DongChao Resumé

neutronest@gmail.com
www.neutronest.moe
github id: neutronest

WORK EXPERIENCE

2015.03-Now Research on NLP (National Lab. of Intelligent Technology & Systems, Tsinghua University)

Research Intern

- Sentiment analysis on microblog data using CRF, HMM model.
- Sentiment analysis on dialogue data preparing for ACL and SIGIR.
- Deep learning approaches and related applications: CNN, RNN, LSTM &etc.
- Probabilistic graphical model: CRF, HMM.

2013-2014 Software Developer (Xiaomi Company)

Working in Miliao team (MiCloud), as a backend developer.

2012.08-2012.09 Software Developer (GeekPark)

PHP intern (one month)

EVENTS

2015-08 Machine Learning Summer School 2015 Kyoto (MLSS2015)

Participant as undergraduate (two weeks)

2014-06 International Conference on Machine Learning (ICML2014)

Participant in ICML2014 (five days)

PROJECTS

Micoblog Analysis (Tsinghua)

- capture the context of microblogs.
- suitable node features and edge features proposed to represent some information between the parent microblog and the children microblogs.
- Use CRF-SVM model to conduct this task.
- about 4% performance increase.
- prepare for SIGIR2016.

Dialogue Act Analysis (Tsinghua)

- Deep Learning approaches on DA classification for sequence data.
- RNN, LSTM and corresponding variants used to capture the context.
- More experiments is currently nin progress.
- Prepare for ACL2016.

Antispam (Xiaomi)

- feature engineering: generate features that identified spam users effectively. performance: 90%~95% accurate.
- user confidence coefficient analysis:

- collect data from distributd log system Scribe.
- store data and index into Hive, HBase.
- use linear models to measure the training data.

Miliao Muc (Xiaomi)

- a group chat subproject in Miliao.
  - features developed with Java and corresponding middlewares: Maven, Zookeeper, RabbitMQ, etc.

EDUCATION

Tianjin University

Management Information System 2009-2013

SKILLS

Python, Java; Linux, Git,

Haskell, OCaml just a little.

INTERESTED

Machine Learning:

```
DeepLearning: sequence models: RNN, LSTM, GRU and other variants.
              computer vision: CNN, cascade CNN, CNN with multi-masks and other variants.
              corresponding applications: sentiment analysis, image classification, etc.

Large Scale ML: GraphLab, Spark etc.

Other Machine Learning topics: high dimensional learning, representation learning, AI&Law, etc.
```

Functional Programming:

```
Programming languages: Haskell, Ocaml with theories related: Lazy Evaluation, Monad,
                       pure functional data sturcture...

Basic theories: lambda calculus, type theory, category theory ...
```