

Yu-Chia Chen

RESEARCH SCIENTIST AT META

Seattle, WA 98109

☎ (208) 329-8707 | ✉ yuchaz@meta.com | 🌐 yuchaz.github.io | 📷 yuchaz | 📄 yuchaz

Summary

- Expertise** Natural Language Processing, Unsupervised Learning, Semi-Supervised Learning, Topological Data Analysis, Network Analysis
- Publications** Published in top-tier Machine Learning conferences and journals (NeurIPS, KDD, and JMLR)
- Programming Languages** Python (Advanced), MATLAB (Advanced), C++ (Intermediate)
- Languages** English (Professional), Mandarin (Native), Taiwanese (Native)

Education

University of Washington

PH.D. IN ELECTRICAL ENGINEERING

- Advisor: Marina Meilă

Seattle, WA

Sep. 2016 - Aug. 2021

National Taiwan University

B.S. IN PHYSICS

Taipei, Taiwan

Sep. 2011 - Jun. 2015

Experience

Meta Platforms, Inc.

RESEARCH SCIENTIST

Seattle, WA

Sep. 2021 - Present

- Developed an end-to-end text variation model to enhance the quality and effectiveness of ad texts for advertisers
 - Fine-tuned the LLM using the *Multi-Text* pipeline with RLHF techniques (PPO + rejection sampling)
 - The proposed framework led to an estimated 2.5% increase in top-line ads-quality metrics, resulting in a revenue boost of millions of dollars
 - News coverage:
 - * <https://www.facebook.com/business/news/introducing-ai-sandbox-and-expanding-meta-advantage-suite>
 - * <https://techcrunch.com/2023/05/11/meta-announces-generative-ai-features-for-advertisers/>
- Investigated automatic text engagement and hallucination scoring models as the evaluators for the generated texts
 - Created a *Multi-Text with Engagement label* pipeline from the impression-level data to rank text performance within the same ads
 - Developed a hallucination detection model using human-annotated texts, distilled datasets from LLaMA, and data augmentation techniques

Geometric Data Analysis Group (prof. Marina Meilă), University of Washington

GRADUATE RESEARCH ASSISTANT

Seattle, WA

Apr. 2017 - Aug. 2021

- Thesis: “Learning Topological Structures and Vector Fields on Manifolds with (Higher-order) Discrete Laplacians” [3]
 - Analyzed the decomposibility of the k^{th} homology embedding (the higher-order generalization of spectral clustering) of the discrete k -Hodge Laplacian under matrix perturbation theory (*NeurIPS’21 oral* [1])
 - Proposed a versatile framework for topological feature discovery and semi-supervised edge flow learning from point cloud data [13]
 - Studied a well-known defect of spectral embedding methods on manifolds with large aspect ratios (*NeurIPS’19* [5])
- Developed a scalable manifold learning python toolkit for millions of points: *megaman* [8] (<https://github.com/mmp2/megaman>)

Facebook

MACHINE LEARNING INTERN

Seattle, WA

Jun. 2020 - Sep. 2020

- Developed deep learning models to optimize the click-through rate (CTR) based recommendation system for search ads placement
- Investigated various modeling techniques, including transfer learning and multi-task learning

Microsoft Research

RESEARCH INTERN

Redmond, WA

Jun. 2018 - Sep. 2018

- Studied the evolution of networks with millions of vertices by a dynamic network model extended from the stochastic block model (*KDD’19* [7])
 - Presented a sampling-based extension to causal impact, which allows the practitioners to determine the significance of the intervention
 - Applied in the social relationships of more than 1000 giraffes in the Tarangire Ecosystem over five years (*Animal Behaviour* [2])

