Certificate I: Understanding AI and Machine Learning in Africa

Course AIML02: AI and Machine Learning in Africa

Module 02: Application Case Studies

Lecture 06: Conservation

Carnegie Mellon University
Africa

Learning Objectives

1. Identify the impact of poaching and explain the importance of wildlife conservation

2. Identify the challenges of modelling the behavior of poachers

3. Explain how machine learning can be used to predict poaching activity in protected parks

4. Explain the importance of addressing prediction uncertainty in machine learning models

Lecture Contents

Poaching

- Illegal wildlife poaching threatens
  - Biodiversity
  - Ecological balance
  - Ecotourism

- Many species are being poached to near-extinction
  - Elephants for ivory tusks
  - Rhino for horn

- Many other animals, such as wild pigs and apes, are hunted for their meat

A tranquilised rhino is dehorned to make it less attractive to poachers

Source: https://www.bbc.com/news/science-environment-20971182

Poaching

"Artificial intelligence frameworks can significantly advance wildlife protection efforts by learning from past poaching activity to prescribe actionable recommendations to park managers."

(Xu et al., 2020)

Poaching

To be successful in

- Assessing risk of attacks by poachers
- Planning patrols

Rangers require knowledge of poachers' behavior

Well-hidden snares
Source: Ugandan Wildlife Authority

Challenges of Machine Learning in Conservation

Machine learning needs the same data, but …

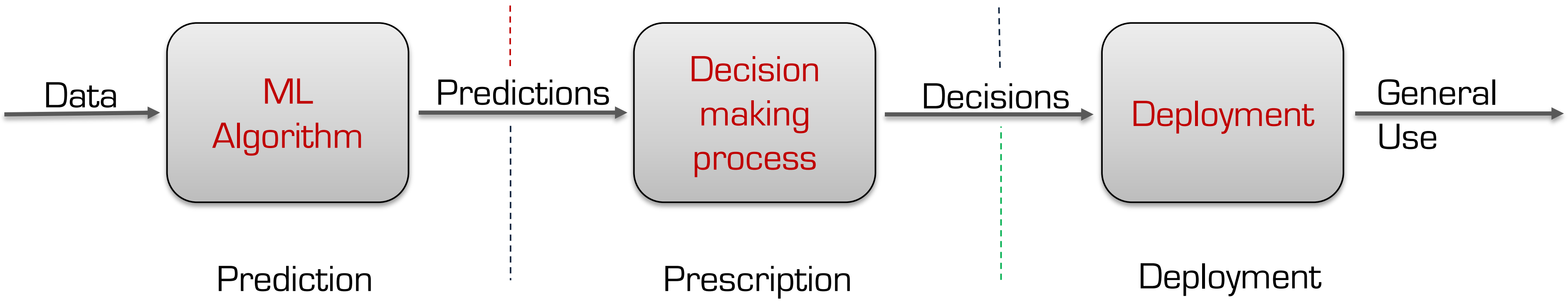
- Data on wildlife crimes are often very imbalanced
- Unreliable negative labels
  (indicating an absence of poaching activity)
- Historical poaching observations are not collected thoroughly and uniformly: biased dataset
- Poaching patterns vary with region

Well-hidden snares
Source: Ugandan Wildlife Authority

Up to 99.6% negative labels (no poaching activity) and 0.4% positive labels (poaching activity)

Protection Assistant for Wildlife Security (PAWS)

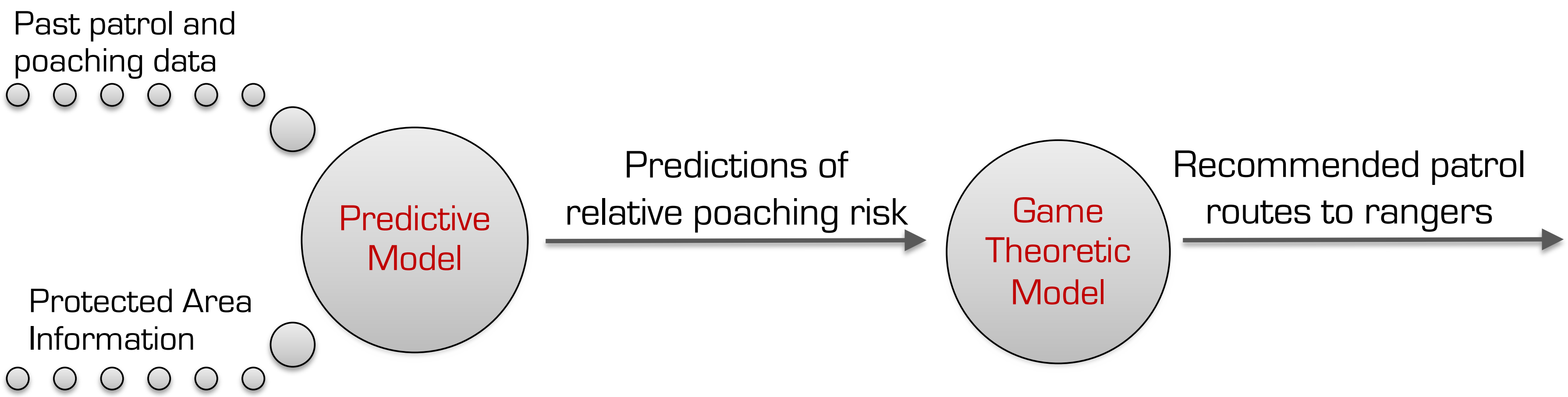
Traditional data-to-deployment pipeline



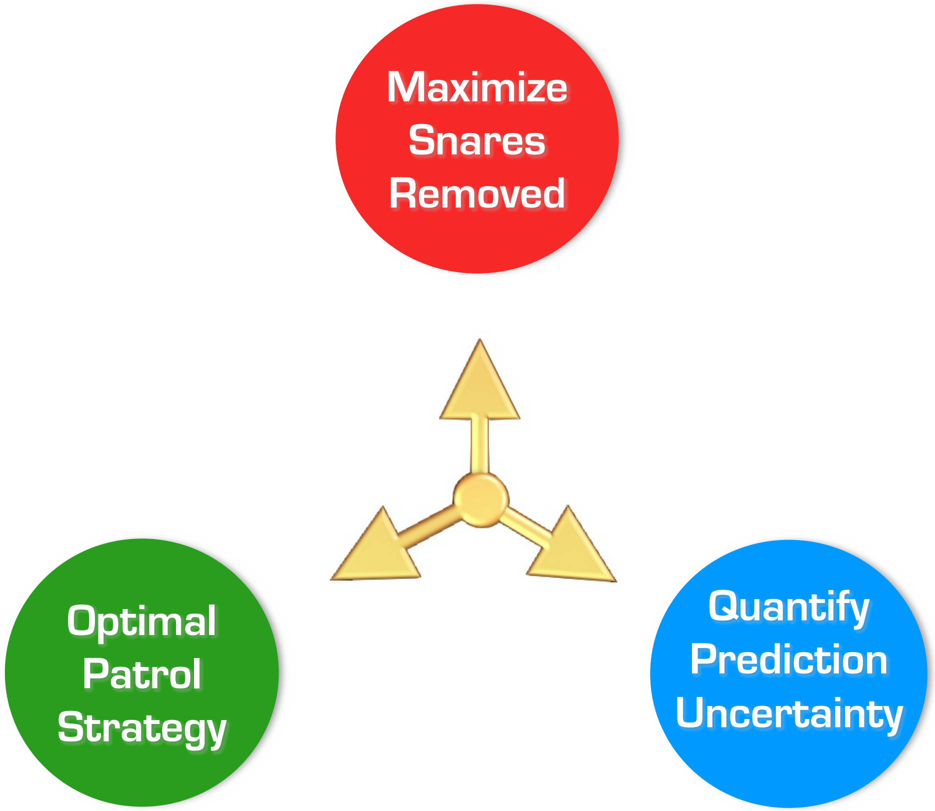
This approach has been used in a previous wildlife conservation applications:
Protection Assistant for Wildlife Security (PAWS)

