Yukun (Edward) Zhang

Website | ykzhang1211@g.ucla.edu | +1 (604) 838-0430 | GitHub

With Microsoft certifications and 7+ years of programming experience in Python and R, I've built a solid foundation in Machine Learning (ML) for business and finance. To excel as the distinguished financial ML engineer in my written statement, I'm dedicated to deepening my understanding of large data information retrieval & ML algorithms by pursuing an advanced engineering degree.

EDUCATION

THE UNIVERSITY OF BRITISH COLUMBIA (UBC) 11/2023
Master of Data Science (MDS) advised by Dr. Varada Kolhatkar ***UBC GPA (cumulative): 4.33 / 4.33***
- Relevant Courses: Advanced Machine Learning & Statistical Modeling, Data Structure, Algorithms, Cloud Computing.

UNIVERSITY OF CALIFORNIA, LOS ANGELES (UCLA) 12/2018
Bachelor of Science in Mathematics/Economics ***UCLA GPA (major): 3.91 / 4.00***
- Cum Laude Latin Honors and the Dean's Honors ***UCLA GPA (cumulative): 3.81 / 4.00***
- Transferred from UNIVERSITY OF CALIFORNIA, SAN DIEGO (UCSD) ***UCSD GPA (major): 4.00 / 4.00***
- Major at UCSD: *Bachelor of Science in Mathematics (Applied)* ***UCSD GPA (cumulative): 3.85 / 4.00***
- Honors at UCSD: Magna Cum Laude Honor List, the Dean's Honors List and 4 Terms Provost Honors.
- Relevant Courses: Linear Algebra, Optimization, Differential Equations, Programming, Data Structure, Algorithms, Statistics.

SELECTED PROFESSIONAL SKILLS

- **GENERAL:** Machine Learning (RNN, GCN and Hidden Markov), Natural Language Processing (NLP), Data Structure, Algorithms, Data Science, Statistical Modeling, Optimization, Project Management, Risk Management.
- **TECHNICAL:** Python (Pytorch, scikit-learn), R, HTML, PostgreSQL & MySQL, MongoDB, QGIS, MATLAB, LaTeX, AWS & Azure Cloud Computing, Git/GitHub, Docker, MS Excel Macros VBA, Stata, Java and C++.
- **CERTIFICATE:** *Microsoft Certified: Azure AI Engineer Associate* ***MICROSOFT*** 07/2022
  Microsoft Certified: Azure Fundamentals ***MICROSOFT*** 03/2022
- **LANGUAGE:** English, Chinese (Mandarin), Japanese & rudimentary Korean and Mongolian.

SELECTED WORK EXPERIENCE

Wells Fargo Bank, Los Angeles, CA, USA
Assistant Vice President (AVP) under the Dept. of Corporate & Investment Banking 06/2020 - 03/2022
Financial Analyst I 01/2019 - 06/2020
- Monitored the Project Management Life Cycle phases and validated the implementation of the Iterative Waterfall framework.
- Queried and researched financial data in the tech industry using MySQL to adjust liquidation and counterparty risks.
- Identified opportunities to enhance the Client Evaluation Program and initiated the Rate Calculation & Comparison Project together with Software Development Life Cycle teams, **increasing client satisfaction by 26% and transaction volume by 21%**.
- Conducted Moody's Analytics pre-screening valuations for comparative & risk analysis on M&A projects in the TMT sector.

UCLA Department of Financial Planning, on-campus, Los Angeles, CA, USA
Junior Analyst 12/2017 - 07/2018
- Provided support for solution implementation at Ronald Reagan UCLA Medical Center, resulting in a **19% increase in revenue.**
- Created projections for three fiscal years using past five-year growth rates estimated by BETA Distribution in Excel data tables.
- Reconciled operating cash flow and assessed permanent endowment funds & real estate investment for fair market value.
- Designed Quartile Comparable Analysis (QCA) benchmarking with USC and other three UC campuses.

Merrill Lynch & Co. - Bank of America, Los Angeles, CA, USA
Financial Analyst Training Development Program 06/2017 - 11/2017
- Analyzed clients' real estate property values using models based on lognormal distributions and Monte Carlo Simulations.
- Performed the Value Chain Analysis on the supply side, and conducted Ratio Analysis & Comparison for the Client Demo Day.

