Jun Zhu
Xi'an Jiaotong University
Xi'an, China 710049

(+86) 187 9270 2157 (cell)
jszhujun2010@126.com
jszhujun2010.farbox.com

Objective

Data mining researcher, Nanjing

Education

2014– **Master student, Computer Science, Xi'an Jiaotong University**

Research field: Automatic program repair

MOOC lover: machine learning foundations, machine learning techniques, algorithms, artificial intelligence, the analytics edge(machine learning, optimization, visualization)

2010–2014 **BE, Information Egineering, Xi'an Jiaotong University**

top 10 student, combined bachelor-master degree programme class

Thesis: Supergraph query algorithm research

Project experience

Network video client

This is a part of a video signal transforming project. The main purpose of the project is to transform the TS stream(on the wall) to network stream through MCU, and then transport it to PC. The MCU acts like a web server, and I made a GUI client on the PC side to commnunicate with server, and receive network stream to display on clients' screen.

It is implemented by Python and GUI is designed through PyQt.

Pacman game

It's actually a project of artificial intelligence course in University California, Berkley. There is a pacman in the maze which is pre-defined, there are some beans and some ghosts in the maze. The ultimate goal is to eat all the beans while not be caught by the ghosts(not hit the maze wall, of course) as fast as possible.

It is implemented by Python and many important search algorithms are involved.

Tianchi competition, 2015

282th/7186

This is a data mining competition held by Alibaba. Customer shopping data from 2014/11/18 to 2014/12/17 are provided as traning data, and the goal is to predict 2014/12/18 shopping data.

The competition has two phases, the first phase training data size is 500M, while TB scale data is invloved in the second phase.

Kaggle competition

186th/2329

It is a text mining project which is held by MIT Edx course, the analytics edge. Given the fully text of blogs, we are asked to predict whether a particular blog will be popular.

The bag of words model is used, and many useful features are extracted from the text. Several classification algorithms are applied and finally ensembled.

Skills

English

- CET4: 601, CET6: 535

Programming languages

- Python, C/C++, Java

Machine learning & Data Mining

- R, Pandas + Scikit-learn, Spark MLLib

Honors and social experience

Siyuan scholarship, 2011, 2012

Volunteer of recruiting office, Xi'an Jiaotong University