Houjun "Jack" **Liu**

houjun@stanford.edu   T: (650) 282-0675   github.com/**Jemoka**   linkedin.com/in/**houjun-liu**   **jemoka**.com

Education

Undergraduate • **Stanford University** Stanford, CA   09/2023 - 06/2027

Upper-School Student • **The Nueva School** San Mateo, CA   UW GPA 4.0   08/2019 - 06/2023
  - Selected awards: "Silver", USA Computing Olympiad; SXSW//EDU Startup Finalist, top 5 out of 2000+
  - Select coursework: Differential Equ., Linear Algebra, Multi-Variable Calculus, Machine Learning, Software Engineering

Skills

  - Data/ML: PyTorch, TensorFlow, NumPy, Pandas
  - Python Full-Stack: Flask, Django
  - {Dev,ML}Ops: Docker, Kubernetes, Kubeflow, CVAT
  - Front End/JS: React, Svelte, Ionic, NextJS
  - iOS Development: AppKit, SwiftUI

  - Math Modeling: R, Mathematica, MATLAB
  - FP/LISP: Clojure, Common Lisp, Haskell
  - C/C++: Eigen, Qt, i2C/UART development
  - Product management: user stories, need-finding
  - Project management: agile, gantt charts, Jira

Significant Endeavors

Lead Research Engineer • **MacWhinney Lab, CMU** talkbank.org   linguistics and speech tech. lab   2022-Present
  - NIH-Funded Consulting Research Associate role
  - Maintained TalkBank, the world's largest auditory language pathology database, in team of 5
  - Developed Batchalign, a speech analysis tool being used by 20+ other speech labs globally

Research Assistant • **UC Davis Dementia Center**   big-data dementia analysis group   2022-Present
  - Led analysis of the large National Alzheimer's Coordinating Center dataset at the group
  - Demonstrated state-of-the-art accuracy in dementia and MCI classification (93% F1)
  - Presented research at National Alzheimer's Disease Research Center

Co-Founder • **The Condution Project** condution.com   open-source task-management app   2019-Present
  - Created task-management app used by 4,000+ active monthly users worldwide
  - Leveraged effective social marketing campaigns for 5,000+ monthly impressions
  - Lead 8-person lean development team with Agile-driven workflows

Project Lead • **MODAP** San Mateo, CA   non-profit firefighting aid R&D   2019-2023
  - Lead 8 engineers in creating a rapid & accurate airborne payload delivery service
  - Trained award-winning in-house CV models on embedded projectile devices
  - Supported by Fire Chiefs at CalFire & Colorado Fire Departments

Consulting ML Engineer • **Dragonfruit AI** dragonfruit.ai   enterprise video analytics startup   2021-2022
  - Hired as one of the youngest ML Engineers, reporting to head of Data Science, developing CV models
  - Pushed 4 major features into production, created training workflow that expedited tuning by 33%

Summer Lecturer • **AIFS AIBridge** aifs.ucdavis.edu   graduate-level machine learning program   2022
  - Developed a comprehensive one-week ML bootcamp hosted in collaboration between UC Davis and AIFS
  - Co-taught course to 22 graduate students in food science across Davis, Berkeley, and Cornell

A curriculum vitae follows with relevant publications and pre-prints.

Publications, Pre-Prints, and Talks

Liu, H., Zhang, J., Liu, X., Weakley, A. *A Transformer Approach for Cognitive Impairment Classification and Prediction* [In Press, Paper Presentation]. The 8th International Workshop in Health Intelligence, AAAI 2024, Vancouver, BC, Canada.

Liu, H., Zhang, J., Liu, X., Weakley, A. (2023, October 12). *Deep-Learning Analysis of Longitudinal Alzheimer's Data* [Conference Presentation]. National Alzheimer's Dementia Research Center, Davis, CA, United States.

Liu, H., Huang, A., (2023). *A System for the Ingestion, Storage, and Selective Recommendation of Textual Data by Size Constrained Textual Computational Systems* (United States Patent Pending. No. 63/514,412). USPTO

Liu, H., MacWhinney, B., Fromm, D., Lanzi, A. (2023). Automation of Language Sample Analysis. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2023_JSLHR-22-00642

Lanzi, A. M., Saylor, A. K., Fromm, D., **Liu, H.**, MacWhinney, B., Cohen, M. L. (2023). DementiaBank: Theoretical Rationale, Protocol, and Illustrative Analyses. *American Journal of Speech-Language Pathology*, 32(2), 426–438. https://doi.org/10.1044/2022_AJSLP-22-00281

Liu, H., Sarboraria, M., Lu, V., Salimi, K., Cho, O., Kuznetsov, F., ... Hauser, P. (2023). R@N: an Asynchronous, Student-Led Program to Train Students to Conduct Independent Research. *Journal of Biological Chemistry*, 299(3), S239.

Krishnan, A., Brown, M., **Liu, H.**, Holmes, G., Hauser, P., De, L. (2023, March 24-28). *Project80 3.0: An Updated Student-Driven Framework for Creating Multimedia Educational Resources from Primary Literature* [Conference Poster]. DiscoverBMB, Am. Society for Biochemistry and Molecular Biology, Seattle, WA, United States.

Liu, H., Sayyah, Z. (2022). ConDef: Automated Context-Aware Lexicography Using Large Online Encyclopedias. In: Arai, K. (eds) Intelligent Computing. SAI 2022. Lecture Notes in Networks and Systems, vol 507. Springer, Cham. https://doi.org/10.1007/978-3-031-10464-0_41

Liu, H. (2022). Encrypted, Anonymized System for Protected Health Information Verification Built via Proof of Stake. arXiv preprint arXiv:2205.02753. https://arxiv.org/abs/2205.02753

Liu, H., Krishnan, A., Brown, M., Hauser, P., De, L. (2022). Project 80: a reproducible, student-driven framework for creating multimedia educational resources from primary literature. *The FASEB Journal*, 36.

Liu, H. (2021). Towards Automated Psychotherapy via Language Modeling. arXiv preprint arXiv:2104.10661. https://arxiv.org/abs/2104.10661

Liu, H. (2020). Byte-Pair and N-Gram Convolutional Methods of Analysing Automatically Disseminated Content on Social Platforms. OpenScienceFoundation https://doi.org/10.20944/preprints202004.0214.v1