Tutorial: Ethics, Bias and Responsible AI: Challenges and Mitigation Strategies

18th International Symposium on Visual Computing

Lake Tahoe, NV, USA

October 16-18, 2023

http://www.isvc.net

Abstract

This tutorial will discuss how deep learning methods can enhance visual knowledge discovery and image processing. The tutorial will further focus on risks of privacy and ethical considerations, with discussing cancellability and de-identification as two of the mechanisms to mitigate the privacy concerns in sharing and storing of visual data. It will focus on technical approaches to addressing bias and unfairness in AI, discussing the ways to identify, explain, mitigate and communicate bias. It will conclude with a demo of some advanced AI tools from Microsoft, Google and IBM for assessing and improving fairness, and mitigating bias.

Description

Computer vision and person authentication are some of the key mechanisms of ensuring secure access to resources and reliable decision making in academic and business sectors. It became an established practice for government, consumer, financial and recreational institutions in modern society. However, with the rise of the technological advancements, such as AI and deep learning, more and more capabilities exist not only to infer private information of individuals, but to use aggregate data mining for commercial or other purposes.

Objectives

The objective of this tutorial is to provide a comprehensive consideration to issues of mitigating risks to data privacy and ethical considerations of using AI. Upon completion of this tutorial, attendees will be equipped with knowledge allowing them to identify and mitigate risks to privacy in data collection and processing, as well as to use the available tools for fair ML and DL.

Introduction	5 minutes
Participant survey on a general understanding of bias in AI (an interactive component)	15 minutes
AI in computer vision and person authentication	30 minutes
Privacy and ethics considerations	20 minutes
Addressing bias and unfairness in AI	20 minutes
Break	20 minutes
Azure Rai Fairness Dashboard, Microsoft Toolkit Fairlearn, What-If tool Google, AIF-360 IBM	30 minutes
Conclusions	10 minutes
Q&A with participants	30 minutes

Intended Audience

This tutorial is suitable for graduate and undergraduate students, researchers, and policy-makers.

Duration

This will be a 3-hour half-day tutorial.

Organizers

Marina L. Gavrilova, Department of Computer Science, University of Calgary, mgavrilo@ucalgary.ca

Marina L. Gavrilova is a Full Professor, an Order of the University of Calgary inductee and a head of the Biometric Technologies and SPARKS Laboratories in the Faculty of Science. Her publications include over 250 refereed articles, edited special issues, books and book chapters in the areas of machine learning, information fusion, knowledge discovery and cybersecurity. She serves as a Founding Editor-in-Chief of Transactions on Computational Science Journal, Springer and an Editor-in-Chief of the International Journal of Digital Human, Inderscience. As a globally renown award-winning researcher and educator, Dr. Gavrilova has given over 50 keynotes, invited lectures and tutorials at major scientific gatherings worldwide, including Stanford University, Purdue University, Fordham University, Microsoft Research USA, Oxford University UK, Samsung Research South Korea and Nanyang Technological University, Singapore. Dr. Gavrilova is a passionate advocate of equity, diversity and inclusion in academia, industry and society.

Date

TBD