Valeriia Cherepanova

College Park MD, USA, +13014010454, vcherepa@umd.edu

Google Scholar

Education

Aug 2018-Present — PhD in Applied Mathematics
University of Maryland, College Park
Advisor: Prof. Tom Goldstein
- *Dean's Fellowship*

Sept 2017-Sept 2018 — MRes in Computational Biology (CoMPLEX)
University College London, London
Advisor: Prof. Alexei Zaikin
- *Distinction*

Sept 2013-Jun 2017 — BSc in Mathematics
National Research University Higher School of Economics, Moscow
Advisor: Prof. Vladimir Poberezhny

Work Experience

June 2022-Aug 2022 — Applied Scientist Intern, **Amazon**
Manager: Prajit Reddy Muppidi
Developed Machine Learning solutions for improving Alexa Voice Search on FireTV.

June 2021-Aug 2021 — Applied Scientist Intern, **Amazon**
Manager: Shrikant Khadilkar
Developed NLP solutions to monitor integrity and transparency of 3P Alexa Skills.

May 2020-Present — Graduate Research Assistant, **University of Maryland**
Advisor: Prof. Tom Goldstein
Work on various research projects in Deep Learning.

Jul 2016-Oct 2016 — Data Scientist Intern, **Teradata (Moscow)**
Advisor: Dr. Sergei Gromov
Worked on educational project in Machine Learning.

Selected Publications

Valeriia Cherepanova, Micah Goldblum, Harrison Foley, Shiyuan Duan, John P Dickerson, Gavin Taylor, Tom Goldstein, *LowKey: Leveraging Adversarial Attacks to Protect Social Media Users from Facial Recognition*, ICLR, 2021,

Roman Levin*, **Valeriia Cherepanova***, Avi Schwarzschild, Arpit Bansal, C. Bayan Bruss, Tom Goldstein, Andrew Gordon Wilson, Micah Goldblum, *Transfer Learning with Deep Tabular Models*, arXiv preprint arXiv:2206.15306,

Valeriia Cherepanova*, Vedant Nanda*, Micah Goldblum, John Dickerson, Tom Goldstein, *Technical Challenges for Training Fair Neural Networks*, ICLR 2021 RAI Workshop, arXiv:2102.06764,

Valeriia Cherepanova*, Steven Reich*, Samuel Dooley, Hossein Souri, Micah Goldblum, Tom Goldstein, *A Deep Dive into Dataset Imbalance and Bias in Face Identification*, arXiv preprint arXiv:2203.08235,

Eitan Borgnia, **Valeriia Cherepanova**, Liam Fowl, Amin Ghiasi, Jonas Geiping, Micah Goldblum, Tom Goldstein, Arjun Gupta[1], *Strong Data Augmentation Sanitizes Poisoning and Backdoor Attacks Without an Accuracy Tradeoff*, ICASSP, 2021,

Micah Goldblum, Steven Reich*, Liam Fowl*, Renkun Ni*, **Valeriia Cherepanova***, Tom Goldstein, *Unraveling Meta-Learning: Understanding Feature Representations for Few-Shot Tasks*, ICML, 2020,

Arpit Bansal, Micah Goldblum, **Valeriia Cherepanova**, Avi Schwarzschild, C. Bayan Bruss, Tom Goldstein, *MetaBalance: High-Performance Neural Networks for Class-Imbalanced Data*, arXiv preprint, arXiv:2106.09643,

Oleg Blyuss, Alexey Zaikin, **Valeriia Cherepanova** et al., *Development of PancRISK, a urine biomarker-based risk score for stratified screening of pancreatic cancer patients*, British Journal of Cancer, 2019, doi:10.1038/s41416-019-0694-0

[1] Authors ordered alphabetically
* Indicates equal contributions

Relevant Coursework

Machine Learning: Deep Learning (HSE), Computer Vision (UMD, CMSC426), Computational Linguistics (UMD, CMSC723), Algorithms in Machine Learning: Guarantees and Convergence (UMD, CMSC828U), Foundations of Deep Learning (UMD, CMSC828W).

Optimization and Numerical Methods: Optimization Methods (HSE), Numerical Methods (HSE), Scientific Computing (UMD, AMSC660, AMSC661), Advanced Optimization (UMD, CMSC764).

Recent Research Projects

Oct 2021 - Present

Tabular Deep Learning
University of Maryland, College Park
As deep learning models challenge the status quo of gradient boosted decision trees in terms of performance, we explore other benefits that deep learning can offer compared to classical methods. In particular, we study advantages of representation learning for domain generalization, transfer-learning, robustness and fairness in tabular data.

Sep 2021 - Dec 2021

A Deep Dive into Dataset Imbalance and Bias in Face Identification
University of Maryland, College Park
Media portrayals often center imbalance as the main source of bias in face recognition. We investigate how different kinds of imbalance in data affect gender bias in face identification. We separately consider imbalance with respect to the number of identities and number of images per identity in the train and gallery sets. Our findings show that each type of imbalance has a distinct effect on a model's performance on each gender presentation.

Aug 2020 - Feb 2021

Technical Challenges for Fairness in Deep Learning
University of Maryland, College Park
We explore how various methods for improving fairness in automated decision making systems work in deep neural networks. We find and explain multiple technical problems associated with applying those tools to deep learning, such as overfitting to fairness objective, fairness gerrymandering and others. The work is published at ICLR 2021 RAI Workshop.

Oct 2020 - Feb 2021

Strong Data Augmentation Sanitizes Poisoning and Backdoor Attacks Without an Accuracy Tradeoff
University of Maryland, College Park
Many previous defenses against poisoning either fail in the face of increasingly strong attacks, or they significantly degrade performance of the models on clean data. In this project we show that strong data augmentations, such as mixup and CutMix, can significantly diminish the threat of poisoning and backdoor attacks without trading off performance. The work is published at ICASSP 2021.

Apr 2020 - Oct 2020

LowKey: Leveraging Adversarial Attacks to Protect Social Media Users from Facial Recognition
University of Maryland, College Park
In this project we develop an efficient adversarial attack against facial recognition systems to protect photos shared online from being used by third-party algorithms to recognize users. We design an attack that could degrade the performance of industrial facial recognition systems (Amazon Rekognition, Microsoft Azure) to below 1%. Finally, we release an easy-to-use webtool for filtering photos with our adversarial attack. The work is published at ICLR 2021.

Conferences and Talks

April 2021

ICLR 2021, RAI Workshop
Technical Challenges for Training Fair Neural Networks

April 2021

ICLR 2021
LowKey: Leveraging Adversarial Attacks to Protect Social Media Users from Facial Recognition

Dec 2020

NeurIPS 2020, Resistance AI Workshop & Workshop on Dataset Curation and Security
LowKey: Leveraging Adversarial Attacks to Protect Social Media Users from Facial Recognition

Professional Service

Reviewer: NeurIPS 2022, ICLR 2022, NeurIPS 2021, ICLR 2021 RAI Workshop, IEEE TPAMI

Computer Skills

Programming: Python (Pytorch, Pyspark), basics of Matlab and R