Towards secure digital farming :

SECURITY MODEL AND RISKS ASSOCIATED TO MACHINE LEARNING

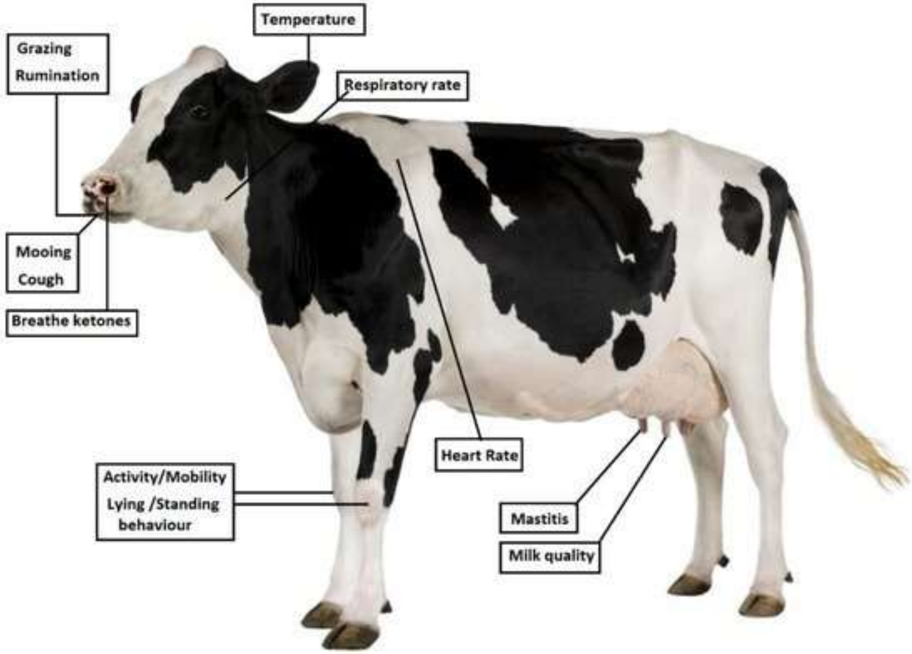
H. Lardé, *S. Gambs, M.O. Killijian, A.B. Diallo*

Université du Québec À Montréal - UQÀM

Introduction

Digital farming relies on data

- Environment
- Production
- Health
- Welfare
- Genomics
- Management



A. Awashti, A. Awashti, D. Riordan, J. Walsh. Non-Invasive Sensor Technology for the Development of a Dairy Cattle Health Monitoring Sys, 2016.

This paves the way for the use of Artificial Intelligence for precision and efficacity

Introduction

Use of Artificial Intelligence and Machine Learning introduces risks to :

Security

Privacy

Farming sector needs to improve cyber security

- U.S. Government Accountability Office (2019)
- Survey conducted by Geil et al. (2018)

Outline

I - Security model for digital farming

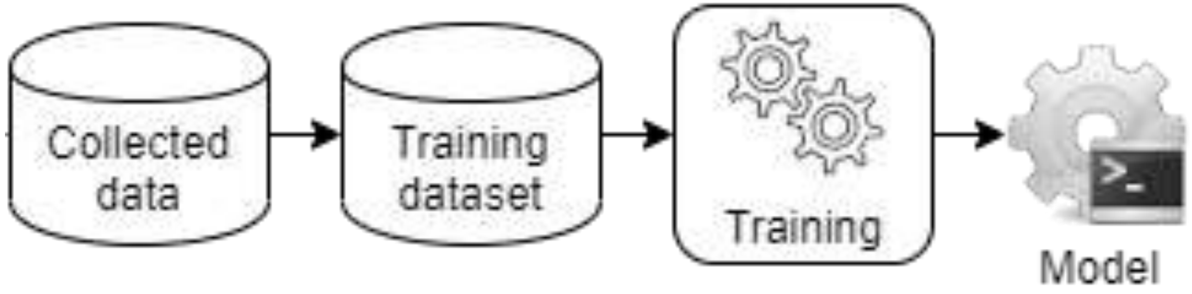
1. Data chain
2. Risk vectors
3. Adversary model

II - Risks to machine learning

1. Privacy of data and model
2. Integrity of model and predictions
3. Means of mitigation

Machine learning flow

1. Data collected : raw data
2. Training dataset: pre-processed data
3. Training: design predictive model
4. Model: query and make predictions