Publications

REFERRED PUBLICATIONS

- [1] **YU-CHIA CHEN** and Marina Meila. The decomposition of the higher-order homology embedding constructed from the k-Laplacian. *Advances in Neural Information Processing Systems*, 34, 2021; **NeurIPS oral presentation** (acceptance rate 1%)
- [2] Juan M. Lavista Ferres, Derek E. Lee, Md Nasir, **YU-CHIA CHEN**, Avleen S. Bijral, Fred B. Bercoivitch, and Monica L. Bond. Social connectedness and movements among communities of giraffes vary by sex and age class. *Animal Behaviour*, 180:315–328, October 2021
- [3] **YU-CHIA CHEN**. *Learning Topological Structures and Vector Fields on Manifolds with (Higher-Order) Discrete Laplacians*. PhD thesis, University of Washington, 2021
- [4] Samson Koelle, Hanyu Zhang, Marina Meila, and **YU-CHIA CHEN**. Manifold Coordinates with Physical Meaning. *arXiv:1811.11891 [stat.ML]*, July 2021. To appear in JMLR.
- [5] **YU-CHIA CHEN** and Marina Meilă. Selecting the independent coordinates of manifolds with large aspect ratios. In *Advances in Neural Information Processing Systems 32*, pages 1086–1095, 2019
- [6] Samson J. Koelle, Hanyu Zhang, Marina Meilă and **YU-CHIA CHEN**. Manifold Coordinates with Physical Meaning. *Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019)*, Vancouver, Canada, December, 2019
- [7] **YU-CHIA CHEN**, Avleen S. Bijral, and Juan Lavista Ferres. On Dynamic Network Models and Application to Causal Impact. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19*, pages 1194–1204, New York, NY, USA, 2019. ACM
- [8] **YU-CHIA CHEN**, Dominique Perrault-Joncas, Marina Meilă, and James McQueen. Improved Graph Laplacian via Geometric Self-Consistency. *NIPS Workshop on NIPS Highlights (MLTrain), Learn How to code a paper with state of the art frameworks*, Long Beach, CA, December 2017
- [9] Peifeng Jing, Kosuke Winston, **YU-CHIA CHEN**, Benjamin S. Freedman, and Lih Y. Lin. Patterning and Colonizing Stem Cells with Optical Trapping. In *Optics in the Life Sciences Congress (2017), Paper OtM4E.2*, page OtM4E.2. Optical Society of America, April 2017
- [10] **YU-CHIA CHEN**, Cih-Su Wang, Tsung-Yuan Chang, Tai-Yuan Lin, Hsiu-Mei Lin, and Yang-Fang Chen. Ultraviolet and visible random lasers assisted by diatom frustules. *Optics Express*, 23(12):16224–16231, June 2015
- [11] Cih-Su Wang, Chi-Shung Liao, Tzu-Ming Sun, **YU-CHIA CHEN**, Tai-Yuan Lin, and Yang-Fang Chen. Biologically inspired band-edge laser action from semiconductor with dipole-forbidden band-gap transition. *Scientific Reports*, 5:8965, March 2015

PREPRINTS/UNDER REVIEW/TECHNICAL REPORTS

- [12] Timothy Siegler, Wiley Dunlap-Shohl, Yuhuan Meng, Wylie Kau, Preetham Sunkari, Chang-En Tsai, Zachary Armstrong, **YU-CHIA CHEN**, David Beck, Marina Meila, and Hugh Hillhouse. Water-Accelerated Photo-oxidation of CH₃NH₃PbI₃ Perovskite: Mechanism, rate orders, and rate constants. *ChemRxiv*, June 2021
- [13] **YU-CHIA CHEN**, Marina Meilă, and Ioannis G. Kevrekidis. Helmholtzian Eigenmap: Topological feature discovery & edge flow learning from point cloud data. *arXiv:2103.07626 [stat.ML]*, March 2021
- [14] **YU-CHIA CHEN**, James McQueen, Samson J. Koelle, Marina Meilă, Stefan Chmiela and Alexandre Tkatchenko. Modern Manifold Learning Methods for MD data – a step by step procedural overview. <http://students.washington.edu/yuchaz/files/2020-md-manifold.pdf>

Other Experience & Projects

Institute for Pure & Applied Mathematics (IPAM), UCLA

Los Angeles, CA

VISITING RESEARCHER

Sep. 2019 - Dec. 2019

- Participated in the *Machine Learning for Physics and the Physics of Learning* long program
- Investigated the intersection of conformal prediction, unsupervised learning, and physical science
- White paper: <https://www.ipam.ucla.edu/news/white-paper-machine-learning-for-physics-and-the-physics-of-learning/>

Department of Electrical & Computer Engineering, University of Washington

Seattle, WA

TEACHING ASSISTANT

Jan. 2017 - Dec. 2017

- Courses: Digital Signal Processing (graduate level), Devices And Circuits, Discrete Time Linear Systems, Fundamentals of Electrical Engineering

