

Yangkai Ren

Data Analyst

(201)4063199 | yren.allen@gmail.com | https://yangkaiaren.github.io/Kai_Profolio/

SUMMARY

Highly motivated professional with experience in Data Analytics and Data Scientists. Passionate about Data Mining, Machine Learning, Big Data, Data Visualization. Love to learn new knowledge and enjoy working under great pressure.

PROFESSIONAL PROJECTS

Used car prices prediction contest project in Tianchi of Alibaba. 2021

- Analysis based on the used car market and requires players to predict the transaction price of used cars. Utilized MAE(Mean Absolute Error) technology to evaluate models.
- Scored 439.83 and Ranked top 3% in 12912 competing teams ($394 / 12912 = 3\%$).
- Used mixed models (LGMB and CatBoost) based on EDA, preprocessing, and feature engineering. Optimized hyperparameter using GridsearchCV to reach the best model.
- The total data volume exceeds 400k and contains 31 columns of variable information, 15 of which are anonymous variables.

Machine Learning Analysis on Census Bureau Dataset. 2021

- Collection Dataset from UCI Machine Learning Repository. Training a binary classifier to predict the income (50k instances and 14 attributes in the dataset).
- Used Logistic Regression classifier to classify the dataset and utilized confusion Matrix and accuracy score to find the accuracy of the model. Finally, achieve 85.2% accuracy.
- Developed an ML pipeline to automate the iterative processes of data transformation and building estimators. The pipelines offer efficient methods to automatically apply to preprocess and feature engineering to numerical and category features.

EDUCATION

Harrisburg University, PA

Master of Analytic

Sep.2021 – present

Master of Computer Information Sciences, GPA 3.7/4.0

Sep.2019 – Aug.2021

Related courses: Data Engineering & Mining, Big Data Architectures, Data Struct & Algorithms, Scientific Computing

School of Visual Arts, NYC

Master of Computer Arts

Sep.2015 – May.2017

University of Electronic Science and Technology, Sichuan, China

Bachelor of Digital Media

Sep.2008 – Jun 2012

SKILLS AND TOOLS

Skills: Data Analytics, Data Mining, Machine Learning, Visualization, Big Data, Data Collection, Statistical Modeling, CI/CD, Reporting.

Tools: Python, R, SQL, C#, Bash, AWS, Hadoop, Spark, Git.

EXPERIENCE

WeDesign (Art and design interactive and experiential education), NYC

Teaching Assistant

Nov.2018 – Jun.2019

- Collecting data and analyzing the entry requirement and history data of each university.
- Mentored students to apply for university like NYU game design program.
- Design a promoted study plan for students one-on-one tutoring. Improved students' learning efficiency by 3 times.

Leshan TV Station (The Biggest TV station of Leshan City in China), Leshan City, Sichuan, China

Journalist

2012 – 2015

- Collecting and assessing all kinds of social information and data for news.
- Reported and written over 400 news. Making the ratings the public livelihood news <News Everyday> to be the top all the year round locally.
- Analyzed the proportion and distribution of local poverty. Planning special reports for poverty alleviation. Helped 3 poor students get funds to continue their studies and many dropout children go back to school.