MACHINE LEARNING

Dr S. Sridhar

Professor,
Department of Information Science and Technology,
College of Engineering, Guindy Campus,
Anna University, Chennai

Dr M. Vijayalakshmi

Associate Professor,
Department of Information Science and Technology,
College of Engineering, Guindy Campus,
Anna University, Chennai

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries.

Published in India by
Oxford University Press
22 Workspace, 2nd Floor, 1/22 Asaf Ali Road, New Delhi 110002

First Edition published in 2021

ISBN-13 (print edition): 978-0-19-012727-5
ISBN-10 (print edition): 0-19-012727-9

Typeset in Palatino Linotype
by B2K-Bytes 2 Knowledge, Tamil Nadu
Printed in India by

Cover image: © whiteMocca/Shutterstock

For product information and current price, please visit www.india.oup.com

Dedicated to my mother, late Mrs Parameswari Sundaramurthy, and my mother-in-law, late Mrs Renuga Nagarajan, whose encouragement and moral support motivated me to start writing this book when they were alive and remained an eternal motivation even after their demise in completing this book.

– Dr S. Sridhar

Dedicated with love and affection to my family, Husband (Mathan), Sons (Sam and Ayden), and Mom (Vennila).

– Dr M. Vijayalakshmi

Preface

Can machines learn like human beings? This question has been posed over many decades and the search for an answer resulted in the domain of artificial intelligence. Informally, learning is nothing but adaptability to the environment. Human beings can adapt to the environment using the learning process. Machine Learning is a branch of artificial intelligence which is defined as "the field of study that gives computers the ability to learn without any explicit programming". Unlike other user-defined programs, machine learning programs try to learn from data automatically. Initially, computer scientists and researchers used automatic learning through logical reasoning. However, much progress could not be made using logic. Eventually, machine learning became popular due to the success of data-driven learning algorithms.

Human beings have always considered themselves as the superior species in object recognition. While machines can crunch numbers in seconds, human beings have shown superiority in recognizing objects. However, recent applications in deep learning show that computers are also good in facial recognition. The recent developments in machine learning such as object recognition, social media analysis like sentiment analysis, recommendation systems including Amazon's book recommendation, innovations, for example driverless cars, and voice assistance systems, which include Amazon's Alexa, Microsoft's Cortana and Google Assistant, have created more awareness about machine learning. The availability of smart phones, IoT, and cloud technology have brought these machine learning technologies to daily life.

Business and government organizations benefit from machine learning technology. These organizations traditionally have a huge amount of data. Social networks such as Twitter, YouTube, and Facebook generate data in terms of Terabytes (TB), Exabytes (EB), and Zettabytes (ZB). In technologies like IoT, sensors generate a huge amount of data independent of any human intervention. A sudden interest has been seen in using machine learning algorithms to extract knowledge from these data automatically. Why? The reason being that extracted knowledge can be useful for prediction and helps in better decision making. This facilitates the development of many knowledge-based and intelligent applications. Therefore, awareness of basic machine learning is a must for students and researchers, computer scientists and professionals, data analysts and data scientists. Historically, these organizations used statistics to analyze these data, but statistics could not be applied on big data. The need to process enormous amount of data poses a challenge as new techniques are required to process this voluminous data, and hence, machine learning is the driving force for many fields such as data science, data analysis, data analytics, data mining, and big data analytics.

Scope of the Book

Our aim has been to provide a first level textbook that follows a simple algorithmic approach and comes with numerical problems to explain the concepts of machine learning. This book stems from the experience of the authors in teaching the students at the Anna University and the National Institute of Technology for over three decades. It targets the undergraduate and post-graduate students in computer science, information technology, and in general, engineering students. This book is also useful for the ones who study data science, data analysis, data analytics, and data mining.

This book comprises chapters covering the concepts of machine learning, a laboratory manual that consists of 25 experiments implemented in Python, and appendices on the basics of Python and Python packages. The theory part includes many numerical problems, review questions and pedagogical support such as crossword and word search. Additional material is available as online resources for better understanding of the subject.

