

Yingkai (Kevin) Xu

Data Scientist

kevinhsu1008@gmail.com

Summary

Personal Website: yingkaixu.com Email: kevinhsu1008(at)gmail.com Yingkai (Kevin) Xu is currently looking for fully funded PhD position on machine learning and computational finance in UK/US. He is also looking for opportunities in data scientist and high frequency algorithmic trading area. Yingkai is a highly proactive, adaptable, diligent and team-orientated professional with serious attention to details, quick learning ability and excellent interpersonal skills who always stays hungry and stays foolish. His industry exposure includes Financial Services, E-commerce, Digital Media and Education.

Experience

Mobile App Designer and Web Developer at Glenmere Technology Ltd

November 2014 - Present (7 months)

- Designed and developed the startup official website (nhkeu.co.uk).
- Setup LNMP (Linux, Nginx, MySQL and PHP) full stack developing environment.
- Used ThinkPHP framework to establish the CRM system.
- Applied AWS, including EC2, Load Balance, S3, RDS, EMR, CloudFront, Route 53 and Elastic Beanstalk, etc. to optimize the performance of server architecture and backend system.
- Improved UI/UX experience of Android and IOS apps based on user testing.

Senior Investment Researcher at Tantinet Capital Ltd

October 2012 - September 2014 (2 years)

- Researched on brokerage service and traded FX, Commodities, Futures and Options.
- Developed quantitative trading algorithms based on VaR, statistical arbitrage, implied volatility, market correlation, etc. Measured risk-return and timing profiles of them.
- Organized Bloomberg UK-China Investment Expo 2013 (ucie.org.uk) on behalf of Tantinet Capital Ltd.
- Invited by Chinese Students and Scholars Association (CSSA) to make presentations on the topic 'Techniques in Trading' in the University of Cambridge and School of Oriental and African Studies.

Part-time Assistant at McKinsey & Company

August 2014 - August 2014 (1 month)

- Identified potential global M&A target companies as a compliment of client's industrial value chain.
- Collected detailed information and analyzed the credit rating of client's counterparties.

Data Analyst at Intelligent Modelling & Analysis Research Group

June 2014 - August 2014 (3 months)

- Worked on big data analysis, knowledge representation, machine learning, decision support algorithm, etc. • Researched on a novel ensemble for colon cancer stage classification based on UK immunological laboratory data.

Intern at ING

February 2011 - February 2011 (1 month)

- Researched on insurance products, insurance development and marketing strategies of mainland China.
- Analyzed the macroeconomy of Asia, including manufacturing competitiveness, financial sustainability and monetary policy.
- Accomplished a business plan and 166 questionnaires on the topic 'Life Insurance in Emerging Market' in one week. Made the presentation to the Honk Kong regional manager and received 90% recognition.

Education

London Graduate School in Mathematical Finance

Preliminary for PhD, Mathematical Finance, 2014 - 2015

UCL

Master of Science (MSc), Computational Statistics and Machine Learning, 2014 - 2015

University of Nottingham

Bachelor of Engineering (BEng), Electrical and Electronic Engineering, 2012 - 2014

Grade: First Class Honours (78%)

3 recommendations available upon request

Nanjing University of Aeronautics and Astronautics

Bachelor of Science (BSc), Information and Computing Science, 2010 - 2012

Grade: Top 2% (3.9/4.0)

The University of Hong Kong

Exchange, Economics, 2011 - 2011

Skills & Expertise

Machine Learning

Quantitative Finance

Statistics

Data Analysis

Data Mining

Quantitative Analytics

FX Trading

Portfolio Management

Algorithmic Trading

Financial Risk Management

Python

Matlab

PHP

Projects

Semi-supervised Recursive Autoencoders for Opinion Detection on Twitter

March 2015 to April 2015

Members: Yazhe Li, Yingkai (Kevin) Xu, Dexuan Zhang

Analyzing the tweets can identify the political opinions expressed by users for a candidate. Platforms like Twitter are very popular for political campaigning. So many people are tweeting about their favorite party or candidate. One could mine this data in interesting ways to extract and summarize political sentiments. The purpose of the project is to build a system that summarizes political sentiment given some tweets.