(Yang et al., 2014)

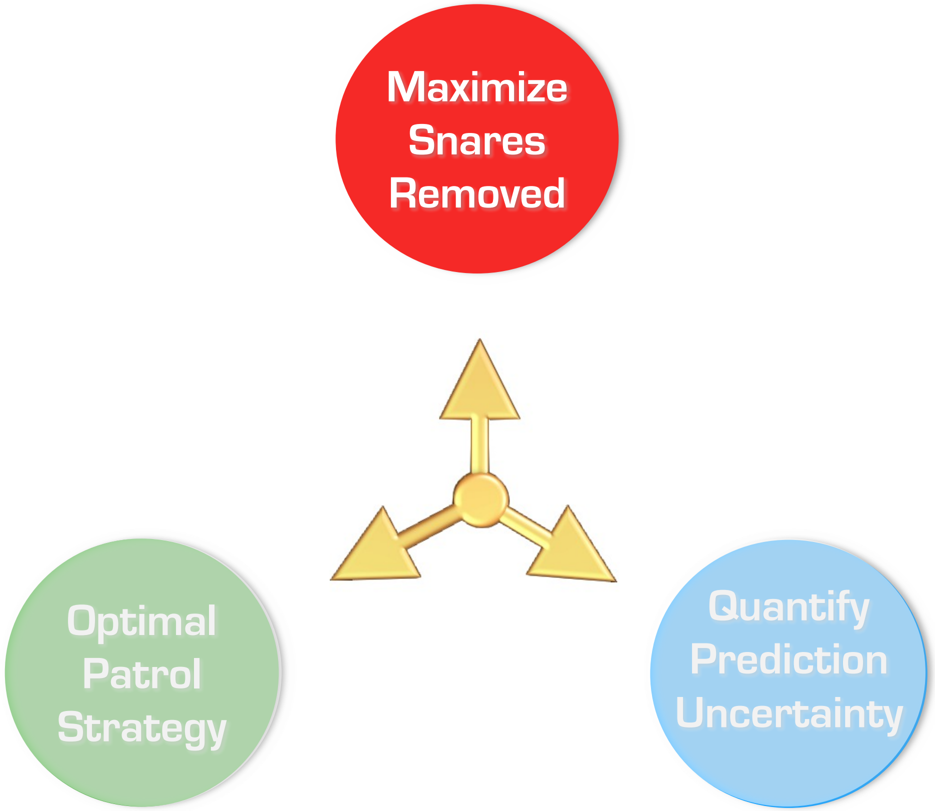
Protection Assistant for Wildlife Security (PAWS)



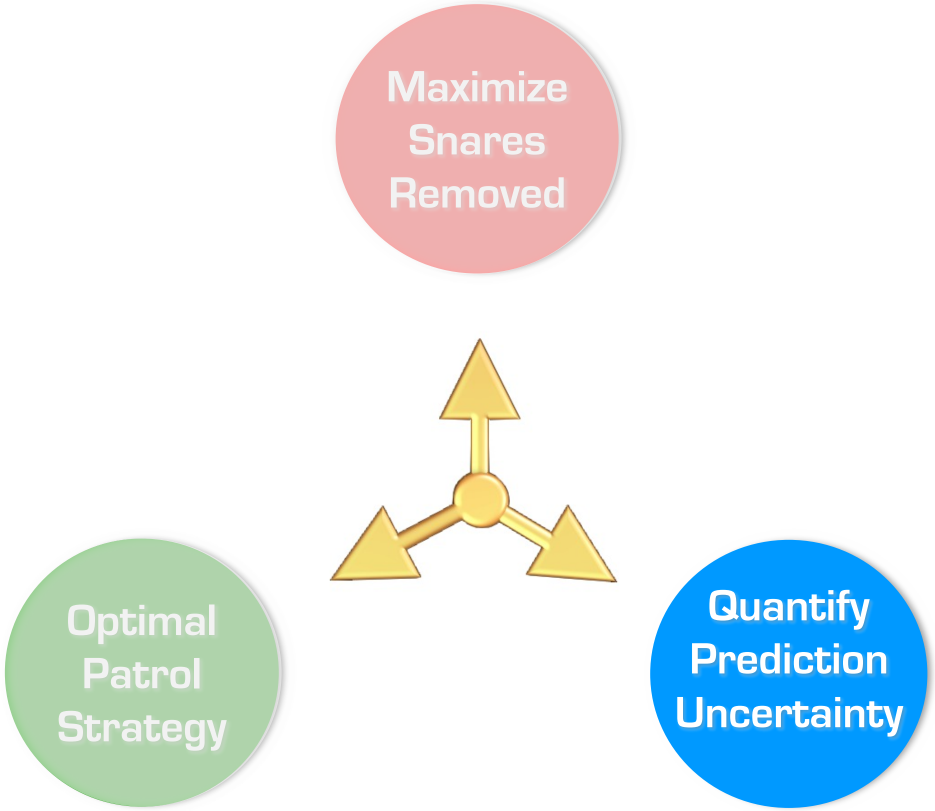
Predictive Modelling with Uncertain Data



Predictive Modelling with Uncertain Data



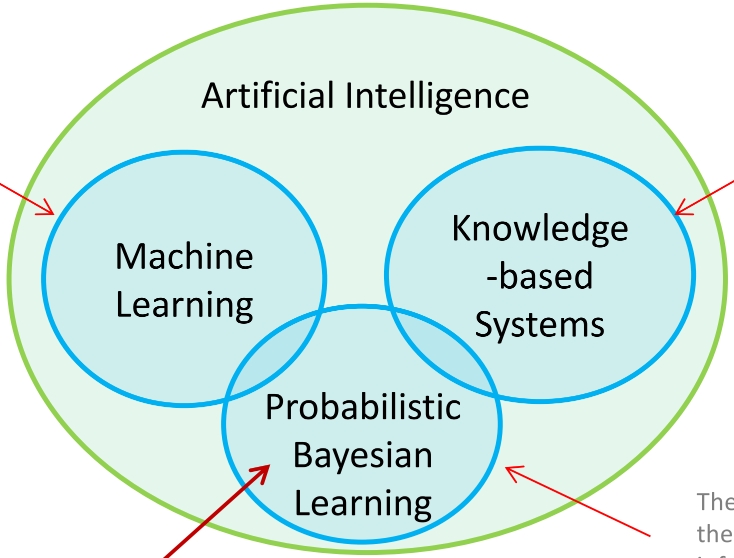
Predictive Modelling with Uncertain Data