Key Features of the Book

- Uses only minimal mathematics to understand the machine learning algorithms covered in the book
- Follows an algorithmic approach to explain the basics of machine learning
- Comes with various numerical problems to emphasize on the important concepts of data analytics
- Includes a laboratory manual for implementing machine learning concepts in Python environment
- Has two appendices covering the basics of Python and Python packages
- Focuses on pedagogy like chapter-end review and numerical questions, crosswords and jumbled word searches
- Illustrates important and latest concepts like deep learning, regression analysis, support vector machines, clustering algorithms, and ensemble learning

Content and Organization

The book is divided into 16 chapters and three appendices. The appendices A, B, and C can be accessed through the QR codes provided in the table of contents.

Chapter 1 introduces the *Basic Concepts of Machine Learning* and explores its relationships with other domains. This chapter also explores the machine learning types and applications.

Chapter 2 of this book is about *Understanding Data*, which is crucial for data-driven machine learning algorithms. The mathematics that is necessary for understanding data such as linear algebra and statistics covering univariate, bivariate and multivariate statistics are introduced in this chapter. This chapter also includes feature engineering and dimensionality reduction techniques.

Chapter 3 covers the *Basic Concepts of Learning*. This chapter discusses about theoretical aspects of learning such as concept learning, version spaces, hypothesis, and hypothesis space. It also introduces learning frameworks like PAC learning, mistake bound model, and VC dimensions.

Chapter 4 is about *Similarity Learning*. It discusses instance-based learning, nearest-neighbor learning, weighted k-nearest algorithms, nearest centroid classifier, and locally weighted regression (LWR) algorithms.

Chapter 5 introduces the basics of *Regression*. The concepts of linear regression and non-linear regression are discussed in this chapter. It also covers logistic regression. Finally, this chapter outlines the recent algorithms like Ridge, Lasso, and Elastic Net regression.

Chapter 6 throws light on the concept of *Decision Tree Learning*. The concept of information theory, entropy, and information gain are discussed in this chapter. The basics of tree construction algorithms like ID3, C4.5, CART, and Regression Trees and its illustration are included in this chapter. The decision tree evaluation is also introduced here.

Chapter 7 discusses *Rule-based Learning*. This chapter illustrates rule generation. The sequential covering algorithms like PRISM and FOIL are introduced here. This chapter also discusses analytical learning, explanation-based learning, and active learning mechanisms. An outline of association rule mining is also provided in this chapter.

Chapter 8 introduces the basics of *Bayesian* model. The chapter covers the concepts of classification using the Bayesian principle. Naïve Bayesian classifier and Continuous Features classification are introduced in this chapter. The variants of Bayesian classifier are also discussed.

Chapter 9 discusses *Probabilistic Graphical Models*. The discussion of the Bayesian Belief network construction and its inference mechanism are included in this chapter. Markov chain and Hidden Markov Model (HMM) are also introduced along with the associated algorithms.

Chapter 10 introduces the basics of *Artificial Neural Networks* (ANN). The chapter introduces the concepts of neural networks such as neurons, activation functions, and ANN types. Perceptron, back-propagation neural networks, Radial Basis Function Neural Network (RBFNN), Self-Organizing Feature Map (SOFM) are covered here. The chapter ends with challenges and some applications of ANN.

Chapter 11 covers *Support Vector Machines* (SVM). This chapter begins with a gentle introduction of linear discriminant analysis and then covers the concepts of SVM such as margins, kernels, and its associated optimization theory. The hard margin and soft margin SVMs are introduced here. Finally, this chapter ends with support vector regression.

Chapter 12 introduces *Ensemble Learning*. It covers meta-classifiers, the concept of voting, bootstrap resampling, bagging, and random forest and stacking algorithms. This chapter ends with the popular AdaBoost algorithm.

