

XU HAN

170C Discovery Learning Center (DLC) ◇ 430 UCB, Boulder, CO, 80309
(+1) 720 3451591 ◇ xuha2442@colorado.edu ◇ <https://knights207210.github.io>

EDUCATION

Ph.D. Student in Computer Science

Sep 2017 - May 2022

University of Colorado Boulder, Boulder, USA

Research Interests: Human Computer Interaction, Data Science, Data Visualization

Advisor: Tom Yeh

GPA: 3.97 /4.0

B.E. in Electronic Engineering

Sep 2013 - Jun 2017

University of Science and Technology of China, Hefei, China

Undergraduate Thesis Advisor: Joern Ostermann, Zhibo Chen

PROFESSIONAL EXPERIENCES

Big Data Architecture Based on Twitter Data

Feb 2019 - Apr 2019

University of Colorado Boulder, USA

- Implemented a big data framework based on twitter data with Kafka (streaming processing), Pyspark (sentimental analysis with NLTK), MongoDB, Elastic Search & Kibana (visualization)
- A dashboard was built for visualization of sentimental data analysis results using Flask

Alexa Skills Evaluation

Jun 2018 - Sep 2018

Sikuli Lab, University of Colorado Boulder, USA

- Developed a voice skill crawler to collect responses data from 45708 Alexa skills (Python)
- Deployed DL models with pre-trained & self-trained embeddings to do Alexa skills' topic classification (PyTorch)
- Analyzed responses data of the most 100 popular Alexa skills and studied these skills' compliance situations of current design guidelines
- Proposed research agenda for future evaluation tool development

Recommendation System Implementation and Visualization Based on Yelp Data

March 2018 - May 2018

University of Colorado Boulder, USA

- Deployed machine learning methods to build a recommendation system based on the prediction of user's rating for business (Python, SQL), the result of Collaborative Filtering algorithm as references
- Blended several machine learning models by assigning weights and achieved best performance (AUC = 0.82, outperforming baseline (AUC by CF = 0.65) by 26.2%)
- Visualized the selected features, feature rankings and model training process (Javascript, HTML & CSS)

Repeat Buyer Prediction Based on User Online Shopping Logs

Oct 2017 - Dec 2017

University of Colorado Boulder, USA

- Deployed machine learning techniques to predict whether a customer will become a repeat buyer (Python, SQL)
- Focused on feature engineering, extracted 129 features (user demographic features, user behavioral features, merchant-related features, user-merchant interaction features, big promotion features, slope features, similarity features, PCA features) and conducted feature ranking analysis
- Ranked top 25 in the competition panel

AWARDS

Graduate Travel Grant, University of Colorado Boulder	2019
CSC(China Scholarship Council) National Scholarship	2017
UTSIP Scholarship, University of Tokyo	2016
Outstanding Student Scholarship, University of Science and Technology of China	2016, 2015, 2014
The Best Technology Award, Robogame2015, University of Science and Technology of China	2015

RELEVANT SKILLS

Languages: C, C++, Python, Javascript, HTML & CSS, SQL, MATLAB
System & Library: PostgreSQL, Git, Linux, OpenGL, D3, Kibana, ElasticSearch, PyTorch
Software: Tableau

PUBLICATIONS

- [1] **Xu Han** and Tom Yeh. 2019. Evaluating Voice Applications by User-Aware Design Guidelines Using an Automatic Voice Crawler. *In Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, March 20, 2019*, 4 pages.
- [2] **Xu, Han**, and Tom Yeh. 2019. Evaluating Voice Skills by Design Guidelines Using an Automatic Voice Crawler. *arXiv [cs.HC]*. *arXiv*. <http://arxiv.org/abs/1906.01122>.
- [3] Shiroma, K., Davis, N., Yeh, T., **Xu, H.**, Sagna, A., & Xie, B. (2019). Co-Designing eHealth tutorials with and for older adults. *To be presented at the Healthier Texas Summit. October 17 - 18, 2019. Austin, Texas.*