

# Xingzhong Du

Location: Melbourne, VIC, Australia

E-mail: domainxz@gmail.com

Web Page: <https://dxz.bitbucket.io>

Linkedin: <https://www.linkedin.com/in/domainxz>

Github: <https://github.com/domainxz>

Kaggle: <https://www.kaggle.com/redflack>

Mobile: +61401271801

## Education

- **University of Queensland** Brisbane, QLD, Australia  
*Ph.D. in Computer Science* 2013 - 2017
  - **Research Topic: Multi-level Video Filtering Using Non-textual Contents**
  - Supervisors: Prof. Xiaofang Zhou & Prof. Yi Yang
- **Nanjing University** Nanjing, Jiangsu, China  
*M.S. in Computer Science* 2010 - 2013
  - **Research Topic: Attribute Entropy Personalized Recommendation**
  - Supervisor: Prof. Zhenyu Chen
- **Nanjing University** Nanjing, Jiangsu, China  
*B.Sc. in Software Engineering* 2006 - 2010

## Working Experience

- **Data Scientist** Melbourne, VIC, Australia  
*Wesfarmers Advanced Analytic Centre* 01/2022 - present
  - Main duties:
    - \* Diagnose and improve customer contact strategy
    - \* Building process to match products from external entities
  - Other duties:
    - \* Sharing knowledge.
- **Data Scientist** Melbourne, VIC, Australia  
*Coles Advanced Analytic Centre* 03/2019 - 01/2022
  - Main duties:
    - \* Improving personalised e-mail marketing and customer segmentation for 6 million flybuys members.
    - \* Building customer segmentations
    - \* Building customer feature table
    - \* Improving store range models based on customer data.
    - \* Designing, developing and operating production jobs.
  - Other duties:
    - \* Mentoring graduates.
    - \* Planning projects
    - \* Sharing knowledge.

- **Tutor** Melbourne, VIC, Australia  
*Monash University* 07/2018 - 02/2019
  - Task: tutoring and grading for three courses: advanced data mining (FIT5142), applied data science (FIT5149) and information retrieval system (FIT5166)
- **Contract Data Scientist** Brisbane, QLD, Australia  
*4impact* 04/2018 - 06/2018
  - Vision: time series analysis and prediction for input temperatures of power stations
  - Task: given a sequence of past temperatures, predicting the temperature per minute next 12 hours
  - Responsibility: data cleaning and feature analysis, data wrangling, predictive model design and parameter tuning, post-process, interface design, integration, code repository maintain
- **Casual Research Assistant** Brisbane, QLD, Australia  
*The University of Queensland* 04/2018 - 06/2018
  - Vision: clothing style mining and recommending using deep learning
  - Task: given several pieces of clothing, predicting the coordination
  - Responsibility: data cleaning, data wrangling, model design and tuning, model evaluation, code repository maintain
- **Contract Data Scientist** Brisbane, QLD, Australia  
*SAP* 03/2018 - 04/2018
  - Vision: learning based custom retention analysis and prediction
  - Task: given customer behaviour data, generating a short list of candidates who will end contracts
  - Responsibility: event exploration, predictive model design and parameter tuning, model ensemble, code repository maintain
- **Research Assistant** Brisbane, QLD, Australia  
*The University of Queensland* 11/2017 - 03/2018
  - Vision: clothing style mining and recommending using deep learning
  - Task: a piece of clothing, predicting the coordination
  - Responsibility: data cleaning, data wrangling, model design and tuning, model evaluation, code repository maintain
- **Research Assistant** Brisbane, QLD, Australia  
*Oracle Labs Australia* 05/2017 - 11/2017
  - Vision: Learning-based Bug Detection Using Code Property Analysis
  - Task: given C or C++ source code, detecting 8 kinds of bugs
  - Responsibility: data cleaning, data wrangling, model design and tuning, model evaluation, report generation, code repository maintain

## Skills

- **Programming Skills**
  - **Experienced:** C, C++, Java, Python, R, SQL, Shell
  - **Familiar:** cuda C, Matlab, SQL, Html, JavaScript, CSS
- **Platform Skills**
  - **Experienced:** Linux, Spark

- **Machine Learning Skills**

- **Classical Model:** K-means clustering, linear and kernel SVM, linear and kernel regression, Weighted Matrix Factorization (WMF) and its variants, PCA, Bayesian personalized ranking (BPR), xgboost tree, random forest, Gaussian mixture model (GMM), latent Dirichlet allocation (LDA)
- **Deep Learning Model:** convolutional neural network (CNN), recurrent neural network (RNN), multiple layer perception (MLP), stacked de-noising auto-encoder (SDAE), variational auto-encoder (VAE)