Chapter 13 discusses *Cluster Analysis*. Hierarchical clustering algorithms and partitional clustering algorithms like *k*-means algorithm are introduced in this chapter. In addition, the outline of density-based, grid-based and probability-based approaches like fuzzy clustering and EM algorithm is provided. This chapter ends with an evaluation of clustering algorithms.

Chapter 14 covers the concept of *Reinforcement Learning*. This chapter starts with the idea of reinforcement learning, multi-arm bandit problem and Markov Decision Process (MDP). It then introduces model-based (passive learning) and model-free methods. The *Q*-Learning and SARSA concepts are also covered in this chapter.

Chapter 15 is about *Genetic Algorithms*. The concepts of genetic algorithms and genetic algorithm components along with simple examples are present in this chapter. Evolutionary computation, like simulated annealing, and genetic programming are outlined at the end of this chapter.

Chapter 16 discusses *Deep Learning*. CNN and RNN are explained in this chapter. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are outlined here. Additional web contents are provided for a thorough understanding of deep learning.

Appendix A discusses Python basics.

Appendix B covers Python packages that are necessary to implement the machine learning algorithms. The packages like Numpy, Pandas, Matplotlib, Scikit-learn and Keras are outlined in this appendix.

Appendix C offers 25 laboratory experiments covering the concepts of the textbook.

Acknowledgments

"The Lord is my strength and my shield; my heart trusted in him, and I am helped: therefore my heart greatly rejoiceth; and with my song will I praise him." **[Psalm 28:7]**

First and foremost, the authors express the gratitude with praises and thanks to **Lord God** who made this thing happen. This book would not have been possible without the help of many friends, students, colleagues, and well-wishers. The authors thank all the students who motivated them by asking them the right questions at appropriate time. The authors also thank the colleagues of the Department of Information Science and Technology, Anna University, SRM University, and National Institute of Technology for supporting them with their constant reviews and suggestions. The authors acknowledge the support of the reviewers for providing critical suggestions to improve the quality of the book.

We thank the editorial team of the Oxford University Press, India for their support in bringing the book to realization. The authors thank the members of the marketing team too for their continuous support and encouragement during the development of this book.

Dr Sridhar thanks his wife, Dr Vasanthy, and daughters, Dr Shobika and Shree Varshika, for being patient and accommodative. He also wants to recollect the contributions made by his mother, Mrs Parameswari Sundaramurthy, and mother-in-law, Mrs Renuga Nagarajan, who provided constant encouragement and remained an eternal motivation in completing the book.

Dr Vijayalakshmi would like to thank her husband, Mr Mathan, Sons, Sam and Ayden, and her mother, Ms Vennila, for supporting her in completing this book successfully.

Any comments and suggestions for further improvement of the book are welcome; please send them at ssridhar2004@gmail.com and his website www.drssridhar.com, or vijim@auist.net.

S. Sridhar
M. Vijayalakshmi

The Oxford University Press, India would like to thank all the reviewers:

1. Dr Shriram K. Vasudevan (K. Ramakrishnan College of Technology)
2. Maniroja M. Edinburgh (EXTC Dept., TSEC)
3. Nanda Dulal Jana (NIT, Durgapur)
4. Dipankar Dutta (NERIM Group of Institutions)
5. Imthias Ahamed T P (TKM College of Engineering)
6. Dinobandhu Bhandari (Heritage Institute of Technology)
7. Prof. BSP Mishra (KIIT, Patia)
8. Mrs Angeline (SRM University)
9. Dr Ram Mohan Dhara (IMT, Ghaziabad)
10. Saurabh Sharma (KIET Group of Institutions, Ghaziabad)
11. B. Lakshmanan (Mepco Schlenk Engg. College)
12. Dr T. Subha (Sairam Engineering College)
13. Dr R. Ramya Devi (Velammal Engineering College)
14. Mr Zatin Kumar (Raj Kumar Goel Institute of Technology, Ghaziabad)
15. M. Sowmiya (Easwari Engg. College)
16. A. Sekar (Knowledge Institute of Tech.)