Forecast Use of a City Bikeshare System

March 2015 to April 2015

Members: Juan Paul Figueroa, Yingkai (Kevin) Xu

Bike sharing systems are a means of renting bicycles where the process of obtaining membership, rental, and bike return is automated via a network of kiosk locations throughout a city. Using these systems, people are able to rent a bike from one location and return it to a different place on an as-needed basis. Currently, there are over 500 bike-sharing programs around the world. Bike sharing systems therefore function as a sensor network, which can be used for studying mobility in a city. The purpose of the project is to combine historical usage patterns with weather data in order to forecast bike rental demand in the Capital Bikeshare program in Washington, D.C. Some features like the duration of travel, departure location, arrival location, and time elapsed are used in the model.

Use Telematic Data to Identify a Driver Signature

March 2015 to Present

Members: Yingkai (Kevin) Xu, Dexuan Zhang

For automobile insurers, telematics represents a growing and valuable way to quantify driver risk. Instead of pricing decisions on vehicle and driver characteristics, telematics gives the opportunity to measure the quantity and quality of a driver's behavior. This can lead to savings for safe or infrequent drivers, and transition the burden to policies that represent increased liability. The purpose of the project is to develop an algorithmic signature of driving type based on a dataset of over 50,000 anonymized driver trips from AXA. We use subsampling trick to transform the unsupervised learning problem into a supervised problem and apply random forest algorithm in the project. It helps to build a telematic fingerprint capable of distinguishing when a trip was driven by a given driver.

Classify the Sentiment of Sentences from the Rotten Tomatoes Dataset

February 2015 to March 2015

Members: Yingkai (Kevin) Xu, Yazhe Li

The Rotten Tomatoes movie review dataset is a corpus of movie reviews used for sentiment analysis. It used Amazon's Mechanical Turk to create fine-grained labels for all parsed phrases in the corpus. The purpose of the project is to benchmark the sentiment-analysis ideas on the Rotten Tomatoes dataset. The labels phrases are on a scale of five values: negative, somewhat negative, neutral, somewhat positive, positive. We solve some obstacles like sentence negation, sarcasm, terseness, language ambiguity, and many others in this practical project.

Predict Whether a Mobile Ad Will Be Clicked

January 2015 to February 2015

Members: Yingkai (Kevin) Xu

Online advertising is a multi-billion dollar industry that has served as one of the great success stories for machine learning. Sponsored search advertising, contextual advertising, display advertising, and real-time bidding auctions have all relied heavily on the ability of learned models to predict ad click-through rates (CTR) accurately, quickly, and reliably. A typical industrial model may provide predictions on billions of events per day, using a correspondingly large feature space, and then learn from the resulting mass of data. Memory savings, performance analysis, confidence in predictions, calibration, and feature management are some majority issues that need to deal with. In online advertising, click-through rate is a very important metric for evaluating ad performance. As a result, click prediction systems are essential and widely used for sponsored search and real-time bidding. The purpose of the project is using 10 days of Avazu click-through data, ordered chronologically, to build and test prediction models. We find a strategy based on logistic regression with L1 and L2 penalty and online learning techniques. It beats original standard classification algorithms.

An Ensemble for Cancer Stage Classification

June 2014 to August 2014

Members: Yingkai (Kevin) Xu

Computers find some data difficult to learn from. This results in classification algorithms performing only slightly better than random guessing. When this is the case, ensemble learning which means combining lots of different classifiers into one is often appropriate. In terms of medical prediction, an ensemble to predict colon cancer based on UK immunological laboratory data is developed to improve the predicting performance by dealing with the collected raw data. Predicting how long potential patients can live and what stage their colon cancers are at is the main purpose. We insert random forest as one of classifiers and evaluate simple imputation, multiple imputation and iterative imputation which can solve unbalanced datasets and missing values issues effectively.

