

# Lizhen Tan

ltg11@nyu.edu | 646.824.6362 | New York, NY

## EDUCATION

### NEW YORK UNIVERSITY

MS IN DATA SCIENCE

May 2017 | New York, NY

### STONY BROOK UNIVERSITY

BS IN ASTRONOMY/PLANETARY SCIENCE

& APPLIED MATHEMATICS AND STATISTICS

May 2013 | Stony Brook, NY

## LINKS

LinkedIn:// [lizhentan](#)

Webpage: [Liz-page](#)

## COURSEWORK

Deep Learning

Machine Learning

Natural Language Processing

Time Series

Causal Inference

Big Data

## SKILLS

### COMPUTER SKILLS/TOOLS

Linux • Python • R

PyTorch •  $\LaTeX$  • GitHub

Amazon AWS • MySQL • TensorFlow

Hadoop

Spark (beginner)

### LANGUAGE

Fluent:

English • Chinese (Cantonese and Mandarin)

### LEADERSHIP

GRADUATE:

Fall 2016 - Spring 2017

Secretary in Leadership Circle of CDS (Center for Data Science)

Spring 2017

Grader of Python class

UNDERGRADUATE:

All semesters

Dean's List

Fall 2012 - Spring 2013

Research intern (Astronomy)

## WORK EXPERIENCE

### CIVITAS LEARNING | ASSOCIATE DATA SCIENTIST (INTERNSHIP)

Summer 2016 | Austin, TX

- Collaborated with another intern on a new project for building predictive models on students' success based on their engagement in Learning Management Systems. (e.g. Blackboard, Oracle, etc.)
- Acquired data from data warehouse by performing SQL queries
- Analyzed top features of models for insights of students' engagement pattern using both unsupervised and supervised learning

### LAWRENCE BERKELEY NATIONAL LABORATORY | RESEARCH INTERN

Aug 2013 - May 2014

- Term 1
  - Utilized existing codes to perform image processing on microscope captured images to reduce aberration and astigmatism (MatLab)
  - Discovered a possible new form of carbon by searching and matching pentagonal and/or heptagonal patterns in the enhanced images
- Term 2
  - Pre-processed agricultural data obtained from GAEZ3 website(R)
  - Performed data analysis on obtained data to investigate interactions between climate science and human activities
  - Employed geo-models to examine impacts on future crop productivity under geographical adaptation and non-adaptation

## ACADEMIC PROJECTS

### BUILDING IMAGE GENERATOR USING CONDITIONAL DCGAN Spring 2017 | New York University

- Controlled image randomness by adding class information to DCGAN model.
- Processed class context information into either one-hot embedding or pre-trained GloVe word embedding.
- Generated images were reasonable by feeding both MNIST and CiFar10 datasets. (better results using one-hot embedding)

### PREDICT AND QUANTIFY DEVELOPMENT INFLUENCE ON REAL ESTATE VALUES Fall 2016 | New York University

- Quantified influence of new transactions of commercial units on a pre-defined neighboring real estate values
- Applied machine learning techniques (Logistic Regression, Random Forest) to further explore the potential building features which had the most impact in the target (whether an influence existed)

### UNDERSTANDING AND LEARNING AN AUTOMATED QUESTION ANSWERING SYSTEM Spring 2016 | New York University

- Applied text analysis on questions asked on Yahoo!Answers : explored several multi-class classification methods to predict the question category (4 categories were used in the project)
- Built a program to output an answer to a newly asked question, where the answer was found from a similar archived question in the predicted category

### ANALYSIS OF INTERACTION AMONG DIFFERENT TAXI MODES IN NYC Spring 2016 | New York University

- Conducted data cleaning and data extraction using Map-Reduce on Hadoop
- Analyzed the interaction between newly introduced taxi modes (Uber, green taxis) and yellow taxis