Mieradilijiang Maimaiti

miradel_51@hotmail.com | +86-13051308938 | Alibaba Xixi Park A, No.2 building, Hangzhou
0000-0002-4231-6519 | miradel51 | https://www.miradeljan.com

Education

09/2007 – 07/2008 **Preparatory, Chinese Learning Intensive Program**
School of Chinese Language and Literature, Xinjiang University

09/2008 – 07/2012 **B.S, Computer Science and Technology**
School of Information Science and Engineering, Xinjiang University
Xinjiang Multi-Lingual Information Technology Key Laboratory

09/2012 – 06/2015 **M.S, Computer Software Engineering**
School of Information Science and Engineering, Xinjiang University
Xinjiang Multi-Lingual Information Technology Key Laboratory

09/2015 – 06/2021 **Ph.D, Computer Science and Technology**
Department of Computer Science and Technology, Tsinghua University
THUNLP Group

Work Experiences

07/2021 – Present **NLP Senior Algorithm Engineer, ALIBABA DAMO ACADEMY**
Multi-lingual Group, Machine Intelligence
- Base: Hangzhou, Zhejiang, China
- Mentor: Luo Si, Fei Huang, Ji Zhang
- Project: Multi-lingual processing for international E-commerce platforms DARAZ, LAZADA and AliExpress

07/2020 – 10/2020 **Research Intern, ALIBABA DAMO ACADEMY**
Multi-lingual Group, Machine Intelligence
- Base: Hangzhou, Zhejiang, China
- Mentor: Haiqing Chen, Ji Zhang
- Project: Urdu-Roman Transliteration
- Contribution: I have presented a hybrid model that can address the main challenges (many-to-one mapping and inconsistent mapping) in this task. Besides, the proposed approach is fused into the dialogue system of international E-commerce platforms DARAZ and LAZADA.

07/2019 – 12/2019 **CTO, Beijing Silk Road Heli Investment Management Group Co.,Ltd**
Information Services Department
- Base: Beijing, China
- Project: Multi-Lingual E-commerce Platform
- Contribution: I had directed our team to develop the multilingual E-commerce platform whose frontend and backend were written by ours rather than using other's frameworks. Meanwhile, it has strong flexibility, and it is an extendable system that can be changed to online teaching or online training system effortlessly.

Selected Projects

03/2011 – 09/2011 | **National Natural Science Foundation**
"Uyghur Hand Writing Identification"
- Base: XJU, Wulumuqi, Xinjiang, China
- Mentor: Engineering Academician Prof. Wushouer Silamu
- Contribution: Developing (C++) and System testing.

04/2012 – 09/2012 | **National Natural Science Foundation**
"Export software development for Central and Western Asia"
- Base: XJU, Wulumuqi, Xinjiang, China
- Mentor: Engineering Academician Prof. Wushouer Silamu
- Contribution: Proofing and collocating the Arabic version of Win7.

07/2013 – 11/2013 | **National Natural Science Foundation**
"Research on Large-scale Software of Uyghur Public Analysis Based on Internent"
- Base: XJU, Wulumuqi, Xinjiang, China
- Mentor: Engineering Academician Prof. Wushouer Silamu
- Contribution: Developed Uyghur Word Segmenter System and Part-of-Speech Labeling Platform.

01/2014 – 04/2014 | **Ministry of Industry and Information Technology Project**
"Uyghur Voice Control System Based on Android"
"Uyghur-Chinese Bi-directional Spoken Translation System Based on Android"
- Base: XJU, Wulumuqi, Xinjiang, China
- Mentor: Engineering Academician Prof. Wushouer Silamu
- Contribution: Developed Android version and built the Uyghur grammar slot file.

11/2016 – 04/2017 | **973 Project**
"Theory and Method of Internet Chinese Information Processing for Ternary Space"
- Base: Beijing, China
- Mentor: Science Academician Prof. Maosong Sun & Prof. Yang Liu
- Contribution: Developed Uyghur-Chinese Bi-directional NMT system and extended to "One Belt One Road" countries.

05/2017 – 01/2018 | **National Natural Science Foundation**
"Research on Basic Theory and Key Technology of Cross-language Public Analysis"
- Base: Beijing, China
- Mentor: Science Academician Prof. Maosong Sun & Prof. Yang Liu
- Contribution: Developed cross-lingual IR system and Crawled Uyghur, Tibet, and Mongolian websites.

04/2018 – 09/2018 | **University - Enterprise Cooperation Projects**
"Research on Uyghur-Chinese Machine Translation Based on Deep Neural Network"
- Base: Beijing, China
- Mentor: Prof. Yang Liu
- Contribution: I have played the technical advisor role and provided some pre/post-processing tools, and some related codes as well.

Computer Patent for Invention

- Authorized
  - **Mieradilijiang Maimaiti**, Yang Liu, Huanbo Luan and Maosong Sun, "Unsupervised Domain Adaptation for Neural Machine Translation", Authorization number: CN 107038159 A; China Patent Application No:

201710139214.0;
  - Maosong Sun, **Mieradilijiang Maimaiti**, Yang Liu and Huanbo Luan, "The Training Method for Neural Machine Translation", Authorization number: CN 109117483 B; China Patent Application No: **2018**10845896.1;
- Under Review
  - Yang Liu, **Mieradilijiang Maimaiti**, Huanbo Luan and Maosong Sun, "The System and Method for Low-Resource Neural Machine Translation Based on Data Augmentation"; China Patent Application No: **2021**108572155;

Selected Publications

2021

1. Maimait, M. *et al.* Segment, Mask, and Predict: Augmenting Chinese Word Segmentation with Self-Supervision. *In International Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2021).
2. Maimaiti, M., Liu, Y., Luan, H., P.Zegao & Sun, M. Improving the Data Augmentation for Low-Resource NMT Guided by POS-Tagging and Paraphrase Embedding. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)* **20,** 1–21 (6 2021).
3. Maimaiti, M., Liu, Y., Luan, H. & Sun, M. Data Augmentation for Low-Resource Languages NMT Guided by Constrained Samplings. *International Journal of Intelligent System (IJIS),* 1–22 (2021).
4. Zheng, Y. *et al.* Self-Supervised Quality Estimation for Machine Translation. *In International Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2021).

