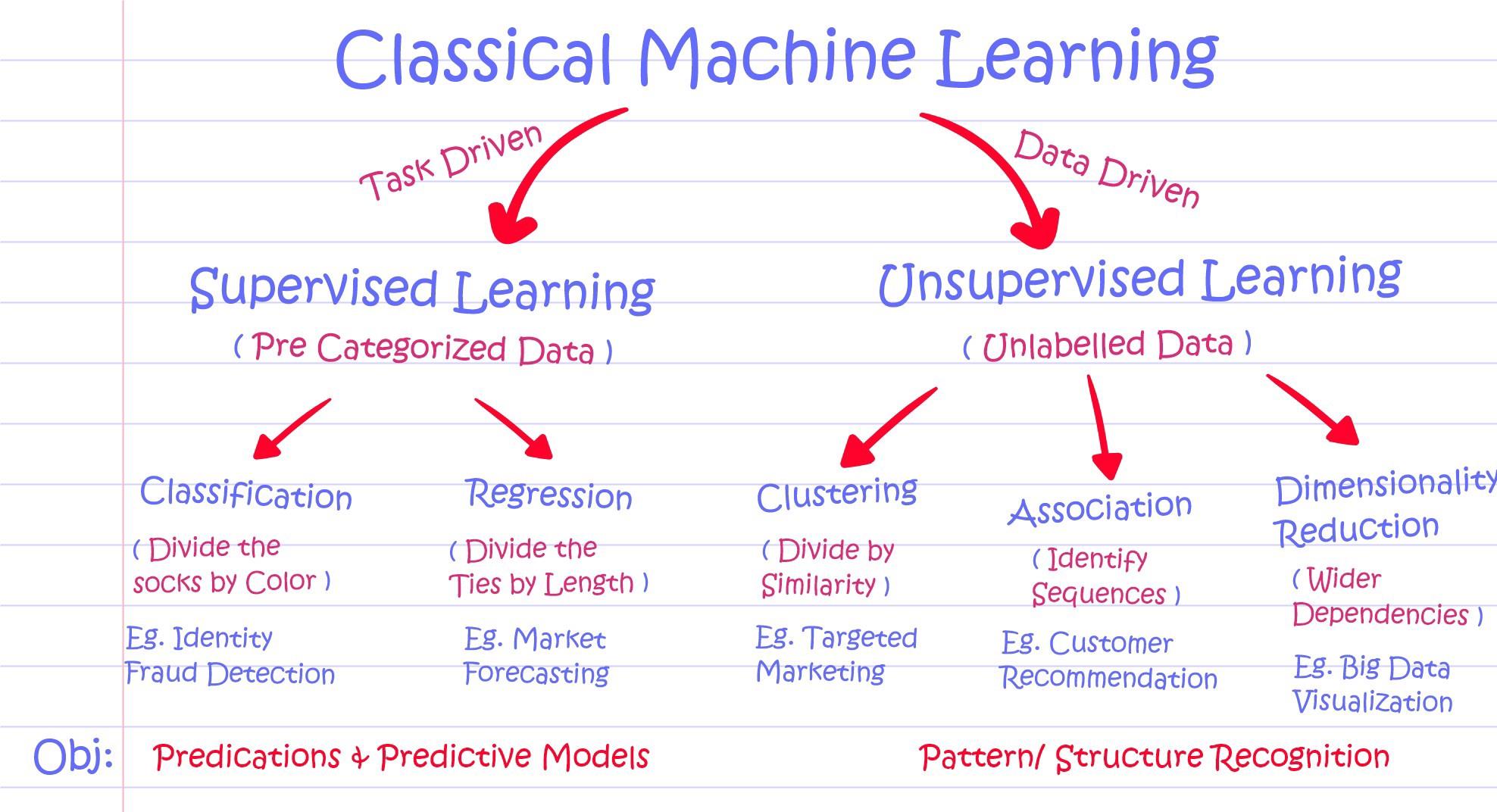
# What is Artificial Intelligence

Artificial intelligence is the branch of engineering that leverages technology in performing complex tasks normally requiring human intelligence. Tasks such as visual perception, speech recognition and informed decision-making are activities requiring high levels of human intelligence. The use of computer programmes and algorithms is progressively making this mammoth task a reality,

