Detailed Contents

Scan for *'Appendix 1 - Python Language Fundamentals'*

Scan for *'Appendix 2 - Python Packages'*

Scan for *'Appendix 3 - Lab Manual with 25 Exercises'*

Chapter 1

Introduction to Machine Learning

"Computers are able to see, hear and learn. Welcome to the future."
— Dave Waters

Machine Learning (ML) is a promising and flourishing field. It can enable top management of an organization to extract the knowledge from the data stored in various archives of the business organizations to facilitate decision making. Such decisions can be useful for organizations to design new products, improve business processes, and to develop decision support systems.

Learning Objectives

- Explore the basics of machine learning
- Introduce types of machine learning
- Provide an overview of machine learning tasks
- State the components of the machine learning algorithm
- Explore the machine learning process
- Survey some machine learning applications

1.1 NEED FOR MACHINE LEARNING

Business organizations use huge amount of data for their daily activities. Earlier, the full potential of this data was not utilized due to two reasons. One reason was data being scattered across different archive systems and organizations not being able to integrate these sources fully. Secondly, the lack of awareness about software tools that could help to unearth the useful information from data. Not anymore! Business organizations have now started to use the latest technology, machine learning, for this purpose.

Machine learning has become so popular because of three reasons:

1. High volume of available data to manage: Big companies such as Facebook, Twitter, and YouTube generate huge amount of data that grows at a phenomenal rate. It is estimated that the data approximately gets doubled every year.

2. Second reason is that the cost of storage has reduced. The hardware cost has also dropped. Therefore, it is easier now to capture, process, store, distribute, and transmit the digital information.
3. Third reason for popularity of machine learning is the availability of complex algorithms now. Especially with the advent of deep learning, many algorithms are available for machine learning.

With the popularity and ready adaption of machine learning by business organizations, it has become a dominant technology trend now. Before starting the machine learning journey, let us establish these terms - data, information, knowledge, intelligence, and wisdom. A knowledge pyramid is shown in Figure 1.1.

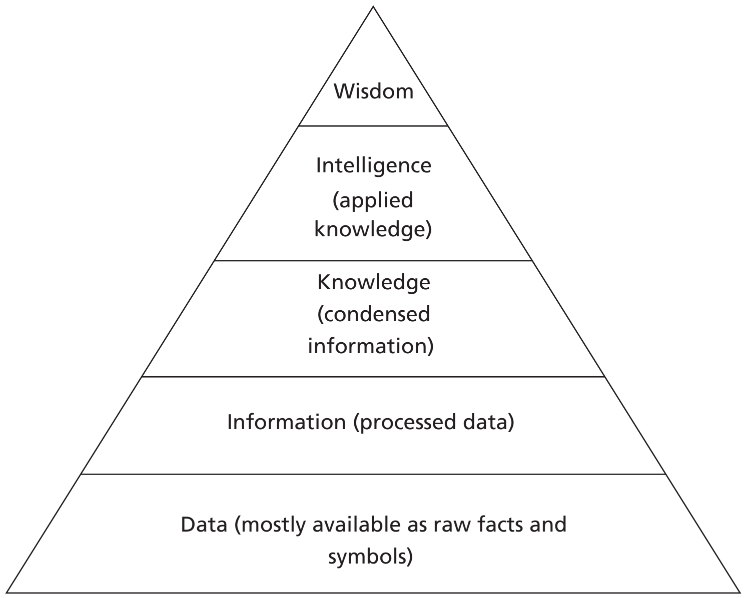

Figure 1.1: The Knowledge Pyramid

What is data? All facts are data. Data can be numbers or text that can be processed by a computer. Today, organizations are accumulating vast and growing amounts of data with data sources such as flat files, databases, or data warehouses in different storage formats.