I - Security model of digital farming

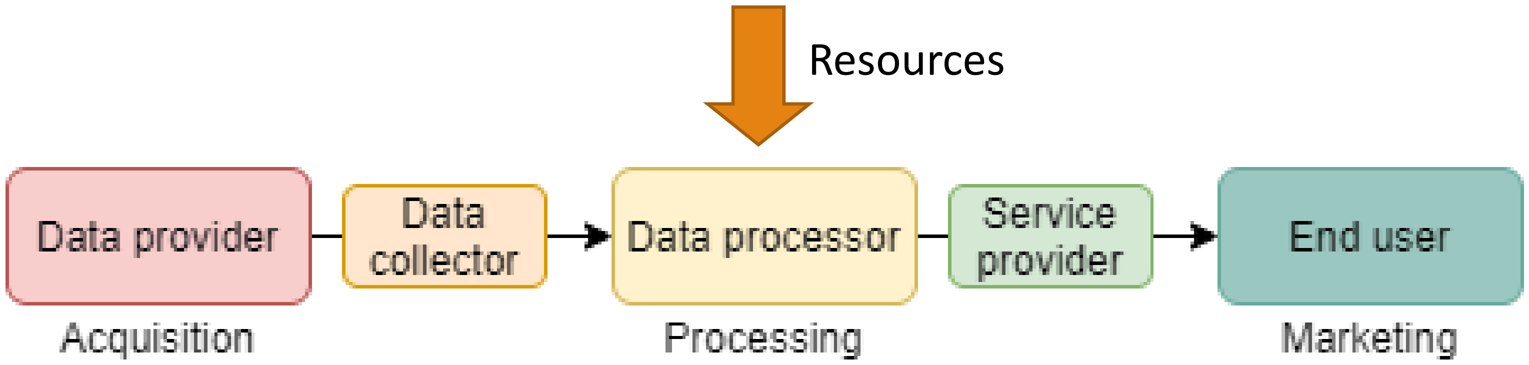
1. Data chain

Data chain describes the data life cycle between resources and actors. Wolfert et al. (2017)

Resources

- Training dataset: confidentiality
- Trained model: confidentiality, integrity
- Predictions: integrity

Actors

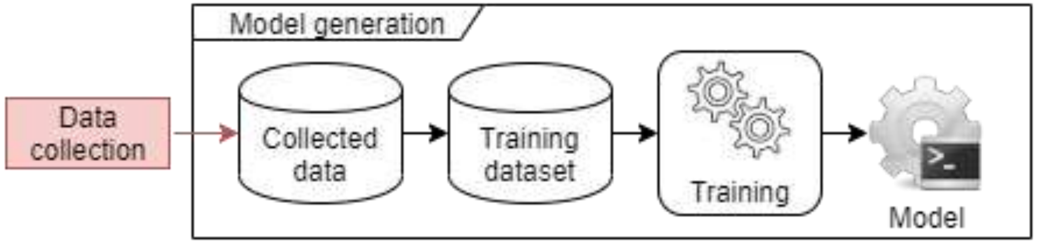


I - Security model of digital farming

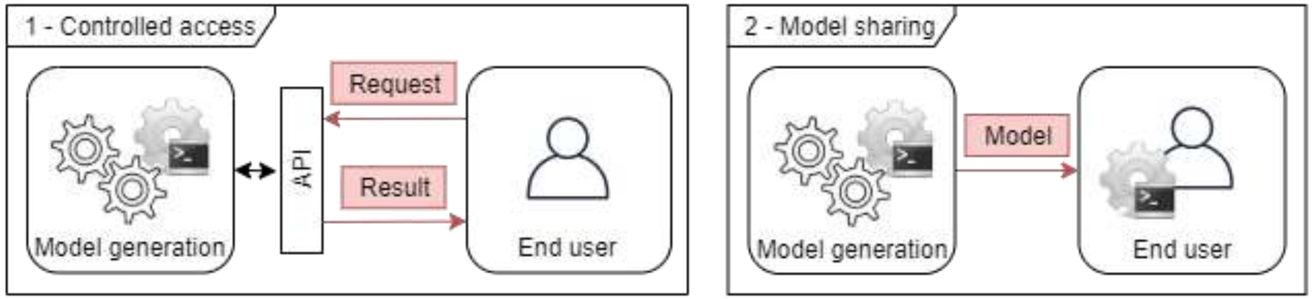
2. Risk vectors

Data processors have two interfaces that are the data collection (upstream) and model predictions (downstream)

