"How it Works"
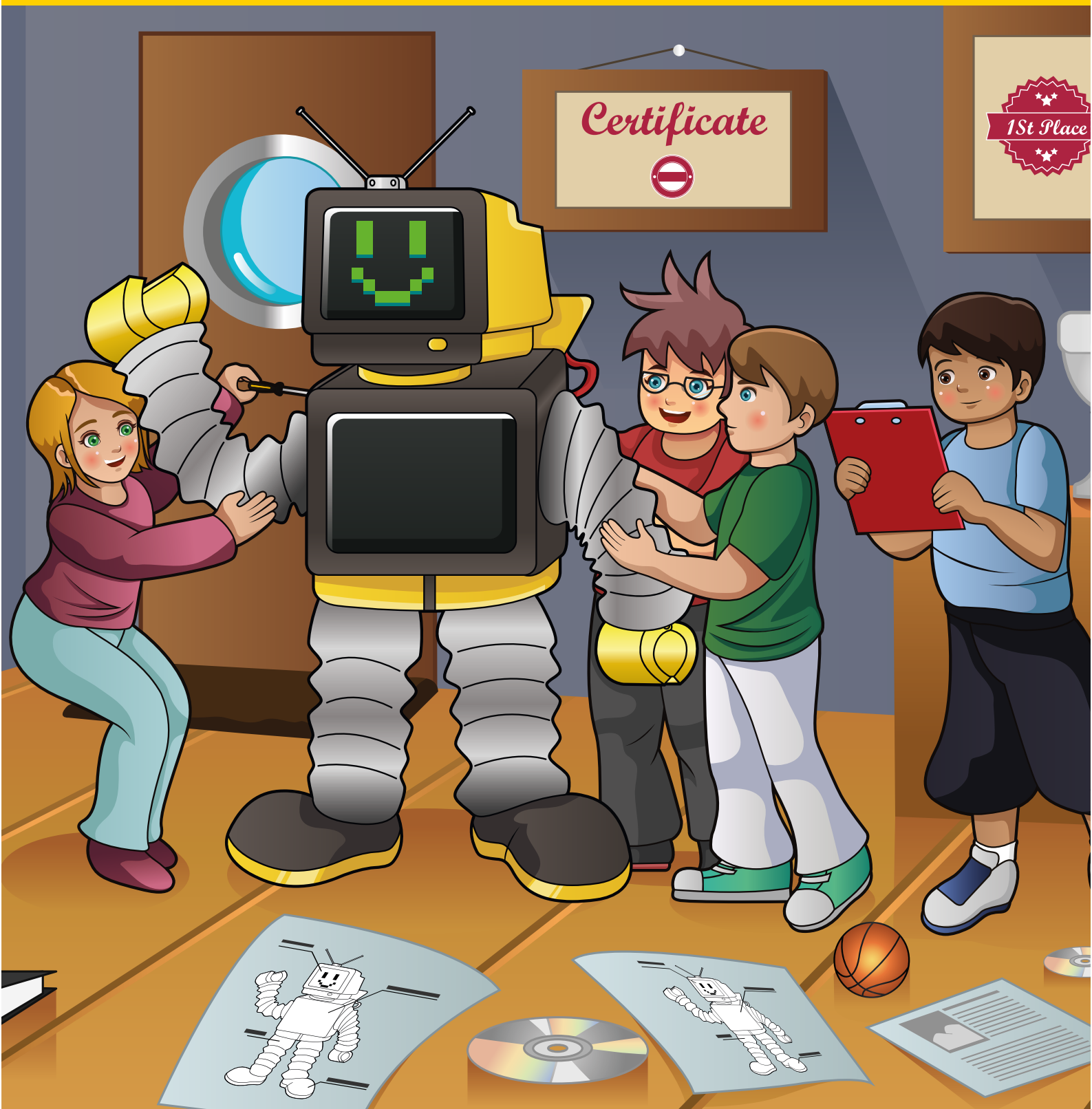# Machine Learning

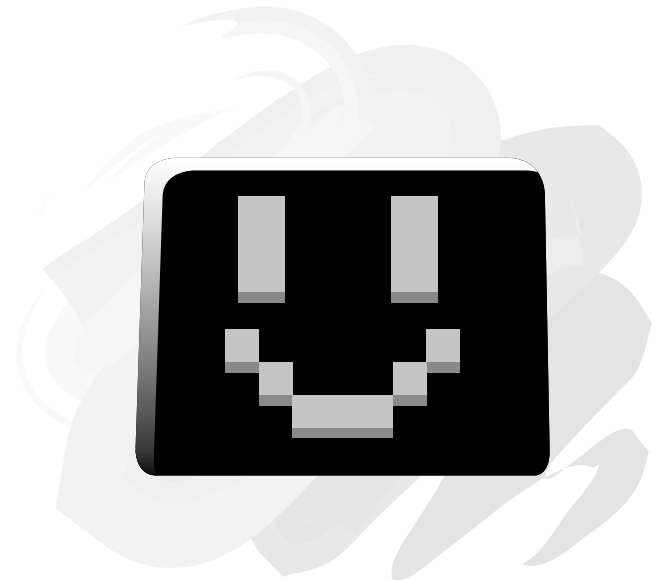# "How it Works"
Machine Learning

Adapted from a variety of sources by **Bob Yelland**

Machine learning – a branch of artificial intelligence - is changing not only how we interact with machines, but how we relate to the world around us. During the past decade, machine learning has given us self-driving cars, speech recognition, effective web search, personalised recommendations, and a vastly improved understanding of the human genome.

The term "machine learning" dates to 1959 when Arthur Samuel, an IBM researcher, defined it as "the ability (for computers) to learn without being explicitly programmed", and the field encompasses a variety of mathematical techniques where computers learn and refine their own solutions based on sample "training" data.