Predictive Modelling with Uncertain Data

Machine learning techniques, such as deep learning, enable AI applications after being trained using data in very large datasets (sometimes referred to as big data)

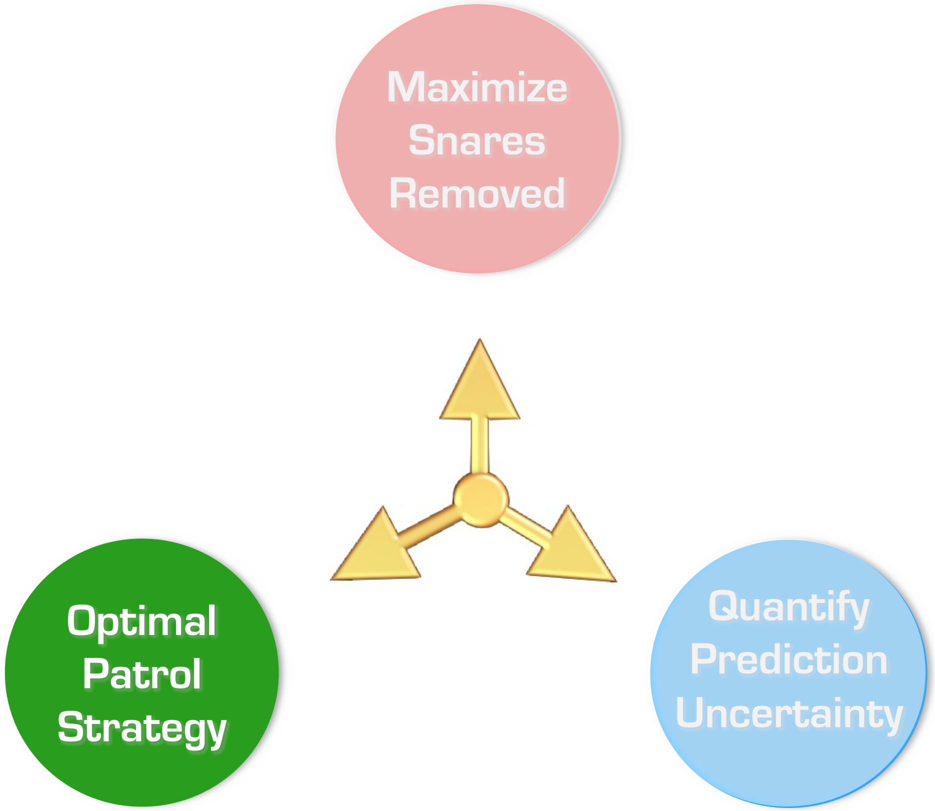
Knowledge-based systems attempt to emulate the problem-solving skills of a human expert by using explicitly encoded knowledge and reasoning to solve problems (they used to be called expert systems)



These provide a powerful way to deal with uncertainty by capturing the statistical relationships among the entities being modeled: inferring the most likely outcomes and drawing conclusions that are the most likely to be correct when solving problems

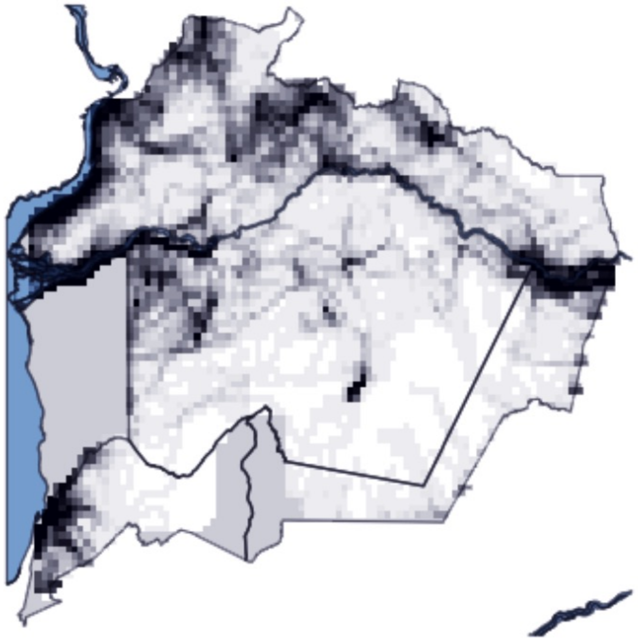
Gaussian processes are used to model uncertainty in this application case study

Predictive Modelling with Uncertain Data

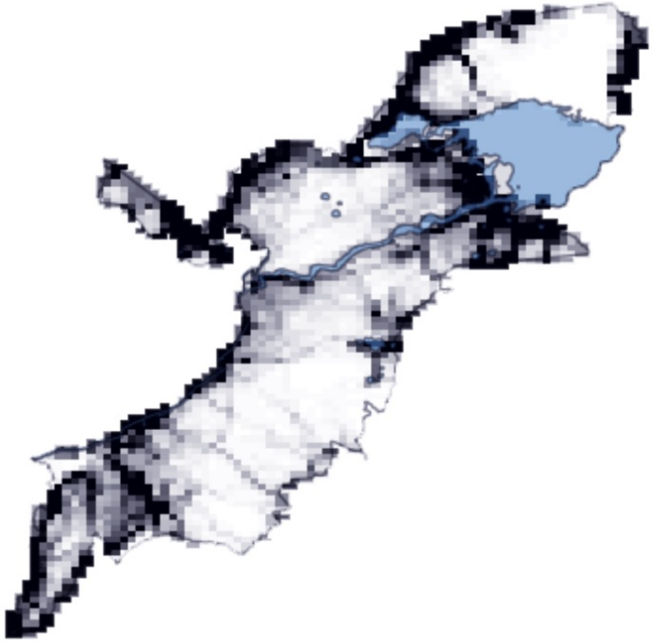


Predictive Modelling with Uncertain Data

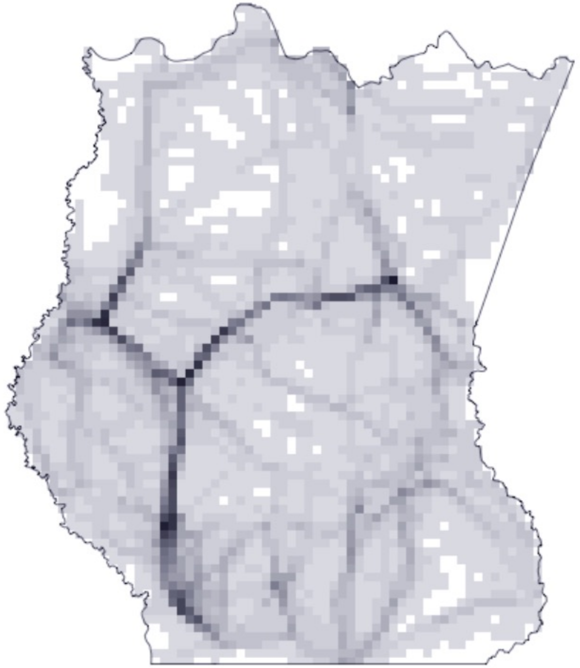
Murchison Falls National Park (MFNP)
Uganda



Queen Elizabeth National Park (QENP)
Uganda



Srepok Wildlife Sanctuary (SWS)
Cambodia



(Xu et al., 2020)

Predictive Modelling with Uncertain Data

01

Geospatial Data

Terrain features, such as elevation maps, rivers, and forest cover;

Landscape features, such as roads, park boundaries, local villages, and patrol posts

Ecological features such as animal density.