QR Code Content Details

Scan the QR codes provided under the sections listed below to access the following topics:

Table of Contents

- Appendix 1 – Python Language Fundamentals
- Appendix 2 – Python Packages
- Appendix 3 – Lab Manual (with 25 exercises)

Chapter 1

- Section 1.4.4 – Important Machine Learning Algorithms

Chapter 2

- Section 2.1 – Machine Learning and Importance of Linear Algebra, Matrices and Tensors, Sampling Techniques, Information Theory, Evaluation of Classifier Model and Additional Examples
- Section 2.5 – Measures of Frequency and Additional Examples
- Exercises – Additional Key Terms, Short Questions and Numerical Problems

Chapter 3

- Section 3.1 – A figure showing Types of Learning
- Section 3.4.4 – Additional Examples on Generalization and Specialization
- Section 3.4.6 – Additional Examples on Version Spaces

Chapter 5

- Section 5.3 – Additional Examples

Chapter 6

- Section 6.2.1 – Additional Examples

Chapter 7

- Section 7.1 – Simple OneR Algorithm
- Section 7.8 – Additional Examples
- Section 7.8 – Additional Examples: Measure 3-Lift
- Section 7.8 – Apriori Algorithm and Frequent Pattern (FP) Growth Algorithm
- Exercises – Additional Questions

Chapter 8

- Section 8.1 – Probability Theory and Additional Examples

Chapter 9

- Section 9.3.1 – Additional Examples

Chapter 10

- Section 10.1 – Convolution Neural Network, Modular Neural Network, and Recurrent Neural Network
- Key Terms – Additional Key Terms

Chapter 11

- Section 11.1 – Decision Functions and Additional Examples

Chapter 12

- Section 12.2 – Illustrations of Types of Ensemble Models, Parallel Ensemble Model, Voting, Bootstrap Resampling, Working Principle of Bagging, Learning Mechanism through Stacking, and Sequential Ensemble Model
- Section 12.4.1 – Gradient Tree Boosting and XGBoost
- Exercises – Additional Review Question

Chapter 13

- Section 13.2 – Additional Information on Proximity Measures
- Section 13.2 – Taxonomy of Clustering Algorithms
- Section 13.3.4 – Additional Examples
- Section 13.4 – Additional Examples
- Section 13.8 – Purity, Evaluation based on Ground Truth, and Similarity-based Measures
- Key Terms – Additional Key Terms
- Numerical Problems – Additional Numerical Problems

Chapter 14

- Section 14.1 – Additional Information
- Section 14.2 – Context of Reinforcement Learning and SoftMax Method
- Section 14.7 – Additional Information on Dynamic Programming
- Section 14.7 – Additional Examples

Chapter 15

- Section 15.2 – Examples of Optimization Problems and Search Spaces
- Section 15.3 – Flowchart of a Typical GA Algorithm
- Section 15.4.6 – Additional Examples and Information on Topics Discussed in Section 15.4
- Section 15.5.1 – Feature Selection using Genetic Algorithms and Additional Content on Genetic Algorithm Classifier
- Section 15.6.2 – Additional Information on Genetic Programming

Chapter 16

- Section 16.2 – Additional Information on Activation Functions
- Section 16.3 – *L*1 and *L*2 Regularizations
- Section 16.4 – Additional Information on Input Layer
- Section 16.4 – Additional Information on Padding
- Section 16.5 – Detailed Information on Transfer Learning
- Section 16.8 – Restricted Boltzmann Machines and Deep Belief Networks, Auto Encoders, and Deep Reinforcement Learning
- Exercises – Additional Questions

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.

