# Tao Sun

Mobile: (219)-386-8179
Email: ts202115@ohio.edu

## Personal statement

In Fall 2016, I was enrolled into the PhD program of Computer Science in Ohio University. In the meantime, I am a second-year candidate of Master of Mathematics (Computer Science track) in the same university. My research interest is in computer vision combing natural language processing methods and theoretical analysis of deep learning training methods. My last degree was Master of Science in Engineering (Physical Electronics) from Huazhong University of Science and Technology in China, which was conferred more than 10 years ago. After graduation, I entered the software industry. As a seasoned software engineer, I had 10 years' successful experience, working for 9Spaces and SAS successively.

## Academic projects

**Final Project of Natural Language Processing**, Spring 2016

**Title**: Acquisition of Topic Signatures with Hypothesis Tests Based on Word Embedding
**Description**: A topic signature is a vector of terms associated with a topic. In this project, a class of adapted hypothesis tests, which leverage on semantic information provided by word embedding, is proposed. Taking self-categorized document sets extracted from Wikipedia as the corpus, topic signatures are acquired with the adatpted methods for each category of documents. For each topic signature, both the terms computed by the adatptions and that computed by general hypothesis tests are searched as keywords in the test corpus through Lucene. $F_{0.5}$-scores for both results are compared to evaluate the adapted methods. The results illustrate that $F_{0.5}$-scores are improved significantly, which indicates that by fully exploiting semantic information the proposed methods at least partly resolve the noise issue existed in the general hypothesis tests.

**Final Project of Machine Learning course**, Fall 2016

**Title**: Classification of Alzheimer's Disease and Normal Congnitive Status with Recurrent Neural Networks in Resting State fMRI
**Description**: A deep learning framework is designed providing assist to the diagnosis of Alzheimer Disease (AD). The framework is a combinatiion of Deep Auto-Encoder (DAE) and Recurrent Neural Networks (CNN). Initially, the resting-state functional Magnetic Resonance Imaging (rs-fMRI) data is preprocessed and time series of Regions of Interest (ROIs) mean values are extracted. Through a DAE the high-dimensional data

is reduced to a lower dimensionality and then is splitted into multiple identical-sized sub-series. After that, a RNN classifier is trained on the sub-series classifying each sub-series as either AD or Normal healthy Control (NC). Finally, the diagnosis suggestion of a subject is made by ensemble of the outputs of the sub-series classification results. However, the experiments on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset shows the framework fails to generalize the trained model to the test data. After analysis, we conclude that the preprocessed rs-fMRI time-series data is not suitable to classify the Alzheimer Disease.

# Work history

**Technical Lead**, Data Visualization Group of SAS Institute, Beijing
April 2015 – July 2015

My role involves leading a team developing and maintaining SAS BI Dashboard.

**UI Common Component Developer**, HTML5 Group of SAS Institute, Beijing
September 2013 – March 2015

As a core team member, I am focusing on the development of Form Control Component.

Form Control, a major HTML5 component in SAS, is designed to collect user-input values. Based on a data model, it generates various UI controls and organizes them into hierarchical forms.

Strengths and achievements:

- Overall object-oriented design of Form Control to meet the requirements, in the meantime considering the extension to Property Sheet, another major SAS HTML5 component.
- Designed and implemented fundamental features of Form Control, including but not limited to responsive web solution, data model and binding and transparent group.
- Fixed most of obstructive bugs.
- Developed other miscellaneous HTML5 components.

The essential supports I provided were crucial to its successful releases. And I was also promoted to a lead position due to my significant contributions to it.

**Software Engineer**, Business Intelligence Group of SAS Institute, Beijing
September 2013 – March 2015

Business Intelligence (BI) Group was responsible for the plugin development of SAS Environment Manager (EV) and SAS Visual Data Builder (VDB).

Strengths and achievements:

- Solved a long blocking bug of Data Spreadsheet, with that safe guarded *File Importer* project for the team and then winning the 3C (Collaboration, Communication and Consistency) Award of SAS 2012.
- Set up the prototype of *File Importer* and led the development of it technically.
- To fix some long-standing bugs, rewrote library plugin of EV.

**Team Lead and Softwar Engineer**, 9Spaces Inc., Beijing and Guangzhou
September 2005 – August 2011

I was focusing on an outsourcing project, whose owner was one of 9Spaces' strategic partners from May 2008 to August 2011.

My responsibilities included designing and implementing applications, collaborating with remote team members to ensure on-time release of high-quality code and project deliverables.

From September 2005 to May 2008, I participated each phase of design and development of 9Spaces.com. In the second half of 2007, I started to lead a team working on payscale.cn, Chinese version of famous salary survey website payscale.com.

## Skills and abilities

**[Development Environment]**
- Unix, Linux
- Windows

**[Programming]**
- Python (Tensorflow, nltk, NumPy, SciPy)
- Java (Java EE, Spring, Lucene, GWT, Heritrix, IBatis etc)
- Javascript (JQuery, Dojo, OpenUI5)
- Actionscript (Flex)
- Ruby

**[Data Science]**
- Deep Learning
- SAS

- Machine Learning
- Natural Language Processing

**[Software Engineering]**
- Object-oriented Analysis and Design (OOAD)
- Project management and agile development (Scrum)

# Education

**Doctor of Computer Science and Master of Mathematics**
Ohio University, Spring 2016 - Now

Course list (GPA 3.934):
*Computational Theory*
*Operating System*
*Natural Language Processing*
*Machine Learning*
*Advanced Image Analysis*

**Master of Science in Engineering**, Physical Electronics
Huazhong University of Science and Technology, September 2002 – June 2005

**Bachelor of Engineering**, Optoelectronics
China Jiliang University, September 1998 – June 2002

# Literature

Sun, T., Zhu, D., Yang, Z., Liu, Z., & Liu, Y. (2006). Theoretical predictions of photonic properties of nanoporous copolymer films as photonic band gap materials using FDTD. *Applied Physics B*, *82*(1), 89-92.

Sun, T., Zhu, D., Yang, Z., Liu, Y., & Wu, F. (2005, January). Analysis of optical waveguiding properties of self-assembled block copolymer films using FDTD method. In *Asia-Pacific Optical Communications* (pp. 208-214). International Society for Optics and Photonics.

Wu, F., Zhu, D., & Sun, T. (2005, January). Analysis and design of low chromatic dispersion in flat-top AWG with parabolic waveguide horn. In *Asia-Pacific Optical Communications* (pp. 605-611). International Society for Optics and Photonics.

# Certifications

**Machine Learning**
Coursera.org
March 2014 – May 2014


# Awards

**SAS 2012 3C (Collaboration, Communication and Consistency) Award**
December 2012