

Yao Lu

<http://yaolu.me>

Room 308, iLab Tongji, School of Software Engineering
4800 Cao'an Road, Tongji University, Shanghai, China
+86 188-1820-6109 | 95luyao@gmail.com

EDUCATION

Tongji University

B.Eng. in Material Science and Engineering

- GPA: 4.3/5.0

Shanghai, China

Sept.2012 - Jun.2017 (Expected)

University of Alberta

Dept. of Computing Science *Visiting Scholar*

- MITACS Scholar (co-funded by China Scholarship Council and Canada Government)
- Research Supervisor: Prof. Randy Goebel and Dr. Mi-Young Kim

Edmonton, Canada

June.2016 - September.2016

RESEARCH EXPERIENCES

Interest: Machine Learning, Natural Language Processing, Graph Analysis, Computer Vision

Salient Fact Mining in Wikipedia

Directed by Prof. Yanghua Xiao, Fudan University

GraphLab@Fudan

Jan.2017 - Present

- This project aimed at finding the salient fact in the knowledge base, for example, Stephen Hawking will have a fact "disability" much more surprise than "physicists". The salient fact is the difference between prior and posterior, the prior can be viewed as the common feature of a specific category. We can find the most salient facts of entities by calculating the distance between prior and posterior.
- Using distant supervision with deep neural models and multi-instance learning to extract information from unstructured category description text. Combining cooccurrence information to compute the semantic similarity between categories. Quantify intrinsic and extrinsic surprise category of entities
- Overcome the link and information missing effect on calculation in comparison to previous methods

Spatial Association Gene Network Analysis

Directed by Prof. Tianwei Yu, Emory University

YuLab@Emory

April.2016 - Present

- A project aimed at integrating expression data with biological networks to find dynamic relationships which has computational hurdles to overcome. The goal of the project was to find vertices around which local expression consistency change significantly between clinical conditions.
- We proposed a new method called DNLC (Differential Network Local Consistency) which can extract significant vertices that are not detected using existing methods.
- Preparing for paper submission & CRAN R Package released

Microsoft Academic Graph Mining

Collaborate with Prof. Randy Goebel, University of Alberta and Dr. Jing Yuan, Microsoft

Microsoft Research Asia

July.2016 - Present

- A project aimed at finding better method to evaluate young researchers in comparison to the h-index and citation number. We use Convolutional Neural Network for the feature extraction for modeling of academic quality. Apply graph embedding to quantify the social impact. We propose a joint learning framework capable of adding more context information for modeling
- The project has been presented at Microsoft Developer Conference

Topic Model Based Microblog Spammer Detection

Collaborate with Prof. Renxian Zhang, Tongji University

iLab&NLP@Tongji University

Oct.2015 - April.2016

- The project aimed at detecting smart spammers. This kind of spammers with profiles and tweets are difficult to identify. We proposed a Topic model based spammer detection approach to detect fake accounts in microblog platform. Our method outperform other methods.
- Co-authored paper accepted by **NAACL-HLT Student Session**, San Diego, 2016

Biologically Inspired Scene Recognition

Directed by Prof. Ye Luo, Tongji University

iLab@Tongji University

Mar.2016 - Present

- Using biologically plausible method of Visual Saliency and Gist Descriptor to recognize scenes
- Modified the feature extraction method proposed by Itti et al.
- New features with Linear SVM achieve 10% higher than previous method

PROJECTS AND COMPETITIONS

Beauty of Programming Competition

May.2016

Microsoft Research Asia

- One of the most famous programming competition in China
- Rank 3 of 1000 teams in the competition
- Project about Microsoft Academic Graph(MAG) search application

Microblog Social Trend Prediction Competition

Aug.2013 - Feb.2014

Alibaba Inc. & Sina Inc.

- Time series analysis to predict the popularity of Microblog posts
- Top 10% among total 2293 teams
- Develop the competition method to academic paper

Open Source Contribution

Differential network local consistency (DNLC)

Authors: **Yao Lu**, Yusheng Ding, Linqing Liu and Tianwei Yu

R Package

Dec.2016

- Project about detection local consistency
- Package available in R-CRAN ([Package Link](#))

PUBLICATIONS

In Progress

Yao Lu, Yusheng Ding, Linqing Liu, Jianwei Lu and Tianwei Yu "Detecting Differential Network Local Consistency" In Progress

Published

Linqing Liu, **Yao Lu**, Ye Luo, Renxian Zhang, Laurent Itti, and Jianwei Lu. "Detecting" Smart" Spammers On Social Network: A Topic Model Approach." In Proceedings of NAACL-HLT, pp. 45-50. 2016.

Mi-Young Kim, Ying Xu, **Yao Lu**, and Randy Goebel. "Legal Question Answering Using Paraphrasing and Entailment Analysis." In Tenth International Workshop on Juris-informatics (JURISIN). 2016.

HONOURS & AWARDS

16' **Top 3 of Beauty of Programming Competition 2016**, Microsoft Inc. (Top 0.3%)

15' **First-class Scholarship**, Tongji University (Top 5%)

First Prize in National Mathematics Competition, Ministry of Education (Top 5%)

14' **First-class Scholarship**, Tongji University (Top 5%)

Outstanding Student Award, Tongji University (Top 2%)

First Prize in Physics Competition, Tongji University (Top 1%)

TECHNICAL STRENGTH

Programming Languages Skills

C/C++, Python, R, Matlab, L^AT_EX, Shell Script

Packages&Tools

Spatial Analysis, Natural Language Processing, Deep Learning
Tensorflow, NLTK, NetworkX, OpenCV, Sci-kit Learn, Git