Privacy in the world of AI and Big data

Xuan-Son Vu

Department of computing science, Umeå University

Abstract

Recent breakthroughs in artificial intelligence (AI) and Big Data have opened many new possibilities in data analytics, autonomous systems, and the like. However, they pose many challenges regarding privacy since big data is completely opposed to the basis of data protection. Similar to the Internet protocol, Machine Learning - which is the core of AI, wasn't designed with privacy in mind. For instance, Support Vector Machines (SVMs) try to solve a quadratic optimization problem by deciding which instances of training dataset are support vectors. This means that the data of people involved in the training process will be also published within the SVM models. Recently, Fredrikson et al. used hill-climbing algorithm on the output probabilities of a computer-vision classifier to reveal individual faces from the training data. Because of all these issues, privacy guarantees must apply to the worst-case outliers and thus will also destroy data utilities. From all above reasons, in this talk, I will walk you through on how to protect privacy when doing machine learning and how to have a good trade-off between data utilities and privacy. Our demonstrated data will be a real social network data, which is important these days because of privacy breaches such as Cambridge Analytical. Hope to see you in the talk.