

# Wei Luo

Technical Expertise Includes: Machine Learning · Computing Advertising · Natural Language Processing.

Age: 29 | Tel: (+86) 18911026978 | Email: vividfree@qq.com | Blog: <http://vividfree.github.io/>

## EDUCATION

- ◆ Master of Science in Computer Application Technology (concentration in natural language processing), Institute of Computing Technology of the Chinese Academy of Sciences (ICT, CAS), 2012. GPA: 86.83.
- ◆ Bachelor of Science in Software Engineering, Huazhong University of Science and Technology (HUST), 2009. GPA: 90.70 (Rank: 2/209)

## MAIN SKILLS

- ◆ Technical expertise includes: Machine Learning, Computing Advertising, Natural Language Processing;
- ◆ Skilled in software development and own good coding style;
- ◆ Familiar with C / C++ / Python / Shell programming language and computing tools such as TensorFlow / Spark / Hadoop / Hive;
- ◆ Experienced in analyzing business requirements from three aspects which are product, algorithm and engineering, and own team cooperation spirit and project management experience;
- ◆ Own good English skills: listening, speaking, reading and writing. IELTS(6.0), CET-6(525), CET-4(580)

## WORK EXPERIENCE

2016.01 - present      360 – business data department – data mining team      senior algorithm engineer

- ◆ **Query suggestion server**      project owner & core member
  - ✧ Object: support requirements from business product, such as 360 DMP<sup>1</sup>, 360 ShangYi<sup>2</sup> and account optimization in 360 DianJing<sup>3</sup>, etc.
  - ✧ Building the query similarity model based on a deep learning model. The accuracy in test set is 95%+.
  - ✧ Launch the query suggestion server mainly based on product quantization algorithm which is a similarity search algorithm.
  - ✧ Now is attempting to optimize the query similarity model by using multi task learning on term weight model and query similarity model.
- ◆ **Look-alike targeting in 360 DMP**      project owner & core member
  - ✧ Object: analyze the known target crowd in order to find more similar target audiences and expand the target marketing coverage.
  - ✧ Mobile look-alike targeting: first use several strategies to mine similar APPs, then use a heuristic algorithm to rank hundreds of millions of 360 users for every app, in order to find more similar target audiences.
  - ✧ PC look-alike targeting: use logistic regression model to train and rank all 360 users for every advertiser in order to find similar target audiences.
  - ✧ By-product: based on the bagging strategy, develop a large-scale word2vec train and inference tool. Compared with word2vec in Spark MLlib, precision@5 increases from 69% to

<sup>1</sup> Homepage of 360 DMP: <http://dmp.360.cn/>

<sup>2</sup> Homepage of 360 ShangYi: <http://shangyi.360.cn/>

<sup>3</sup> Homepage of 360 DianJing: <http://e.360.cn/>

81% in similar app mining.

- ◆ **Demography Prediction for feature targeting in 360 DMP** project owner & core member
  - ✧ Mobile demography includes dimensions such as gender, life stage and life expert. First use several strategies to mine similar APPs, and construct dictionaries for the mentioned dimension, then use a heuristic algorithm to rank hundreds of millions of 360 users for every tag in every dimension.
  - ✧ PC demography includes dimensions such as gender, age and purchase ability, etc. Use Max Entropy model and Naïve Bayes model to train and rank all 360 users for every tag in every dimension. The precision for gender, age and purchase ability is 90.6%, 78.3% and 63.2% respectively, the AUC is 0.948, 0.813 and 0.740 respectively.

**2013.12 – 2015.12 360 – business product department – data mining team algorithm engineer**

- ◆ **Trigger and rank module of high commercial query in ad network** core member
  - ✧ High commercial query trigger: in charge of content targeting and query targeting strategy.
  - ✧ Responsible for ranking high commercial query.
  - ✧ Achieve the results that CTR1 increases by 180% and CPM1 increases by 120% among high UV sites.
  - ✧ By-product: large-scale sparse matrix multiplication based on MapReduce.
- ◆ **Trigger module in display ad** project owner & core member
  - ✧ Object: the old ad engine built reverse index from tag to ad, and this scheme suffered from long reverse index, the engine needed to do hard pruning when searching which hurt CTR. Because of this problem, we planned to build reverse index from conjunction to ad.
  - ✧ Method: first compute statistical CTR score for ad and conjunction pair, then compute rank score for ad and conjunction pair concerning bid price.
  - ✧ Achieve the results that CTR1 increases by 30% and CPM1 increases by 10%.

**2012.7 – 2013.11 People Search – public sentiment department – data mining team algorithm engineer**

- ◆ For purpose of monitor public sentiment in microblog data, be responsible for several strategy modules such as area detection, positive and negative emotion classification and bloggers' industry detection and one engineering module which is streaming based microblog data processing.
- ◆ In charge of designing the social business product, its main functions include social media management, social monitoring, social account analysis and potential customer mining.

**2011.7-2011.9 Hulu Corporation (Beijing) – video search team software engineer (intern)**

- ◆ Optimized the query understanding module and spell correction module. The accuracy of spell correction is improved from 75% to 78%.

## **RESEARCH EXPERIENCE**

**2011.9-2012.5 Research on fast incremental training algorithm for word alignment**

- ◆ Word alignment is the bottleneck in building the translation model. As for unsupervised learning based word alignment, when new training data is coming, use online EM algorithm to replace batch EM algorithm because online learning algorithm could help the convergence of model more fast.
- ◆ Experiment results: The proposed method is 5 times faster than batch EM and would not hurt the quality of word alignment and translation.

**2010.5-2010.9 International Workshop on Spoken Language Translation Evaluation (IWSLT 2010)**

- ◆ Implemented rule-based Named Entity translation systems for Chinese-to-English and English-to-Chinese respectively. Optimized the hierarchical phrase-based statistical MT decoder.

- ◆ **Campaign result:** Ranked top 2 in various tracks among 11 world research teams.

#### **PUBLICATIONS & PATENT**

- ◆ During working at 360: 16 pieces of patent are submitted as the first author, including one piece whose legal status is "entitled - authorized", 3 pieces as "trial - actual trial" and 5 pieces as "trial - public". (By the end of January 2018)
- ◆ **Wei Luo.** Research on Fast Incremental Training Algorithm for Word Alignment. Acta Scientiarum Naturalum Universitatis Pekinesis, 49(1).
- ◆ Hao Xiong, Jun Xie, Hui Yu, Kai Liu, **Wei Luo**, Haitao Mi, Yang Liu, Yajuan Lv and Qun Liu. 2010. The ICT Statistical Machine Translation Systems for the IWSLT 2010. In Proceedings of IWSLT 2010, Paris, France, December.
- ◆ **Wei Luo**, Zongcheng Ji, Yajuan Lv, Qun Liu. A Novel Approach to Improve Word Alignment, In YWCL2010, Wuhan, Oct.2010.

#### **HONORS & AWARDS**

##### **During working at 360:**

- ◆ Outstanding employee of 2017, star of the first quarter in 2016;
- ◆ Performance review: rank top 5% in 2017, several times rank top 20%.

##### **During studying at ICT, CAS:**

- ◆ All-round excellent student (three times).

##### **During studying at HUST:**

- ◆ Outstanding thesis of bachelor of Hubei province (2009), outstanding graduate of HUST (2009);
- ◆ National Scholarship (once), all-round excellent student of HUST (twice);
- ◆ Outstanding student cadre (once), excellent student for academic performance (once).