# Machine Learning

Machine learning is a way of improving accuracy in a computers ability to imitate the way humans learn and perform activities. Through computer programmes and algorithms, a computer makes sense of the data inputs by performing pattern recognitions and trend analysis, it does this in several ways:



**Supervised Learning**

is one method that relies on labelled input data to learn a function that produces an appropriate output when given new unlabelled data. This type of Machine Learning occurs when the training comes from a set of data with pre-defined solutions. The Machine uses these desired outcomes to make decisions and draw conclusions about new data by recognising features and assigning labels.

**Unsupervised Learning**

This type of machine learning attempts to make sense of a vast array of data without any labels. It learns from a training data set that does not include desired solutions. One way this happens is through clustering and association, in other words the algorithm attempts to group data items based on similarities.

**Semi Supervised**

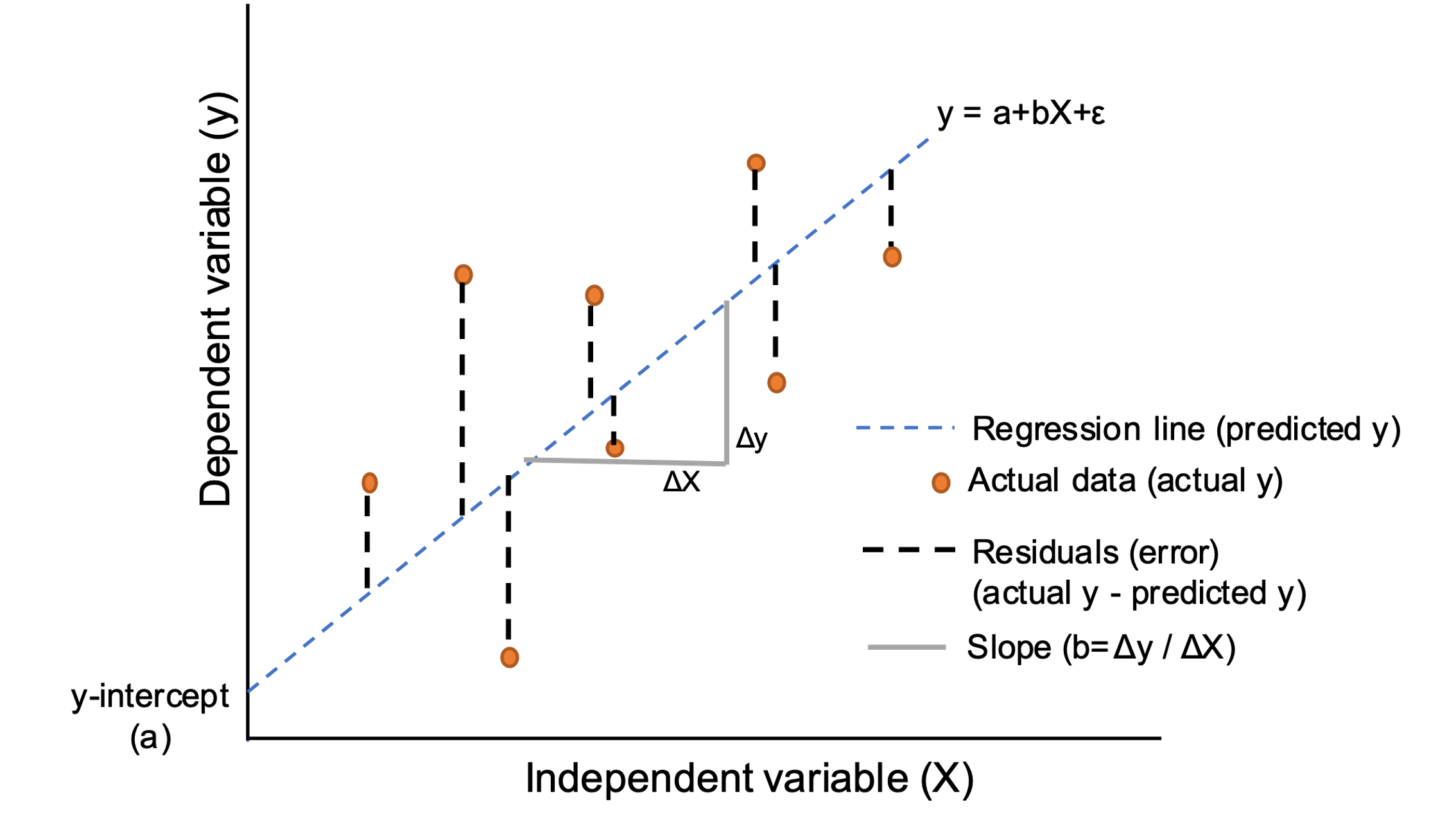
Utilises both labelled and unlabelled data. A small amount of labelled data is used to train the algorithm and a large amount of unlabelled data is cross referenced with desirable outcomes. It attempts to find trends, patterns, groups, and then tries to label them.

