10. Data Analytics What Can New Analyses Techniques Bring for Better Farm Results 2

Title presentation

Secure precision farming : a security model and a security risk assessment of machine learning & genomics

Author(s)

H. Lardé , S. Gambs, M.O. Killijian & A.B.Diallo

Institution for which the first author of this abstract is working

Université du Québec à Montréal - UQAM, P.O. Box 8888, Station Centre-ville, Montreal, Quebec H3C 3P8, CANADA

Abstract

In precision agriculture, machine learning approaches are now being used to address research issues related to genomic data. These approaches are, for example, used for forecasting within farms. It has often been highlighted that genomic data are extremely sensitive from the point of view of computer security in general. This is essentially due to the massive amount of precise information they embed. In addition, the increasing use of machine learning methods has led to the discovery of numerous confidentiality and integrity breaches both in the models and in the data they represent. This work is a position paper that allows the field of precision agriculture to be confronted with possible security issues that could arise from the synergistic use of machine learning techniques with genomic data. In this paper, we first propose a security model dedicated to the specific settings and threats of the precision agriculture domain. In this model, we identify and classify the resources at risk, define the different classes of actors, determine the risk vectors and propose some realistic attack scenarios. Two versions of the model are proposed in order to take into account the potential presence of a trusted third party to regulate data access and machine learning model management. On one hand, this third party is usual in this setting, particularly with the objective of enforcing some security policies, but on the other hand we also identify some potential threats and risks introduced by this third party. We then use this model to conduct a security risk assessment for the domains of precision agriculture and machine learning. During the study, we assess the security impact of genomic data usage during both training and prediction stages. The considered attacks encompass model theft, model inversion, membership inference, data poisoning and antagonistic examples. For each type of attack, we study the impact of the presence of a trusted third party for training and access to model predictions. Finally, we propose some adapted mitigation means, such as differential privacy, robust models or removal of confidence index from predictions and discuss their effectiveness.