Mortar Element Method for the Two Dimensional Transverse Electric Boundary Integral Equation

September 2013 to June 2014

Members: Yingkai (Kevin) Xu

Scattering of transverse electric polarized frequency dependent electromagnetic waves by two dimensional objects can be modelled by electric boundary integral equations. In recent years, the algorithms based on these equations have been parallelized almost completely. Only the preprocessing and further handling of the geometry remains sequential. This was no problem up to now because the memory required for geometry

storage was negligible compared to the rest of the data. Eventually, however, this will pose a limit on the size of the problems that can be simulated. Mortar element method is convenient for parallel computing. Roughly said, such a method allows for the solution of the problem on separate parts of the geometry, after which the partial solutions are glued or mortared to construct the global solution. To some extent, mortar element method in this project is based on boundary element method. This method facilitates parallelisation, which is important since single computing nodes have stopped becoming faster. Further, more than one PC can be used to tackle the large-scale problems. In this project, I researched on the implementation of classical numerical method and mortar boundary element method with respect to a perfectly conducting cylinder and validate the solutions of scattering fields by analytical results. In the project, both transverse magnetic and transverse electric problems will be discussed. For the parallelization purpose, mortar boundary element method for the two dimensional scattering of fields and waves will be finally developed.

Campus Seller

February 2014 to May 2014

Members: Ilona Brodovska, Yingkai (Kevin) Xu, Kunlun Mao, Ayesha Taylor

Campus Seller is a new and local service within campuses which allows students and staff to dispose any goods that they no longer want and wish to sell. The aim is to maximize the value of second hand goods and build a sustainable and recycling model. The project obtained £4,000 grants from Student Venture Challenge (SVC) and UnLtd to launch the business and won a finalist place in National Apprentice Challenge (NAC) which is the largest live competition for entrepreneurial students across UK. Knowledge Exchange and Innovation Awards also nominated this project in 2014.

Inovanao-Hipark Operation Centre

November 2011 to April 2012

Members: Yingkai (Kevin) Xu, Leon Zhou, Chao Huang, Binbin Wang

Inovanao focuses on marketing and commercialisation of scientific and technological achievements. It provides entrepreneurial practice opportunities for over 160 college students who are willing to access to real business in the industry. Inovanao assisted the incubation team Yomoo to register as a limited company and promote market-oriented products, which achieved approximated #500,000 (£50,000) profit in 2012. The project received #300,000 (£30,000) seed funding from BHC venture capital and Jiangsu Province Technology Department.

Humanoid Gait Planning and Stability of Biped Robot

January 2011 to March 2012

Members: Boyun Tang, Yingkai (Kevin) Xu, Xinye Zhang

Through researches on dynamically self-tuning walking control, a reinforcement learning algorithm was designed based on Zero Moment Point (ZMP) and Poincare Map to control eight degrees of freedom in the parallel process. The learning mechanism and adaptation capability of dynamic walking on a complex terrain were effectively realized. Practically, it optimized falling motion and gravity balance for the humanoid robot.

Volunteer Experience

Machine Learning Fellow at Startup.ML

April 2015 - Present (2 months)

- Empowered the business model with accurate and fast machine learning systems for startups.

Manager of London Branch at ISIG Capital

March 2014 - September 2014 (7 months)

- Facilitated investment communication, practices and educational panels among college students and financial professionals in London.

Core Member of Electricity Team at NUstream

November 2012 - July 2013 (9 months)

- Installed a small scale hydropower scheme to provide sustainable electricity to Nyumba ya Masimbiro community centre, health centre and primary schools in rural Malawi where only 2% of residents have access to electricity.
- Collected environmental information and designed hydropower transmission system.

