Haojie Ma

18151681382 · 1137633684@qq.com · angelhunt ·

Information

1994/07/09 Job intention: Algorithm engineer Work experience: near 1 year

Education

- Master, Nanjing University, Computer Science and Technology, 2017.9~2020.7
- Bachelor, Nanjing University, Computer Science and Technology, 2013.9~2017.7

Experience

- **Microsoft China, m365 department, 2020.6~2021.4**
  Big data platform user/cluster anomaly detection
  Data lineage system based on static analysis of scope syntax tree

Research

- Two CCF-A conference papers
- Unsupervised Human Activity Representation Learning with Multi-task Deep Clustering (**UbiComp 2020**)
  - Constructed multi-task unsupervised deep clustering for human activity recognition
  - The model obtains hidden layer representations through joint learning of the three tasks of clustering, classification and auto-encoding
  - The model can surpass mainstream unsupervised methods without labels, and the performance of supervised algorithms can be approximated by fine-tuning with a small number of labels
- AttnSense: Multi-level Attention Mechanism For Multimodal Human Activity Recognition (**Ijcai 2019**)
  - Human behavior recognition in multi-modal scenes
  - Dynamic fusion of multi-modal sensor signals is carried out through the self attention model

Project

- Big data cluster anomaly detection
  - Policy-based detection and alarm, alarm merging mechanism
  - PCA-based user behavior anomaly detection and blackout
  - Cluster node abnormality detection, such as task blocking, job failure rate is too high
- Accelerate the performance of spark MLlib algorithm based on Intel MKL (undergraduate completion)
  - Implemented PCA, ALS, Logistic Regression and other algorithms through MKL
- NJU-ProjectN
  - Implement a complete pipeline CPU, c language subset compiler, MIPS simulator and simple operating system
  - Run the MIPS CPU, x86 hardware simulator, operating system, and compiler implemented by yourself on FPGA, and compile and run Legend of Sword and Fairy
- Some small project
  - Machine learning algorithms implemented in python, such as decision trees, naive Bayes
  - Parallel machine learning algorithms implemented by libraries such as MPI and cuda
  - High-performance KV storage, lsm tree based storage engine
  - Use Nlp technology to analyze Jin Yong's novels, implement a multi-label clustering algorithm, and display the relationship diagram of Jin Yong's novels

Award

- 2015-2017 Nanjing University People's Scholarship
- 2016-2017 Second Prize of the Second Program Design Competition of Nanjing University
- 2019-2020 National Scholarship