Upstream



Downstream



I - Security model of digital farming

3. Adversary model

Goals

- Financial gain: model privacy
- Information leak: training dataset privacy
- Disruption: model and prediction integrity

Capabilities

- Insider: specific knowledge
- Outsider: large cyber resources and advanced skills

II - Risks to ML in digital farming

1. Confidentiality of data and model

Membership inference

- Determine if a data point is part of the training set
- Ex: On hospital discharge dataset, Shokri et al. (2017)

Model inversion

- Gain knowledge about the training dataset
- Ex: Reconstruct unknown features of patient with warfarin dosing system, Fredrikson et al. (2014)

Model theft

- Steal the model parameters or extract model behavior
- Ex: Steal model for vendor (Machine Learning as a Service), Tramèr et al. (2016)

II - Risks to ML in digital farming

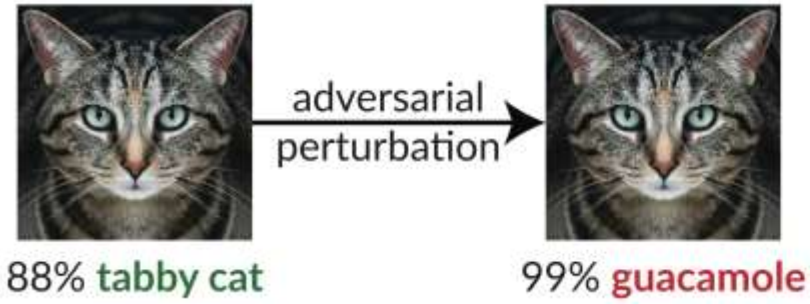
2. Integrity of model and predictions

Data poisoning

- Inject malicious data point in training set to compromise the model
- Ex: Backdoor on authentication system, Chen et al. (2017)

Adversarial example

- Craft malicious request to compromise the prediction
- Ex: Evade malware detection system, Al-Dujaili et al. (2018)



Adversarial example on InceptionV3 classifier
Retrieved from https://github.com/anishathalye/obfuscated-gradients/blob/master/example.png

II - Risks to ML in digital farming
3. Means of mitigation

Differential privacy
- Increase privacy of each element in dataset by adding small noise to without affecting utility
- Membership inference, model inversion

Query auditing
- Analyse queries or filter results to prevent attacks
- Membership inference, model inversion

Robust model
- Use training techniques that increase model resilience
- Data poisoning, adversarial example

Conclusion

Digital farming must improve cyber security

ML research exposes new practical risks to security and privacy

I. We developed a security model for digital farming
II. We investigated risk to machine learning and practical means of mitigation

Opportunity to increase resilience of digital farming

References

- Priority open recommendations: U.S. Department of Agriculture, Apr 2019. Available at https://www.hsdl.org/?view&did=824065.
- Andrew Geil, Glen Sagers, Aslihan D. Spaulding, and James R. Wolf. Cyber security on the farm: An assessment of cyber security practices in the united states agricultural industry. International Food and Agribusiness Management Review, (1030-2018-1811), Feb 2018
- Sjaak Wolfert, Lan Ge, Cor Verdouw, and Marc-Jeroen Bogaardt. Big data in smart farming – a review. Agricultural Systems, 153:69 – 80, 2017.
- Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. Membership inference attacks against machine learning models. 2017 IEEE Symposium on Security and Privacy(SP), May 2017.
- Matthew Fredrikson, Eric Lantz, Somesh Jha, Simon Lin, David Page, and Thomas Ristenpart. Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing. In23rd USENIX Security Symposium (USENIX Security 14), pages 17–32, San Diego, CA, aug 2014. USENIX Association.
- Florian Tramèr, Fan Zhang, Ari Juels, Michael K. Reiter, and Thomas Ristenpart. Stealing machine learning models via prediction apis, 2016.
- Xinyun Chen, Chang Liu, Bo Li, Kimberly Lu, and Dawn Song. Targeted backdoor attacks on deep learning systems using data poisoning, 2017.
- Abdullah Al-Dujaili, Alex Huang, Erik Hemberg, and Una-May OReilly. Adversarial deep learning for robust detection of binary encoded malware. 2018 IEEE Security and Privacy Workshops (SPW), May 2018