SELECTED PROJECTS & DATA RESEARCH EXPERIENCE

I. Business Oriented Projects:

UBC Product Knowledge Graph NLP Project | Project Page **Python, R, AWS, Neo4j, Dash, Git**
Capstone Research Project under Dr. Varada Kolhatkar, UBC and partners from Cymax Group Inc. 05/2023 - 07/2023
- Managed and monitored project workflows on a weekly basis, adhering to the Agile development framework.
- Extracted product-entity relations from 23,000 furniture descriptions using the bert uncased model, finalized as the best for custom Named-entity Recognition fine-tuning, achieving a **0.8 exact match F1 test score** and a **31% improvement** over the benchmark).
- Adapted Spotify's ANNOY model to establish product-product similarity and built automated pipelines using Python Makefile.
- Collaborated in building a Neo4j product graph database and deploying a visualization Dash App on Render for e-commerce clients.

NYC Restaurants Health Inspection Analysis Project | Project Page **Python, R, Docker, Git**
Class Final Project for DSCI 522, UBC 2022 W I

- Designed the down-sampling approach for binary classifications to mitigate imbalanced data, aligning with our research objectives.
- Evaluated the performance of logistic regression and SVMs (balanced vs. imbalanced) with hyperparameter tuning and classification report analysis (including PR and ROC curves), leveraging data from NYC Open Data (achieving a **0.975 F1 test score**).
- Collaborated in developing automated pipelines in both Docker image and Python Makefile (with a full dependency diagram).

ML Benchmarking Models Prediction Project – The Epic Games Ratings | Project Page **Python, R, MongoDB, LaTeX, Git**
Research Assistant at Epic Games, Inc. 11/2021 - 03/2022

- Categorized users in restricted game data as node classifications on graphs and trained Graph Convolutional Networks (GCNs) and Multi-layer Perceptron (MLP) models using Nvidia RAPIDS cuML to predict user age categories (Minors, Adults, Middle-aged).
- Trained a stacking ensemble of KNN and bootstrap random forest algorithms in Python scikit-learn with finely tuned hyperparameters, achieving better game rating predictions with **0.87 RMSE** and a **12% improvement** over the benchmark model.
- Implemented a gaming recommendation system that leverages users' game-playing histories, utilizing a weighted average of hybrid approaches combining unsupervised collaborative filtering and supervised content-based filtering.

QGIS LA Local Market Analysis Project for Starbucks | Project Page **Python, R, SQL, QGIS, LaTeX, Git**
Class Final Project under Dr. Michael Shin, UCLA and the Undergraduate Business Society, UCLA 2017 Fall

- Analyzed the effects of population density, the presence of colleges and shopping centers within a 1-mile buffer, and income distribution on the QGIS Geo-distribution of Starbucks in LA, using the Python SQLite database from the US Census Bureau.
- Designed a linear mixed-effects model to analyze Starbucks revenues in ten US cities, estimating both nationwide and local effects.
- Developed a multi-layer Artificial Neural Network ANN model from scratch, **exclusively using basic NumPy and SciPy only**, for multi-class classification (achieving **0.97 test accuracy**).

II. Finance Oriented Projects:

BlackRock Troubled Waters ESG Investment Research Project | Project Page **Python, R, SQL, QGIS, LaTeX, Git**
Research Assistant at BlackRock, Inc. under Dr. Yuxi Suo 03/2020 - 07/2020

- Utilized Python and QGIS to develop a dynamic, multivariate 3D mapping for visualizing 2014-2020 global water shortages.
- Conducted a three-level quantile analysis on US REITs with low, middle, and high returns in 2014, using water allocation and scarcity stress panel data from Bloomberg, WRI, the United Nations World Water Report, and the SNL database for REIT properties.
- Investigated the implications of multivariate time series predictions for 2030 on the US REITs market and water stress level proportions, based on sensitivity benchmark results from deep learning memoryless 2D CNN, short-memory RNN, and LSTM.

US Finance Studies - The Effectiveness of the Big Four Auditors on Restatement Errors | Project Page **Python, R, LaTeX**
R Workshop Project under Dr. Ronghui (Lily) Xu, UCSD 2017 Spring

- Manipulated and classified the 2008-2016 internal audit data and US SEC multivariate raw panel data by firm industrial sector types and restatement error types, and applied STL decomposition and SARIMA to extract patterns of seasonality and autocorrelation.
- Confirmed the positive impact of the Big Four auditors in reducing the likelihood of committing restatement errors for their client firms, using binary logistic regression to compare firms with Big Four versus non-Big Four auditors.

US Stock Portfolio Return Analysis on SIC Industry and Market Capitalization | Project Page **R, Pine Script, LaTeX**
Class Final Project under Dr. Gordon B. Dahl, UCSD 2016 Spring

- Visualized time series US stock trends and indicators from 2008 to 2015 by implementing GMMA, Bollinger Bands, moving average systems, and the reversal "9" indicator from DeMARK Analytics in Pine Script, coded from scratch on TradingView.
- Calculated and reported the five-year annualized compounded returns of three selected stock portfolios from 2011 to 2015 in R, based on different SIC first digits, volatility, and market capitalization rankings.

