

Neng Wan

1120 Karby Ter, Unit 204, Sunnyvale, CA, 94089

☎ (669)-577-4582 | ✉ wanneng19940620@gmail.com

I'm currently a master's student in NLP at UCSC, recently graduated with a master in linguistics, I'm passionate about building real world language applications using cutting edge technology.

Education

University of California, Santa Cruz

Santa Clara, CA

Master of Science in Natural Language Processing

Sep. 2023 - Current

- Courses: Natural Language Processing, Deep Learning for NLP, Data Science & Machine Learning Foundation

University of Edinburgh

Edinburgh, UK

Master of Science in Linguistics

Sep. 2021 - May. 2022

- Courses: Intro to Syntax, Syntax Theory, Semantic Theory, Intro to Phonology, Phonological Theory, etc.

University of Turin

Turin, Italy

Bachelor of Arts in Languages and Cultures of Asia and Africa, Final grades: 105/110

Sep. 2014 - Jan. 2021

- Courses: General Linguistics, Linguistics (Advanced Course), Theory and History of Grammars, African Linguistics, etc.

Technical Skills

Coursera Certifications Mathematics for Machine Learning, Introduction to Statistics, Python for Everybody, Python 3 Programming, Discrete Mathematics for Computer Science, Accelerated Computer Science Fundamentals, Machine Learning, Deep Learning, Natural Language Processing

Programming Languages Python, C++

Tools scikit-learn, TensorFlow, Trax

Languages

Native: Mandarin Chinese

Full professional proficiency: English, Italian

Limited working proficiency: Japanese, Spanish, French, Ancient Hebrew

Work Experience

Instructor

University of Turin, Italy

Project: Language awareness: Our Languages and Us - (Noi e Le Nostre Lingue)

Feb. 2016 - Feb. 2021

- Designed and prepared various comprehensive and fun Chinese Linguistics workshops for elementary school students between 10-12 year-old.
- Lectured in more than ten public schools.

Project Experience

Chatbot

using reformer with transfer learning

Winter 2022-2023

- Built a chatbot with reformer. Implemented functionalities to get conversations, tokenization, batching and bucketing for pre-processing, reversible layer for decoder block, and utilized a pre-trained model with transfer learning to generate a dialog.

Text summarization

using transformer

Winter 2022-2023

- Built a text summarization system with transformer from scratch. Implemented tokenization, de-tokenization, concatenation, batching and bucketing for pre-processing, causal attention and transformer decoder block for the model, and greedy decoder for token generation. Implemented training, evaluation and tested the model with customized input.

Neural machine translation

using LSTM with attention

Winter 2022-2023

- Built a NMT model using LSTM with attention. Implemented tokenization, formatting and padding for pre-processing, built the model using LSTM as encoder and decoder, attention with dense layer and LSTM for output, implemented token generation, training and evaluation.

Part of speech tagging

using hidden markov model with Viterbi algorithm

Winter 2022-2023

- Implemented a POS tagging algorithm for WSJ articles. performed data pre-processing to obtain transition and emission matrices, implemented forward and backward algorithm for Viterbi decoder using dynamic programming.