

Leon Zhang

Phone (US): (206) 487-6780 | E-Mail: lz198@duke.edu | Address: Durham, NC ; Vancouver, BC | [Website](#) | [Linkedin](#) | [GitHub](#)

Summary of Qualifications

Programming Languages: Python, Java, R, SQL, MATLAB, C++, JavaScript, HTML, CSS.

Software Tools and Skills: Git (GitHub, GitLab), Containers (Docker), Spread Sheets (Excel, Google Sheet), Cloud Computing (AWS, Azure, GCP), ML Libraries (Sklearn, PyTorch, TensorFlow, Keras), Continuous Integration & Deployment (CI/CD), Flask Application, Data Visualization Tools (Grafana).

Relevant Coursework: Data Structures, Algorithms, Database Management Systems, Probability, Statistical Modeling, Machine Learning, Natural Language Processing (NLP), Data Engineering Systems, Linear Algebra, A/B Testing.

Certifications: AWS Certified Solution Architect – Associate [\[Credentials\]](#)

Education

Duke University, Durham, NC	Aug. 2020 – Apr. 2022
Master of Science, Data Science (MIDS)	Overall GPA: 3.78/4.00
University of Washington, Seattle, WA	Sep. 2016 – Jun. 2020
Bachelor of Science, Chemical Engineering	Overall GPA: 3.55/4.00 Computer Science GPA: 3.76/4.00

Professional Experiences

Data Science Intern, Windstream. **May. 2021 – Present**

- Built, trained, and packaged NLP transformer model in PyTorch to score customer sentiment in chatbot systems to facilitate proactive live customer services.
- Discovered data insights on system message errors and created visualization dashboards for business reference.

ML Software Programmer, Duke Health System **Jan. 2021 – Present**

- Develop Bi-Clustering algorithms for R package to match patient demographics with cancer symptoms to assist doctors by finding treatment solutions more effectively.

Research Assistant, University of Washington Jim Pfaendtner's Research Group **Jan. 2019 – Jun. 2020**

- Devised a deep learning model, variational autoencoder, with the research team to explore chemical reaction pathways and predict intermediate chemical species using TensorFlow.
- Built a training dataset by computing all possible chemical species from reaction pathways using Python libraries including NumPy, Pandas, etc.
- Implemented molecular rotational techniques through the use of quaternion coordinate system in Python and C++ to aid researchers visualize molecule movement in space and study interactions with different interfaces.

Pre-Clinical Research Intern, Chinese Academy of Sciences **Aug. 2018 – Sep. 2018**

- Synthesized tumor targeting Nano drugs for photo-thermal cancer therapy with research group using high-end scientific instruments such as electron microscope.
- Improved drug yield over 300% by reducing clumping, making the synthesis process realistic for large-scale testing.

Projects & Competitions

2020 Duke Datathon – 1st Place [\[Link\]](#) **Oct. 2020**

- Collaborated in a team of four and achieved 1st place in finding out the economic impact of COVID-19 across the world.
- Designed a comprehensive metric using PCA that reflects the economic condition of a country over time, aggregating multiple economic indicators including GDP and stock market indices to perform analysis and modeling.
- Presented analysis of what countries are likely to be impacted by the pandemic and made suggestions to help relieve the economic impact using regression modeling and time series forecasting in R and Python.

Movie Recommendation Web Application [\[Link\]](#) **Aug. 2020**

- Designed a visually appealing, scalable web application to provide users with movie recommendations using Flask, Python, JavaScript, HTML, and CSS.
- Integrated Cloud tools from GCP into development cycle to configure continuous integration and deployment for automated code testing and hosting of website.