Predictive Modelling with Uncertain Data

01

Geospatial Data

Terrain features, such as elevation maps, rivers, and forest cover;

Landscape features, such as roads, park boundaries, local villages, and patrol posts

Ecological features such as animal density.

Poaching & Patrol Data

GPS location, date, time of observations: animals / humans spotted; indications of illegal or poaching activity (campsites, cut trees, firearms, bullet cartridges, snares, or slain animals).

02

Predictive Modelling with Uncertain Data

01

Geospatial Data

Terrain features, such as elevation maps, rivers, and forest cover;

Landscape features, such as roads, park boundaries, local villages, and patrol posts

Ecological features such as animal density.

02

Poaching & Patrol Data

GPS location, date, time of observations: animals / humans spotted; indications of illegal or poaching activity (campsites, cut trees, firearms, bullet cartridges, snares, or slain animals).

03

Data Discretization

Construct a 1 km x 1km discrete spatio-temporal map of the area with N cells, each cell comprising a sequence of T three-month time intervals, recording k features, both static geospatial features and one time-variant feature identifying the amount of patrol coverage in the previous time step.

Predictive Modelling with Uncertain Data

Dataset

Data
matrix

Observation
vector of labels

$$\mathcal{D} = (\mathbf{X}, \mathbf{y})$$

Predictive Modelling with Uncertain Data

Dataset

Data matrix

Observation vector of labels

$$\mathcal{D} = (\mathbf{X}, \mathbf{y})$$

Number of time intervals

Number of locations

Number of features

$$\mathbf{X} \in \mathbb{R}^{T \times N \times k}$$

One of these features is the time-variant patrol effort c

Predictive Modelling with Uncertain Data

Dataset

Data matrix

Observation vector of labels

$$\mathcal{D} = (\mathbf{X}, \mathbf{y})$$

Number of time intervals

Number of locations

Number of features

$$\mathbf{X} \in \mathbb{R}^{T \times N \times k}$$

One of these features is the time-variant patrol effort c

No poaching activity

Poaching activity

$$\mathbf{y} \in \{0, 1\}^{T \times N}$$

Each time interval in each location is labelled

Predictive Modelling with Uncertain Data

- The success with which rangers detect poaching activity in a given 1 km x 1 km cell depends on the amount of effort they exert in patrolling that cell

- Positive instances are reliable, irrespective of the amount of patrol effort.

Cells at a given time that are labelled poaching activity

- Why? If rangers find a snare in a cell, poaching occurred with certainty

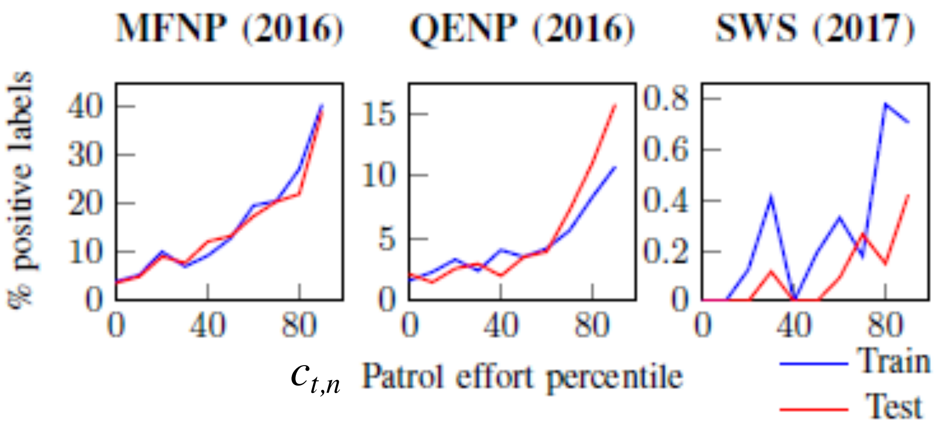
- Negative instances have different levels of uncertainty that depend on the patrol effort $c_{t,n}$ exerted in cell n during time t

Predictive Modelling with Uncertain Data

The percentage of poaching activity detected within each cell increases proportionally with the patrol effort exerted within that cell

To quantify the uncertainty of negative instances, apply a threshold θ of patrol effort (the distance patrolled)

- Patrol effort $c_{t,n} \geq \theta \rightarrow$ more reliable (less uncertain)
- Patrol effort $c_{t,n} < \theta \rightarrow$ less reliable (more uncertain)



Percentage of positive labels at different thresholds of patrol effort
Year in parentheses used for test set (Xu et al., 2020)

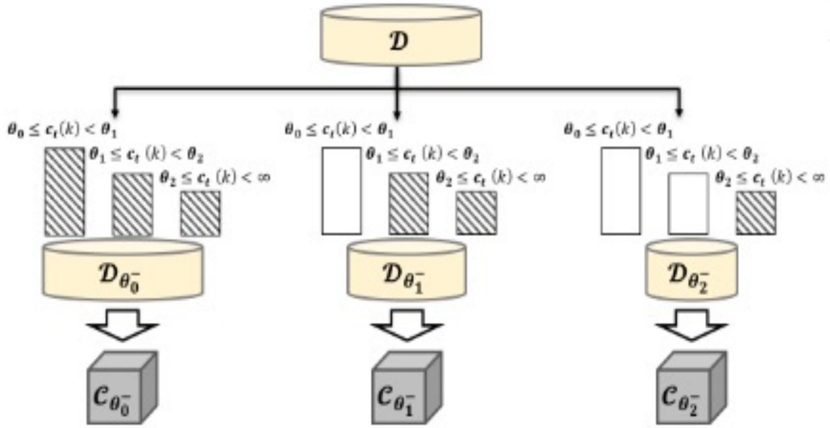
Predictive Modelling with Uncertain Data

The imperfect observation aWare Ensemble (iWare-E)

Uses a bagging ensemble of weak classifiers

- Decision Trees, or
- Support Vector Machines (SVMs)

to generate a strong classifier



(Xu et al., 2020)