2020

5. Li, Z. *et al.* An Empirical Study on Deep Neural Network Models for Chinese Dialogue Generation. *Symmetry* **12,** 1756 (2020).
6. Maimaiti, M., Liu, Y., Luan, H. & Sun, M. Enriching the Transfer Learning with Pre-Trained Lexicon Embedding for Low-Resource Neural Machine Translation. *Tsinghua Science & Technology (TST)* **27,** 150–163 (1 2020).

2019

7. Maimaiti, M., Liu, Y., Luan, H. & Sun, M. Multi-Round Transfer Learning for Low-Resource NMT Using Multiple High-Resource Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)* **18,** 1–26 (2019).

2018

8. Maimaiti, M., Zou, S., Wang, X. & Zou, X. *How to Understand: Three Types of Bilingual Information Processing?* in *ICCSIP* (2018).
9. Maimaiti, M. & Zou, X. *Discussion on Bilingual Cognition in International Exchange Activities* in *IFIP TC12 ICIS* (2018).

Previous works which were published from 2014 to 2016 please refer to my home page.

Academic Service

- Reviewer in International Conferences
  - ACL2016, AAAI2018, COLING2018, NAACL2019, PACLIC2021
- Reviewer in International Transactions
  - International Journal of Intelligent System 2021

Technical Skills

- Natural Language Processing: Machine Translation (related techniques), Language Model, Information Retrieval, Pre-processing
- Deep Learning: Architectures (RNN (GRU & LSTM), CNN), Tools (Tensorflow, Pytorch)
- Machine Learning: Model, Strategy, Algorithm, Supervised Application, Statistical Learning, Unsupervised Learning,
- Speech Signal Processing (Chinese, Uyghur): ASR, TTS, Spoken Translation
- Computer Programming: Fundamentals, Data Structure, Algorithm, Mathematics, WEB development, Programming

Language (C, C++, C#, Java, Python (ML & DL toolkits)), Database, MVC Architecture, Smart Mobile
- Other Skills: Graphic Design, Project Management, Organization Skill

Selected Softwares

- Developed Cooperatively
  - Product Management System for Windows Mobile, C#, 2009
  - Large-scale Training School Information Management System, C#, 2012
  - Uyghur TTS Text Message Reader System, Android, 2013
  - Uyghur GIS system, VB, 2013
  - Uyghur-Chinese SMT system Based on Moses and Microsoft Translator Hub, C++/JavaScript/HTML, 2014
  - Chinese Voice Input Method, Java & Android, 2014
  - Uyghur-Chinese Bi-directional Spoken MT System, C++ & Android, 2014
- Developed by Own
  - Library Information Management System, C#, 2009
  - University Exam Information Management System, C#, 2009
  - University Dormitory Information Management System, C#, 2010
  - University Exam Proctor Arrangement System, C#, 2011
  - Chinese Learning Platform for Android, Java & Android, 2012
  - Uyghur, Kazakh, and Kirghiz Online Text Collection and Processing Software, C# & ASP.net, 2013
  - Uyghur Multi-script Unicode Converter, C#, 2013
  - Uyghur Word-segmenter and Pos-tagging Pre-processing Tools, C++ & C#, 2014
  - Uyghur Automatic Pos-tagging System, C++ & C#, 2014
  - Uyghur Automatic Word-segmenter System, C++ & C#, 2015
  - Uyghur, Tibet, Mongolian - Chinese Bi-directional SMT System, C++/JavaScript/HTML, 2016
  - Uyghur - Chinese Bi-directional NMT System, C++ (Tensorflow) /JavaScript/HTML, 2016
  - Cross-lingual Information Retrieval System, C++/JavaScript/HTML, 2017
  - Low-Resource NMT System, C++ (Tensorflow) /JavaScript/HTML, 2018
  - Multi-lingual NMT System, C++ (Tensorflow) /JavaScript/HTML, 2019

Hobbies

- Reading
  - **Psychology**
  - Science Fiction
  - Linguistics
- Movie
  - **Acrobatic Fighting**
  - Science Fiction Film
  - American Drama
- Sports
  - **Body Building**, swimming
  - Badminton, football, **billiards**
  - **Hiking**, skiing
- Travelling
  - Scenic area
  - **Old Buildings**
  - Museum

Honors

- Outstanding student cadre of XJU, 2007-2008

- National Encouragement Scholarship, 2008-2009
- Excellent league member of XJU, 2008-2009
- National Encouragement Scholarship, 2009-2010
- Advanced Individual in Social Practice Intern of XJU, 10/2010
- Second prize in the program design competition of XJU, 06/2011
- National Scholarship (Undergraduate), 2010-2011
- Outstanding Student, 2010-2011
- Outstanding Graduates (Undergraduate), 06/2012
- Second-class scholarship for postgraduate, 09/2012, 09/2013, 09/2014
- National Scholarship (Postgraduate), 12/2014
- Third-class Social Practice Scholarship for Postgraduates of THU, 12/2016