tsg insights

November 2021

Testing of artificial intelligence (AI) and machine learning (ML) – unsupervised learning

This whitepaper is the fourth in a series which act as companion pieces:

- An overview of artificial intelligence (AI) and machine learning (ML).
- Introduction to the testing of AI and ML.
- Testing of AI (artificial intelligence) and ML (machine learning) – supervised learning.
- Testing of AI (artificial intelligence) and ML (machine learning) – unsupervised learning.
- Testing of AI (artificial intelligence) and ML (machine learning) – reinforcement learning.
- AI and machine learning – algorithmic bias – the cruel mirror AI and ML reflects back at us.

In this paper, we look at the approaches to and challenges regarding unsupervised learning.

What is unsupervised learning?

Unsupervised learning is where a machine learns for itself. It has to identify naturally occurring patterns within data sets on its own – there are no pre-assigned labels or scores to help it. The conclusions it draws are based on statistical methods such as clustering and modelling.

Main advantages:

- While being fundamentally backwards-looking, it can result in new and unexpected conclusions in the data and hidden structures being revealed.
- Unlabelled data is faster and easier to obtain than the labelled variety and is quicker to input as well. Yet it can still be used to undertake complex tasks.
- It reduces the likelihood of mistakes – labelling takes, training, time and effort, and can suffer from human error.
- It's useful for testing the efficacy of AI in a given situation.

Main disadvantages:

- It generally results in less precise and slower training compared with supervised learning.
- Its very unpredictability can mean the creation of unwanted categories and conclusions, resulting in confusion and error.
- There's a risk of learned data bias, which can become normalised over time – it can take a while for errors and unwanted categories and results to come to light.
- This means that while less time may be spent on input, more may have to be spent on interpretating the results.

Sometimes, the line between supervised and unsupervised learning is blurred by semi-supervised learning. With this, a part of the given input data has been labelled. This can reduce unpredictability, but at the expense of greater cost and time input.

How unsupervised learning works

Unsupervised learning is most often focused on clustering. This is where groups of objects or data points are similar to each other, but dissimilar to objects or data points in other clusters.

For clustering, various algorithms are used, which are categorised according to how they work. For example, exclusive, overlapping, hierarchical, density-based, and probabilistic clustering.

One of the most popular and straightforward algorithms used in unsupervised learning is k-means clustering. This is the algorithm that defines which features are present in the dataset and then groups data with common elements into clusters. This makes it essential for data mining, particularly across documents and images. There's also a subset of this, fuzzy k-means clustering, which allows data points to apply to more than one cluster.

The Hidden Markov model

This is a more specialised unsupervised learning approach which makes use of an elaborate ML algorithm to analyse and group data.

Real-life applications include:

- Optical character and handwriting recognition – used for validating documents and for wider security features.
- Speech recognition and synthesis – this is an important part of conversational user interfaces and features in virtual assistant technology such as Siri.
- Text classification – often featuring parts-of-speech tagging, this is used in opinion mining, and across medical documents.
- Text Translation – employed by multi-national organisations and governments, as well as in Google Translate and other language tools. It's also increasingly used across social media.
- Data analytics – this includes specialised applications such as cryptanalysis and time series analysis, as well as the more broadly-based computational finance.