

Maochen Guan

444 Washington Blvd, Apt 4417, Jersey City, NJ 07310
Phone: +1 917-684-1317

E-Mail: mg3364@nyu.edu
Website: www.maochen.org

EDUCATION

M.S., Computer Science, **New York University**, NY, USA (2011 to 2013) **GPA: 3.80/4**
B.S., Electrical Engineering, **Donghua University**, Shanghai, China (2005 to 2009)

PROFESSIONAL EXPERIENCE

IPsoft Inc. New York, USA 11/2013 - Current

- **Lead R&D Engineer (Specialized in Natural Language Processing)**

Constructing a complex Sentence Similarity computing system using word vector representation.
Developed rule based co-referencing recognizing in Semantic Role Understanding System.
Implemented an adapter layer for Stanford Parser to accept CoNLL-U format training data.
Recognize and auto combine proper nouns.
Trained sentence type recognizer by using MaxEnt supervised model.
Trained dialog act recognizer by using LibLinear model.
~30% accuracy improvement on original dialogue scripts for 82 related topics.
Implementing Named Entity Tagger to AI Interactive system.
Developed and optimized IPsoft specific model for Stanford PCFG and shift-reduced Parser.

ByteConsulting Inc. New York, USA 06/2013 – 09/2013

- **Application Developer**

Developed PhoneGap E-Market financial mobile application with REST-ful web service provider.

Industrial and Commercial Bank of China Shanghai, China 06/2009 to 04/2010

- **Software Development Engineer**

Built currency recognition tool for foreign transaction system.
Concurrent threads optimization for 3rd party's transaction systems.
Enhanced framework of HR platform to separate data structure and function operations.

PROJECT EXPERIENCE

- **Feature Extraction Optimization for Multicore Architecture** 12/2012

Concurrency Programming: Parallel feature extraction process for NLP. (Java 1.7.0)
Data level parallelism in training process (Thread Pool Model).
Profiling and optimize the CKY for parsing process.
TDD + Unit test case covered.

- **Natural Language Processing: Sentence Sentimental Analysis** 04/2012 to 05/2012

Supervised Machine Learning Classification System. (Java 1.6.0)
Target: Classify tweet's sentiment extracted from twitter.com into Positive, Negative or Neutral.
Approach: Maximum Entropy classifier (Grammatical Model) + Bigram (Lexical Model backup) for prediction.
Self-defined feature extraction.
F1-Score: 88% (~5000 Training Samples, Cross Validation).
Website: <http://www.tweetemotion.com>

Personal Interests

- Word2vec with syntax feature

Working and developing word2vec model with syntax role similarity using dependency parse tree.

Ex: Given “Apple”, instead returning “cellphone”, “ios”, “California”, it tries to return “HTC”, “Samsung”, “Microsoft” etc.

COMPUTER SKILLS

- Java (proficiency), C#/ASP.NET, C/C++.
- Natural Language Processing, Statistical Machine Learning, Regex.
- OO Design Pattern, Agile Development, TDD.
- Git, Subversion.

GitHub

<https://github.com/maochen>