**Reinforcement Learning**

The aim of reinforcement learning is to maximise reward in each scenario, the premise is that there’s an agent and a reward with hurdles in between. The agent can observe/interact with an environment, select, and perform actions which in return are rewarded or penalised, the feedback from performing these actions is used to build on the experience of the agent to get progressively better at a given task. The learning happens through experience, which enables the agent to find its own strategy, which is called a policy, to maximise reward over time.

**Linear Regression**

A type of machine learning algorithm based on supervised learning. It’s a predictive modelling technique that attempts to find a relationship between an Input and an Output. A model is trained to predict the behaviour of data based on some variables. The data is modelled using a straight-line and is appropriate for analysing continuous values.

Used in trend forecasting, for example, can be used to study the impact on sales as a result of increasing the price of a product. Consumer behaviour can be modelled from this information therefore future pricing decisions can be made using this insight.

**Logistic Regression**

Is a supervised classification algorithm, the target variable/ output can only take discrete values for any given set of features/inputs. This algorithm can be used in the medical field to classify the progression of cancer i.e. is a tumour malignant or benign.

Chart, line chart

Description automatically generated

**Decision Tree**

Is a supervised learning algorithm that can be used for solving regression and classification problems by learning a few simple decision rules. It depends on building a model using a training data set and is then able to predict the class or value of a desired variable by using these decision rules on new data making it a Supervised learning algorithm. It graphically represents all the possible outcomes for the decision.

Diagram

Description automatically generated

**Random Forest**

Is a supervised learning algorithm that works for Regression and Classification. Random Forest consists of many decision trees. In regression, the average prediction of each individual tree is returned.

Radar chart

Description automatically generated

**SVM (Support Vector Machine)**

Supervised machine learning algorithm that can be used for classification, regression, and outliers’ detection. It transforms the data using the Kernel trick and then finds optimal boundaries based on these transformations

**Naive Bayes**

Supervised machine learning algorithm based on bayes theorem and is used for solving classification problems. Particularly used on large volumes of data and is rooted in conditional probability. By definition conditional probability is the likelihood of an event occurring given that another event has already happened. Therefore, prior knowledge is an important step making this a supervised learning algorithm. It is a very fast and highly extensible algorithm that can be used for both binary and multiclass classification.

Applications include spam email classification, real-time Prediction being a fast learning

algorithm can also be used to make predictions in real-time.

**KNN (K- Nearest Neighbours)**

Supervised machine learning algorithm that can be used for classification and regression problems. KNN works by finding the distances between a query and all the examples in the data, selecting the specified number of examples (**K**) closest to the query, then votes for the most frequent label (in the case of classification) or averages the labels (in the case of regression)

Advantages of KNN lie in its simplicity to implement, Flexibility to feature/distance choice, naturally handles multi-class cases and can do well in practice with enough representative data

The downsides however include the need to determine the value of parameter K (number of nearest neighbours) which is often trial and error. Also, the Computation cost can be quite high as a result of computing the distance of each query instance to all training samples.

**K-Means**

Unsupervised learning algorithm that works through clustering, it relies on grouping similar data sets together to discover underlying patterns. Clustering refers to a collection of data points aggregated together because of certain similarities

Chart, scatter chart

Description automatically generated