

KAIRUI WU

Current Address: 736 Bay Street, Toronto, M5G 2M4

Email: kairuiwu@yeah.net **Mobile:** (437)-983-3399

October 4th, 2018

Department of Human Resource
Capital One

Position: Data Scientist, Full-Time 2019

Dear HR Manager,

I'm writing to express my strong interest in the position of Data Scientist at Capital One. The explosion of data makes me realize that only by obtaining advanced skills which enable us to keep pace with the coming of 'feedback economy', will we win in the treasure hunting game in big data era. This becomes my motivation for arming myself with the knowledge and skills of data analytics in my previous study and starting my career as a data scientist. The area of banking and financial service is one of the best platforms on which the technology of data analytics could be applied to make a difference in light of the huge amount of data produced and obtained in this industry. More importantly, Capital One is a famous bank holding company with a high reputation in the world and stays ahead in the industry through a great innovation ecosystem. Therefore, I strongly believe the excellent platform of Capital One is the best place for me to apply my knowledge and skills to make data speak.

In my study at University of Toronto and Carnegie Mellon University, I finished various courses that were closely related with data analytics, machine learning and computer programming with good performances. Through these courses, I have developed a solid background on the knowledge of exploratory data analysis and predictive data analytics. Besides, I participated in various projects through Kaggle competition and Udacity Nanodegree Programs to improve the integrity of my knowledge set and skills. These projects also enhanced my practical skills of machine learning and data analytics using Python, R and Tableau. These previous experiences brought myself a systematical knowledge and skill frame on the whole chain of data gathering, data wrangling, data analytics and data visualization as well as in-depth understanding on machine learning techniques. In addition, the 8-month co-op experience at RBC as a risk analyst in machine learning architecture armed me with good insights into the industrial applications of machine learning and AI techniques and strong team-work sense.

I feel confident that I can achieve a good performance and well applied my knowledge and skills of data analytics in this position and am very eager to grow professionally with Capital One. Your time and consideration is greatly appreciated.

Sincerely,
Kairui Wu

KAIRUI WU

Current Address: 736 Bay Street, Toronto, M5G 2M4

Email: karrywu.uot@gmail.com **Mobile:** (437)-983-3399

OBJECTIVE	To obtain a full-time job to pursue a career in the area of machine learning and data analytics	
EDUCATION	University of Toronto , Toronto, Canada	Sep 2017 – Apr 2019
	<i>Master of Information, Information System Design</i>	GPA: 3.81
	Related Courses: Data Analytics; Machine Learning; Programming for Information System; Data Modeling and Database Design; System Analysis & Innovation; Data Structure and Algorithm; Natural Language Computing	
	Carnegie Mellon University , Pittsburgh, the United States	Sep 2014 – Dec 2015
	<i>Master of Science in Engineering</i>	GPA: 3.75
WORKING EXPERIENCE	Related Courses: Data Mining; Data Management; Engineering Computation	
	Shanghai Jiao Tong University , Shanghai, China	Sep 2010 – Jun 2014
	<i>Bachelor of Science in Engineering</i>	GPA: 3.3(5/30)
	Udacity.com , Online Education	Oct 2017 – Mar 2018
	<i>Nanodegree of Advanced Data Analyst</i> <i>Nanodegree of Machine Learning Engineer</i>	
WORKING EXPERIENCE	Coursera.com , Online Education	May 2018 – Sep 2018
	<i>Neural Networks and Deep Learning, Data Structures and Performance, Object Oriented Programming in Java</i>	
	Royal Bank of Canada , Toronto	May 2018 – Dec 2018
	<i>Enterprise Risk Analyst Coop (Machine Learning Architecture), Group Risk Management</i>	
	<ul style="list-style-type: none">• Made proof of concepts on innovative ideas via researches to transferred business use cases to machine learning and AI applications• Collected datasets using SQL query from various databases in RBC and implemented data wrangling works and visualization• Implemented machine learning algorithms using Python	
WORKING EXPERIENCE	Cushman & Wakefield Consulting Company , China	Dec 2015 – Jul 2017
	<i>Senior Analyst</i>	
	<ul style="list-style-type: none">• Explored customers' preferences and the market vacancy by conducting exploratory data analysis using Excel• Collected and extracted the information of the policy and market through in-depth interviews to experts• Made development strategies or investment strategies for clients based on the result of data analytics and the information obtained by interviews	