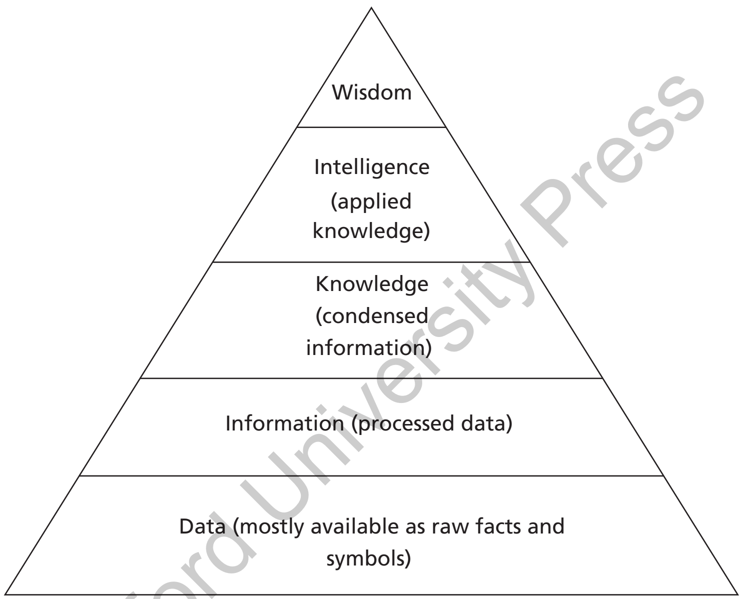


Figure 1.1: The Knowledge Pyramid

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.

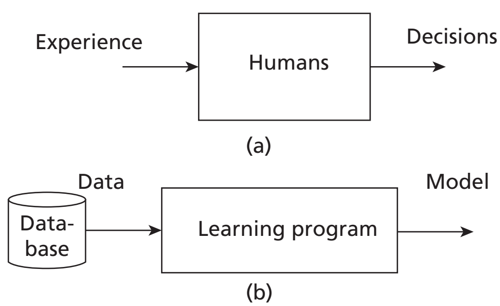


Figure 1.2: (a) A Learning System for Humans (b) A Learning System for Machine Learning

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.

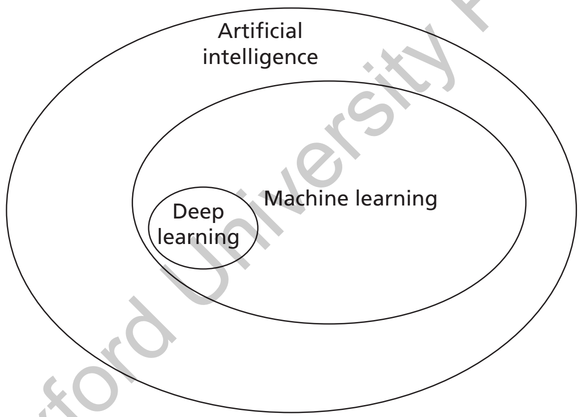


Figure 1.3: Relationship of AI with Machine Learning

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:

Big Data Data science concerns about collection of data. Big data is a field of data science that deals with data's following characteristics:

1. Volume: Huge amount of data is generated by big companies like Facebook, Twitter, YouTube.
2. Variety: Data is available in variety of forms like images, videos, and in different formats.
3. Velocity: It refers to the speed at which the data is generated and processed.

Big data is used by many machine learning algorithms for applications such as language translation and image recognition. Big data influences the growth of subjects like Deep learning. Deep learning is a branch of machine learning that deals with constructing models using neural networks.

Data Mining Data mining's original genesis is in the business. Like while mining the earth one gets into precious resources, it is often believed that unearthing of the data produces hidden information that otherwise would have eluded the attention of the management. Nowadays, many consider that data mining and machine learning are same. There is no difference between these fields except that data mining aims to extract the hidden patterns that are present in the data, whereas, machine learning aims to use it for prediction.

Data Analytics Another branch of data science is data analytics. It aims to extract useful knowledge from crude data. There are different types of analytics. Predictive data analytics is used for making predictions. Machine learning is closely related to this branch of analytics and shares almost all algorithms.

