

# 张 侗冬

+61 420 264 362

[allenzdd@gmail.com](mailto:allenzdd@gmail.com)

GitHub: <https://github.com/allenzdd>

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

9/48-50 Premier st. Kogarah, NSW, 2217, Australia



## 教育背景

2017 <sub>03</sub> – 2019 <sub>03</sub>	悉尼大学 (USNews#34)	计算机 (数据分析与管理)	工学硕士
	主要课程: 机器学习, 深度学习, 云计算, 计算统计方法, 数据库管理, 多媒体检索, 项目管理, 数据分析与商业智能		AVG: 75 / 100
2011 <sub>09</sub> – 2015 <sub>06</sub>	江南大学 (教育部直属 211)	信息与计算科学 (数学类)	理学学士
	主要课程: 数学分析, 高等代数, 概率论, 数学模型		AVG: 85 / 100

## 技能掌握

- 编程能力: Python, R, C++, SQL, LaTeX
- 方法: 机器学习, 神经网络: TensorFlow/Keras, 云计算: Hadoop, MapReduce, Spark
- 计算机软件: MS Office: Excel, PowerPoint, Word, SAP, Cognos
- 语言能力: 英语, 普通话

## 硕士经历

2018.08 – 至今	TED 演讲者表情分析	毕业设计项目
	• 正在进行中, 目的是实现分析一场 TED 中, 演讲者的感情倾向 (表情与语义), 目前研究属于物体识别、分割与面部表情分析, 使用 YOLO 做人体和脸部识别, 用 deep-sort 去做人体的追踪, 最终做成了能实现实时追踪识别的应用。	
2018.03 – 2018.06	Chars74K 图像数据集分类竞赛	深度学习项目
	• 使用 TensorFlow 底层 API 搭建 CNN 模型, 并加入多种优化算法 (BatchNormalization, Dropout);	
	• 研究了多种复杂的网络模型如 DenseNet, ResNet, VGG 等, 最终采用了 bagging 的融合模型去进行分类, 正确率达到了 90%, 并且获得了前五的排名。	
2018.03 – 2018.06	搭建神经网络框架	深度学习项目
	• 在 Python3 下使用面向对象编程 (OO) 不基于 API (TensorFlow, sklearn) 去搭建神经网络框架,	
	• 在框架下增加了多种激活函数, SGD, momentum, dropout, softmax and cross-entropy loss, mini-batch, batch normalization, weight normalization, weight decay 等等。	
2018.03 – 2018.06	Spark 框架下实现 kNN 图像分类	云计算项目
	• 用 Spark 和 Spark.sql 框架去搭建 kNN 模型, 并用其将 MNIST 数据集进行分类;	
	• 分析 I/O 消耗和运行时间, 在集群内, 最终的模型用时为 1 分钟内, shuffle 接近 0;	
	• 用 Spark.ml 框架去分析每个机器学习算法的优劣, 并得出结论。	
2017.08 – 2017.11	糖尿病胰岛蛋白分类	计算统计方法项目
	• 基于蛋白质磷酸酯数据, 通过动态规划, 在线性规划中, 求得所有已知分类的固定时间变量的 R 值差别最大的线性组合形式建立新属性, 对基因序列使用 PPSM 方式建立新属性, 实用工具为 Rstudio;	
	• 以随机部分样本方式优化 SVM 算法, 并以优化后的算法分类 Akt 和 mTOR 特性的蛋白质, 工具为 R。	
2017.03 – 2017.07	图像拼接及碎纸片复原	多媒体检索项目
	• 图像拼接中, 通过 Python3, 用 SIFT 算法得到图片特征点, kNN 方式去拼接, 最后用 RANSAC 法映射图片;	
	• 碎纸片复原中, 使用动态规划及皮尔逊相关系数实现复原, 工具为 Python3。	
2018.03 – 2018.06	新南威尔士州交通数据分析	悉尼大学 OLAP 数据分析竞赛
	• 基于 2011-2018 年新州交通数据, 分析地区以及驾驶员的危险驾驶倾向性, 并得出相关的建议;	
	• 报告获得了第一名, 并获得 1000 澳元的奖励。	
2018.04 – 至今	参与多个 Kaggle 竞赛	Kaggle
	• 用多种方法对数据进行预处理 (除噪、填补缺省值、抓取特征点等), 并经常使用融合模型去进行数据的预测与分类, 其中竞赛项目最好的名次是 TOP10;	

## 工作经验

2016.02 – 2017.01	精硕科技 AdMaster	数据分析师
<ul style="list-style-type: none"><li>监测每个数字广告（视频贴片，网页 Banner 等）项目的数据并分析流量提供观点与解决方案；</li><li>对广告引流的网页进行监测代码添加及流量监测，并分析数据提供观点与解决方案；</li><li>通过社交媒体（微博为主）数据分析舆情情况，并提供观点与解决方案；</li><li>主要客户：资生堂中国，汉莎航空，可口可乐。</li></ul>		
2016.11 – 至今	领英 LinkedIn	江南大学主页管理员
<ul style="list-style-type: none"><li>管理江南大学领英主页并定期与领英官方沟通。</li></ul>		
2014.06 – 2014.09	乐必扬 DigitasLBi (4A)	数据分析实习
<ul style="list-style-type: none"><li>社交媒体数据进行舆情分析，客户为盖茨基金会和必胜客中国。</li></ul>		

