

Xiangli Zhang (Justin)

Email: justinzhang.xl@gmail.com | **Cell:** 505-7098187 | **Github:** [justin-mbca/ML](https://github.com/justin-mbca/ML) | **LinkedIn:** [in/justinzhang/](https://www.linkedin.com/in/justinzhang/) | **Publication**

SUMMARY

Data Scientist with 15 years of experience in diverse data analysis projects, including data preprocessing, cleaning, exploration, visualization, feature selection/engineering, model training/evaluation, and system biology and network analysis. Skilled in machine learning, statistics, programming, and data science tools. Experienced in biotechnology, pharmaceuticals, research labs, medical research institutes, universities, and public health. Strong track record of delivering high-quality analysis reports within tight deadlines.

SKILLS

Machine Learning: Deep Learning (CNN, RNN, Transformers), Supervised (Logistic Regression, Random Forest, SVM, etc.) and Unsupervised Learning (Clustering, Dimension Reduction, etc.), Model Evaluation (Cross-Validation, Accuracy, F1, ROC etc.), Validation and Optimization (tuning hyperparameters), Computer Vision, NLP, OpenAI GPT, Langchain, HuggingFace

Statistics: Parametric and Nonparametric Hypothesis Testing, Bayesian Statistics, Linear Regression, Generalized Nonlinear Regression, Testing Correction

Programming: Python (Scikit-learn, TensorFlow, PyTorch, Keras), R, SAS, Perl, VB.NET, SQL, C/C++, HTML, JavaScript, Agile, Git, Docker, Linux

Data Science Tools: Tableau, Power BI, VS Code, Jupyter Notebook, R Shiny, Apache Spark, Apache Nifi

EDUCATION

- MicroMasters Statistics and Data Science, Massachusetts Institute of Technology, 2020-present
- IBM Data Science Professional Certificate, IBM, 2020
- B.Sc. Bioinformatics, Simon Fraser University, 2006
- B.Eng. Chemical Engineering, Tsinghua University, 1992

CURRENT ACTIVITY

- Leading a team of 5 to participate AI competition organized by [DFW Chinese IT Association](#)
- Developed health care recommendation systems for diabetes including chatbot; collected, processed public data; Implemented with Python Flask, JavaScript, HTML etc.
- Applied cutting edge ML technologies, such as OpenAI GPT, Langchain, Pinecone etc.

WORK EXPERIENCE

Data Scientist | Freelancer | Winnipeg, Canada | 05/2023-present

- Worked on diverse data analysis projects using programming, statistics, and machine learning
- Developed product recommendation systems with KNN model and item similarity based method
- Performed multicollinearity analysis for shipping industry investment
- Developed hotel booking cancellation prediction model using Random Forest
- Developed Python application with GUI for RNASeq alignment
- Analyzed clinical patient encounter data using R for data exploration and visualization

Senior Bioinformatics Programmer | Daiichi Sankyo, Inc. | Basking Ridge, NJ | 10/2019-01/2023

- Supported clinical trial study of antibody conjugated oncology drugs
- Established bioinformatics data pipeline for biomarker omics data profiling
- Analyzed oncology multiomics data and generated visualization figures
- Experienced with Jira, Confluence, AWS, S3 bucket

Pharmacometrics R Programmer | ADC Therapeutics | New Providence, NJ | 07/2022-01/2023

- Supported clinical trial study of antibody conjugated oncology drugs
- Updated and maintained analysis pipeline of pharmacological PKPD analysis in R
- Developed R shiny app for data exploration, visualization, and analysis
- Performed modeling and analysis of drug efficacy, safety, and overall survival

Data Science/Bioinformatics Lead | The Bill & Melinda Gates Medical Research Institute | Boston, MA | 01/2022-07/2022

- Supported TB clinical trial development
- Developed R shiny app for data exploration, visualization, and analysis of TB patient data
- Performed statistics analyses on microbiological assay data

Bioinformatician | Los Alamos National Laboratory | Los Alamos, NM | 10/2017-10/2019

- Developed software application for NGS data simulation and genome annotation pipeline
- Published sequencing data and contributed to publications
- Experienced with Agile, Daily Scrum, Code review

Bioinformatics Analyst | UT Southwestern Medical Center | Dallas, TX | 1/2016-5/2017

- Provided sequencing and bioinformatics service to UTSW researchers
- Developed analysis pipelines for diverse sequencing projects
- Evaluated genome structural variant tools and established human whole genome analysis pipeline

Bioinformatician | University of Manitoba | Winnipeg, MB, Canada | 7/2010-11/2015

- Characterized biofuel potential bacteria strains using NGS sequencing
- Established NGS pipelines and performed analyses for microbial genomes
- Analyzed gene expression microarray data and performed pathway annotation
- Experienced with SAS programming for clinical research projects

Bioinformatics Analyst | Public Health Agency of Canada | Canada | 07/2009-05/2010

- Developed gene association application and analyzed microarray data
- Analyzed sequence data on comparative genomics

Programmer | Canadian Grain Commission | Winnipeg, MB, Canada | 02/2007-05/2009

- Developed grain quality control software applications
- Experienced in VB 6.0, VB.NET, and Microsoft SQL