Pattern Recognition It is an engineering field. It uses machine learning algorithms to extract the features for pattern analysis and pattern classification. One can view pattern recognition as a specific application of machine learning.

These relations are summarized in Figure 1.4.

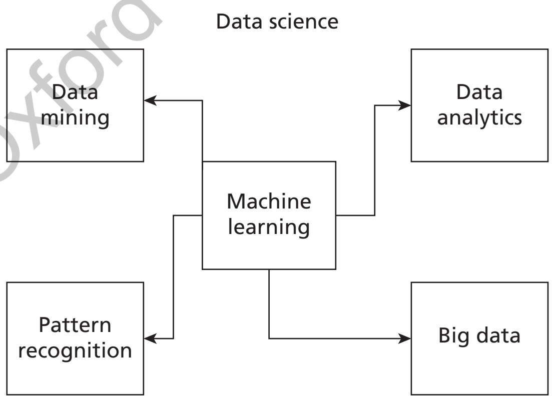


Figure 1.4: Relationship of Machine Learning with Other Major Fields

1.3.3 Machine Learning and Statistics

Statistics is a branch of mathematics that has a solid theoretical foundation regarding statistical learning. Like machine learning (ML), it can learn from data. But the difference between statistics and ML is that statistical methods look for regularity in data called patterns. Initially, statistics sets a hypothesis and performs experiments to verify and validate the hypothesis in order to find relationships among data.

Statistics requires knowledge of the statistical procedures and the guidance of a good statistician. It is mathematics intensive and models are often complicated equations and involve many assumptions. Statistical methods are developed in relation to the data being analysed. In addition, statistical methods are coherent and rigorous. It has strong theoretical foundations and interpretations that require a strong statistical knowledge.

Machine learning, comparatively, has less assumptions and requires less statistical knowledge. But, it often requires interaction with various tools to automate the process of learning.

Nevertheless, there is a school of thought that machine learning is just the latest version of 'old Statistics' and hence this relationship should be recognized.

1.4 TYPES OF MACHINE LEARNING

What does the word 'learn' mean? Learning, like adaptation, occurs as the result of interaction of the program with its environment. It can be compared with the interaction between a teacher and a student. There are four types of machine learning as shown in Figure 1.5.

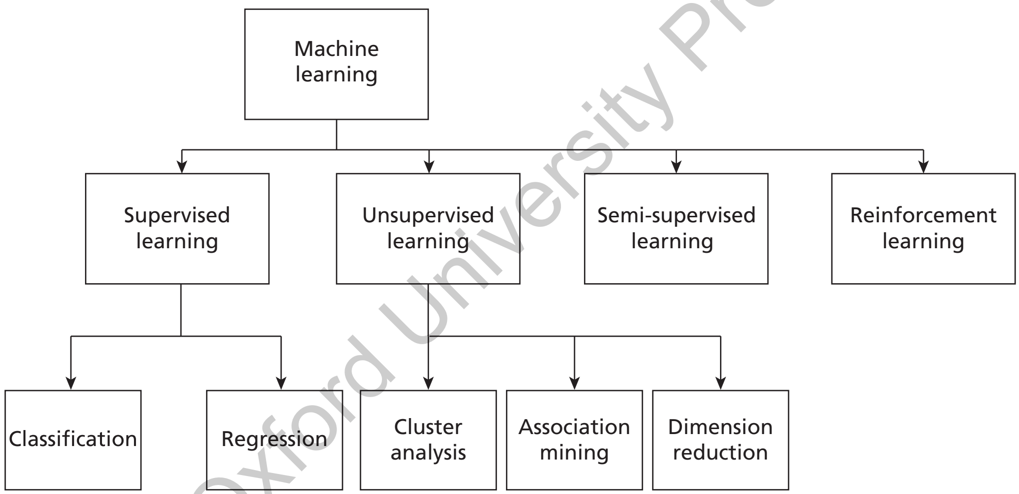


Figure 1.5: Types of Machine Learning

Before discussing the types of learning, it is necessary to discuss about data.