- **Project Management and Collaboration Skills**

- **Experienced:** Jira, Bitbucket, GitLab, Slack, Microsoft Teams, Commitizen, Auto-changelog, Git flow, R markdown, Jupyter notebook and lab

- **Certifications**

- IBM Certified for Advanced Electronic-Business Engineering
- Qualification Certificate of Software Designer, China

## Selected Data Science Competition Experience

- **Google Cloud & YouTube-8M Video Understanding Challenge @ Kaggle**

*Solo*

2017

- Task: Predict the tags of the videos
- Position: 98<sup>th</sup> of 650, top 15%
- Tools: python, tensorflow

- **State Farm Distracted Driver Detection @ Kaggle**

*Solo*

2016

- Task: Identify drivers' actions from images
- Position: 106<sup>th</sup> of 1440, top 8%
- Tools: python, caffe, keras, pandas, tensorflow, and torch

- **Santander Customer Satisfaction @ Kaggle**

*Solo*

2016

- Task: Predict which customer will break contract in next month
- Position: 660<sup>th</sup> of 5123, top 13%
- Tools: xgboost, R

- **WARD-CMU @ TRECVID 2015**

*Group Member*

2015

- Primary member in surveillance event detection (SED)
- Final position in automatic detection task: 1st
- PersonRun event has achieved the highest accuracy (around 50%) until now.
- Tools: sklearn, opencv, ffmpeg, python

- **Infomedia@TRECVID 2014 MED, MER and SED**

*Group Member*

2014

- Primary member in surveillance event detection (SED)
- Final position in automatic detection task: 2nd

- Final position in interactive detection task: 1st
- Tools: sklearn, opencv, ffmpeg, python

## Selected Publications

1. **Xingzhong Du**, Hongzhi Yin, Ling Chen, Yang Wang, Yi Yang, and Xiaofang Zhou, “Personalized Video Recommendation Using Rich Contents from Videos”, TKDE, 2018
2. **Xingzhong Du**, Feiping Nie, Weiqing Wang, Yi Yang, and Xiaofang Zhou, “Exploiting Combination Effect for Unsupervised Feature Selection by  $\ell_{2,0}$  Norm”, TNNLS, 2018
3. **Xingzhong Du**, Hongzhi Yin, Zi Huang, Yi Yang, Xiaofang Zhou, “Exploiting Detected Visual Objects for Frame-level Video Filtering”, WWW journal, 2018
4. Weiqing Wang, Hongzhi Yin, Zi Huang, Qinyong Wang, **Xingzhong Du**, Quoc Viet Hung Nguyen, “Streaming Ranking Based Recommender Systems”, SIGIR, 2018
5. Luyao Liu, **Xingzhong Du**, Lei Zhu, Fumin Shen and Zi Huang, “Discrete Binary Hashing Towards Efficient Fashion Recommendation”, DASFAA, 2018
6. Yan Zhang, Hongzhi Yin, Zi Huang, **Xingzhong Du**, Guowu Yang, Defu Lian, “Discrete Deep Learning for Fast Content-Aware Recommendation”, WSDM, 2018
7. Zhicheng Zhao, Xuanchong Li, **Xingzhong Du**, Qi Chen, Yanyun Zhao, Fei Su, Xiaojun Chang, Alexander G.Hauptmann, “A unified framework with a benchmark dataset for surveillance event detection”, Neurocomputing, 2017
8. Hongzhi Yin, Liang Chen, Weiqing Wang, **Xingzhong Du**, Quoc Viet Hung Nguyen, Xiaofang Zhou, “Mobi-SAGE: A Sparse Additive Generative Model for Mobile App Recommendation”, ICDE, 2017
9. **Xingzhong Du**, Hongzhi Yin, Zi Huang, Yi Yang, Xiaofang Zhou, “Using Detected Visual Objects to Index Video Database” (**Best Paper Award**), ADC, page 333–345, 2016
10. **Xingzhong Du**, Yan Yan, Pingbo Pan, Guodong Long, Lei Zhao, “Multiple Graph Unsupervised Feature Selection”, Signal Processing, Elsevier, page 754–760, 2016
11. **Xingzhong Du**, Xuanchong Li, Xiaofang Zhou, Alexander Hauptmann, “WARD-CMU @ TRECVID 2015”, Proceedings of TRECVID, 2015
12. **Xingzhong Du**, Tieke He, Zhenyu Chen, Jia Liu, Chengfeng Hui, “ABEY: An Incremental Personalized Method Based on Attribute Entropy for Recommender Systems”, SEKE, page 318–321, 2013