Selfie Sensei: Convolutional Neural Network based selfie instructor

Seattle, WA

COURSE PROJECT

Apr. 2017 - Jun 2017

- Trained the Google Inception-v3 model on 40 thousand selfies collected from Twitter with hashtag *#selfie*

Large scale medical subject heading (MeSH) term indexing

Seattle, WA

COURSE PROJECT

Jan. 2017 - Mar. 2017

- Constructed a Convolutional Neural Network trained with *skipgram* word2vec embedding to annotate 27k MeSH terms on 12M academic articles

Photonics Lab, University of Washington

Seattle, WA

GRADUATE RESEARCH ASSISTANT

Sep. 2016 - Dec. 2016

- Investigated high accuracy mass sensing using Nanostructure-enhanced laser tweezers and its application to stem cell patterning [9]

Psychological Warfare Group, Ministry of National Defense

Taipei, Taiwan

FRONT-END SOFTWARE ENGINEER (COMPULSORY MILITARY SERVICE)

Aug. 2015 - Jul. 2016

- Lead engineer on cloud-based file exchanging platform, which enables user to search, view, and share streaming media
- Technology used: JavaScript (*react.js*), HTML/CSS

- Investigated bio-photonics devices with wide spectrum range [10]
- Studied Perovskite and CdTe core shell quantum dots assisted random laser in bio-inspired materials [11]

Honors & Awards

2019	Student Travel Award , NeurIPS 2019	Vancouver, Canada
2019	Student Travel Award , KDD 2019	Anchorage, AK
2019	Travel Grant , UW Department of Electrical & Computer Engineering	Seattle, WA
2013	Scholarship , Taipower Academic Scholarship	Taipei, Taiwan
2012	Scholarship , Taipower Academic Scholarship	Taipei, Taiwan
2010	Second prizes , Physics Scholastic Ability Contest	Kaohsiung, Taiwan

Selected Talks

Dec. 2021	Seminar Talk , NeurIPS'21 oral, <i>The decomposition of the higher-order homology embedding</i> [1].	Seattle, WA
Nov. 2020	Seminar Talk , UW ML retreat, <i>Higher-order topological feature discovery and edge flow learning</i>	Seattle, WA
Feb. 2020	Seminar Talk , UW Geometric Data Analysis Group, <i>Hodge Laplacians on graphs</i>	Seattle, WA
Dec. 2019	Poster Presentation , NeurIPS'19, <i>Selecting the Independent Coordinates of Manifolds with Large Aspect Ratios</i>	Vancouver, Canada
Oct. 2019	Seminar Talk , IPAM, <i>Selecting the Independent Coordinates of Manifolds with Large Aspect Ratios</i>	Los Angeles, CA
Aug. 2019	Poster Presentation , KDD'19, <i>On Dynamic Network Models and Application to Causal Impact</i>	Anchorage, AK
Sep. 2018	Seminar Talk , Microsoft, <i>On Dynamic Network Models and Application to Causal Impact</i>	Redmond, WA
Jan. 2018	Seminar Talk , UW Geometric Data Analysis Group, <i>Improved Graph Laplacian via geometric self-consistency</i>	Seattle, WA

References

Marina Meilă

PROFESSOR, DEPARTMENT OF STATISTICS, UNIVERSITY OF WASHINGTON

mmp@stat.washington.edu

Juan Lavista Ferres

VP, CHIEF DATA SCIENTIST, MICROSOFT CORPORATION

jlavista@microsoft.com

Minh Phuong Nguyen

STAFF MACHINE LEARNING ENGINEER, META PLATFORMS INC.

phuong88@meta.com

Ioannis G. Kevrekidis

BLOOMBERG DISTINGUISHED PROFESSOR, DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS, JOHNS HOPKINS UNIVERSITY

yannisk@jhu.edu

Avleen S. Bijral

APPLIED SCIENTIST, MICROSOFT CORPORATION

avbijral@microsoft.com

Yang-Fang Chen

CHAIR PROFESSOR, DEPARTMENT OF PHYSICS, NATIONAL TAIWAN UNIVERSITY

yfchen@phys.ntu.edu.tw