

# ZHANG, ZIHAN

488 Xiangyang Road, Xinxiang City, Henan Province, 453000, P. R. China

Tel: +86 188 1307 2016    ◇    Email: s1719186 @ ed.ac.uk

## INTRODUCTION

---

I am a postgraduate student in University of Edinburgh. I have completed all the courses and the dissertation of Data Science programme, and will graduate in Nov. 2018. Prior to that, I studied in a National Mathematical Training Base at Shandong University, where I built strong mathematical and statistical skills. My research interest is Machine Learning & Natural Language Processing.

## EDUCATION

---

**Master of Science with Distinction @ University of Edinburgh (UK)** Sep. 2017 - Nov. 2018  
Data Science @ School of Informatics **GPA: 78/100 Thesis: 72/100**

Core Courses: Machine Learning and Pattern Recognition (69), Machine Learning Practical (84), Probabilistic Modelling and Reasoning (90), Data Mining and Exploration (80), etc.

**Bachelor of Science @ Shandong University (985 & 211 project, China)** Sep. 2013 - Jun. 2017  
Mathematics and Applied Mathematics @ School of Mathematics **GPA: 85/100 or 4.3/5.0**

Studied at *Hua Luogeng Class*, a National Mathematical Training Base, ranked top 3 in China by specialization. Related Courses: Probability Theory (89), Mathematical Statistics (95), Applied Statistics (88), etc.

## SKILLS

---

<b>Advanced</b>	Python (TensorFlow, Keras, Pandas, sk-learn, etc.), Linux, Hadoop
<b>Beginner</b>	C, C++, Java, SAS, SQL, Matlab, R

## RESEARCH & PROJECTS

---

**TensorFlow Applied to Neural Language Models** Jun. 2018 - Aug. 2018  
*MSc thesis, supervised by Dr. Ben Allison (Amazon) and Dr. Tania Bakhos (Amazon).* [pdf](#) and [code](#)

- Implement Neural Language Models with Mixture of Softmaxes using TensorFlow to solve the Softmax bottleneck.
- Optimize hyperparameters mainly by specific Random Search; explore the influence of different LSTM implementations and the number of Softmax components.
- Speed up training by: training models in parallel on multiple GPUs; weight tying; weight matrix factorization; Gradient Clipping.
- The final model has lower validation and test perplexities as well as higher speed, compared with the original model.
- Compare the final TensorFlow model with the same model implemented using PyTorch.

**Marketing: Predicting Customer Behaviors** Mar. 2018 - Apr. 2018  
*Coursework, with Yi Wei, Wen Jia and Jiayu Li.* [pdf](#)

- Preprocess the large and dirty dataset provided by KDD Cup 2009 competition using pandas, making it clean enough to feed neural networks.
- Clip variables with large amounts of missing data; fill in the missing data; encode categorical data by one hot; reduce dimensions by linear or unlinear methods including PCA, Isomap LLE.
- Predict behaviors by deep models; optimize hyperparameters; evaluate the result by AUC scores.
- In terms of customer churn, one of the three behaviors, the prediction AUC exceeds the first place of the competition.

## Instance Segmentation of Nucleus Images

Feb. 2018 - Apr. 2018

*Coursework, with Yi Wei and Zhengjun Yue, supervised by Prof. Steve Renals.*

[pdf](#)

- Segment nucleus instances from nucleus images acquired in different microscopy systems, which is provided by Kaggle 2018 Data Science Bowl.
- In the baseline experiments, use fully convolutional networks with VGG16 backbone to semantically segment nuclei from the background, and use open operation to separate individual nuclei.
- In the contrast experiments, apply Mask RCNN to direct segment nucleus instances. Extract features by ResNet101 with Feature Pyramid Networks; select regions of interest by Region Proposal Networks; semantically segment the nuclei in all regions by fully convolutional networks.
- Implement the two models and pre-train them on COCO datasets; implement evaluation criteria by mean average precision; carefully tune hyperparameters. Mask RCNN significantly outperforms the baseline model.

## AWARDS & HONORS

---

Informatics International Masters Scholarship	Nov. 2017
Hua Luogeng Scholarship	Dec. 2016
Honorable Mention @ Interdisciplinary Contest in Modeling (US)	Apr. 2016
First Prize @ Shandong Province Mathematics Competitions (China)	Dec. 2015
Third Prize @ Chinese Mathematics Competitions	Nov. 2015
Second Prize @ China Undergraduate Mathematical Contest in Modelling (Shandong Province)	Oct. 2015
Meritorious Winner @ Certificate Authority Cup International Mathematical Contest in Modelling	Jan. 2015
Merit (96/100) @ National Computer Rank Examination (Level 2) with specialization in C language	Apr. 2014

## EXPERIENCE

---

Volunteer @ the 22 <sup>nd</sup> International Congress of Historical Sciences	Aug. 2015
Volunteer @ the 13 <sup>th</sup> National Key Middle School Principals Forum	Apr. 2015
Volunteer @ the 11 <sup>th</sup> Congress of National Federation of Industry and Commerce	Dec. 2014
Manager @ Billiards Association of School of Management in Shandong University	Oct. 2014 - Jun. 2015
Director @ Shandong University Student Union Office	Sep. 2014 - Jun. 2015

## HOBBIES

---

I am interested in table tennis, billiards, board games and movies; I love travelling and visited more than ten European countries in the last two years.