

Haomin Lin

humaslin97@gatech.edu
<https://github.com/HumasLin>

EDUCATION	M.S. in Computational Science and Engineering Georgia Institute of Technology, Atlanta, GA (exp.) May 2021 <ul style="list-style-type: none">GPA: 4.0/4.0Selected Courses: Intro Database Systems; Statistical Machine Learning; Data & Visual Analytics; Web Search & Text Mining B.Eng. in Optical Engineering (joint with Nankai University) Tianjin University, Tianjin, China July 2019 <ul style="list-style-type: none">GPA: 3.6/4.0Selected Courses: Fundamentals of Computer Software; Probability Theory and Mathematics Statistics; Linear Algebra
RESEARCH EXPERIENCE	Research Assistant , Indiana University May 2020 - Now Supervised by Xiaojing Liao <ul style="list-style-type: none">Scraped text data from an advisory report website and pre-processed it as pure text files.Extracted key information from structured/unstructured text data with regular expressions.Transformed information to more general values that represent the features of vulnerabilities. Research Assistant , Purdue University July 2018 - Sept 2018 Supervised by Meng Cui <ul style="list-style-type: none">Designed a component to hold the lens and install the motor then produced it using 3D printing.Created a GUI program in Python to interface with the motor that adjusts lens in optical experiments manually and automatically.Studied and implemented Structured Illumination Microscopy algorithm in Python by reading and summarizing research papers. Research Assistant , Nankai University Feb 2017 - Mar 2018 Supervised by Yange Liu <ul style="list-style-type: none">Analyzed the result of energy absorption of the fiber at different frequencies from experiments on different fiber structures.Conducted several experiments to collect data and made visualizations.Compared the performance of different fiber structures to adjust experiment plans, finding the best one with great absorption rate at 980nm. Research Assistant , Huaqiao University July 2017 - Aug 2017 Supervised by Jixiong Pu <ul style="list-style-type: none">Implemented Genetic algorithm to produce diffracted light field with a laser system.Collected data and produced visualization of key parts in the light field to adjust the parameters to generate a good focusing pattern.Finished a project of constructing multi-focused light field that includes multiple high resolution focusing points.
PROJECTS	Uber saver <ul style="list-style-type: none">Conduct data wrangling on large-scale raw data and integrate data from different Uber services using SQL and Pandas.Build a software to predict surge multiplier for riders with Random Forest classifier using Python SciKit-Learn with accuracy over 80%.Design a visualization of prediction results shown as a heatmap with interactive features using D3.js.

Simulation of rumor spreading in social network

- Create a simulation software to simulate rumor spreading in a social network with Discrete Event Simulation.
- Validate the output of simulation system through statistical analysis and create visualization with Matplotlib.

Community & Digital Archives Project

- Identify negative comments on our website using sentiment analysis in PHP with the help of VADER lexicon.
- Train a Machine Learning model to recognize different kinds of maps with accuracy over 90% based on Transfer Learning using Tensorflow.

Revealing Gendered Language in Job Descriptions

- Crawl around half a million job postings from Indeed.com and then aggregate them into one dataset using Pandas.
- Make prediction on salary levels of job postings that don't have salary information with Ensemble learning to get more data, achieving 87% accuracy.
- Classify job descriptions into 16 industries by counting the frequencies of keywords from each industry and analyze the gender scores in each industry.

PROFESSIONAL EXPERIENCE **Research Technician Intern**, HC SemiTek Corporation, Yiwu, China Jan 2018 - Feb 2018

- Worked out a new design with a 10-people team by studying novel research results and running simulations.
- Analyzed customer behavior and predict product features worth improving using NumPy and Pandas.
- Contributed to the increase by 10,000+ in sales of the new product.

COMPETENCES **Techniques** Python, SQL, JavaScript, D3.js, Tableau, SciKit-Learn, NumPy, Hadoop