PROJECTS	Data Quality Assessment by Anomaly Detection Technique, RBC May 2018 – Present <ul style="list-style-type: none"> • Made the proof of concept to the machine learning application on client data quality assessment and improvement • Implemented unsupervised learning algorithms such as Isolation Forest and K-prototypes to detect the data quality issue and anomalies via anomaly detection technique using Python
	Home Credit Default Risk, Kaggle.com Competition Jun 2018 – Sep 2018 <ul style="list-style-type: none"> • Conducted data wrangling works and feature engineering using Python • Applied the model stacking technique on LightGBM, Random Forest, Neural Network and Logistic Regression to build a model with an AUC score of 0.798 (Ranking: Top 11%)
	TalkingData Fraud Detection, Kaggle.com Competition Feb 2018 – Apr 2018 <ul style="list-style-type: none"> • Conducted exploratory data analysis and feature engineering using Python • Built a model with an AUC score of 0.981 using the algorithm of XGBoost (Ranking: Top 13%)
	Web Scraping and Auto-filling for 10-K Financial Reports, RBC May 2018 – Jun 2018 <ul style="list-style-type: none"> • Built a web scraping program to extract the 10-K financial reports from Edgar.com • Built an automation program to parse the 10-K financial reports which are in XBRL format and auto-fill the information to RBC's Newton Template of financial report
	Customer Segmentation by Clustering Algorithm, Udacity Jan 2018 – Mar 2018 <ul style="list-style-type: none"> • Applied PCA transformation on datasets to reduce the dimensionality • Applied Gaussian Mixture Model to do clustering on the dataset of customers and found two clusters of customers
	Research on Sensitivities of Machine Learning Classifiers on Imbalanced Datasets, University of Toronto <i>Team Leader</i> Sep 2017 – Dec 2017 <ul style="list-style-type: none"> • Applied SMOTE method on the dataset of loan defaults to change the imbalanced degree of the dataset • Tested sensitivities of different classification algorithms such as Logistic Regression, Decision Tree, Random Forest, SVM, KNN, Gaussian Bayes on datasets with different imbalanced degrees by monitoring prediction performances of models
HONORS	<i>Excellent Employee</i> of Cushman & Wakefield Jun 2017 <i>Excellent Graduates</i> of Shanghai Jiao Tong University Jun 2014 <i>Huashan Scholarship</i> from Shanghai Jiao Tong University Dec 2013 <i>Three Good Student Prize</i> from Shanghai Jiao Tong University Jun 2013 <i>Academic Fellowship</i> from National Education Department Sep 2010 - May 2013
SKILLS	Proficient: Python, Java, R, C++, SQL, Excel, Powerpoint Familiar: Tableau, Azure, Matlab
LANGUAGE	English, Chinese



Complete Academic History

Kairui Wu

Accurate as of: September 16
2018

This is not an official transcript.

Registration History

2017 Fall-2019 Winter: School of Graduate Studies

School of Graduate Studies

2017 Fall - Master of Information - Faculty of Information

Crs Code	Title	Wgt	Mrk	Grd	CrsAvg
CSC2515H	Introduction to Machine Learning	0.50		A-	
INF1340H	Programming for Information Systems	0.50		A+	
INF1341H	Systems Analysis and Process Innovation	0.50		A-	
INF2190H	Data Analytics: Introduction, Methods and Practical Approaches	0.50		A+	

Credits Earned: 2.00

2018 Winter - Master of Information - Faculty of Information

Crs Code	Title	Wgt	Mrk	Grd	CrsAvg
INF1342H	System Requirements and Architectural Design	0.50		A-	
INF1343H	Data Modeling and Database Design	0.50		A+	
INF2040H	Project Management	0.50		A-	
INF3900H	Workplace Integrated Learning	0.50		A-	

Credits Earned: 2.00**2018 Summer - Master of Information - Faculty of Information**

Crs Code	Title	Wgt	Mrk	Grd	CrsAvg
INF3902H	Co-operative Workplace Placement II	0.50		SDF	

Credits Earned: 0.00**2018 Fall - Master of Information - Faculty of Information**

Crs Code	Title	Wgt	Mrk	Grd	CrsAvg
ECE1762H	Algorithms and Data Structures	0.50		IPR	
INF2177H	Information Management and Systems	0.50		IPR	

Credits Earned: 0.00**2019 Winter - Master of Information - Faculty of Information**

Crs Code	Title	Wgt	Mrk	Grd	CrsAvg
----------	-------	-----	-----	-----	--------

INF1005H	Information Workshop I Futures Thinking	0.25	IPR
INF1344H	Introduction to Statistics for Data Science	0.50	IPR

Credits Earned: 0.00

This is not an official transcript.