**Four Types of Machine Learning Algorithms Explained**

Machine learning is increasingly becoming more important to the everyday function of the modern world. [Machine learning](https://www.seldon.io/what-is-machine-learning/) algorithms are behind a range of technologies, whether providing predictive analytics to businesses or powering the decision-making of driverless cars. There are distinct approaches to machine learning which change how these systems learn from data.

In general, there are four main types of machine learning algorithms. Each one takes a different approach to how a machine will learn from data and are positioned to solve unique problems and challenges. They also differ on the level of input required from developers and the condition of the datasets they learn from.

The four common machine learning algorithm types are:

1. Supervised machine learning algorithms
2. Unsupervised machine learning algorithms
3. Semi-supervised machine learning algorithms
4. Reinforcement machine learning algorithms

This guide will explore and explain the different types of machine learning algorithms, how they differ, and what they’re used for.

**Supervised machine learning algorithms**

Supervised machine learning algorithms are reliant on accurately labelled data and oversight from a developer or programmer. The algorithm is fed data which includes input and desired output defined by the developer. The system then learns from the relationship between the input and output training data to build the model. The model maps input data to the desired output and is trained until the model reaches a high level of accuracy.

This type of machine learning algorithm requires direct supervision from a developer. Training data must be accurately labelled and the algorithm boundaries should be set too. All supervised machine learning algorithms are trained exclusively on labelled data.

The subsequent models can be used to predict outcomes of future data and trends. They can also be used to classify new data against rules set by analysing the training data.

**How are supervised machine learning algorithms used?**

Supervised machine learning algorithms are generally used to either categorise data against a model or predict continuous outcomes of new data. In the first instance, the algorithm will be trained to identify and categorise objects using training data. The system learns how to understand the relationship between data points from the training data. Once deployed, new data can then be fed into the model to be classified.

The second common usage is in the prediction of continuous outcomes. A model is trained to identify patterns and trends in a training dataset and can then apply this to new live data once deployed.

Supervised machine learning algorithms can be used to:

* Classify new data into established groups and categories.
* Forecast emerging and future trends based on predictive models.
* Help businesses in campaign and acquisition projects by forecasting changes in value.

**Classifying data**

Supervised machine learning algorithms will often be trained to classify datasets.  The models will be trained on labelled datasets on how to recognise objects and their classifications. Models can be trained to classify a range of data types, such as images, text or audio. The process is supervised, as the parameters of each classification must be set by the developer. The process is known as classification.

Once the model is trained, it will be able to recognise and classify new data and objects. The model can be used to identify specific subjects within images for example.

**Predicting continuous outcomes**

Another common use of supervised machine learning algorithms is the prediction of outcomes. The model is trained to identify patterns within a training dataset, which may relate to their values or label groupings. Once the model understands the relationship between each label and the expected outcomes, new data can be fed into it when deployed. It can then be used to make calculated predictions from the data, for example identifying seasonal changes in sales data. The process is known as regression.

These types of machine learning algorithms are key elements of predictive analytics tools. Regression machine learning use cases may include:

* Price prediction models to project retail sales or stock trading outcomes.
* Predictive analytics in a variety of sectors such as education or healthcare.
* Marketing and advertising campaign planning, to forecast the value of prospective ad space.

**Unsupervised machine learning algorithms**

Unsupervised machine learning algorithms are not directly controlled by a developer and will be trained on datasets with no labels. Unsupervised machine learning algorithms are used to identify patterns, trends or grouping in a dataset where these elements are unknown. This type of machine learning can identify the relationship between different data points and be used to segment similar data.

Models will be fed huge datasets to understand the underlying patterns and structure of the data. It can be used to identify trends or to categorise unlabelled datasets. New and unforeseen trends can be discovered using this technique, as the algorithm detects patterns within the dataset with no direct human intervention.

Common uses include the detection of rules that govern unlabelled data, and the segmentation of data into groups. Examples of use include automatic customer segmentation in the sales and marketing sectors.

**How are unsupervised machine learning algorithms used?**

Unsupervised machine learning algorithms are used to gain insight into unlabelled and unknown datasets. A vast amount of data held by organisations and businesses is unlabelled. In some cases there is a lack of resources to clean and label this data correctly. Labelling data can take a huge amount of resources depending on the file type, for example labelling images or rich text can be intensive. Unsupervised machine learning algorithms give businesses insight into these unmapped datasets

The technique is used in many data management platforms as a way to automatically group audiences and customers. It can also be used to identify trends in user data for more efficient marketing campaigns or personalised content. Trend analysis is driven by the data itself instead of a supervising developer.

Unsupervised machine learning algorithms are used to:

* Segment audience data for more personalised marketing campaigns.
* Highlight unknown customer trends to finetune services and products.
* Deliver personalised content based on interests on online music and film streaming services.
* Automate systems for recommending items to users on ecommerce and retail websites.
* Automatically improve efficiency of systems and processes.

**Clustering datasets**

One of the main uses of unsupervised machine learning algorithms is making sense of unlabelled data. The algorithm will cluster or segment data into categories depending on the relationship between each data point. Unlike supervised machine learning algorithms which require labelled training data, unsupervised algorithms will segment data based on trends it picks up from the unlabelled data.

