DIWEI "DAVID" ZHU

277 Gold Street, Brooklyn, NY 11201 | 917-864-6724 | dz1397@nyu.edu | https://chwdy.github.io/

SUMMARY

- Current NYU MS student, focused on Business data analysis, with GPA 3.87/4.0, majored in Industrial Engineering
- 3 years of experience on python, including Sklearn, NumPy, Pandas, seaborn, bs4, projects on machine learning
- 1 year experience on R, working on 3+ course projects, answering questions and grading in R as Teaching assistant
- Solid background on visualization, 4 years exp on making PowerPoint slides, Tableau Desktop specialist
- Experienced in business case analysis, communicating and providing solutions through presentations
- Fast learner, problem solver, coding enthusiasts, 10+ course project experience on ReactJS, Nodejs, Tableau

EDUCATION

New York University, Tandon School of Engineering, Brooklyn, New York **Sep 2019 - May 2021(Expected)**
MS in Industrial Engineering (focused on Business Data Analyst) GPA:3.87/4.00
Relevant Coursework: Machine Learning, Programming for BI & Analytics, Business Analytics, Project Management

Northeastern University, Shenyang, China **Sep 2015 - Jul 2019**
Bachelor of Business Administration, Industrial Engineering**,** GPA:3.50
Relevant Coursework: Quality Control & Reliability, System Engineering, Production Planning & Control

TECHNICAL SKILLS

- **Programming Languages:** Python (MatplotLib, Numpy, Sklearn), R (ggplot2, caret), Java(basic), C++, JavaScript
- **Certificate:** Tableau Desktop Specialist
- **Other Tools:** Git/GitHub, MySQL, MongoDB, Jupyter Notebook, AWS, R studio, NodeJS, ReactJS, Premiere Pro

PROFESSIONAL EXPERIENCE

SAP Implementation Project Intern, Capgemini Consulting, Guangzhou, China **Feb 2019 – May 2019**
Worked with middleware development team to build web app providing easy access to the ERP system on cellphone

- Generated testing environment datasets with SQL, using Navicat software to manipulate MySQL database
- Reported and followed up on 90% of bug cases with the software, assisted developers in reproducing issues
- Assisted leaders in presenting plans and progress to clients by creating concise slides and visuals on weekly basis
- Supported employee training by producing instructional video series and detailed, role-based manuals

Industrial Engineer Intern, Avic Shenyang Aircraft Co., Ltd., Shenyang, China **Jul 2018 – Aug 2018**
Focused on production line optimization and assist with the paperwork in the department

- Devised a research paper focus on high temperature and noise pollution in the factory
- Developed 2 solutions for these problems, including using electric power, air conditioners
- Monitored and tracked factory environment on temperature and noise using thermometer and decibel meter

Management Assistant, Metro Construction Business Center, Guangzhou, China **Jan 2018 – Feb 2018**

- Demonstrated company's accomplishments by producing a 4-minute video presentation
- Organized the annual meeting venue for 200 people by managing the audio and video setup
- Streamlined the reporting process on the database of supplier project progress information

PROJECT EXPERIENCE

YouTube trending video analysis (R, Git, ETL, Text analysis) **NYU - Dec 2020**

- Used R to provide 8 suggestions, 3 channel templates based on industry insights with genres, content and tags
- Performed text analysis; removed stop words, computed word recurrence and visualized via wordcloud
- Developed "engagement score" model to evaluate video's engagement with the audience and plotted in ggplots2

Computing System Architecture Course Project (C++) **NYU - Dec 2020**

- Implemented 5-stage pipeline of MIPS architecture with branch prediction and data forwarding
- Simulated a 2-bit saturated branch predictor indexed by the Least significant bit (LSB)
- Designed a 2-level cache that handles data miss by round-robin eviction and configurable blocks and way sizes

Project management course project (Gantt chart, Bootstrap) **NYU - Dec 2020**

- Piloted a meditation guide with step-by-step instructions for beginners to start meditation in a team of 5
- Reduced the time spent on scheduling by modelling a formula to automate time allocation calculation on excel
- Designed a website using bootstrap to showcase audio meditation guide and demonstrated it in presentation

Home credit default risk project (R, Caret, ggplot2, LGBM, Random Forest) **NYU - Oct 2020**

- Predicted the clients' probability of loan repayment with loan history and credit card usage using R in team of 5
- Balanced the dataset by oversampling and performed grid search for machine learning models using caret
- Visualized the feature engineering process using ggplot2 and demonstrated the result and analysis to our audience
- Compared models' performance and chose gradient boosting tree because of its high accuracy of 73%

