YUTONG YANG

1015 Barrington Street | Halifax, NS

YutongYang@dal.ca | 902-399-5288 | www.linkedin.com/in/yutong998

EDUCATION

Master of Digital Innovation (Data Science Certificate) 09/2022-12/2023 (Expected)
Dalhousie University, Halifax NS

- **Relevant coursework:** Data Analysis, Machine Learning, Process of Data Science

B.S. Econometrics & Quantitative Economics, GPA 3.94/4 05/2022
University of Illinois at Urbana-Champaign, Champaign IL, USA

- **Relevant coursework:** Applied Statistics, Applied Regression and Design, Statistics Programming Methods

SKILLS

Certificates: Microsoft Azure Fundamentals, Tableau Desktop Specialist, Coursera SQL for Data Science
Programming languages: Python, R, MySQL, HTML, CSS, JavaScript, PHP
Software/Packages: Tableau**,** Stata, MATLAB, SAP, NumPy, Matplotlib, Pandas, Seaborn, TensorFlow
General: LaTeX, GitHub, French (B1), English (Fluent), Chinese (Native)
Non-Technical Skills: Teamwork, Public Speaking, Critical Thinking, Self-starter, Growth Mindset

PROFESSIONAL EXPERIENCE

Part-time Analyst - China Securities Co., Ltd. *03/2020-08/2020*

- Managed transaction data in Chinese real estate market using ***MySQL***, generated monthly balance sheet and income statement, improved the overall data accessibility for the group.
- Collaborated with a team of 5 employees, researched business progress in the security industry, tracked industry trends under data scientist supervision.
- Created detailed visualization of numeric and categorical data using ***SAP analytics cloud*** and delivered 30-minute weekly analysis reports.

PROJECT EXPERIENCE

Zillow Housing Price Extractor | Python, requests, OpenStreetMap, Matplotlib, Pandas, Tableau *12/22–01/23*

- Extracted housing prices in U.S. from Zillow using ***Python requests***, cleaned data using ***Pandas.***
- Converted addresses into latitude and longitude using ***GeoPandas***, exported price and geospatial into CSV.
- Visualized median house price for each state by a ***Tableau*** map with color marks, applied states as filter to generate a list of top-20 listing price for each state with the house location.

Stroke Risk Prediction Using ML | Python, XGBoost, scikit-learn, TensorFlow, Seaborn 09/22–*12/22*

- Selected 12 variables with ***principal component analysis***, used ***Seaborn*** to visualize the latent space.
- Constructed 3 ML models, ***logistic regression, random forest*** and ***voting classifier***, to predict stroke risk.
- Used grid-search ***hyper-parameter optimization*** to improve model accuracy, best weighted f1 score of >0.9.

Personal Portfolio Website | HTML, CSS, JavaScript *08/22–09/22*

- Built a personal website with ***HTML*** and ***CSS*** for layouts, added interactive features using ***JavaScript*** functions, and utilized ***Google Sheets API*** for message board, published the website through ***GitHub.***
- Used as a platform to demonstrate my data analysis skill, constant updates with recently completed projects.

Amazon Price Tracker & Predictor | Python, Beautiful Soup, R, ggplot2, forecast *03/22–05/22*

- Scraped the pricing information automatically using ***Beautiful Soup*** given a URL for any Amazon product, exported the price and date information into a CSV file.
- Applied ***ARIMA time-series forecasting*** in R, captured the trend and seasonality of price fluctuation, achieved statistically significance with p-value <0.05.
- Visualized price prediction in one month using ***ggplot2***, delivered a 20-minute presentation in class.