

QIUSHI YAN

qiushi.yan@vanderbilt.edu ◇ 629 239-9151

[Website](#) ◇ [Github](#) ◇ [Linkedin](#)

EDUCATION

Vanderbilt University - Data Science Institute
Master of Science in Data Science

Nashville, TN
Expected May 2023

Communication University of China
Bachelor of Arts in Data Journalism

Beijing, China
September 2017 - June 2021

EXPERIENCE

RStudio - Open Source Intern

May 2022 - August 2022

- Developed the R package [agua](#) for integrating the H2O.ai machine learning platform with tidymodels.
- Contributed bug fixes and features to other R packages in tidymodels: parsnip, tune, usemodels, bundle.

51Talk - Analytics Engineer Intern

September 2020 - June 2021

- Modelled key variables in teaching performance and ROI with regression and hypothesis testing.
- Developed a dashboard for monitoring key metrics identified above with Django and Vue.

DiDi - Data Science Intern

March 2019 - August 2019

- Created data cleaning pipelines with PySpark and Shiny dashboards for visualizing AB testing results.

DATA SCIENCE PROJECTS

[agua](#) - Tidymodels Integration with H2O

[R](#)

- Developed an R package providing tidymodels interface to the H2O sever, implemented a set of consistent APIs for model fitting, tuning, parallel processing and interpretation.

Anime Recommendation - A recommendation system for anime series

[Python](#), [Airflow](#), [Neo4j](#)

- Developed a recommendation system for anime streaming website, including Airflow ETL tasks with PySpark, training pipelines with Node2Vec models with Neo4j, microservice GraphQL APIs with FastAPI.

[nyclodging](#) - A machine learning web app for exploring Airbnb listings

[R](#), [JavaScript](#)

- Developed an interactive website which predicts price ranges for a Airbnb listing using word embeddings and multinomial regression. Provides various customizing options for crafting user-defined geospatial visualizations.

SOFTWARE PROJECTS

pylubridate - A Python library of datetime utilities

[Python](#), [Rust](#)

- A Python library for working with datetime data with a Rust backend. Implemented parsing of ISO strings into date objects, common year, month, day extractors, and date arithmetics.

[qlang](#) - A toy programming language with a combination of R and Python

[Go](#)

- A interpreter written in Go with a mix of features in R, Python and JavaScript. Provided native support for vectorization, easier syntax for creating arrays and hashmaps, and first-class functions.

For more projects see [projects.qiushiyan.dev](#).

ADDITIONAL

Frameworks Tidyverse, Shiny, FastAPI, Airflow, Vue, React
Certification [AWS Certified Developer](#)