Prospective Student Persona Modelling (R, ETL, Dplyr) **NYU - Oct 2020**

- Derived 5 personas of potential students using R for Marketing purpose with database on Rivier University
- Developed pipelining function and conducted grid search to find compelling subgroup with feature combinations
- Generated 5 key features for marketing department and focus on Persona that covered 54.3% of overall population

Economic effect of COVID-19 (Python, Sklearn, K-mean) **NYU - May 2020**

- Collaborated with a team of 5 to collect data and analyze the relationship between COVID and GDP growth.
- Developed a linear regression model to predict the economic effects based on education and economics
- Performed data clustering using K Means and preprocessed the data by handling missing values and joining data
- Visualized the results and clusters using Matplotlib and presented the results to the class of 30

Analysis of Volatility and Fractal Dimension (Python) **NYU - May 2020**

- Interpreted the Volatility and Hurst coefficient of 5 currencies to define its correlation with time
- Formulated algorithms to calculate volatility, using the Hurst Coefficients, using 'compute_HC' package
- Devised regression model prediction, grid search, visualization and evaluated using the r squared values

Student Performance Prediction (Python, ETL, Machine learning) **NYU – Apr 2020**

- Predicted student exam scores based on their background using student performance dataset
- Preprocessed data by variance threshold, and one-hot encoder to explore data relationship by the correlation matrix.
- Clustered the performance by Kmeans and selected features by RFE (recursive feature elimination).
- Trained and evaluated supervised models included logistic regression, naïve Bayes and decision tree for comparison

ADDITIONAL COURSEWORK

Cloth retail analysis (Tableau) **Jul 2020**

- Rectified unnecessary data using data interpreter to map it into quantified pivot tables
- Applied trend lines to predict future values using Time-based Linear Regression
- Demonstrated the best locations for cloth sales in Australia using the "story" feature

Coffee store expansion analysis (Tableau) **Jul 2020**

- Navigated the store expansion candidates by deriving their potential based on the location and sales data
- Developed a successful model that allows customizations and automation for calculating potential candidates

Coal terminal machine failure prediction (Tableau) **Jun 2020**

- Visualized a forecast of potential failure and idle capacity of coal machine using the trend line graphics in Tableau
- Built an efficiency plan via showcasing a validated analysis of weekly calculations to render idle times

FRONT END

Yelpcamp (Node.js) **Jul 2020**

- Perfected a secure website to share campus information using Express and MongoDB
- Deployed user authorization and authentication with express-route middleware and "passport" module
- Embellished the website with the grid system and styles in Bootstrap 4 as well as icons from font-awesome

TODO list app (Node.js, RESTful) **Jul 2020**

- Built a single page to-do list app based on Express framework with RESTful design
- Deployed Mongo DB and connected to Nodejs through the mongoose package
- Animated buttons with CSS3 animation along with fonts and colors on the webpage

Warbler (Node.js, Express, React, Redux) **Aug 2020**

- Built a message board app with the client-server structure using express and mongo DB
- Deployed the app on Heroku and made it capable of switching mongo DB connections based on the environment.
- Developed Restful API interface on the server and applied middleware using web token for user identification
- Implemented React - Redux structure for clients, capable for fast-paced development and easy data access.

IE-RELATED PROJECTS

Production science individual project **NYU - Mar 2020**

- Supervised simulations and inspection of the production line for 500 days using Excel data

- Verified the system's stability by applying the Little's Law on input rate, WIP, variation of output.
- Mapped the improvement plans with comparison on performance by changing the throughput and variance.
- Automated simulation run by VBA script in Excel, capable of recording the data of 10 runs in one click.

HAL factory production line analysis project **NYU - May 2020**

- Upgraded daily production from 1200 to 3000 with HAL factory's production data in team of 4
- Boosted the analyses process by setting up Excel formula and determining the bottlenecks
- Forged long term improvement plan on labor, machine, and timetable aspects with visualized graphs

Factory simulation final project (Witness) **NYU - Nov 2019**

- Refined the Steel Nugget manufacturing process to improve production capacity using Witness software
- Modeled a restructure report after simulating effects by labor, delivery speed and pallet count
- Customized line performance with zero-spending option, changed priority of products and its maintenance

Inventory simulation (Python) **NYU - Oct 2019**

- Crafted warehouse inventory system in Python, using given data of lead time and demand distribution data
- Simplified the searching and reviewing process by using the Prettytable package to generate tables
- Compared and analyzed inventory and backorder level under different pairs of order-level and reorder line.

EXTRACURRICULAR ACTIVITIES

Team member, Honorable Mention, *2018 MCM/ICM Contest,* **Feb 2018**

- Predicted countries fragile index by constructing Linear Regression models on MATLAB in a team of 3
- Verified the model by segmenting the three countries and predicting their indexes
- Performed data cleanup by filling in the missing values and encoding data types on Excel