## 学士经历

2012.09 – 2014.03	江南大学理学院学生会	主席
<ul style="list-style-type: none"><li>成功完成常规任务，同时规划并进行了精简部门、工作标准化、资源开放化等院会改革；</li><li>成功组织了两届“笃学江南---数学竞赛”，每届近 2000 人参加。</li></ul>		
2012.09 – 2013.09	江南大学科学技术协会	教育培训中心部长
<ul style="list-style-type: none"><li>为学校学生组织了许多次的培训，包括 IT 培训、礼仪培训、求职培训等；</li><li>成功邀请了韩寒一个团队、延参法师、刘同等人演讲。</li></ul>		
2012.10 – 2013.10	江苏省大学生创新创业项目---苏州市发展分析	数据分析负责人
<ul style="list-style-type: none"><li>基于统计年鉴和调研数据，分析苏州城市化进程中的问题以及观点和未来展望；</li><li>已经成功发表论文。</li></ul>		
2012.08 – 2013.08	全球创业周中国站	校园伙伴江苏站站长
<ul style="list-style-type: none"><li>以校园代表身份（全国共 60 个大学代表）被组委会邀请参会，并在会上当选江苏站站长。</li></ul>		

## 奖项

- 2011.09 江南大学艺术团成员
- 2011.11 江南大学理学院冠军辩论队队长
- 2011.12.31 元旦钟声晚会演出者---手风琴
- 2012.12 首届黑苹果青年汇成员
- 2013.05 2012-2013 年江南大学优秀学生干部
- 2013.08 江南大学数学建模队成员
- 2013.10 2013 年暑期实践校级优胜奖项目团队负责人
- 2015.07 2015 年江南大学校级优秀毕业论文
- 2016.08 通过 AdMaster 第四届社交舆情数据分析培训认证
- 2016.08 通过 AdMaster 第四届广告监测数据分析培训认证
- 2018.07 澳大利亚新南威尔士州交通数据竞赛冠军

## 自我评价

- 悉尼大学（世界 top50）数据科学硕士应届生，本科江南大学（211 工程）数学系毕业。
- 擅长 Python、R、SQL 编程，
- 熟练掌握各种数据处理方法；
- 精通机器学习各类算法（sklearn）；
- 深度学习各类算法（TensorFlow）；
- 云计算（Spark、MapReduce、Hadoop）；
- 本科期间，为学生干部，多次参加校际活动，有较强的社交与组织能力。

# Dongdong ZHANG

+61 420 264 362

[allenzdd@gmail.com](mailto:allenzdd@gmail.com)

GitHub: <https://github.com/allenzdd>

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

9/48-50 Premier st. Kogarah, NSW, 2217, Australia



## Education

2017 – 2019	University of Sydney	Master of Information and Technology and Management (Data Science)	
		Main Courses: Machine Learning and Data Mining, Deep Learning, Computational Statistical Method, Cloud Computing, Database and Management	AVG: 75 / 100
2011 – 2015	Jiangnan University	Bachelor of Science (Mathematics: Information and Computing Science)	
		Main Courses: Mathematical Analysis, Probability Theory, Mathematical Modelling	AVG: 85 / 100

## Advantage

- **Programming:** Python, R, C++, SQL, LaTeX
- **Methodology:** Any **Machine Learning** Method, Any **Neural Network** Method, Tensorflow, **Hadoop**, **MapReduce**, **Spark**
- **Software:** MS office: Excel, PowerPoint, Word, SAP, Cognos
- **Language:** English, Mandarin

