

Honglin Yu

School of Computer Science, Australian National University, NICTA
Homepage : <http://yuhonglin.github.io>

Email: honglin.yu@anu.edu.au
Mobile: +61 406502886

Research Interests Data Science, Machine Learning, Large-scale Online Media Analysis

Education **Australian National University** Aug. 2011 - Aug. 2015
Ph.D. in Computer Science

- Thesis submitted: Understanding the Popularity Evolution of Online Media: A Case Study on YouTube Videos
- Affiliated with Statistical Machine Learning group of NICTA

Southeast University, China Aug. 2009 - Jul. 2011
Finished master course in Automation

Southeast University, China Aug. 2005 - June 2009
Bachelor of Automation

- Rank: 4/140
- Best Graduation Thesis (2 out of 140)

Skill Highlights

- Experienced in **Data Science** and **Statistical Analysis**
 - Classification, regression, clustering, time series analysis
 - Data exploration and visualization (matplotlib, R, d3.js)
 - Working on large scale **real data** for 4 years
- Broad knowledge in **Machine Learning**
 - Deep understanding of main-stream algorithms (SVM, Random Forest, ANN etc.)
 - Experienced in carrying out experiments and **feature engineering**
- Firm programming skills, especially in **C/C++** and **python**
 - Also familiar with SQL, Javascript, R, Julia
 - Using Linux everyday for more than 6 years, skilled emacs user
- Experienced in collecting and analyzing large scale Internet data
 - Skilled in implementing various online data **crawlers**
- Domain knowledge related to Computational Social Science Research
 - YouTube video popularity analysis
 - Twitter user network analysis
- Basic controller design and control theory

Research Projects Please see <http://yuhonglin.github.io/project/> for more details

- Analyzing Twitter-driven YouTube Views
 - Predicting videos' viewcount with Twitter's feed
 - Contributions: Most of the ideas; All the Software implementation, data collection, analysis and prediction
- Data Mining on YouTube Video Shot Sequences
 - Mining frequent remixed shots from large number of videos
 - Contributions: Part of the ideas and data collection; All the implementation, data preprocessing and analysis
- Exploring the Phases of Popularity Evolution of YouTube Videos
 - Understanding viewcount dynamics by novel time series segmentation techniques
 - Contributions: Most of the ideas; All the software implementation, data collection, analysis and prediction

Software Highlights	<ul style="list-style-type: none"> • YTCrawl: a YouTube video history viewcount crawler • SegFit : a time series segmentation algorithm written in C++ • Shotdetect : a video shot detection program written in C++ 	
Media Coverage	<ul style="list-style-type: none"> • ANU Reporter <i>How the viral video star is born.</i> August 2015. http://www.anu.edu.au/news/all-news/how-the-viral-video-star-is-born • NCI Research News <i>Predicting popularity.</i> September 2015. http://nci.org.au/2015/09/30/predicting-popularity/ 	
Papers	<ul style="list-style-type: none"> • Honglin Yu, LEXING XIE AND SCOTT SANNER, Exploring the Popularity Phases of YouTube Videos: Observations, Insights, and Prediction. In proceedings of International AAAI Conference on Web and Social Media (ICWSM) 2015. • Honglin Yu, LEXING XIE AND SCOTT SANNER, Twitter-driven YouTube Views: Beyond Individual Influencers. In proceedings of ACM Multimedia Conference (ACMM) 2014. • Honglin Yu, LEXING XIE AND SCOTT SANNER, Phase-sketch Clustering on Large Scale Time Series Dataset. (<i>Working paper</i>). • Honglin Yu, Supplementary Observations on YouTube Video Popularity Evolution: Shocks, Periodicity and Correlation with Twitter. (<i>Working paper</i>). • Marian-Andrei Rizoiu, Lexing Xie, Scott Sanner, Manuel Cebrian, Honglin Yu, and Pascal Van Hentenryck. Can This Video be Promoted? – Linking Endogenous and Exogenous Processes to Forecast Popularity on Youtube. (<i>Working paper</i>). 	
Teaching Experience	<ul style="list-style-type: none"> • Research Group Tutorial on Random Forest 2014 Slide: http://www.slideshare.net/yuhonglin/main-33219459 • Research Group Tutorial on Support Vector Machine 2013 Slide: http://www.slideshare.net/yuhonglin/main-44194421 • Computational Social Science Summer School 2013 Website: http://users.cecs.anu.edu.au/~xlx/teaching/css2013/index.html Teaching Assistant • Computer Network 2011 Teaching Assistant, Southeast University, China 	
Awards	<ul style="list-style-type: none"> • Excellent Graduation Dissertation of Southeast University (2 out of 140) 2009 • Champion of the 3D Soccer Simulation League, Robocup Worldcup 2008 • National Mathematical Modeling Contest Third Prize of Jiangsu Province 2008 • Ailipu Scholarship 2008 • Higher Mathematics Competition of Jiangsu Province, Third Prize 2006 	
Review Experience	<ul style="list-style-type: none"> • International World Wide Web Conference (WWW) • ACM International Conference on Web Search and Data Mining (WSDM) • International AAAI Conference on Web and Social Media (ICWSM) 	
Language	Chinese (native)	English (fluent)