Labelled and Unlabelled Data Data is a raw fact. Normally, data is represented in the form of a table. Data also can be referred to as a data point, sample, or an example. Each row of the table represents a data point. Features are attributes or characteristics of an object. Normally, the columns of the table are attributes. Out of all attributes, one attribute is important and is called a label. Label is the feature that we aim to predict. Thus, there are two types of data – labelled and unlabelled.

Labelled Data To illustrate labelled data, let us take one example dataset called Iris flower dataset or Fisher's Iris dataset. The dataset has 50 samples of Iris – with four attributes, length and width of sepals and petals. The target variable is called class. There are three classes – Iris setosa, Iris virginica, and Iris versicolor.

The partial data of Iris dataset is shown in Table 1.1.

Table 1.1: Iris Flower Dataset

S.No.	Length of Petal	Width of Petal	Length of Sepal	Width of Sepal	Class
1.	5.5	4.2	1.4	0.2	Setosa
2.	7	3.2	4.7	1.4	Versicolor
3.	7.3	2.9	6.3	1.8	Virginica

A dataset need not be always numbers. It can be images or video frames. Deep neural networks can handle images with labels. In the following Figure 1.6, the deep neural network takes images of dogs and cats with labels for classification.

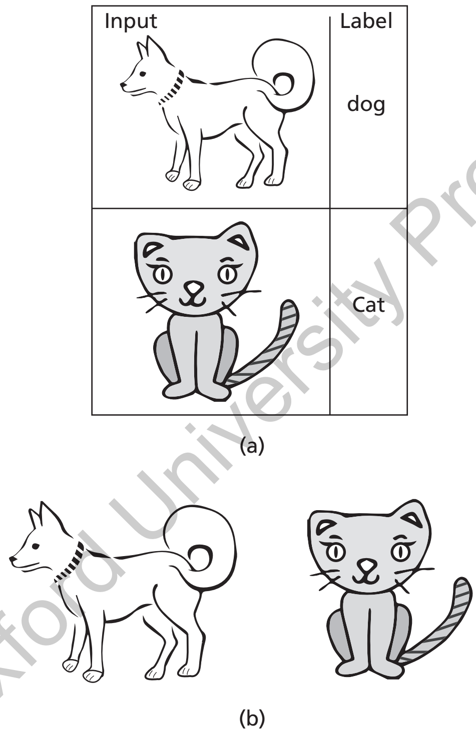


Figure 1.6: (a) Labelled Dataset (b) Unlabelled Dataset

In unlabelled data, there are no labels in the dataset.

1.4.1 Supervised Learning

Supervised algorithms use labelled dataset. As the name suggests, there is a supervisor or teacher component in supervised learning. A supervisor provides labelled data so that the model is constructed and generates test data.

In supervised learning algorithms, learning takes place in two stages. In layman terms, during the first stage, the teacher communicates the information to the student that the student is supposed to master. The student receives the information and understands it. During this stage, the teacher has no knowledge of whether the information is grasped by the student.

This leads to the second stage of learning. The teacher then asks the student a set of questions to find out how much information has been grasped by the student. Based on these questions, the student is tested, and the teacher informs the student about his assessment. This kind of learning is typically called supervised learning.

Supervised learning has two methods:

1. Classification
2. Regression

Classification

Classification is a supervised learning method. The input attributes of the classification algorithms are called independent variables. The target attribute is called label or dependent variable. The relationship between the input and target variable is represented in the form of a structure which is called a classification model. So, the focus of classification is to predict the 'label' that is in a discrete form (a value from the set of finite values). An example is shown in Figure 1.7 where a classification algorithm takes a set of labelled data images such as dogs and cats to construct a model that can later be used to classify an unknown test image data.

