Anastasia Razdaibiedina

arazd.github.io ⋄ anastasia.razdaibiedina@mail.utoronto.ca

EDUCATION

University of Toronto 2018 - Present
Ph.D. Machine Learning & Computational Biology (GPA: 4.0)
<u>Research areas</u>: Representation Learning, Natural Language Processing, Biomedical Machine Learning, Continual Learning

Kyiv National University of Taras Shevchenko 2013 - 2017
B.S. Applied Mathematics with Honors (GPA: 4.0)

RESEARCH EXPERIENCE

Meta (Facebook) AI, Seattle, US Jun 2022 - Present
Research Intern, Natural Language Processing and Machine Learning
- Developed a continual learning method for language models based on prompt tuning, which trains $< 0.1\%$ of total parameters.
- Validated on T5 and BERT models, achieved $+22\%$ improvement over previous state-of-the-art on a standard NLP benchmark.
- Submitted a research paper to ICLR 2023, gave talk at FAIR.
  Technologies: pytorch, huggingface, numpy, pandas, jupyter, matplotlib, bash, slurm, git

Amazon Research, remote May - Oct 2021
Applied Scientist Intern, Natural Language Processing
- Developed a regularization method for language models fine-tuning, which avoids representation collapse.
- Achieved $+2.6$ points improvement over standard fine-tuning on 13 NLP tasks with BERT model.
- Submitted a research paper to EACL 2023, gave talk at Amazon Science.
  Technologies: tensorflow, huggingface, numpy, pandas, scikit learn, bash, aws cloud computing

Recursion Pharmaceuticals, remote Jun - Sep 2020
Data Science Intern
- Developed computer vision algorithms for drug discovery using ResNet and DenseNet architectures.
- Improved existing CNN pipeline accuracy by 8% via data augmentation techniques (CutMix and MixUp).
  Technologies: pytorch, determined, scikit learn, numpy, pandas, seaborn, plotly, jupyter notebook, bash, git, google cloud

University of Toronto / Vector Institute Jan 2018 - Present
Graduate Student Researcher
- Developed a self-supervised method to predict protein function from microscopy data, and validated on a novel dataset of 3,000,000 single-cell images. Discovered aging-related functions of 7 previously unknown proteins.
- Developed an image restoration method for microscopy data using GANs (Generative Adversarial Networks), effective in few-shot conditions. Published in NeurIPS 2019 medical imaging workshop and rated in top-15 submissions.

National Academy of Sciences of Ukraine, Kyiv, Ukraine Nov 2016 - Mar 2017
Bioinformatics Research Intern
- Performed statistical analysis on patient data using R, identified 2 single-nucleotide polymorphisms associated with high miscarriage risk. Published findings in Obstetrics and Gynaecology International journal.

PUBLICATIONS

1. A. Razdaibiedina et al. Progressive Prompts: continual learning for language models without forgetting. Submitted to **ICLR, 2023**. OpenReview link

2. A. Razdaibiedina et al. PIFiA: a self-supervised approach for protein functional annotation from single-cell imaging data. In submission to **Nature Methods**.

3. A. Razdaibiedina et al. Improving language models fine-tuning with representation consistency targets. Submitted to **EACL**, 2022. https://arxiv.org/abs/2205.11603

4. A. Razdaibiedina et al. Learning multi-scale functional representations of proteins from single-cell microscopy data. In **ICLR**, 2022, MLDD workshop. https://arxiv.org/abs/2205.11676

5. A. Razdaibiedina et al. Multi-defect microscopy image restoration under limited data conditions. In **NeurIPS**, 2019, Medical Imaging workshop (rated in top-15). https://arxiv.org/abs/1910.14207

6. A. Razdaibiedina et al. Effects of single-nucleotide polymorphisms in cytokine, toll-like receptor, and progesterone receptor genes on risk of miscarriage. In **Obstetrics and Gynecology International**, 2018. https://pubmed.ncbi.nlm.nih.gov/30116270/

7. A. Razdaibiedina et al. Biomolecular modeling on iOS devices: review and software comparison. In **RJBCS**, 2016. ResearchGate Link

TEACHING EXPERIENCE

1. ***Bias and Fairness in Machine Learning, Vector Institute*** (Winter 2022). Developed and conducted two tutorials on bias and fairness in pre-trained language models (BERT, T5).

2. ***CSC384: Intro to Artificial Intelligence*** (Winter 2021). Developed 5 quizzes and conducted 3 tutorials on knowledge representation and deep learning fundamentals.

3. ***Deep Learning 2, Vector Institute*** (Fall 2020). Developed two machine learning assignments on computer vision (segmentation with UNET, object detection with YOLO) and interpretability (CNNs, Gradient Maps, SmoothGrad) in Google Colab.

4. ***CSC311: Intro to Machine Learning*** (Fall 2019) Developed and conducted two tutorials that cover linear algebra basics and singular value decomposition, corrected homeworks.

INVITED TALKS

Facebook AI Research talk	Sep 2022
Continual learning for language models without forgetting	
Toronto Bioinformatics User Group	Mar 2022
Self-supervised approach for protein functional annotation from single-cell imaging data	
Temerty Center for AI Research and Education in Medicine	Aug 2021
Discovering gene-disease relationships with Deep Learning	
York University $\times$ Vector Institute invited panelist	Oct 2019
Panel discussion: AI in Healthcare and Future	

HONORS AND AWARDS

School of Graduate Studies travel award (1000$)	2022
Ontario Graduate Scholarship (15,000$)	2021-2022
Vector Institute PGA Fellowship (6,000$ $\times$ 3)	2021, 2020, 2019
NeurIPS travel award (500$)	2019
School of Graduate Studies travel award (1000$)	2019
NVIDIA GPU grant program	2018
Ph.D. Merit entrance scholarship (2000$)	2017
Augmented academic merit scholarship (top-10% of the class)	2015-17

Total: 40,000$