# Andrew Li

206-390-3060 | ali@oberlin.edu | www.andrewyinli.com | AndrewYinLi | andrew-yin-li

## Education

**Oberlin College** *August 2016 to May 2020*

**B.A. in Computer Science, minor in Mathematics** *3.62 Major GPA*

**Coursework:** Data Structures, Algorithms, Machine Learning, Artificial Intelligence, Human Computer Interaction, Computer Architecture, Systems Programming, Programming Abstractions, Privacy and Social Networks, Computer and Information Security, Calculus I to III, Discrete Mathematics, Machine Learning Research

**Activities:** Oberlin College Computer Science Majors Committee Officer, ACM Programming Contest Regionals

## Experience

**Incoming Spring 2019 Data Scientist Intern** *January 14th 2019 to May 3rd 2019*

**NASA Ames Research Center** *Mountain View, CA*

- Will spearhead development of a data pipeline using R, Python, and C++ for analysis of satellite data for changes in vegetation productivity and land surface temperature across the United States.

**Deep Learning Research Intern** *May 2018 to August 2018*

**Massachusetts Institute of Technology Lincoln Laboratory** *Lexington, MA*

- Assessed classification capabilities of a dynamic deep neural network utilizing TensorFlow and expedited testing process to rapidly test classification of specific features.
- Designed preprocessor using Python and the Arcpy and PDAL libraries to label and voxelize LIDAR point clouds based on geographic truth data from Google Maps and government databases.
- Leveraged ArcGIS to program various geospatial and LIDAR exploitation and post-processing scripts for geographic surveying and analysis.

**Undergraduate Researcher** *February 2018 to Present*

**Oberlin College** *Oberlin, OH*

- Currently conducting research on neural networks with U-Net and VGG architectures using Keras to identify all windmills in the state of Iowa due to the unknown quantity and implications on green energy.
- Conducted research on predicting the outcome of at-bats in Major League Baseball (listed under 'Projects').
- Wrote a Python script using SQL queries to scrape the MLB Statcast database and used the scikit-learn and Pandas libraries to model the data with random decision forests.

**Computer Science Teaching Assistant** *August 2017 to May 2018*

**Oberlin College** *Oberlin, OH*

- Graded weekly labs for "CS 241: Systems Programming" taught by Professor Roberto Hoyle.
- Utilized a core understanding of C, shell scripting, and Unix in order to assist students and grade weekly labs.

## Skills

Java, Python, C#, C, C++, SQL, Javascript, Racket, Bash, HTML, CSS, Linux, Git

## Projects

**Intelligent Spam Filter Comparison**

- Compared the spam detection capabilities of Long Short-term Memory neural network using Keras to Multinomial Naive Bayes using scitkit-learn.
- Wrote Python preprocessor for Spam Assassin's email corpus to stem words in an appropriate format for machine learning.
- Created a Naive Bayes spam filter as a baseline for comparison. Assisted with creation and testing of LSTM neural network.
- Analysis and comparison of both methods consolidated in a *paper* .

**MLB: Machine Learning Baseball**

- Research project intended to predict the outcomes of every at-bat during the 2016 Major League Baseball season.
- Wrote scraper and preprocessor in Python using Pandas and SQLite3 libraries to scrape the MLB Statcast database for at-bat data.
- Used scikit-learn library and Jupyter notebook to model random decision forests fed various combinations of sabermetrics for each at-bat.