

Boxin Zhao

Nankai University, Tianjin, China

zbx1995@126.com

Education Background

Nankai University, School of Mathematical Sciences

Tianjin, China

- Bachelor of Science in Statistics, **GPA: 92/100 (rank 1/82)**

Sep 2014 – Jun 2018

- Math & Statistics Courses: Mathematical Statistics, Multivariate Statistical Analysis, Statistical Computing (Statistical Algorithms, e.g. Bootstrap), Machine Learning (Coursera), Linear Regression, Bayes Statistics, Probability, Stochastic Process, Numerical Analysis, Real Analysis, Complex Analysis, Mathematical Analysis, Advanced Algebra.
- Programming Courses: C++, Data Structures and Algorithms, R, Database (MySQL), MATLAB, Operating Systems and Network
- Honors: National Scholarship (top 1%), Tianjin Municipal People's Government Scholarship (top 3%), Zhide Scholarship (top 3%), First-Class Scholarship (top 5%)

Working Experience

Broad Measure Asset Management Company

Beijing, China

Summer Quantitative Analyst

Jun 2017 – Aug 2017

- Constructed a multi-factor alpha model with value factors, quality factors and momentum factors; used PCA (Principal Component Analysis) as a major method to lower the dimension of IC (Information Coefficient) vector and compute correlation matrix.
- Developed a risk-adjusted IC based portfolio construction method; risk-adjusted IC was designed as an alternate for traditional IC to achieve a more robust prediction performance as risk-adjusted IC takes the risk exposure in alpha factors into account.
- Established risk models based on the BARRA risk model as a tool to control portfolio risk exposure.
- Created an optimal execution process model using calculus of variations; typically focused on Euler-Lagrange Equation.
- Constructed a performance measurement schema with considerations of cumulative returns, information ratio, information coefficient and other technical indicators to reflect the performance of portfolio.
- Developed a quantitative portfolio management system, with risk-adjusted IC based multifactor model as alpha model, BARRA based risk model as risk-control tool, an optimal execution process to implement trading, and a measurement schema to monitor performance.

Hyde Renaissance Capital

Beijing, China

Quant Analyst

Sep 2017 – Oct 2017

- Constructed factor driven strategies; stocks were ranked according to their weekly factor scores and were assigned into five quantile portfolios with equal stock weights; quality factors, value factors, and momentum factors were considered.
- Developed a backtest model covering multiple indicators such as annualized returns, annualized volatility, coverage, factor turnovers, information ratio, portfolio returns, quantile returns, serial correlation, spearman IC, etc.
- Implemented backtest model using Python with its packages, including but not limited to ipython, numpy, pandas and matplotlib.

Research Experience

“A feature-based improvement of computational protein function prediction using machine learning”

Mar 2017 - Present

Jianyi Yang, Boxin Zhao, Bowen Wang

- Predicted multiple protein features by implementing feature prediction software on Linux platform; features we predicted and software we implemented included *transmembrane features* with MEMSAT-SVM, *Psipred-helices* with PSIPRED 3.3, etc.
- Implemented the classical FFPred model, a svm model with Gaussian kernel to predict Gene Ontology term, with the help of scikit-learn; approximately 90 features calculated in previous steps and 1,000 proteins were used for training.
- Implemented a newly proposed method called GOLabeler, which utilized Learning to Rank framework to combine component protein function prediction methods. This method achieved the best result in latest CAFA competition.
- We proposed a BLAST based ensemble method. We transfer the best method between similar protein sequences. Our method achieves a better result than GOLabeler on our test set.
- Python, along with its packages like ipython, numpy, pandas, was used as major tools for data analysis in the process; we also wrote shell scripts to deal with simple tasks.

“Handwriting recognition with neural network”

Apr 2017

Course Projects

- Constructed a neural network with one hidden layer comprised by 25 units; data set contained 5,000 handwriting pictures where each training example was a 400-dimensional vector unrolled by grey scale pixel images.
- Adopted sigmoid activation function and regularized cost function; learned parameters by backpropagation algorithm.
- Visualized the hidden layer by choosing input vector that gave the highest activation for the hidden unit; found hidden units correspond roughly to detectors that looked for strokes and other patterns in the input.

Programming Skills

- C++, Python, R, MATLAB, MySQL, Linux Shell