Processed data is called information. This includes patterns, associations, or relationships among data. For example, sales data can be analyzed to extract information like which is the fast selling product. Condensed information is called knowledge. For example, the historical patterns and future trends obtained in the above sales data can be called knowledge. Unless knowledge is extracted, data is of no use. Similarly, knowledge is not useful unless it is put into action. Intelligence is the applied knowledge for actions. An actionable form of knowledge is called intelligence. Computer systems have been successful till this stage. The ultimate objective of knowledge pyramid is wisdom that represents the maturity of mind that is, so far, exhibited only by humans.

Here comes the need for machine learning. The objective of machine learning is to process these archival data for organizations to take better decisions to design new products, improve the business processes, and to develop effective decision support systems.

1.2 MACHINE LEARNING EXPLAINED

Machine learning is an important sub-branch of Artificial Intelligence (AI). A frequently quoted definition of machine learning was by Arthur Samuel, one of the pioneers of Artificial Intelligence. He stated that "*Machine learning is the field of study that gives the computers ability to learn without being explicitly programmed*."

The key to this definition is that the systems should learn by itself without explicit programming. How is it possible? It is widely known that to perform a computation, one needs to write programs that teach the computers how to do that computation.

In conventional programming, after understanding the problem, a detailed design of the program such as a flowchart or an algorithm needs to be created and converted into programs using a suitable programming language. This approach could be difficult for many real-world problems such as puzzles, games, and complex image recognition applications. Initially, artificial intelligence aims to understand these problems and develop general purpose rules manually. Then, these rules are formulated into logic and implemented in a program to create intelligent systems. This idea of developing intelligent systems by using logic and reasoning by converting an expert's knowledge into a set of rules and programs is called an expert system. An expert system like MYCIN was designed for medical diagnosis after converting the expert knowledge of many doctors into a system. However, this approach did not progress much as programs lacked real intelligence. The word MYCIN is derived from the fact that most of the antibiotics' names end with 'mycin'.

The above approach was impractical in many domains as programs still depended on human expertise and hence did not truly exhibit intelligence. Then, the momentum shifted to machine learning in the form of data driven systems. The focus of AI is to develop intelligent systems by using data-driven approach, where data is used as an input to develop intelligent models. The models can then be used to predict new inputs. Thus, the aim of machine learning is to learn a model or set of rules from the given dataset automatically so that it can predict the unknown data correctly.

As humans take decisions based on an experience, computers make models based on extracted patterns in the input data and then use these data-filled models for prediction and to take decisions. For computers, the learnt model is equivalent to human experience. This is shown in Figure 1.2.

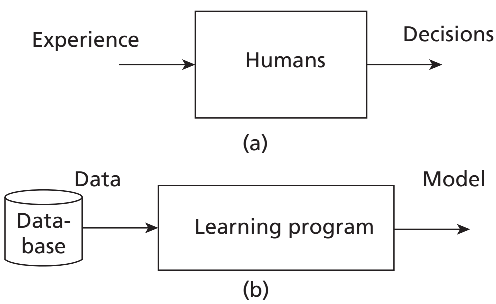

Figure 1.2: (a) A Learning System for Humans (b) A Learning System for Machine Learning

Often, the quality of data determines the quality of experience and, therefore, the quality of the learning system. In statistical learning, the relationship between the input x and output y is modeled as a function in the form $y = f(x)$. Here, f is the learning function that maps the input x to output y. Learning of function f is the crucial aspect of forming a model in statistical learning. In machine learning, this is simply called mapping of input to output.

The learning program summarizes the raw data in a model. Formally stated, a model is an explicit description of patterns within the data in the form of:

1. Mathematical equation
2. Relational diagrams like trees/graphs
3. Logical if/else rules, or
4. Groupings called clusters

In summary, a model can be a formula, procedure or representation that can generate data decisions. The difference between pattern and model is that the former is local and applicable only to certain attributes but the latter is global and fits the entire dataset. For example, a model can be helpful to examine whether a given email is spam or not. The point is that the model is generated automatically from the given data.

Another pioneer of AI, Tom Mitchell's definition of machine learning states that, "*A computer program is said to learn from experience E, with respect to task T and some performance measure P, if its performance on T measured by P improves with experience E*." The important components of this definition are experience *E*, task *T*, and performance measure *P*.