The algorithm will identify the similarities and differences between each data point then map the dataset into segments. This process is useful for highlighting unseen trends in large, unlabelled datasets. Common usage of this technique is the automatic segmentation of audience or customer data in digital marketing and sales environments.

One of the most common techniques to cluster datasets through unsupervised machine learning algorithms is K-means Clustering. It works by assigning data to the nearest cluster point. The target number of clusters can be defined by the user.

**Data visualisation**

Data visualisation models created from unsupervised machine learning algorithms can create charts, diagrams and graphs from unlabelled data. The process can take complex and unlabelled data sets and quickly plots visualisations to provide insight. Data visualisation overlaps with clustering, as the technique visualises the different clusters of data plotted across two or three dimensions. Data visualisation makes observing and understanding the grouping of complex data more straightforward.

An example of a machine learning technique for visualisation is t-Stochastic Neighbor Embedding (t-SNE). The method models similar and dissimilar data points across two or three dimensions, helping to visualise the distribution of data clusters.

**Reducing dimensionality**

Reducing the dimensions of a sample of unlabelled data will help to refine the groups and clusters. By reducing the number of variables in the model, the data trends are simplified and the overall processing can be more efficient. This technique will be used in instances where too many dimensions are clouding the resulting insights. It will often simplify the data, improving performance and speed of analysis.

One of the most-used techniques for dimensionality reduction is Principal Component Analysis (PCA). It works by establishing the principal components which govern the relationship between each data point, before simplifying to use only the main principal components. The technique maintains the variety of data grouping but streamlines the number of separate groups.

**Semi-supervised machine learning**

As the name suggests, semi-supervised machine learning is a blend of supervised and unsupervised approaches. It combines elements of both types of machine learning algorithms. It is used with datasets that have only a portion of data accurately labelled.

Semi-supervised machine learning algorithms are trained on the subset of correctly labelled data. The model then uses this training to label the remaining unlabelled data in the sample.

Semi-supervised machine learning uses the classification process from supervised machine learning to understand the desired relationships between data points. It then uses the clustering process from other unsupervised machine learning algorithms to group the remaining unlabelled data.

**How are semi-supervised machine learning algorithms used?**

Semi-supervised machine learning algorithms are used to process and understand data which is only partially labelled. The presence of unlabelled data is common in large or complex datasets. The labelling of such a large dataset may be too resource-intensive or difficult to be done manually by data analysts. For example, the labelling of large text documents would be incredibly labour-intensive if done by a human.

The technique is often used in image analysis, with the model trained on a subset of clearly labelled images. The model can then cluster unlabelled images along the parameters of the learnt rules.

Usage of semi-supervised machine learning algorithms include:

* Processing and categorising audio or image files when a sample of accurately transcribed or labelled data is available.
* As a cost-effective way of categorising data which would be resource-intensive if labelled by human data specialists.
* Grouping large text documents such as scanned books and documents.

**Reinforcement machine learning**

Reinforcement machine learning allows a system to learn and improve the performance of a function through trial and error. The model will find the best solution to a problem in a specific environment by learning from past actions. The process is a feedback loop in which successful actions are rewarded and reinforced. Training generally consists of a system performing an action in a specific environment whilst receiving continuous feedback. Reward signals are released when the action is successful.

The technique iteratively improves the algorithm through positive and negative reward signals. A successful action will receive positive reward signals, whereas a failed action will cause a negative reward signal. The feedback loop will become more complex as the complexity of the task increases. Reinforcement machine learning algorithms are deployed when an action is too complex for a static human-written algorithm. This could be because the challenges faced by the system are too fluid or unpredictable.

The development of driverless cars is a well-known example of reinforcement machine learning. The system learns from interacting with the environment to decipher the best course of action in a given scenario. It’s similar to how an intelligent being will learn from interacting with its environment and learning from past experiences. The idea is for a system to train itself once the parameters of the action are defined.

A popular example of a reinforcement machine learning model is the Markov Decision Process (MDP). The technique uses reward and decision processes to ensure a model takes the optimal action in its current state.

**How are reinforcement machine learning algorithms used?**

Reinforcement machine learning algorithms are used when systems are required to perform complex actions relevant to a specific scenario. The situation or environment may be random or hard to predict. The model can form its approach to a problem or process itself and in a flexible way.

It’s a useful technique when the approach to a problem needs to be reactive or flexible. For example, when a static algorithm written by a human developer would not cover all the variables of a situation. Reinforcement models are reactive to incoming data, so can make decisions based on a changing environment.

Examples of where reinforcement machine learning is used includes:

* Non-player characters in video games are trained to react to human player interactions.
* Development of artificial intelligence.
* The systems behind driverless cars.
* Improving natural language processing for chatbots and virtual assistants.
* Teaching a model to play a game like chess against a human opponent.
* Incentivising a system to make efficiency savings through a reward system.

s

# References

Seldon. (2021, June 28). *Four Types of Machine Learning Algorithms Explained*. Retrieved from Four Types of Machine Learning Algorithms Explained: https://www.seldon.io/four-types-of-machine-learning-algorithms-explained