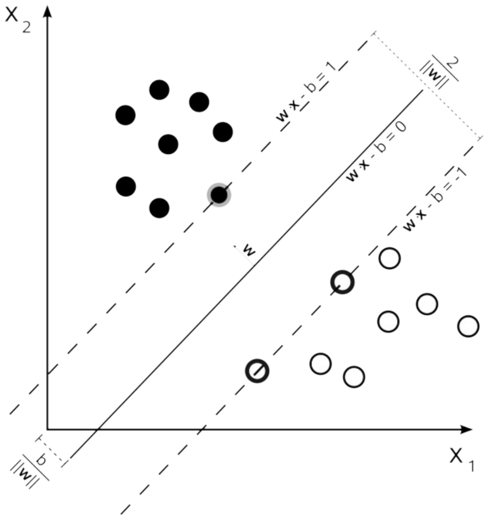
Predictive Modelling with Uncertain Data

The imperfect observation aWare Ensemble (iWare-E)

Uses a bagging ensemble of weak classifiers

- Decision Trees, or
- Support Vector Machines (SVMs)

to generate a strong classifier



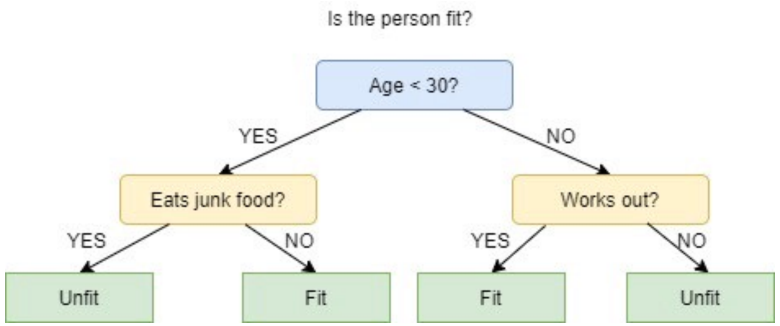
Predictive Modelling with Uncertain Data

The imperfect observation aWare Ensemble (iWare-E)

Uses a bagging ensemble of weak classifiers

- Decision Trees, or
- Support Vector Machines (SVMs)

to generate a strong classifier



https://towardsdatascience.com/decision-trees-for-classification-id3-algorithm-explained-89df76e72df1

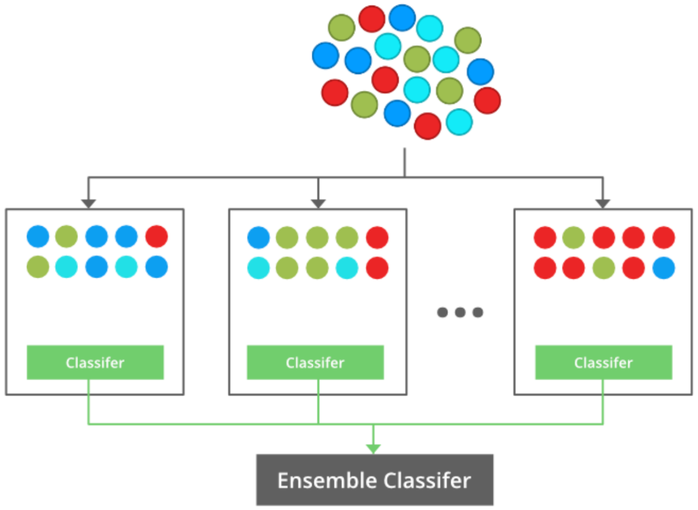
Predictive Modelling with Uncertain Data

The imperfect observation aWare Ensemble (iWare-E)

Uses a bagging ensemble of weak classifiers

- Decision Trees, or
- Support Vector Machines (SVMs)

to generate a strong classifier



https://www.geeksforgeeks.org/bagging-vs-boosting-in-machine-learning/

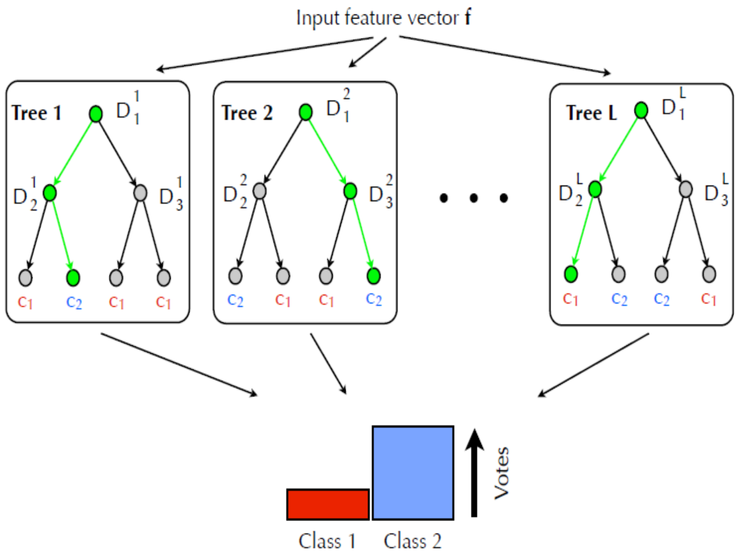
Predictive Modelling with Uncertain Data

The imperfect observation aWare Ensemble (iWare-E)

Uses a bagging ensemble of weak classifiers

- Decision Trees, or
- Support Vector Machines (SVMs)

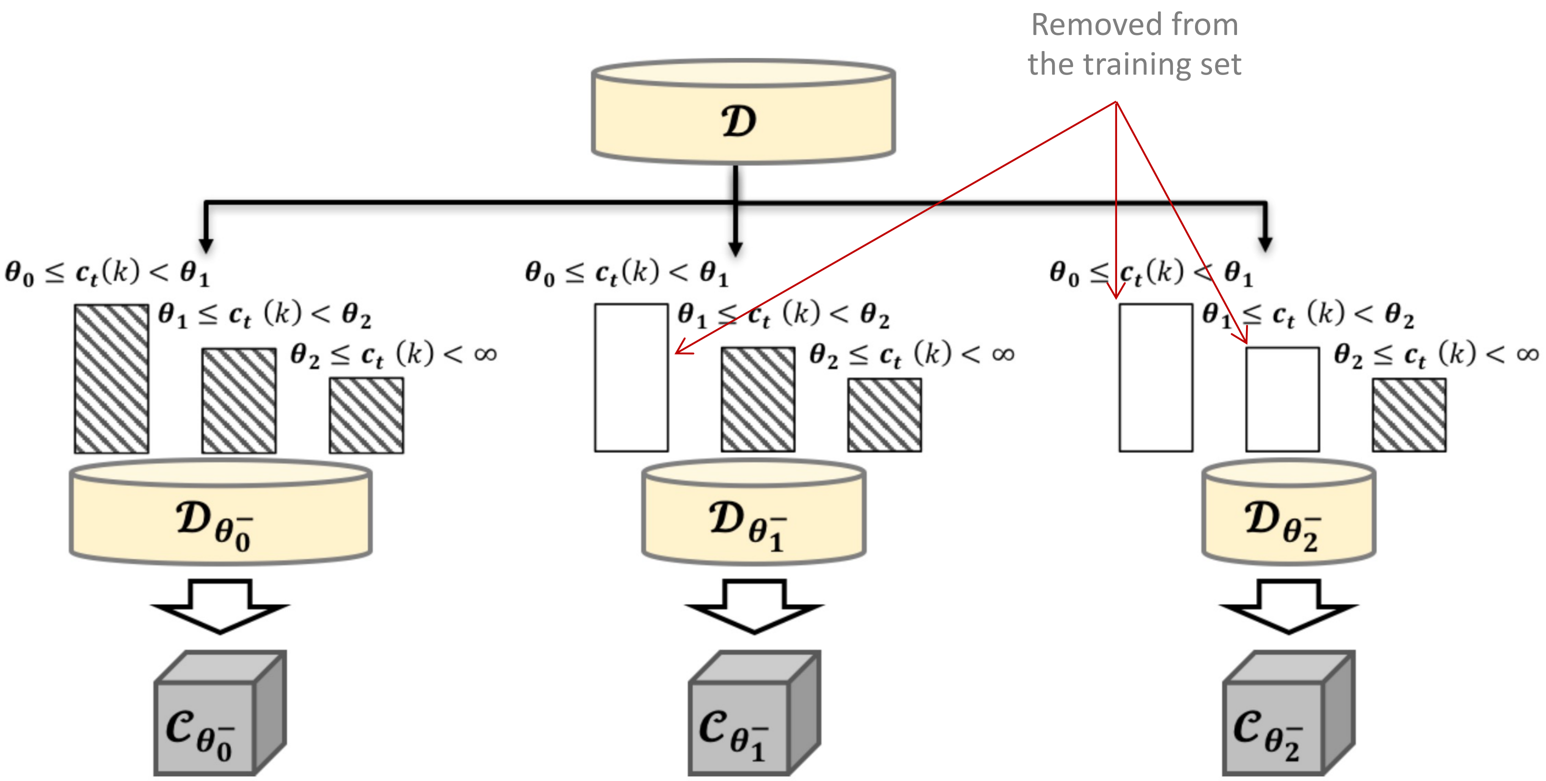
to generate a strong classifier



ing/

Credit: Toby Breckon, Durham University

Predictive Modelling with Uncertain Data



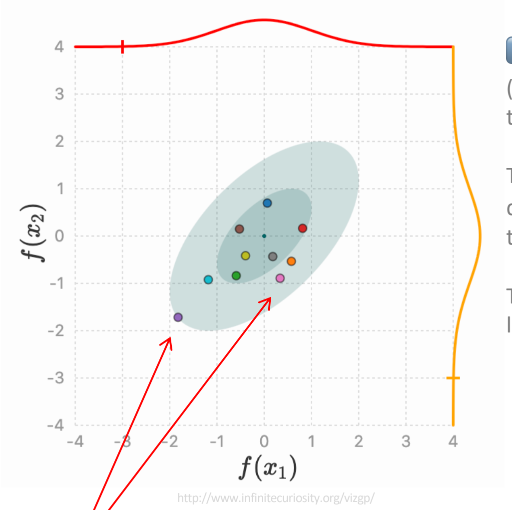
(Xu et al., 2020)

Predictive Modelling with Uncertain Data

iWare-E +

1. Optimally weight classifiers to allow training using the entire dataset instead of a subset
2. Select patrol effort threholds θ_i using percentiles and characteristics of the data instead of using equally-distanced values
3. Using Gaussian process classifiers as weak learners to quantify the uncertainty of the predictions

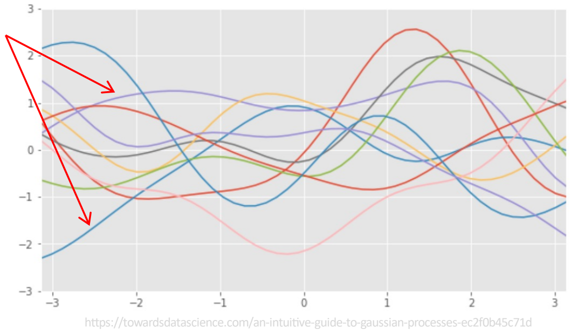
Predictive Modelling with Uncertain Data

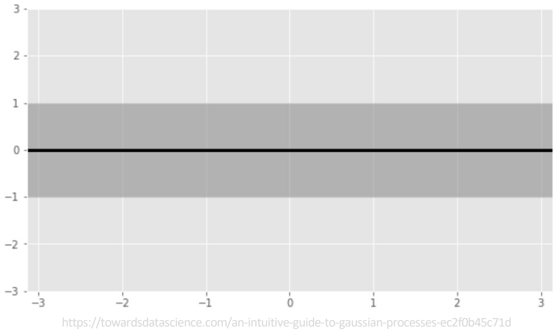


2 Each sample produces a function (this function is a random variable that predicts some outcome).

The shape of the function is also determined by covariance matrix and the kernel function.

The variety of functions reflects the lack of specific data.



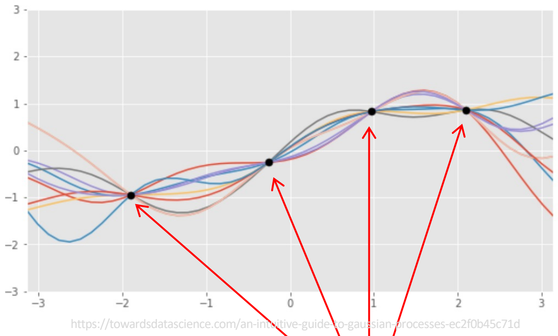


3 The lack of information about the function we want to predict is reflected by the zero mean and large variance (or uncertainty) about any prediction.

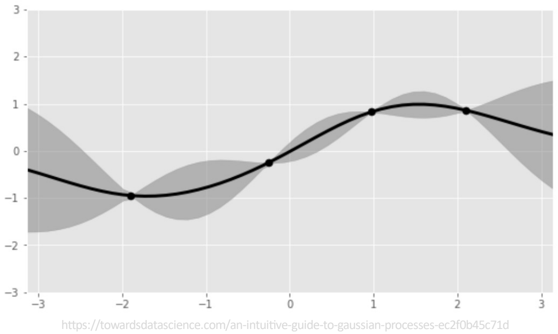
1 Ten samples from a zero-mean 2D Gaussian with a positive covariance for x_1 and x_2 (meaning a change in x_1 will have some corresponding change in x_2, and vice versa).

The shape of the Gaussian is governed by the covariance matrix, which, in turn, is determined by the so-called kernel function.

This shape of the Gaussian reflects prior belief about the nature of process being modelled (e.g., smooth, periodic, or linear) but without any specific data yet to determine the process we are modelling.



4 The introduction of four data points alters the Gaussian distribution (by introducing information into the covariance matrix) and, thus, alters the associated sampled functions which now intersect each data point. This introduction of data allows us to model the posterior Gaussian distribution.



5 The mean of these sampled functions, now constrained by the data, is the function we wish to model (the black curve) and it is the basis for predicting values that are not in the (four point) training data set. The uncertainty associated with these predictions at each point of the curve, i.e., the shaded areas, is also altered: it is zero at the training data points (because there is no uncertainty there) and increases as we move away from them.

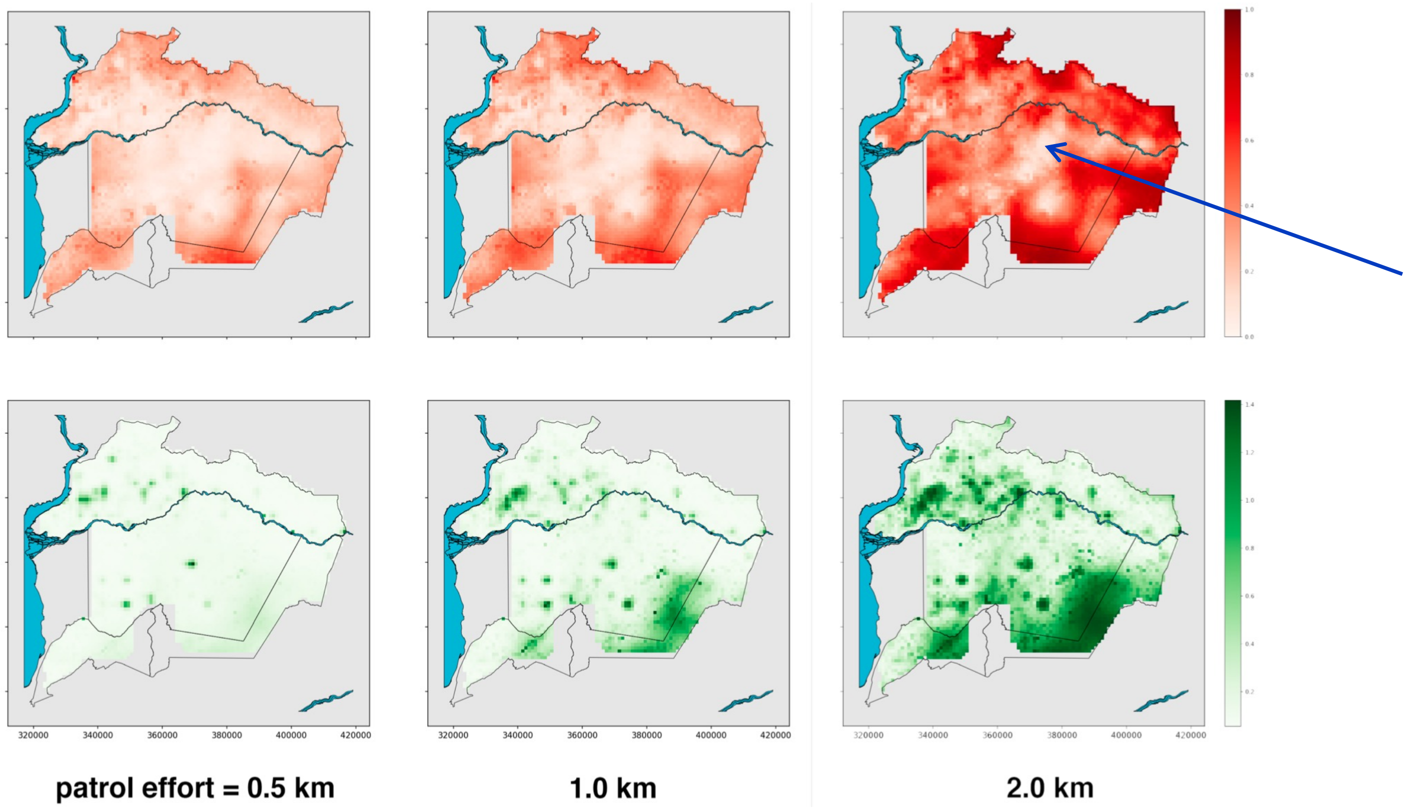
Predictive Modelling with Uncertain Data

iWare-E model	Without the iWare-E model (Baseline model)
Bagging ensembles of SVMs (SVB-iW)	Bagging ensembles of SVMs (SVB)
Bagging ensembles of Decision Trees (DTB-iW)	Bagging ensembles of Decision Trees (DTB)
Bagging ensembles of Gaussian Process Classifiers (GPB-iW)	Bagging ensembles of Gaussian Process Classifiers (GPB)

Predictive Modelling with Uncertain Data

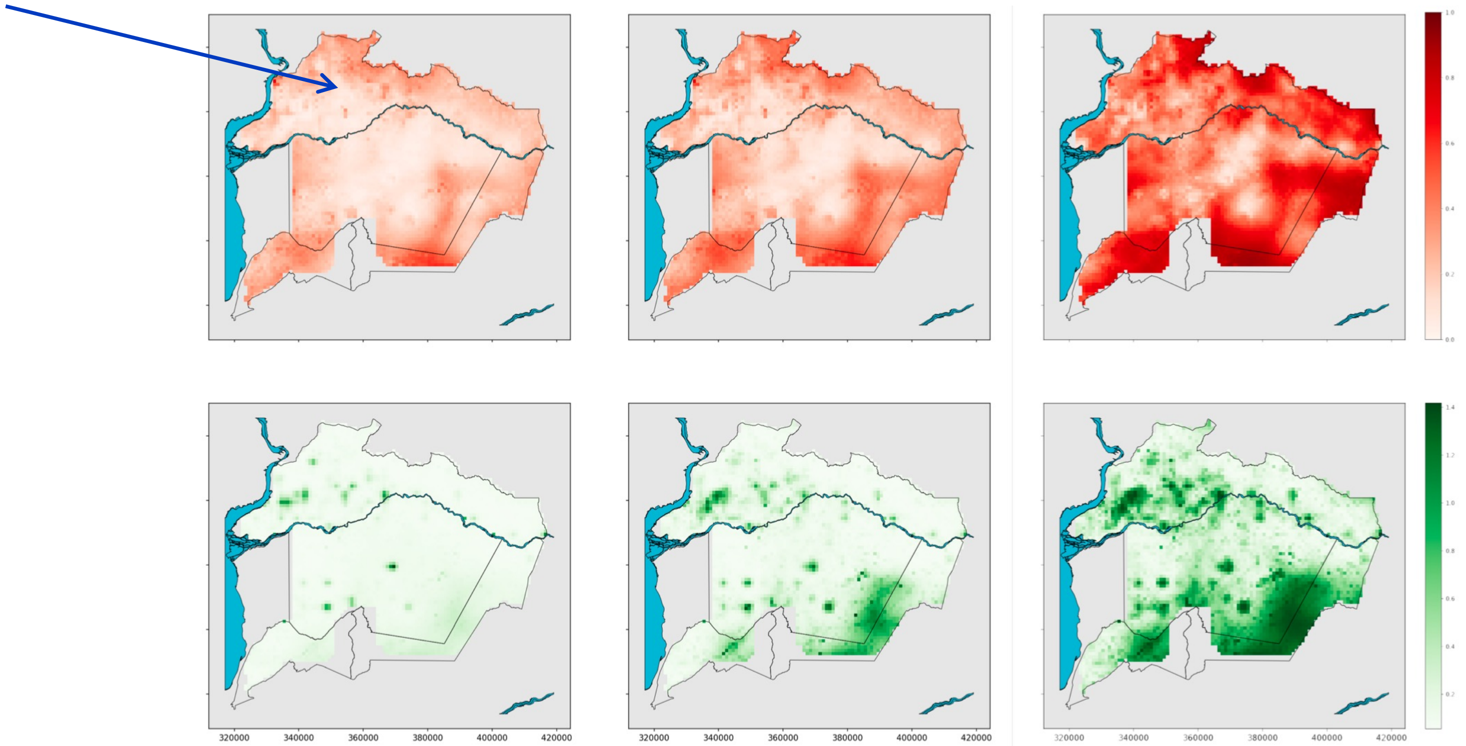
		without iWare-E			with iWare-E		
		SVB	**DTB**	**GPB**	**SVB**	**DTB**	**GPB**
MFNP	**2014**	0.518	0.587	0.626	0.695	**0.711**	0.685
	2015	0.504	0.620	0.670	0.683	0.706	**0.726**
	2016	0.513	0.589	0.603	0.672	0.680	**0.681**
	Avg	0.512	0.599	0.633	0.683	**0.699**	0.697
QENP	**2014**	0.505	0.654	0.693	0.619	**0.735**	0.717
	2015	0.501	0.589	0.600	0.632	0.696	**0.713**
	2016	0.502	0.635	0.611	0.644	0.728	**0.733**
	Avg	0.503	0.626	0.635	0.632	0.720	**0.721**

Predictive Modelling with Uncertain Data



The heatmaps in red at the top show the predicted probability of detecting of poaching activity in Murchison Falls National Park, Uganda, in 2017 based on the based on the Gaussian Process iWare-E model. The associated uncertainty of the predictions are shown in the green heatmaps below.

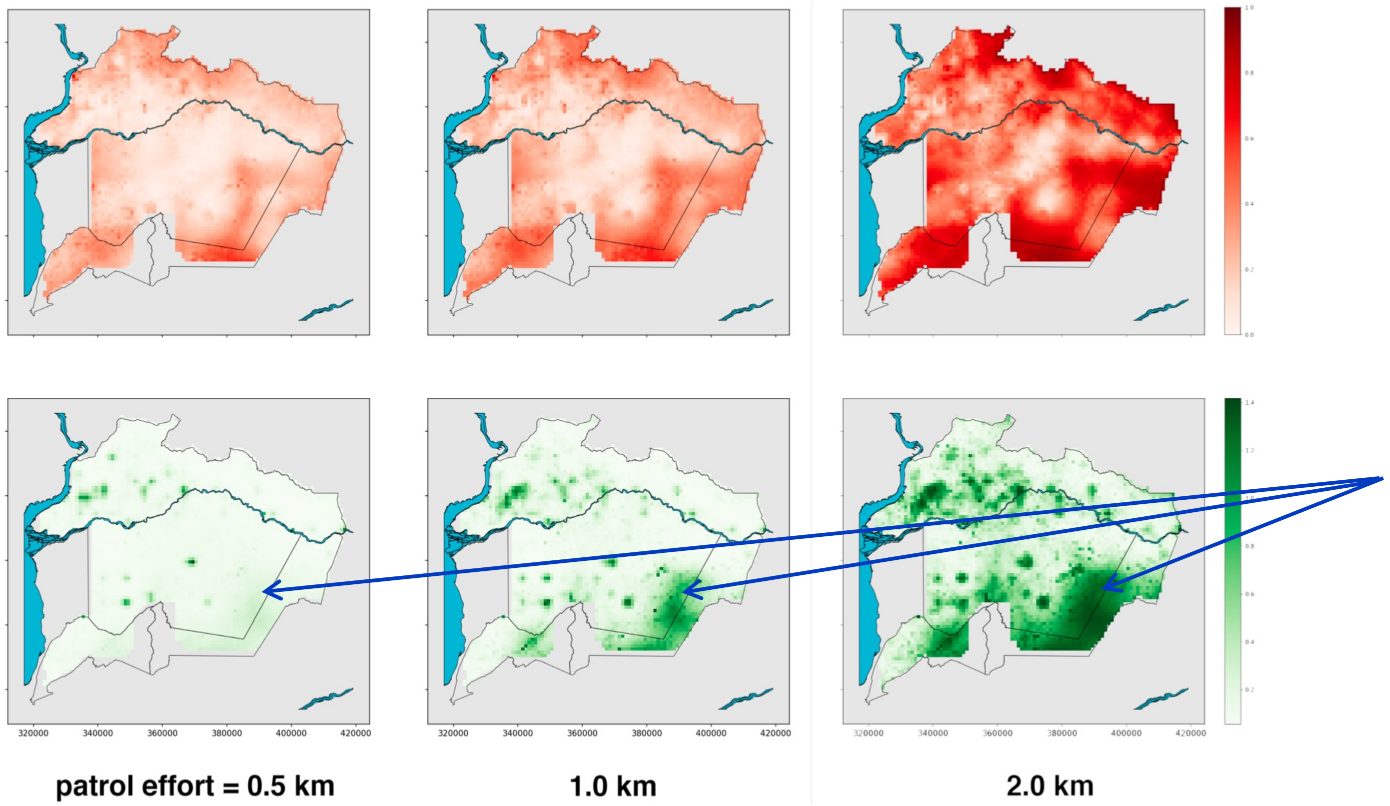
Predictive Modelling with Uncertain Data



patrol effort = 0.5 km **1.0 km** **2.0 km**

The heatmaps in red at the top show the predicted probability of detecting of poaching activity in Murchison Falls National Park, Uganda, in 2017 based on the based on the Gaussian Process iWare-E model. The associated uncertainty of the predictions are shown in the green heatmaps below.

Predictive Modelling with Uncertain Data



The heatmaps in red at the top show the predicted probability of detecting of poaching activity in Murchison Falls National Park, Uganda, in 2017 based on the based on the Gaussian Process iWare-E model. The associated uncertainty of the predictions are shown in the green heatmaps below.

Lecture Summary

1. Ecosystems are put in peril and endangered species are driven to extinction by illegal wildlife hunting and the limited resources to enforce wildlife conservation

2. PAWS is a machine learning pipeline created as a data-driven strategy to identify regions in protected areas that are at high risk of poaching and compute the best patrol routes to prevent poaching

3. The use of Gaussian processes as weak learners improves the existing iWare-E model in PAWS by modelling the uncertainty of the predictions of detecting poaching attacks

Recommended Reading

Xu, L., Gholami, S., McCarthy, S., Dilkina, B., Plumptre, A., Tambe, M., . . . Enyel, E. (2020). Stay ahead of poachers: Illegal wildlife poaching prediction and patrol planning under uncertainty with field test evaluations (short version). In 2020 IEEE 36th international conference on data engineering (ICDE). IEEE.

https://arxiv.org/abs/1711.06323

References

Yang, R., Ford, B., Tambe, M., & Lemieux, A. (2014). Adaptive resource allocation for wildlife protection against illegal poachers. In Proceedings of the 2014 international conference on autonomous agents and multi-agent systems (pp. 453–460). Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

https://dl.acm.org/doi/10.5555/2615731.2615805