Documented report

Machine learning is a field of artificial intelligence that employs algorithms to learn from data make a prediction. These algorithms fall in to two categories being supervised and unsupervised algorithm or methods.

There are five main types of machine learning algorithms that are employed that fall under the category of supervised and unsupervised;

Supervised category: (machines are taught by examples)

* Classification algorithm, an example here is the spam filter in the email box
* Regression algorithm; liner regression, e.g., predicting a salary based on work experience
* Forecasting, weather forecasting is an example

Unsupervised category: (machines studies data to identify pattern)

* Clustering: this involves grouping sets of similar data, based on a specific criterion, it helps to segment data into groups and perform analysis on each dataset to find patterns
* Dimension reduction: this reduces the number of variables considered in order to find the exact information wanted

Reinforcement learning: an algorithm that allows AI to learn in a similar trial and error method like humans to learn and achieve the desired result.

Time series is a machine learning technique that forecasts target value based solely on a known history of target values. It is a specialized form of regression, known in the literature as auto-regressive modelling. The input to time series analysis is a sequence of target values.

Deep learning algorithm: a type of ml algorithm that uses artificial neural networks to perform sophisticated computations on large amount of data, working based on the structure and function of the human body.

* CNNs = convolutional neural network – processes structured grid data like images, e.g., image classification, object detection and facial recognition tasks
* Recurrent neural network – design to track or identify patterns in data sequences
* Long short-term memory networks, e.g., speech recognition, time series.

Common machine learning algorithms that are employed

* **Naïve Bayes Classifier Algorithm (Supervised Learning - Classification)**  
  The Naïve Bayes classifier is based on Bayes’ theorem and classifies every value as independent of any other value. It allows us to predict a class/category, based on a given set of features, using probability.  
    
  Despite its simplicity, the classifier does surprisingly well and is often used due to the fact it outperforms more sophisticated classification methods.
* **K Means Clustering Algorithm (Unsupervised Learning - Clustering)**  
  The K Means Clustering algorithm is a type of unsupervised learning, which is used to categorise unlabelled data, i.e., data without defined categories or groups. The algorithm works by finding groups within the data, with the number of groups represented by the variable K. It then works iteratively to assign each data point to one of K groups based on the features provided.
* **Support Vector Machine Algorithm (Supervised Learning - Classification)**  
  Support Vector Machine algorithms are supervised learning models that analyse data used for classification and regression analysis. They essentially filter data into categories, which is achieved by providing a set of training examples, each set marked as belonging to one or the other of the two categories. The algorithm then works to build a model that assigns new values to one category or the other.
* **Linear Regression (Supervised Learning/Regression)**  
  Linear regression is the most basic type of regression. Simple linear regression allows us to understand the relationships between two continuous variables.
* **Logistic Regression (Supervised learning – Classification)**  
  Logistic regression focuses on estimating the probability of an event occurring based on the previous data provided. It is used to cover a binary dependent variable, that is where only two values, 0 and 1, represent outcomes.
* **Artificial Neural Networks (Reinforcement Learning)**  
  An artificial neural network (ANN) comprises ‘units’ arranged in a series of layers, each of which connects to layers on either side. ANNs are inspired by biological systems, such as the brain, and how they process information. ANNs are essentially a large number of interconnected processing elements, working in unison to solve specific problems.  
    
  ANNs also learn by example and through experience, and they are extremely useful for modelling non-linear relationships in high-dimensional data or where the relationship amongst the input variables is difficult to understand.
* **Decision Trees (Supervised Learning – Classification/Regression)**  
  A decision tree is a flow-chart-like tree structure that uses a branching method to illustrate every possible outcome of a decision. Each node within the tree represents a test on a specific variable – and each branch is the outcome of that test.
* **Random Forests (Supervised Learning – Classification/Regression)**  
  Random forests or ‘random decision forests’ is an ensemble learning method, combining multiple algorithms to generate better results for classification, regression and other tasks. Each individual classifier is weak, but when combined with others, can produce excellent results. The algorithm starts with a ‘decision tree’ (a tree-like graph or model of decisions) and an input is entered at the top. It then travels down the tree, with data being segmented into smaller and smaller sets, based on specific variables.
* **Nearest Neighbours (Supervised Learning)**  
  The K-Nearest-Neighbour algorithm estimates how likely a data point is to be a member of one group or another. It essentially looks at the data points around a single data point to determine what group it is actually in. For example, if one point is on a grid and the algorithm is trying to determine what group that data point is in (Group A or Group B, for example) it would look at the data points near it to see what group the majority of the points are in. for example, Forecasting stock market: Predict the price of a stock, based on company performance measures and economic data

Machine learning in healthcare/fitness

Machine learning can be employed in healthcare to analyse patients and predict any possibility of disease, sickness, etc. in our situation we are using machine learning algorithms to predict the hydration level of a person.

K – MEANS CLUSTERING ALGORITHMS

This type of algorithm divides a population or set of data into a number of groups so that the data in each group are comparable to one another, and different from other groups. In short words, it is the grouping of data based on how similar they are.

Algorithms used in health/fitness technology:

* Apple watches employs a brain (artificial intelligence) neural network in their smartwatches, that is able to process data and give predictions in a similar way that the human brain does, it is also a type of machine learning called deep learning, that uses interconnected nodes/neurons in a layered format that resembles the brain. In this machine learning, each neurons receive data, applies a process to them and release the outputs. What makes Apple’s ECG technology so effective is that they have worked with researchers to gather extensive data sets from patients that allow them to train deep convolutional neural networks. Data was gathered from healthy patients and patients suffering from Atrial fibrillation. The deep neural network was trained extensively with this data so that it can distinguish between healthy and unhealthy ECG patterns.
* Deep learning algorithms is one of the main, if not the most commonly used in wearables to analyse data.
* Fitbit employs six algorithms for their wearables; decision trees, Random Forest, XGBoost, k-nearest neighbour, fully-connected neural network, and LTSMs
* Some machine learning algorithms that are used in hydration monitoring are;
  + Support vector machine algorithms to track drinking
  + Decision tree for activity recognition