

LUWEI (CAROLINE) LEI

1860 N Scott St. Apt 544 Arlington, VA 22209
(814) 862-8939 • leiluwei4221@gmail.com

EDUCATION

Georgetown University

M.S. in Analytics, Data Science Track

Relevant Coursework: Optimization, Statistical Computing, Massive Data Fundamentals, Data Analytics

Expected Graduation: May 2021

GPA: 4.00/4.00

The Pennsylvania State University

B.S. in Computational Data Science

B.S. in System Analysis Mathematics

Relevant Coursework: Stochastic modeling, Data Structure and Algorithms, Computer Vision, Artificial Intelligence

Aug 2015 - May 2019

GPA: 3.61/4.00, Dean's List

SKILLS

- **Programming:** Python, R, SQL, MATLAB, Java, HTML
- **Data Science:** Scala, Spark, Machine Learning, Deep Learning, Natural Language Processing
- **Business Intelligence:** Tableau, Amazon Web Service (AWS), Adobe InDesign

PROFESSIONAL EXPERIENCE

Software Engineer

SenseTime Artificial Intelligence Company

May 2019 - Aug 2019

Shanghai China

- Led the team to design and implement a Python Educational game and successfully promoted to 20+ middle schools in China
- Tested and optimized voice recognition and multi-armed bandit game performances via Pytorch on AWS platform
- Participated in an algorithm textbook editing project and wrote sorting algorithm, which was used by 2000+ middle school students
- Constructed image classification algorithm using machine learning models in Pytorch and generated user-ended API to over 20+ educational organizations

Research Assistant

The Pennsylvania State University

May 2018 - Aug 2018

State College, PA

- Performed ETL on U.S. government petition dataset and tokenized over 7,000 petitions into sentences using NLTK package
- Extracted and integrated natural language processing (NLP) features of the sentences including word embedding, dependency parsing and POS tag on Python NLTK package
- Implemented a convolutional neural networks (CNN) using Tensorflow to classify sentences based on semantic meanings and structures, and improved the model performance to 92%

Research Assistant

The Pennsylvania State University

Jan 2018 - May 2018

State College, PA

- Cleaned and analyzed large datasets to access the factors of federal employment decision making dataset
- Established multiple SVM and neural network models with different sensitive attributes to predict hiring decisions using R
- Summarized underlying patterns of hiring decisions generated by the model and compared with true decisions to conclude the biases of decision making

UNIVERSITY PROJECTS

Social and Economic Impacts of Broadway Shows

Aug 2019 - Present

- Scrapped over 50,000 records of weekly gross data and textual reviews of Broadway shows in history, and evaluated data cleanliness by self-defined quality metrics using Python Urllib and Beautiful Soup packages
- Constructed predictive models including random forest, neural networks to predict sales using social media statistics
- Performed topic modeling via Python Gensim package to unveil the critics' reviews

Discovering Climate Changes in Tree Ring and PDSI Data

Jan 2019 - May 2019

- Adapted time series to analyze different species and geolocation of tree ring width in past thousands of years with Python
- Evaluated the drought conditions using quantitative analysis, and predicted long-term climate changes based on evaluations

AWARDS & INVOLVEMENT

- Member of Association of Analytics for Non-profit Aug 2019 - Present
- Winner of 2019 ASA DataFest at Penn State University April 2019
 - Analyzed the factors to improve Canadian Rugby team game performances and delivered valuable insights
- Third Prize of HackPSU March 2019
 - Built a website to help people understand legal documents
- Shortlist of Nittany AI Challenge February 2019
 - Received \$500 grand for proposing ideas to improve PSU e-library system
- Best Visualization of 2018 ASA DataFest at Penn State University April 2018
 - Analyzed Indeed.com job data and created interactive visualizations