III. Miscellaneous Projects:

US Labor Mobility Patterns Analysis | Project Page **Python, R, LaTeX**
Class Final Project under Dr. Abel Valenzuela and Dr. Kent Wong, UCLA 2018 Fall

- Conducted exploratory data analysis (EDA) using lagged variables in Python and R to compare movements across different regions and urban areas, utilizing panel data from the WBIS database and the US Bureau of Labor Statistics NLSY 79 survey.
- Analyzed and estimated the wage gap between movers and stayers across four selected regions and in urban/non-urban areas using plots and OLS regressions in R, controlling for factors such as age, employment, education, and gender.

US GDP Stata Modeling Analysis **Stata, R, LaTeX**
Class Final Project under Dr. Carroll B. Foster and Dr. James D. Hamilton, UCSD 2016 Summer

- Tested both the Lagged Regressor Model and the Lagged Dependent Variable Model in Stata and R, conducting annual GDP regressions on the money supply, labor force, and government purchases using 1973-2013 data from the US Federal Reserve.

UCSD Healthcare Labor Studies on Overtime Work Analysis for ER Physicians | Project Page **Python, R, LaTeX**
Class Final Project under Dr. Yuan Emily Tang, UCSD 2016 Spring

- Inspected data entry errors and preprocessed the data using POSIXct standard time format, focusing on patient flow and physician shift schedules in the emergency departments (ED) of UC San Diego Health Center.
- Analyzed ggplots depicting hourly patterns of patient arrivals and patient severity levels, and estimated the likelihood of ED physicians working overtime using binary logistic regression with four regressors.

TEACHING EXPERIENCE

Supplemental Student Self-Organized Workshop for DSCI 525 Web and Cloud Computing, UBC
Google Cloud Platform Workshop & Demo (80%+ of MDS cohort RSVP) 2022 W II

Student Services Workshop for Statistics, UBC (medium post last updated on 10/15/2023)
Instrumental Variables Estimation Workshop (150+ students RSVP at UBC) 2022 W I

Math 115A Linear Algebra, UCLA
Grading Reader (5.00/5.00 assessment review from the TAs and the professor) 2018 Fall

Coding in Finance Workshop Series hosted by the Undergraduate Business Society, UCLA
Python Algorithms Workshop Series (120+ students enrolled at UCLA) 2018 Spring

Math 33B Differential Equations, UCLA
Math Tutor (97.8% student recommendation) 2018 Winter

Math 181A Introduction to Mathematical Statistics I, UCSD
R Code Workshops (90%+ of Math 181A students enrolled) 2017 Spring

Math 170A Introduction to Numerical Analysis, UCSD
MATLAB Tutor (93.7% student recommendation) 2017 Winter

Econ 120B Econometrics B, UCSD
Stata Tutor (96.1% student recommendation) 2016 Fall

Math 20D Introduction to Differential Equations, UCSD
MATLAB Tutor (91.6% student recommendation) 2016 Summer I

LEADERSHIP EXPERIENCE & ACTIVITIES

Volunteer at the PwC x Microsoft Transformational AI Conference 06/22/2023

Contributor of the Women in Data Science Conference, Vancouver 2023 hosted by Microsoft & DataCan 04/14/2023
- Interacted closely with guest speakers and sponsors from companies such as Microsoft, Nvidia, and Google.
- Collaborated in organizing Google's IamRemarkable movement event and contributed to its extension to SheIsRemarkable.

Volunteer at the Pacific Conference on Artificial Intelligence at UBC 04/02/2023

Member of the DataCan Network, Vancouver 08/2022 - Present

Member of the UBC Japan Association 08/2022 - Present

Contributor of the Wells Fargo Women in Leadership Event, San Francisco 09/03/2021

Committee of the Undergraduate Business Society, UCLA 09/2017 - 12/2018
- Interacted closely with US domestic sponsors and global clients from EY, Cornerstone Research, Accenture, etc.
- Led the runner-up team "Bruin Ace" in the Merrill Lynch Private Equity LBO & Merger Model Case Competition.
- Managed a team to organize the UCLA 2018 Financial Services Night career fair.

Semi-marathon Runner 01/2019 - Present

Member of the Japanese Student Association, UCLA 09/2017 - 12/2018

Member of Local Badminton Club 09/2015 - Present