For example, the task *T* could be detecting an object in an image. The machine can gain the knowledge of object using training dataset of thousands of images. This is called experience *E*. So, the focus is to use this experience *E* for this task of object detection *T*. The ability of the system to detect the object is measured by performance measures like precision and recall. Based on the performance measures, course correction can be done to improve the performance of the system.

Models of computer systems are equivalent to human experience. Experience is based on data. Humans gain experience by various means. They gain knowledge by rote learning. They observe others and imitate it. Humans gain a lot of knowledge from teachers and books. We learn many things by trial and error. Once the knowledge is gained, when a new problem is encountered, humans search for similar past situations and then formulate the heuristics and use that for prediction. But, in systems, experience is gathered by these steps:

1. Collection of data
2. Once data is gathered, abstract concepts are formed out of that data. Abstraction is used to generate concepts. This is equivalent to humans' idea of objects, for example, we have some idea about how an elephant looks like.
3. Generalization converts the abstraction into an actionable form of intelligence. It can be viewed as ordering of all possible concepts. So, generalization involves ranking of concepts, inferencing from them and formation of heuristics, an actionable aspect of intelligence. Heuristics are educated guesses for all tasks. For example, if one runs or encounters a danger, it is the resultant of human experience or his heuristics formation. In machines, it happens the same way.
4. Heuristics normally works! But, occasionally, it may fail too. It is not the fault of heuristics as it is just a 'rule of thumb'. The course correction is done by taking evaluation measures. Evaluation checks the thoroughness of the models and to-do course correction, if necessary, to generate better formulations.

1.3 MACHINE LEARNING IN RELATION TO OTHER FIELDS

Machine learning uses the concepts of Artificial Intelligence, Data Science, and Statistics primarily. It is the resultant of combined ideas of diverse fields.

1.3.1 Machine Learning and Artificial Intelligence

Machine learning is an important branch of AI, which is a much broader subject. The aim of AI is to develop intelligent agents. An agent can be a robot, humans, or any autonomous systems. Initially, the idea of AI was ambitious, that is, to develop intelligent systems like human beings. The focus was on logic and logical inferences. It had seen many ups and downs. These down periods were called AI winters.

The resurgence in AI happened due to development of data driven systems. The aim is to find relations and regularities present in the data. Machine learning is the subbranch of AI, whose aim is to extract the patterns for prediction. It is a broad field that includes learning from examples and other areas like reinforcement learning. The relationship of AI and machine learning is shown in Figure 1.3. The model can take an unknown instance and generate results.

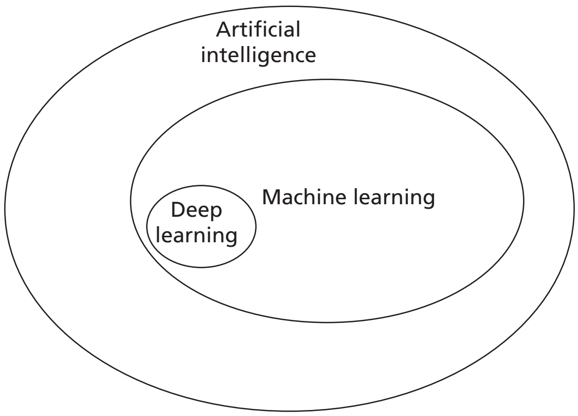

Figure 1.3: Relationship of AI with Machine Learning

Deep learning is a subbranch of machine learning. In deep learning, the models are constructed using neural network technology. Neural networks are based on the human neuron models. Many neurons form a network connected with the activation functions that trigger further neurons to perform tasks.

1.3.2 Machine Learning, Data Science, Data Mining, and Data Analytics

Data science is an 'Umbrella' term that encompasses many fields. Machine learning starts with data. Therefore, data science and machine learning are interlinked. Machine learning is a branch of data science. Data science deals with gathering of data for analysis. It is a broad field that includes: