

Machine Learning And Neural Networks (CM3015)

Course Notes

Felipe Balbi

April 20, 2021

Contents

Week 1	3
1.101 Applications of Machine Learning	3
1.102 Types of ML	4
Week 2	5
1.202 Further reading	5
Week 3	6
2.101 Introduction to supervised Learning	6
2.201 K-Nearest Neighbours Classification	6
2.301 Decision tree	6
Week 4	8
2.501 Classifier evaluation	8
2.602 Further reading	9

Week 1

Key Concepts

- explain the concepts of clustering and dimensionality reduction
- Describe various types of machine learning problem
- Describe various applications of machine learning

1.101 Applications of Machine Learning

Machine Learning is a branch of artificial intelligence that enables machines to learn by example. Carried by the increase in data availability and computational power, we can already experiences applications of machine learning in our everyday lives: mobile phones, personal assistants, language translators, etc.

One application of machine learning are the e-passport gates at some airports which rely on face recognition to identify passengers with high probability.

Computer Vision systems can also be used to detect and classify human posture and facial expressions. Machine Learning can also be applied to other types of data such as text (handwriting recognition) or audio (speech recognition).

These systems collect and process vast amounts of data and the issue of privacy arises. We must be conscious about what data has been recorded, who has access to it, and how it can be used.

Autonomous Vehicles are a focus in machine learning research. They pose interesting and complex challenges both technically and ethically. Vehicles need to be able to detect and avoid pedestrians and other objects on the road. In the case of the accident, who's to blame? The owner of the vehicle? The company who made the car? The software engineers who built the system?

Another common system in our daily lives are recommender systems. We encounter them in streaming services, online shopping experiences, MOOC education providers, and many more. The main goal of these systems is to recommend other items similar to what we have already *consumed*. Because these systems are also used to suggest similar content to what we already watch, they may end up skewing our view of the world.

Generative Machine Learning are models that can generate new data based on a sample, for example given a sample of someone's handwriting, we want to produce more text in the same style.

Another application is related to Sensor-based Activity Recognition. Here the goal is to detect what activity the user is executing (sitting, walking, running, playing football) based on the data from sensors the user's wearing.

1.102 Types of ML

Machine Learning is used when we want to learn from data rather than hardcode a solution. There are two types of machine learning

Supervised Learning in supervised learning, the label y is associated with every sample x . We're trying to learning mapping from x to y

Unsupervised Learning here the goal is usually about clustering data in subgroups. For example, given a dataset containing pictures of animals, separate the images by animal.

We can use the decision tree depicted in figure 1 to decide which type of Machine Learning application to apply:

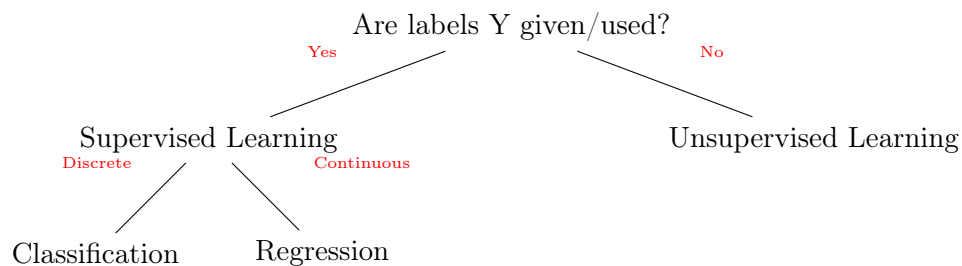


Figure 1: Decision Tree

One final type is Reinforcement Learning where the goal is to learn a sequence of actions that entail some reward. This can be used to teach a machine how to play a specific video game.

Week 2

Key Concepts

- explain the concepts of clustering and dimensionality reduction
- Describe various types of machine learning problem
- Describe various applications of machine learning

1.202 Further reading

- Chapter 1, sections 1.1 to 1.2 of the course textbook (Chollet).
- Chapter 1 of Alpaydin, E. Introduction to machine learning. (Cambridge, MA: MIT Press, 2014) 3rd edition [ISBN 9780262028189].
- Chapter 1, Introduction, up to and including section 1.3, of the following textbook gives a good introduction to the topic of ML: Murphy, K. Machine learning: a probabilistic perspective. (Cambridge, MA: MIT Press, 2012) [ISBN 9780262018029]

Week 3

Key Concepts

- Explain how a simple nearest neighbour algorithm works
- Describe the Decision Tree Classifier
- Evaluate a supervised classification algorithm on a dataset

2.101 Introduction to supervised Learning

Classification is a type of supervised learning where the labels on a data are discrete and categorical.

