

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

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

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

## Professional Experiences

---

**Data Science Intern, Windstream – Durham, NC**

**May. 2021 – Aug. 2021**

- Facilitated proactive live customer services through sentiment analysis by building, training, and packaging NLP transformer models for chatbot system in PyTorch.
- Accelerate customer digital adoption by creating insight dashboards on user activities, remedy tickets, system message errors, and customer satisfaction using Snowflake, Python, REST API, RapidMiner, and Grafana.

**ML Software Programmer, Duke Health System – Durham, NC**

**Jan. 2021 – Present**

- Assist doctors to find effective treatment solutions through developing Bi-Clustering algorithms in R and Python, which match patient demographics with cancer symptoms.

**Research Assistant, University of Washington – Seattle, WA**

**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 with quaternion coordinate system in Python and C++ to visualize molecule movement in space and study interactions with different interfaces.

## 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

## Projects & Competitions

---

**2020 Duke Datathon – 1<sup>st</sup> Place** [\[Link\]](#)

**Oct. 2020**

- Collaborated in a team of four and achieved 1<sup>st</sup> 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 by aggregating multiple economic indicators to perform modeling.
- Presented insights 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 movie recommendations using Flask, Python, JavaScript, HTML, and CSS.
- Adapted continuous integration and deployment for automated code test and production using cloud services from GCP.