

# WEI XU

✉ readonlyfile@hotmail.com · ☎ (+86) 1800-451-5446 · 🔗 <http://memeda.github.io>

## 🎓 EDUCATION

---

**HARBIN INSTITUTE OF TECHNOLOGY**, Harbin, China Sep. 2015 -- Present

*Master student* in Computer Science (CS), expected Jul. 2017

**HARBIN INSTITUTE OF TECHNOLOGY**, Harbin, China Aug. 2011 -- Jul. 2015

*B.S.* in Computer Science (CS), grads among top 10%

## 👤 EXPERIENCE

---

**Baidu Inc.** Beijing, China Jun. 2014 -- Sep. 2014

*C/C++ , Python , shell , MapReduce* NLP Intern

Deployment for knowledge base extraction procedure In Hadoop cluster

- Verify kernel procedure's correctness . Promote shell program performance .
- Deploy them in 5 Hadoop clusters and process data from 20K web database .
- 200+ TB knowledge obtained after continues 1 month .

**The Design and Implementation for Classification System on unbalanced datasets**

Mar. 2015 -- Jun. 2015

*Python* Graduation Design

Two systems has been build , one for good generalization , the other for good classification result .

- Totally experiments for combination of different features , classifiers on balanced and unbalanced datasets
- Different methods have been demonstrated for unbalanced datasets to improve classification effect

**Yinshi map**

Oct. 2014 -- Present

*Python , MySQL* Extending and Maintenance

Calculate relations among various foods , regions and others from Weibo data in the backend . Display it in frontend , <http://ys.8wss.com/> .

- Make sense of the frontend quickly , and fix the BUG .
- Extend backend function , modify and deploy relation extraction procedure .

## ⚙️ SKILLS

---

- Programming Languages: C/C++ is preference , familiar to Python , have experience in Shell , JS , Java
- Development: experience in Scripts , Web , little for QT , MFC , Java Application
- NLP fields : familiar to Text Classification , Chinese Word Segmentation , knowing some about sequence labeling by Neural Network

## ♥️ HONORS AND AWARDS

---

**National Scholarship**

Oct. 2014

## 📖 MISCELLANEOUS

---

- Techblog: <http://blog.csdn.net/readonlyfile>
- GitHub: <https://github.com/memeda>
- Foreign Language: English - CET-6