But how does machine learning actually work?
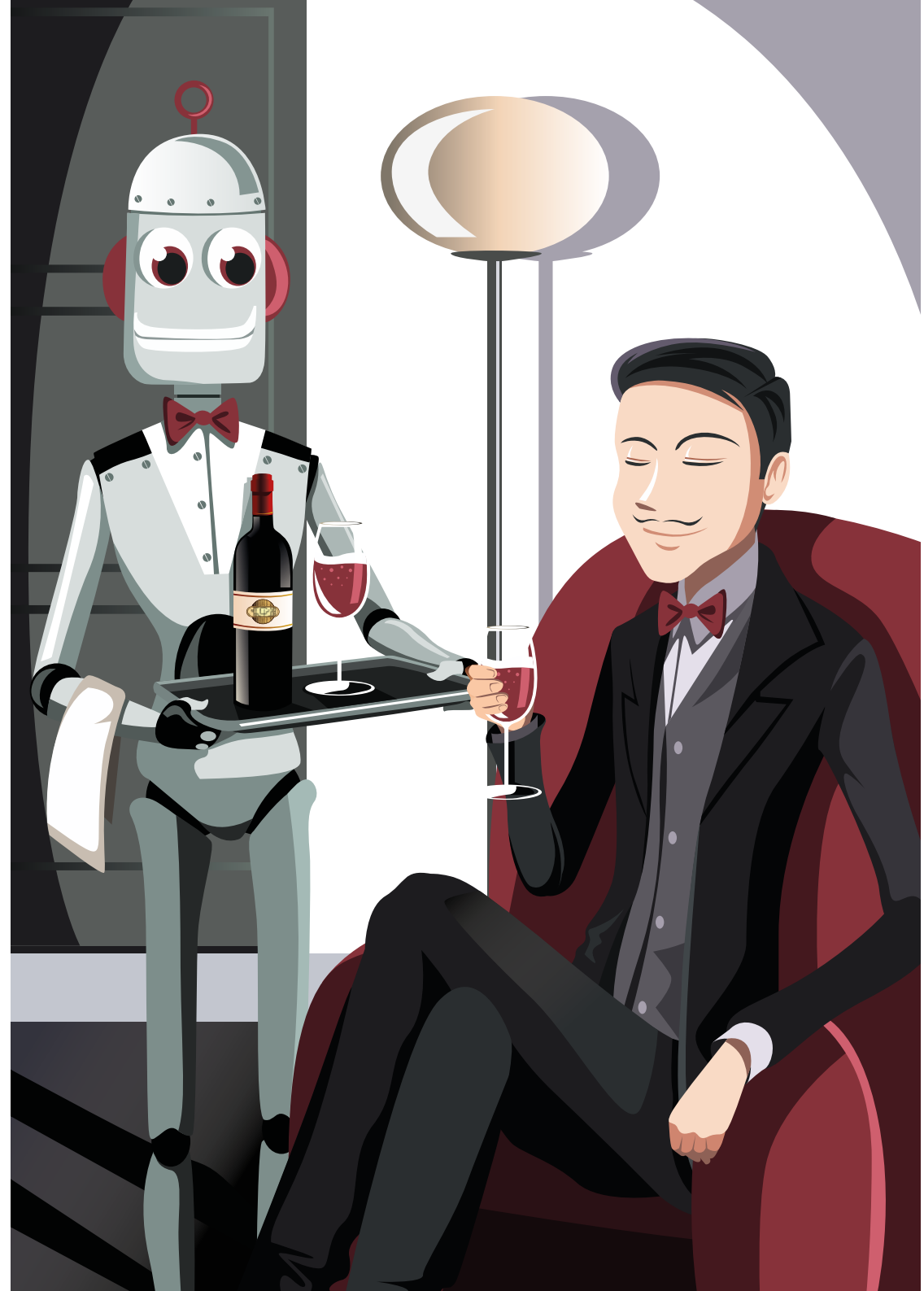
Computers were originally designed to follow algorithms. An algorithm is simply a series of steps coded in a computer language. Skilled computer programmers would liaise with process experts to map business operations into a flowchart diagram which could then be implemented as a computer program.

A flowchart explicitly positions the tasks that should be performed, in the order that they need to be executed, together with any decisions that need to be made along the way. Flowcharts are great at modelling repetitive, predictable processes where decisions are made on unambiguous data. These systems are said to be deterministic.

But not all processes follow clear, unchanging rules, and most decisions in the real world do not lead to a single unambiguous answer. Machine Learning systems are probabilistic: tasks are executed and decisions are made on incomplete information and outcomes are assigned probabilities of being correct.

Machine learning approaches are typically classified into three broad categories:

Supervised learning: The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs. The training process continues until the model achieves a desired level of accuracy on the training data.

Unsupervised learning: No explicit guidance is given to the learning algorithm, as the desired outcome may not be known, leaving the computer on its own to deduce structure or discover patterns from the input data.

Reinforcement learning: A computer program interacts with a dynamic environment in which it must perform a certain task (such as driving a vehicle or playing a game), learning through trial and error as it seeks to achieve its goal.

Machine learning is suited to problems involving classification (dividing objects into two or more classes), regression (discovering relationships between variables) and clustering (grouping objects by similar characteristics). This leads to uses such as:

Recognising patterns:

- Objects in real scenes
- Facial identities or expressions
- Spoken words

Extracting insight:

- From free-format text, audio, or video
- Spotting spam email

Discovering anomalies:

- Unusual sequences of financial transactions
- Unusual patterns of sensor readings

Making predictions:

- Future stock prices or currency exchange rates
- Which movies will a person like?

Many industries have recognised the value of machine learning technology:

Financial services: Highlighting new investment opportunities, or helping investors know when to trade. Preventing fraud by identifying clients with high-risk profiles, or spotting anomalous behaviour.

