Qin Wu

E-mail: wuqin@wustl.edu | Address: Gainesville, FL

SUMMARY

Information System Management graduate student seeks data scientist opportunities, with sufficient project experience in machine learning and data analytics. Solid skills in Python, SQL, Tableau. Have a complete understanding of Hadoop framework.

EDUCATION BACKGROUND

Washington University in St. Louis   Saint Louis, Missouri
Master of Information system management   08/2019 - 01/2021
- GPA: 3.87/4.00
- Major Courses: Data Mining, Optimization, Applied Data Science for Practitioner

Sichuan Agricultural University   Ya'an, China
Bachelor of Engineering in Computer Science and Technology   09/2015 - 06/2019
- Average Scores: 87.8/100.0
- Major Courses: Data Structures and Algorithms, Principle and Application of Database, Computer Network, Computer Composition Principle, Probability Theory and Mathematical Statistics

COMPUTER SKILLS

- Proficient in: Python, SQL, Tableau, Linux, MATLAB, C, JAVA

RESEARCH EXPERIENCE

MINST Handwritten dataset   10/2019 - 11/2019
- Apply Keras package to deploy the convolutional neural network; customize a VAE model; Calculate the mean and variance of each number k; Construct normal distributions and perform sampling reconstruction.
- Customize a convolutional layer to calculate the loss between the input image and the decoded image.
- Apply cosine similarity comparison method to compare the similarity of the images; Modify the number of dimensions and layers to generate the best parameters of the model.
- Plot 3-D normal distributions to show both Vk (V0) and Vk' (the difference between V0~9 and V1~9).
- Change the mean value by increasing the i-fold delta value, and compare the similarity of the two sets of data again to determine the best parameters.

Regression and classification analysis on UCI wine dataset   03/2020 - 04/2020
- Preprocess dataset by data cleaning, categorical feature transformation, missing value filling and standardization, etc.
- Visually analyze the data to determine the analysis steps and methods.
- Execute linear regression, Lasso regression, and random forest to analyze the 14 chemical information variables of wine, obtain the model to calculate the predicted value; evaluate the model through the average absolute error.
- Implement logistic regression, K-nearest neighbor algorithm and self-built model to learn the training set and evaluate the model through the average absolute error.
- Refine the model by finding the best parameters using grid search (GridSearchCV). The final accuracy rate reached 99.8%.

Data visualization and predictive analysis of electronic game sales   10/2020 - 11/2020
- Preprocess dataset by missing value filling and standardization. Execute exploratory data analysis.
- Visualize data to show the commercial value of the data, such as obtaining the most popular game types in the world, the publisher with the most games and the game platform with the most sales.
- Based on the information of game type, publisher and platform, implementing one-hot coding and establish a linear regression model to predict the sales volume of the game.

Leadership Experience

Director of Propaganda Dept., Association of SAU Clubs   05/2016 - 06/2017
- Responsible for the new club member recruitment in the 2016 Hundred Regiment Campaign. Recruit nearly 10,000 participants. Design and issue 1,000 brochures for the activity.
- Participate in discussions and decision-making with the instructor and the Bureau of the Social Federation, and responsible for the overall arrangement of departmental work.
- Propose and implement the WeChat Official Account for the Association; Take charge of daily management of the account and gain 2000 follows during one-year's operations.