## Master Experience

2018.08 – now	Facial Expression Recognition	Graduation Project
	<ul style="list-style-type: none"><li>• Just start, purpose is analyzing the TEDer facial expression in the TED, it belongs to object detection and emotion analysis in real-time, used YOLO to detect people and face, and used deep-sort to achieve people tracking.</li></ul>	
2018.03 – 2018.06	Chars74K dataset classification	Deep Learning Assignment
	<ul style="list-style-type: none"><li>• Used tensorflow framework to create a CNN with Batch Normalization, dropout and others optimize function;</li><li>• Tried many complex deep neural networks, like DenseNet, ResNet, VGG etc., finally used bagging model got 90% accuracy, and the rank is TOP5.</li></ul>	
2018.03 – 2018.06	Multiple Layer Perception Structure	Deep Learning Assignment
	<ul style="list-style-type: none"><li>• No used API (tensorflow and ski-learn) to build the MLP framework on Python3 with OO programming;</li><li>• Added multiple activation function, SGD, momentum, dropout, softmax and cross-entropy loss, mini-batch, batch normalization, weight normalization, weight decay etc.</li></ul>	
2018.03 – 2018.06	Images Classification with kNN by Spark	Cloud Computing Assignment
	<ul style="list-style-type: none"><li>• Used spark and spark.sql framework to create a kNN algorithm then classify MNIST data in the Hadoop cluster;</li><li>• Analyzed I/O cost and execution time, finally used less 1 min to classify 70,000 image data, and near 0 shuffle data;</li><li>• Using spark.ml framework to analysis any machine learning method I/O cost and execution time.</li></ul>	
2017.08 – 2017.11	Classification for insulin	Computational Statistical Method Assignment
	<ul style="list-style-type: none"><li>• Based on the insulin phosphoproteome data, created each insulin's fixed variables linear regression r-squared feature by dynamic programming, and created gene sequence by PPSM method, using RStudio;</li><li>• Implemented SVM classifier by repeated random sampling unlabeled data to classify Akt or mTOR, using R.</li></ul>	
2017.03 – 2017.07	Application of Picture Stitching	Multimedia Retrieval Assignment
	<ul style="list-style-type: none"><li>• Got feature point by SIFT, then matched by kNN, finally wrapped transformation by RANSAC, using Python3;</li><li>• In shreds recovering, implemented dynamic programming to select the biggest sum of Pearson's r of each edge.</li></ul>	
2018.03 – 2018.06	NSW Traffic Data Analysis	USYD OLAP Analysis Competition
	<ul style="list-style-type: none"><li>• Based on 2011-2018 traffic data of New South Wales Australia, analyzed the location and motorists risky driving tendency then provided some credible, justified, data-driven suggestions;</li><li>• My report won 1<sup>st</sup> place and awarded 1000AUD prize.</li></ul>	
2018.04 – Now	Many Kaggle Projects	Kaggle
	<ul style="list-style-type: none"><li>• Implementation of many preprocessing method for data, and frequently used ensemble learning to classify or predict data, and one project get TOP10.</li></ul>	

## Working Experience

2016.02 – 2017.01	AdMaster	Data Analyst
<ul style="list-style-type: none"><li>Tracking data of digital adverts, analysis the traffic to provide insights and solutions for each ad project;</li><li>Setting code on the website and tracking, analysis the traffic to provide insights for each website;</li><li>Getting data form social media web by crawling, then analysis it to provide insights and solution for each project;</li><li>My main client: Shiseido China, Coca-Cola, Lufthansa Inc.</li></ul>		
2016.11 – Now	LinkedIn	Administrator of Jiangnan University
<ul style="list-style-type: none"><li>Management for Jiangnan University LinkedIn Page.</li></ul>		
2014.06 – 2014.09	DigitasLBi (Publicis Group)	Data Analyst Intern
<ul style="list-style-type: none"><li>Social media data analysis for Gates Foundation and Pizza hut Inc.</li></ul>		

## Bachelor Experience

2012.09 – 2014.03	The Student Union of School of Science	President
<ul style="list-style-type: none"><li>Had done every basic works successfully, and leaded reform movement about working normalization;</li><li>Successfully organized Mathematical Competition (more than 2000 participated) twice.</li></ul>		
2012.09 – 2013.09	Science and Technology Association	Minister of Education center
<ul style="list-style-type: none"><li>Organized many technology trainings for students, including IT, etiquette, career etc.;</li><li>Successfully invited and organized some Chinese famous persons' presentation.</li></ul>		
2012.10 – 2013.10	Jiangsu Province Students' Project (City analysis for Suzhou)	Data Analyst Member
<ul style="list-style-type: none"><li>Based on the statistical yearbook and survey data, provided some insights of Suzhou city innovation;</li><li>Has published paper successfully.</li></ul>		
2012.08 – 2013.08	The Head of Global Entrepreneurship Week China	Campus Partners' Jiangsu Leader
<ul style="list-style-type: none"><li>Was invited the convention (President Xi lettered the convention), and conversation with TOP 60 Chinese universities' representatives at Shanghai and was elected Campus Partner's Jiangsu Province Region Leader.</li></ul>		

## Award

- 2011.09 The member of Jiangnan University art troupe
- 2011.11 The captain of the debate champion team in School of Science in JNU
- 2011.12.31 Performer of Jiangnan University New Year's Day Party
- 2012.12 Member of first of Blackapple Youth Engagement
- 2013.05 Outstanding student leader of Student Union in 2012-2013
- 2013.08 Mathematical modeling team member of Jiangnan University
- 2013.10 University summer project excellence award (Group Leader)
- 2015.07 Thesis Award for Outstanding Undergraduate Graduation Thesis (Jiangnan University)
- 2016.08 Awarded Certificate of 4<sup>th</sup> round hard skill training-Track Participation
- 2016.08 Awarded Certificate of 4<sup>th</sup> round hard skill training-Social Participation
- 2018.07 Champion of New South Wales State Traffic Data Analysis Competition

## Self-Introduction

- University of Sydney Master of Data Science , Jiangnan University Bachelor of Mathematics.
- Skilled at Python, R, SQL, C++,
- Experience at handling any data processing method,
- Knowledge at any machine learning and deep learning algorithm and Tensorflow,
- Good at Cloud Computing (Spark, MapReduce, Hadoop).