Vice President of Nanjing University of Aeronautics and Astronautics at Enactus

October 2010 - April 2012 (1 year 7 months)

- Proposed and implemented the project 'Craft in Permanence' to solve the obstacles of the heritage of Chinese folk hand craftsmanship and draw society attention to those fading art treasures. The new E-commerce model and social media saved plentiful human resources and the income of 30 beneficiaries achieved an average 20% increase per year.
- Led the team to win the 3rd place in 2012 Enactus China Final Competition held in Beijing International Convention Centre and granted Elite Student Certificate signed by Enactus global president and CEO.

Publications

The MTSP Optimal Model Suited for Local-cross Repeat Path

Chinese Control and Decision Conference (CCDC) May 25, 2013

Authors: Yingkai (Kevin) Xu, Shan Kang, Hui Lin

Although many people have paid attention to the conventional MTSP problem and doing research on it, its constraint condition "only one travelling salesman can pass by in each city" suffered from many limitations in reality. On the basis of conventional MTSP model, this article promotes it in adding the "material transportation capacity" constraint condition and builds a new MTSP material transportation optimal model permitting the existence of local-cross repeat path. We use lingo software to solve the problem combined with an example. The result indicates that the new optimal model accords with fact well. Another advantage of this new model can also avoid the cumbersome process of the GA (Genetic Algorithm) and it is much easier to operate for researchers. Meanwhile, we can bring in other variants and constraint conditions on the basis of this new model and extend to build other models suited for specific needs.

Location of the Heavy Metal Pollution Source in the Soil Based on GA and BP Artificial Neural Network

Hubei Agricultural Sciences March 1, 2013

Authors: Yuwen Wu, Yingkai (Kevin) Xu, Tong Liu, Hui Lin

A model based on Genetic Algorithm and Back Propagation Artificial Neural Network is developed in this paper to effectively depict the distribution of heavy metal in certain area and accurately locate the pollution source by picking out the coordinate that contents the maximum quantity of heavy metal according to the sampled data. This paper takes Qingdao for instance. It detects the distribution of heavy metal in the soil of different areas in Qingdao, thus locating the possible pollution source.

Face Orientation Recognition Based on Multiple Facial Feature Triangles

International Conference on Control Engineering and Communication Technology (ICCECT) December 7, 2012

Authors: Linlin Gao, Yingkai (Kevin) Xu

This paper proposes a new method that combining multiple feature triangles with BP neural network, to improve the efficiency and accuracy of face orientation recognition. Based on the traditional index the inverted triangle formed by pupils and nasal tip, we find another feature triangle formed by nasal tip and corners of mouth. First we do image preprocessing which includes smoothing linear filter, edge detection and so on. Then both rough and precise detection of feature points are done. Next we extract feature triangle based on two-dimensional plane. Finally BP neural network is used for face orientation recognition. Experimental results show that an approximately 90% success rate is achieved. They also reveal that our new method improves the recognition effect.

The SCM Control Electric Circuit Systems of a Drum Robot Based on Technique of Multithreading

Machinery August 1, 2011

Authors: Yingkai (Kevin) Xu, Hui Lin

A drum hitting robot was designed and manufactured. It is functioned with 5 degrees of freedom, it is driven by 3 motors. The structure of the robot arm is made by Aluminum alloy, and the base of the robot is made by complex chemical material. There are two control electric circuit systems, the major system is the SCM and its functioning circuit, and the other one is sound and touch detector system. It is controlled by the SCM, and it runs according to the pre-set program automatically, it also runs according to what human teaches at spot. The technique of multithreading in the SCM control electric circuit systems of a drum hitting robot was applied. It cherishes the following specialties including entertaining, easy using, and good for showing.

