Yi-Wei **Chen**
PH.D., SOFTWARE ENGINEER, MICROSOFT

2808 Calder Ave NE Unit 6224, Redmond, WA, 98052, U.S.A.

☐ (+1) 979-267-0792 | ✉ yiwei_chen@tamu.edu | 🏠 yiwei-chen.github.io | ⌗ YIWEI-CHEN | in yiwei-chen-50568092

Research Focus

Retrieval augmented generation, Recommendation system, Defect detection, hyperparameter optimization, early stopping, neural architecture search, object detection, video generation, distributed optimization, label noise, Bayesian optimization, distributed system, embedded system.

Work Experience

Microsoft Corporation
Redmond, U.S.A.

SOFTWARE ENGINEER II — *Jan. 2023 -*

- Built recommendation systems of management actions in Microsoft Admin Center, 3× higher CTR than fixed ones.
- Constructed a RAG-based LLM with Microsoft Teams Documents, generating 18% more precise responses.
- Customized Microsoft Copilots to accelerate routine data analysis in Azure Storage and Azure Machine Learning.
- Deployed a GPU cluster to finetune small LLMs (GPT-2, GPT-Neo, GPT-J, and Llama) for text classification tasks, 30% improvement on F1-scores.

Data Analytics at Texas A&M (DATA) Lab
College Station, U.S.A.

RESEARCH ASSISTANT — *Aug. 2018 - Dec. 2022*

- Reduced 4× network parameters of auto-encoders for defection detection by evolutionary search.
- Deployed tiny YOLOv1 (432KB) to MAXIM78000 edge device with Camera input and LCD output.
- Invented a model parallelism for neural architecture search with 1.18× speedup.
- Improved 3% accuracy under high-level symmetric label noise by a new neural architecture search.
- Wrote a survey of Automated Machine Learning, including automated feature engineering, hyperparameter tuning, and neural architecture search.

Amazon.com
Seattle, U.S.A.

APPLIED SCIENTIST INTERN — *Sep. 2021 - Jan. 2022*

- Built an auto video generation framework to animate image highlights for hundreds of miscellaneous product categories.
- Established the image filter using ResNet to exclude text-intensive images.
- Designed text and geometric object detectors using Amazon Rekognition and DeepLab to capture headlines, descriptions, and rectangular subjects.
- Utilized Navier-Stoker inpainter and OpenShot animator to animate captured subjects in product images.
- Accelerated the video generation from 30 minutes (human) to 3 minutes (our framework), getting 10× speedup.

Microsoft Corporation
Redmond, U.S.A.

RESEARCH INTERN — *May 2021 - Aug. 2021*

- Proposed adaptive constraint early stopping (ACE) algorithm to find optimal hyperparameters under specific constraints.
- Modeled the expected evaluation cost to tailor the constraint interval to different constraint cost.
- Developed a new early stopping by using constraint violations.
- Improved feasible AUC of LGBM model from 0.801 to 0.852, getting 6% improvement.

Trend Micro
Taipei, Taiwan

SOFTWARE ENGINEER — *Feb. 2016 - Aug. 2018*

- Customized rules in YARA format for scanning suspicious objects
- Interacted with Microsoft Active Directory (AD) via LDAPv3 protocol, including password authentication, searching AD users and groups, and querying parent groups of an AD user and total members of an AD group.
- Fulfilled secure connection with CheckPoint Firewalls via SSL certificate and CheckPoint SDK OPSEC.
- Integrated Palo Alto Panorama and virtual systems of Palo Alto Firewalls via PAN XAPI.

Education

Texas A&M University
College Station, U.S.A.

PH.D. IN COMPUTER SCIENCE AND ENGINEERING — *Aug. 2018 - Dec. 2022*

- GPA: 4.0/4.0
- Dissertation: Automated Machine Learning with Constraints and Imperfect Data

National Taiwan University (NTU)   *Taipei, Taiwan*
M.S. in Computer Science and Information Engineering   *Feb. 2014 - Sep. 2015*
- GPA: 4.04/4.30
- Thesis: Virtual Hadoop: MapReduce over Docker Containers with an Auto-Scaling Mechanism for Heterogeneous Environments

National Taiwan University (NTU)   *Taipei, Taiwan*
B.S. in Electrical Engineering   *Sep. 2008 - Jun. 2012*
- GPA: 3.79/4.00 (84.25/100)

Publications

JOURNAL PAPER
- **Yi-Wei Chen**, "Automated Machine Learning with Constraints and Imperfect Data", Texas A&M University Dissertation, Dec. 2022
- **Yi-Wei Chen**, Qingquan Song, Xia Hu, "Techniques for Automated Machine Learning", ACM SIGKDD Explorations Newsletter, Dec. 2021
- **Yi-Wei Chen**, Qingquan Song, Xi Liu, P.S. Sastry, Xia Hu, "On Robustness of Neural Architecture Search under Label Noise", Frontiers in Big Data, Jan. 2020

CONFERENCE PAPER
- **Yi-Wei Chen**, Chi Wang, Amin Saied, Rui Zhuang, "ACE: Adaptive Constraint-aware Early Stopping in Hyperparameter Optimization", ACM SIGKDD AutoML Workshop, Washington, DC, USA, Aug. 2022
- **Yi-Wei Chen**, Shih-Hao Hung, Chia-Heng Tu, Chih Wei Yeh, "Virtual Hadoop: MapReduce over Docker Containers with an Auto-Scaling Mechanism for Heterogeneous Environments", 2016 Research in Adaptive and Convergent Systems (RACS), Odense, Denmark, Oct. 2016
- Guanchu Wang, Zaid Pervaiz Bhat, Zhimeng Jiang, **Yi-Wei Chen**, Daochen Zha, Alfredo Costilla Reyes, Afshin Niktash, Gorkem Ulkar, Erman Okman, Xuanting Cai, Xia Hu, "BED: A Real-Time Object Detection System for Edge Devices", Conference on Information and Knowledge Management (CIKM), Atlanta, Georgia, U.S.A., Oct. 2022
- Daochen Zha, Zaid Pervaiz Bhat, **Yi-Wei Chen**, Yicheng Wang, Sirui Ding, Jiaben Chen, Kwei-Herng Lai, Mohammad Qazim Bhat, Anmoll Kumar Jain, Alfredo Costilla Reyes, Na Zou, Xia Hu, "AutoVideo: An Automated Video Action Recognition System", International Joint Conference on Artificial Intelligence (ICJAI), Messe Wien, Vienna, Austria, Jul. 2022

Technical Skills

Languages	Python, C, Bash, Java, C, C++, Ruby, HTML, PHP, Verilog, MATLAB
Tools	Microsoft Copilot, FAISS, LlamaIndex, OpenAI, PyTorch, Tensorflow, HuggingFace, RAY, FLAML, AWS, Azure, Scikit-Learn, Git, Perforce, Trello, JIRA, Latex, Docker, Flask, Django, Ruby-on-Rails, Hadoop