2.201 K-Nearest Neighbours Classification

The k-NN algorithm works on the premise that things are similar if they are closer together. Essentially, we measure the distance from a *test* sample X^* to every sample in the *training* set X , in other words, we compute $(X^* - X_i)^2$.

We can employ Euclidean distance $(\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2})$ or the Manhattan distance $(|(x_2 - x_1)| + |(y_2 - y_1)|)$ for this computation.

The two parameters for the algorithm are the number k and the distance algorithm used.

K-Nearest Neighbours is known as a lazy learning algorithm, which means that we don't generalize on the training dataset until we want to make a query.

2.301 Decision tree

Decision Tree is a very versatile machine learning algorithm, because they are capable of handling both classification and regressions tasks. They can also handle non-linear datasets.

This algorithm can be used as the basic classifier in Random Forests, which is among the most powerful class of machine learning algorithm.

To illustrate how Decision Tree algorithm, we will look at how we can classify Hares vs Rabbits. Figure 1 shows an example of how the decision tree could look like. When applying the algorithm to a new data to be classified, we descend through the tree until we get to a leaf node.

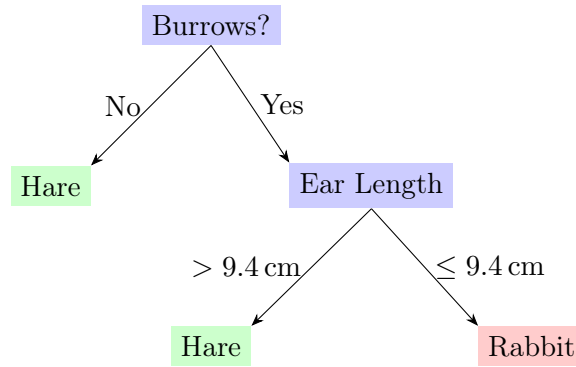


Figure 2: Decision Tree

Decision Trees are referred to as *White Box* models, this means they are easy to interpret, because of the hierarchical nature of the classification rules which is easy to visualize. Conversely, *Black Box* models make decisions in a more opaque process.

There are several types of Decision Tree algorithms, one of which is known as Classification And Regression Tree, or CART. It's a binary tree which can be used, as the name suggests, for both classification and regression.

A decision tree is prone to overfitting, which is when the model fits the training dataset perfectly; this makes it harder to generalize the trained model to other sets of data. Finding the optimal split and feature combination is an NP-complete problem.

Week 4

Key Concepts

- Explain how a simple nearest neighbour algorithm works
- Describe the Decision Tree Classifier
- Evaluate a supervised classification algorithm on a dataset

2.501 Classifier evaluation

One way of evaluating the performance of a classifier is to measure its accuracy, this, however, is not always a good measure of the quality of a classifier.

Another approach to measuring quality of the classifier is to employ a Confusion Matrix. This matrix lets us compare a true condition vs a predicted condition, Like shown in table 1.

Table 1: Confusion Matrix

	Condition Positive	Condition Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

The *True Positive Rate*, also known as Recall or Sensitivity, tells us how likely the model is to predict the correct value. It's computed as shown below:

$$TPR = \frac{TP}{TP + FN}$$

The *Precision* or *Positive Predictive Value* tells us how likely the prediction is to be correct, given a positive prediction. It's computed as shown below:

$$PPV = \frac{TP}{TP + FP}$$

The *True Negative Rate*, also known as Specificity or Selectivity, is computed as:

$$TNR = \frac{TN}{TN + FP}$$

The *False Negative Rate*, also known as Miss Rate, is computed as:

$$FNR = \frac{FN}{FN + TP}$$

The *False Positive Rate*, also known as Fall-out is computed as:

$$FPR = \frac{FP}{FP + TN}$$

2.602 Further reading

- <http://scikit-learn.org/stable/modules/neighbors.html>
- <https://scikit-learn.org/stable/modules/tree.html>
- Chapter 1, section 1.2 of the course textbook (Chollet), also briefly mentions decision tree classifiers.
- Sections 2.1 and 2.5 of Ethem Alpaydin's book provide a good overview of supervised classification: Alpaydin, E. Introduction to machine learning. (Cambridge, MA: MIT Press, 2014) 3rd edition [ISBN 9780262028189].
- Alpaydin's book also discusses decision trees in depth in Chapter 9, sections 9.1 to 9.3.