Languages

English

(Professional working proficiency)

Mandarin

(Native or bilingual proficiency)

German

(Elementary proficiency)

Honors and Awards

Knowledge Exchange Innovation Awards

The University of Nottingham

May 2014

Finalist in National Apprentice Challenge (NAC)

National Apprentice Challenge (NAC)

April 2014

Entrepreneur Grant of Student Venture Challenge (SVC)

Haydn Green Institute for Innovation and Entrepreneurship (HGI)

March 2014

Meritorious Winner in Mathematical Contest in Modelling (MCM)

Consortium for Mathematics and Its Application (COMAP)

February 2011

Second Runner in BIP Business Plan Competition

Intentional Netherland Group (ING)

February 2011

Courses

Master of Science (MSc), Computational Statistics and Machine Learning

UCL

Supervised Learning

Probabilistic and Unsupervised Learning

Information Retrieval and Data Mining

Stochastic Methods in Finance I

Stochastic Methods in Finance II

Applied Bayesian Methods

Applied Machine Learning

Statistical Models and Data Analysis

Bachelor of Engineering (BEng), Electrical and Electronic Engineering

University of Nottingham

Machine Learning

Software Engineering Design

VLSI Design

Control System Design

Mathematical Techniques for Electrical and Electronic Engineers

Field Waves and Antennas

Power Electronic Design

Business Planning for Engineers

Energy Conversion for Motor and Generator Drives

Electronic Construction Project

Power Supply Electronics

Signal Processing and Control Engineering

Professional Skills for Electrical and Electronic Engineers

Electrical Engineering Design Project
Electronic Engineering
Telecommunications
IT Infrastructure
Electronic Design

Exchange, Economics

The University of Hong Kong
Leadership in the Knowledge Economy
Marketing Research and Strategies
International Business Strategic Planning
Accounting and Financial Management

Bachelor of Science (BSc), Information and Computing Science

Nanjing University of Aeronautics and Astronautics
Mathematical Analysis
Advanced Algebra
Numerical Algebra
Probability Theory and Mathematical Statistics
Analytic Geometry
Ordinary Differential Equations
Discrete Mathematics
Complex Function
Fundamental of Applied Software
Modern Control Theory
Analog Electronic Technology
Applied SQL Server Technology
Programming Language
Data Structure
Algorithm Design and Analysis

Preliminary for PhD, Mathematical Finance

London Graduate School in Mathematical Finance
Computational Finance
Information and Finance: Filtration Modelling,
Stochastic Filtering and Asset Pricing

Certifications

National Level 3 Human Resources Management Division (CHN)

Interests

Reading, Music, Traveling, Photography, Psychology and MOOCs.

Yingkai (Kevin) Xu

Data Scientist

kevinhsu1008@gmail.com



3 people have recommended Yingkai (Kevin)

"When I first met Yingkai in National Apprentice Challenge, I was impressed with not only his extraordinary intelligence but also his strong sense of leadership that helped his team get into the final of this competition. There is no need to spell out his outstanding academic and professional performance. However, previous success does not make him arrogant. In contrast, I found him extremely modest, fond of learning and considerate. Therefore, I believe he is a promising candidate for any team who is pursuing success. In short, Yingkai is a great person both in life and work and undoubtedly will go far. "

— **Ivy Yiqing Wang**, was with another company when working with Yingkai (Kevin)

"Kevin is one of the co-founders of the Campus Seller. He is very hardworking, dedicated and creative person. He knows how to solve problems quickly and efficiently. He has strong entrepreneurial skills, which helped our team to get to the finals for the National Apprentice Challenge. He is a great in working within the team! He is able to compromise and negotiate with people. Well done Kevin!!"

— **Ilona Brodovska**, worked directly with Yingkai (Kevin)

"I have been working with Mr Yingkai Xu in the framework of his final year project (3rd year Beng) here at the University Of Nottingham. Mr Xu is without a very strong final year project student. He is a very able and motivated student and a very ambitious young man; he presented me with an impressive curriculum vitae for someone at such an early stage in his career. Mr Xu is a pleasure to work with, not only because of his extensive background knowledge and his ability to learn new skills and techniques, but also because of his polite and pleasant manner of interaction."

— **Kristof Cools**, *Lecturer in Computational Electromagnetics, University of Nottingham*, advised Yingkai (Kevin) at University of Nottingham

[Contact Yingkai \(Kevin\) on LinkedIn](#)