

RESEARCH INTERESTS

- **Machine Learning:** sequence learning, multimodal LLMs, self-supervised learning, representation learning
- **Computational Biology:** proteomics, protein design, regulatory genomics, biomedical data science

EDUCATION

University of Washington

Ph.D. in Computer Science

Co-advisors: William Noble, Sewoong Oh

Seattle, WA

Sep 2020–Current

Koc University

B.S. in Electrical and Electronics Engineering

GPA: 4.00/4.00, Ranked 1st in the class

Istanbul, Turkey

2015–2020

EXPERIENCE

University of Washington

Ph.D. Candidate, (Supervisors: William Noble, Sewoong Oh)

Seattle, WA

Sep 2020–Current

- My current research builds deep learning methods to analyze mass spectrometry data in proteomics
- Formulated *de novo* peptide sequencing as a supervised seq2seq learning task and trained a transformer model advancing the state of the art [1-4]. Exploring self-supervised learning tasks to train foundation models for mass spectra on repository-scale data sets

Amazon

Applied Scientist Intern, (Supervisors: Bella Dubrov, Tommi Jaakkola)

Seattle, WA

Summer 2024

- Developed novel multimodal protein language models for a confidential project within the Grand Challenge organization

Alphabet (Calico Life Sciences)

Machine Learning Research Intern, (Supervisor: David Kelley)

South San Francisco, CA

Summer 2023

- Explored model ensembling and knowledge distillation strategies to improve gene expression prediction with state-of-the-art DNA sequence models

Nucleate

Managing Director, Seattle

Seattle, WA

Apr 2022 - Sep 2023

- Founded Pacific Northwest chapter of [Nucleate](#), a global trainee-led nonprofit supporting emerging entrepreneurs in biotech and helping them spin out of academic labs (as featured in [GeekWire](#))
- Recruited and led a team of 12 PhD/MBA/MD students and postdocs running the [Activator](#) incubator program to identify and support venture teams with early-stage technologies

Novo Nordisk

Machine Learning Research Intern, (Supervisors: Per Greisen, Kristine Deibler)

Seattle, WA

Summer 2022

- Developed a novel approach combining protein language models and molecular graph neural networks to learn representations for modified peptide drug candidates with the Computational Drug Discovery group

- Modeled post-chemotherapy patient reported outcomes and electronic health records (EHRs) for cancer patients
- Performed trajectory clustering and risk group classification to identify vulnerable patient populations [5]

PUBLICATIONS

- [1] **M. Yilmaz**, W. Fondrie, W. Bittremieux, C. Melendez, R. Nelson, V. Ananth, S. Oh, and W. Noble, “[Sequence-to-sequence translation from mass spectra to peptides with a transformer model](#)”, *Nature Communications*, 2024, [code].
- [2] C. Melendez, J. Sanders, **M. Yilmaz**, W. Bittremieux, W. Fondrie, S. Oh, and W. Noble, “[Accounting for digestion enzyme bias in Casanovo](#)”, *Journal of Proteome Research*, 2024.
- [3] V. Ananth, J. Sanders, **M. Yilmaz**, B. Wen, S. Oh, and W. Noble, “[A learned score function improves the power of mass spectrometry database search](#)”, *Bioinformatics*, 2024.
- [4] **M. Yilmaz**, W. Fondrie, W. Bittremieux, S. Oh, and W. Noble, “[De Novo Mass Spectrometry Peptide Sequencing with a Transformer Model](#)”, *International Conference on Machine Learning (ICML)*, 2022, [code][talk].
- [5] A. Azad, **M. Yilmaz**, S. Bozkurt, J. Brooks, D. Blayney, and T. Hernandez-Boussard, “[Diverse Patient Trajectories during Cytotoxic Chemotherapy: Capturing Longitudinal Patient Reported Outcomes](#)”, *Cancer Medicine*, 2021.

CONFERENCE AND INVITED TALKS

- **Translating from mass spectra to peptides with a transformer model** (slides)
 - ASMS 2023 (Oral Presentation and Evening Workshop)
- **De Novo Peptide Sequencing Transformer** (slides)
 - ISMB CompMS 2022
 - Novo Nordisk Research Center Seattle 2022

SCHOLARSHIPS AND AWARDS

- **Paul G. Allen School First-Year Ph.D. Fellowship** 2020
- **Monbukagakusho (MEXT) Scholarship in Science** 2016
 - Awarded by Japanese Ministry of Education, Culture, Sports, Science and Technology
- **Turkish Government High Honour Scholarship** 2015
 - Awarded based on ranking (14th out of 1.8 million students) in National University Entrance Exam
- **Koc University Suna Kirac Scholarship** 2015
 - Full tuition waiver and stipend during the B.Sc. based on National University Entrance Exam rank

SERVICE

- **Program Committee:** CompMS (ISMB 2024) • **Reviewer:** *Genome Research*, *PLOS Computational Biology*, *Journal of the American Society for Mass Spectrometry*, ISMB (2024), RECOMB (2022, 2024) • **UW CSE:** PhD Admission Committee Member (2021)

SKILLS

- **Programming Languages and Tools:**
 - Python, R, SQL, Julia, MATLAB, Java, C, C++
- **Libraries:**
 - PyTorch, Hugging Face, Keras, NumPy, Pandas, Scikit-Learn

SKILLS

- **Programming Languages and Tools:**
 - Python, R, SQL, Julia, MATLAB, Java, C, C++
- **Libraries:**
 - PyTorch, Hugging Face, Keras, NumPy, Pandas, Scikit-Learn

LANGUAGES

- **Turkish:** Native
- **Japanese:** Advanced (JLPT N1)
- **French:** Intermediate
- **Spanish:** Elementary