Melih Yilmaz

Email: melih@cs.washington.edu
Website: melihyilmaz.github.io

RESEARCH INTERESTS

- **Machine Learning:** sequence learning, self-supervised learning, representation learning

- **Computational Biology:** proteomics, mass spectrometry, regulatory genomics, biomedical data science

EDUCATION

University of Washington — Seattle, WA
Ph.D. in Computer Science — Sep 2020–Current
Co-advisors: William Noble, Sewoong Oh

Koc University — Istanbul, Turkey
B.S. in Electrical and Electronics Engineering — 2015–2020
GPA: 4.00/4.00, Ranked 1st in the class

RESEARCH EXPERIENCE

University of Washington — Seattle, WA
Ph.D. Candidate, (Supervisors: William Noble, Sewoong Oh) — Sep 2020–Current

- Focusing on proteomics, my current research builds deep learning methods to analyze mass spectrometry data

- Formulated *de novo* peptide sequencing as a supervised seq2seq learning task and trained a transformer model advancing the state of the art [1,2]. Exploring self-supervised learning tasks to train foundation models for mass spectra on repository-scale data sets

Calico Life Sciences — South San Francisco, CA
Machine Learning Research Intern, (Supervisor: David Kelley) — Summer 2023

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

Novo Nordisk — Seattle, WA
Machine Learning Research Intern, (Supervisors: Per Greisen, Kristine Deibler) — 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

Stanford University — Stanford, CA
Research Intern, (Supervisor: Tina Hernandez-Boussard) — Summer 2019

- 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 [3]

PREPRINTS AND PUBLICATIONS

[1] M. Yilmaz*, W. Fondrie*, W. Bittremieux, R. Nelson, V. Ananth, S. Oh, and W. Noble, "Sequence-to-sequence translation from mass spectra to peptides with a transformer model", *bioRxiv*, 2023, [code].

[2] **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*, 2022, [code][talk].

[3]   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

LEADERSHIP EXPERIENCE

Nucleate   Seattle, WA
Managing Director, Seattle   Apr 2022 - Current

- 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*)
- Headed a team of 10+ PhD/MBA/MD students and postdocs running the Activator incubator program to identify and support venture teams with early-stage technologies from research institutions across PNW.

PROFESSIONAL SERVICES

- **Reviewer**: *Genome Research*, RECOMB 2022 • **UW CSE**: PhD Admission Committee Member (2021)

SKILLS

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

LANGUAGES

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