Chan-Jan (Jeff) Hsu
Taipei, Taiwan | +886 912354485
a24998667@gmail.com | r09946011@ntu.edu.tw
🐙 github | 🔗 linkedin | leetcode | kaggle

SKILLS

RESEARCH	ML/DL/AI	Natural Language Processing	Multimodal Representation Learning
PROGRAMMING LANGUAGES	*Experienced:* Python	*Familiar:* C++
FRAMEWORKS & LIBRARIES	PyTorch	Tensorflow	Jax	Scikit-learn	Pandas	Numpy	Matplotlib	Huggingface
LANGUAGES	*Native:* Chinese	*Fluent:* English (TOEFL 112/120)	*Beginner:* German (A2)

EDUCATION

NATIONAL TAIWAN UNIVERSITY   *2016-2020*

B.S IN MATERIAL SCIENCE AND ENGINEERING

• GPA:3.88/4.30, 16th/53, 143+6 (Exempted) credits

NATIONAL TAIWAN UNIVERSITY   *2020-2022*

M.S IN DATA SCIENCE DEGREE PROGRAM, COLLEGE OF EECS ASSOCIATED WITH ACADEMIA SINICA

• GPA:4.27*/4.30, 1st/21, Graduate Representative for Data Science Degree Program

EXPERIENCE

LINEAR ALGEBRA (EE1002) TEACHING ASSISTANT	*2020Fall / NTU*
MACHINE LEARNING (EE5184) TEACHING ASSISTANT	*2021Spring / NTU*
MACHINE LEARNING (EE5184) TEACHING ASSISTANT	*2022Spring / NTU*
MEDIATEK RESEARCH RESEARCH SCIENTIST	*2022-Present/ Taiwan*

• Extended Pretraining of Language Models to support Traditional Chinese in a Research Group of 5+

• Analyzed FishLeg's Performance on Language Benchmarks such as the SMC Bench in a Research Group of 6+

• Technical Advisor of "ClairAudience: Enhancing ASR Systems with Internal Language Model Adaptation"

THESIS

ENHANCING SELF-SUPERVISED LEARNING BY MULTIMODAL INFORMATION

SUPERVISED BY PROFESSOR HUNG-YI LEE AND DR. YU TSAO

RESEARCH PUBLICATIONS

ANALYZING THE ROBUSTNESS OF UNSUPERVISED SPEECH RECOGNITION   *2022 ICASSP*

CHAN-JAN HSU*, LIN, GUAN-TING*, DA-RONG LIU, HUNG-YI LEE, AND YU TSAO

• First Co-author

XDBERT: DISTILLING VISUAL INFORMATION TO BERT FROM CROSS-MODAL SYSTEMS TO IMPROVE NATURAL LANGAUGE UNDERSTANDING   *2022 ACL*

CHAN-JAN HSU, HUNG-YI LEE, AND YU TSAO

T5LEPHONE: BRIDGING SPEECH AND TEXT SELF-SUPERVIED MODELS FOR SPOKEN LANGUAGE UNDERSTANDING VIA PHONEME LEVEL T5   *2023 ICASSP*

Chan-Jan Hsu, Hung-Yi Lee, and Yu Tsao

Bridging Speech and Textual Pre-trained Models with Unsupervised ASR   *2023 ICASSP*

Jiatong Shi, Chan-Jan Hsu, Holam Chung, Dongji Gao, Paola Garcia, Shinji Watanabe, Ann Lee, Hung-yi Lee

Extending the Pre-Training of BLOOM for Improved Support of Traditional Chinese Models, Methods and Results   *Arxiv*

Philipp Ennen, Po-Chun Hsu, Chan-Jan Hsu, Chang-Le Liu, Yen-Chen Wu Et al.

Service to Research

Served as a reviewer for AACL 2022   *2022 AACL*

Academic Highlights

Coding	**Data Structure and Programming:** A+ | **Algorithms:** A+ | **Discrete Mathematics:** A
ML/DL/AI	**Machine Learning:** A+ | **Deep Learning for Human Language Processing:** A+
	Applied Deep Learning: A+ | **Deep Learning for Computer Vision:** A+
NLP Methods	**Web Retrieval and Mining:** A+ | **Corpus Linguistics:** A+ | **Introduction to Linguistics (1)(2):** A+
Other	**Computation in Data Science:** A+ | **Advanced Computer Architectures:** A+
	Computer Gaming Theory: A+ | **High-Performance Big Data and Artificial Intelligence Systems:** A

Achivements

Leetcode (Competitive Coding): Highest Rating 2324 (top 0.53%)

Kaggle (Machine Learning Contest): Highest Rank 973 (top 1%)

PetFinder.my - Pawpularity Contest : 24th/3537 (top1%)

• Image Regression

Lux AI : 50th/1178 (top5%)

• Multi-variable Optimization Agent Design

Santa 2021 - The Merry Movie Montage: 92nd /867 (top 11%)

• Travelling Salesman Problem Optimization

Stable Diffusion - Image to Prompts (top 5%)

• Reverse the Typical Direction of a Generative Text-to-Image Model