Government: Helping citizens to navigate government services. Proactively identifying citizens in need of services. Identifying fraudulent claims. Linking up multiple government services for greater insight and improved planning.

Healthcare: Analysing data from wearable devices and sensors to assess a patient's health. Spotting patterns that may lead to improved diagnoses and treatment. Discovering medical articles and research studies relevant to a patient's condition.

Retail: Recommendations based on previous purchases. Personalised shopping experiences. Improved campaign targeting. Diagnosing reasons for cart abandonment.

Oil and gas: Finding new well sites based on environmental data analysis. Predicting refinery sensor failure. Streamlining energy distribution and consumption to make it more efficient and cost-effective.

Transportation: Making routes more efficient by predicting demand patterns. Modelling the impacts of weather or equipment failure and possible solutions.

Manufacturing: Improving maintenance, repair and overhaul performance with greater predictive accuracy to the component and part level. Optimising supply chains and just-in-time production.

Media: Predicting likely consumer interest and response before significant new content is commissioned. Targeting consumers based on known likes and dislikes. Reducing piracy.

There are many mathematical techniques underpinning machine learning; the key ones are:

Linear and polynomial regression

Regression is concerned with modelling the relationship between numeric variables that is iteratively refined using a measure of error in the predictions made by the model. The basic assumption is that the output variable (a numeric value) can be expressed as a combination (weighted sum) of a set of numeric input variables.

Decision trees

These tree-like flowcharts use branching to illustrate every possible outcome of a decision. Most decision trees use binary branching (two options) based on actual values or attributes in the data. For large amounts of data, a random forest of multiple decision trees can be created, which together form a consensus decision on the output. Decision trees can be used for both classification and regression problems.

Neural networks

This concept is inspired by the way biological nervous systems, such as the brain, process information. A large number of highly interconnected processing elements work in unison to solve specific problems, usually classification or pattern-matching problems. Each neuron 'votes' on the decision outcome, this might then trigger other neurons to vote, and the votes are tallied creating a ranking of the outcomes depending on the support each has received.

Bayesian networks

These graphical structures, also known as belief networks, are used to represent knowledge about an uncertain domain. The graph is a probabilistic map of causes and effects where each node represents a random variable, while the edges between the nodes represent probabilistic dependencies. For example, "red sky at night" might lead to a 75% chance of "good weather". These conditional dependencies are often estimated by using statistical and computational methods.

Support vector clustering

The objective of clustering is to partition a data set into two or more groups to organise the data into more meaningful collections. For example, to segment customers by similar buying behaviour. This is achieved by identifying the smallest sphere that encloses the data points of one type.

Markov chains

These are mathematical systems that hop from one "state" (a situation or set of values) to another: it is assumed that future states depend only on the present state and not on the sequence of events that precede it. For example, if you made a Markov chain model of a baby's behaviour, you might include "playing", "eating", "sleeping", and "crying" as states, which together with other behaviours could form a "state space": a list of all possible states. In addition, a Markov chain tells you the probability of hopping, or "transitioning", from one state to any other state – e.g., the chance that a baby currently playing will fall asleep in the next five minutes without crying first.

The plethora of algorithms available for machine learning has restricted this technology to those companies with the requisite expertise to select the right tool for the project in hand. For machine learning to gain wider adoption, these technologies need to be simplified and delivered as a service. This is our aim with the IBM Watson Machine Learning Service.

Built on Apache Spark, Watson Machine Learning intelligently and automatically builds models using open machine learning libraries and the most comprehensive set of algorithms in the industry. Its patented Cognitive Assistance for Data Science technology scores each machine-learning algorithm against the data provided to recommend the best match for the need. IBM Watson Machine Learning Service can be accessed through the Watson Data Platform, as an API on IBM Bluemix or on z/OS.

Why not explore Machine Learning further with IBM?

IBM.