Dian Chen

Final Year (Ph.D candidate)
Machine Learning & Data Mining Group,
Institute of Computing Technology (ICT),
Chinese Academy of Sciences (UCAS)

UCAS, Haidian District, Beijing
Mob.: +86-18883383965
Email.:okcd00@qq.com
Web.:http://www.okcd00.tech/about

Homepage

Linkedin.com/in/**okcd00**
Blog.csdn.net/**okcd00**
Github.com/**okcd00**
- **Contributed** to 3 repos (748 stars).

Skills

MAJOR
NLU/NLP, CSC, Deep Learning,
Information Retrieval/Extraction;

PROGRAMMING ENVS
VSCode, PyCharm, DevC++;
JupyterLab, SecureCRT, Termius;

LANGUAGES
Python, C++, CET-6, Mandarin-2A;

FRAMEWORK
Pytorch, Tensorflow 1.x

Research

Entity/Relation Recognition
NER (IPMC Journal Pending)
RE (FCS Journal Accepted)
RE (CIKM Conference Accepted)
RE (NLPCC Conference Accepted)
Chinese Spelling Correction
CSC (EMNLP Conference Pending)
CSC (Outstanding Project with **HKEX**)
Text to SQL Query
Text2SQL (CIKM Conference Pending)
Text2SQL (Awarded at **SZSE TechConf**)

Education

2016-2022
PH.D IN CS
Institute of Computing Technology,
Chinese Academy of Sciences (CAS)

2012-2016
COLLEGE
ChongQing University, CS, IOT
GPA: 3.60/4.00 (Rank 2)

2009-2012
HIGH SCHOOL
High School Affiliated To Nanjing Normal University
Score: 359/480 (1st-class line: 340)

Experience

2017-NOW **PaodingAI** **Reseaching Group & Algorithm Engineer**

AutoDoc : The project focused on extracting various types of information from financial documents, proofreading for correctness (faulty values and typos), and then giving advice on changes.
⇒ I designed the models, rules and KBs for the NER, RE and CSC tasks.
Foundry : A Natural language processing AI platform, through the whole process of annotation, training, and prediction, integrated processing of text semantics. The platform is used flexibly for a variety of basic information extraction tasks.
⇒ I was involved in the design of different tasks on this platform.

PyTorch-Lightning, Tensorflow, Transformers, Requests, PostgreSQL
Projects: **Glazer** , **AutoDoc** , **Foundry** , *Mantra, P5 (more details in* **linkedin** *)*

2015-2017 **Big Data Lab (BDL), Baidu** **Research & Data Intern**

Implement methods to automatically and incrementally crawl large amounts of data, both platform-based and stand-alone. Write algorithms and models for processing and mining valuable information.

PySpider, BeautifulSoup4, ProxyPool, Map-Reduce, Sklearn, Theano
Projects: Sentiment Polarity Analysis, Recommendation for offline retailing

Achievements/Awards

2016 **CCF Outstanding Student of the Year Award** **Top of the college**

Received the CCF Annual Outstanding Student Award and was invited to CNCC 2016 to receive the award.

2016 **CCF-CSP Certification** **Top 5% of all-time rankings**

CCF-CSP certification is a software professional competency certification program held by CCF for college students.

2015 **National Scholarship** **Top of the class**

Scholarships awarded by the Ministry of Education of the People's Republic of China.

2014-2016 **ACM-ICPC Awards** **1 Gold, 3 Silver and 2 Bronze medals**

1 Silver medal (Shanghai) and 2 Bronze Medals (Beijing, Anshan) in ACM-ICPC Asia Regional Competitions. 1 Gold medal (Chongqing) and 2 Silver medals (Sichuan) in ACM Provincial Competitions.

Publications

[C] **Explicit Modeling the Context for Chinese NER** **CCF-A (Pending)**

Dian Chen , Yixuan Cao, Ping Luo: Explicit Modeling the Context for Chinese NER. *ACL-ARR: Pending.

[C] **Nested Relation Extraction with Iterative Neural Network** **CCF-B**

Yixuan Cao, **Dian Chen** , Hongwei Li, Ping Luo: Nested Relation Extraction with Iterative Neural Network. CIKM 2019: 1001-1010

[C] **Nested Causality Mining on Financial Statements** **CCF-C**

Dian Chen , Yixuan Cao, Ping Luo: Pairwise Causality Structure Towards Nested Causality Mining on Financial Statements. NLPCC 2020: 725-737