENCE

DataDay Design | *Co-Founder*   Jun 2019 – Feb 2020
- Founded digital marketing agency that develops integrated digital solutions and strategy
- Designed, implemented, and launched WordPress websites and marketing material for start-ups and SMEs
- Configured SEO and created Google Analytics reports to improve customer conversion rates for clients

Huge, Brooklyn, NY | *Data Science Intern*   Feb 2019 – Aug 2019
- Designed and built interactive R Shiny NLP analytics dashboard to analyze thousands of mobile app reviews using neural networks, leading to improved customer experience insights
- Developed neural network model using Python Keras API to classify mobile app reviews
- Developed feature in Java for R Shiny analytics dashboard that generated customer insight reports for stakeholders

Regeneron Pharmaceuticals, Tarrytown, NY | *Clinical Informatics Intern*   May 2018 – Aug 2018
- Created an interactive data visualization tool using JavaScript, D3.js, HTML and CSS that utilizes network graphs to visualize and analyze electronic health record data, quantifying the relationships between genes and diseases

Sapience Therapeutics, Harrison, NY | *Data Science Intern*   Jun 2017 – Jan 2018
- Utilized Python/Jupyter notebook to analyze gene expression datasets and patterns in cancer gene expression levels
- Generated data visualizations in Matplotlib of gene expression data to characterize the genetic profile of select cancers

Mount Sinai Hospital, New York, NY | *Clinical Research Database Intern*   Jun 2017 – Aug 2017
- Analyzed cancer patient cohorts from EPIC database to investigate the reliability of PSA, a prostate cancer biomarker
- Managed QA database projects to maintain accuracy of patient records

PROJECTS

Mood Over Time: Characterizing how an Artist's Music Changes During their Career   Fall 2019
- Analyzed song lyrics from a music artist's discography using sentiment analysis and KL divergence
- Utilized sentiment analysis to quantify relationship between an artist's most popular Spotify songs and its sentiment

Customizable Task Tracking App   Fall 2018
- Designed and developed a customizable task/expense tracking progressive web application using Vue.js framework
- Conducted user interviews and designed mid-fidelity prototypes

Health Services Recommendation App   Spring 2018
- Developed and deployed web application that recommends health services using NLP/IR techniques for a queried symptom based on geographic location using Yelp reviews
- Implemented recommendation system using cosine similarity, sentiment analysis and LDA topic modeling to classify and retrieve relevant reviews

Undergraduate Research Assistant | *Social Dynamics Lab*   Sep 2018 – Dec 2018
- Designed Postgres database and transferred Reddit data into database using SQL and Python in order to apply neural networks on Reddit mental health subreddit data, classifying a Reddit post's presentation of psychiatric conditions
- Used Python and Zillow API to match voter registration records with geotagged Twitter   Oct 2017 – May 2018 messages in order to analyze sociodemographic characteristics and income homophily of Twitter users