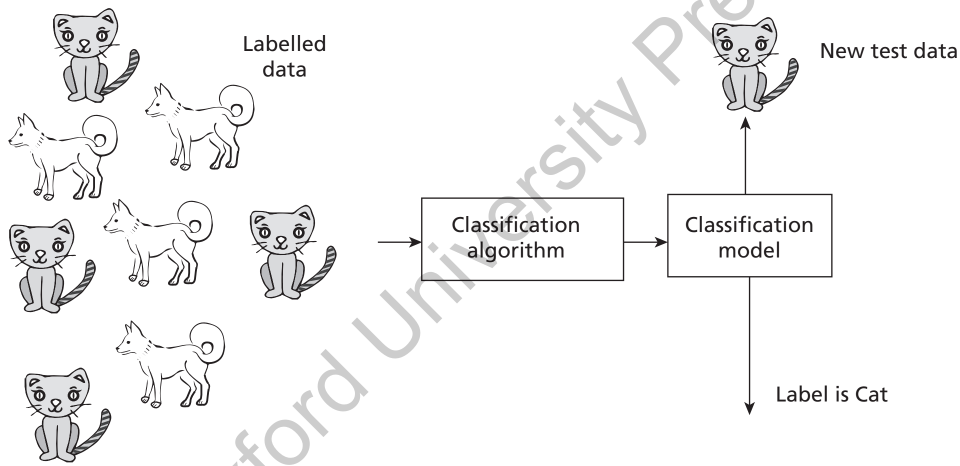


Figure 1.7: An Example Classification System

In classification, learning takes place in two stages. During the first stage, called training stage, the learning algorithm takes a labelled dataset and starts learning. After the training set, samples are processed and the model is generated. In the second stage, the constructed model is tested with test or unknown sample and assigned a label. This is the classification process.

This is illustrated in the above Figure 1.7. Initially, the classification learning algorithm learns with the collection of labelled data and constructs the model. Then, a test case is selected, and the model assigns a label.

Similarly, in the case of Iris dataset, if the test is given as (6.3, 2.9, 5.6, 1.8, ?), the classification will generate the label for this. This is called classification. One of the examples of classification is – Image recognition, which includes classification of diseases like cancer, classification of plants, etc.

The classification models can be categorized based on the implementation technology like decision trees, probabilistic methods, distance measures, and soft computing methods. Classification models can also be classified as generative models and discriminative models. Generative models deal with the process of data generation and its distribution. Probabilistic models are examples of generative models. Discriminative models do not care about the generation of data. Instead, they simply concentrate on classifying the given data.

Some of the key algorithms of classification are:

- Decision Tree
- Random Forest
- Support Vector Machines
- Naïve Bayes
- Artificial Neural Network and Deep Learning networks like CNN

Regression Models

Regression models, unlike classification algorithms, predict continuous variables like price. In other words, it is a number. A fitted regression model is shown in Figure 1.8 for a dataset that represent weeks input x and product sales y.

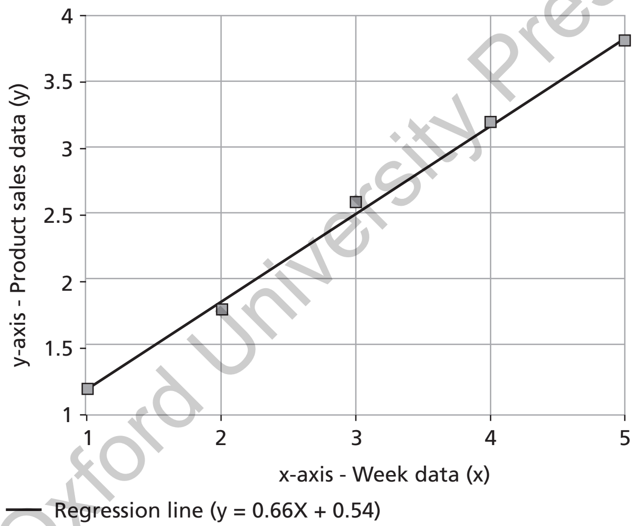


Figure 1.8: A Regression Model of the Form $y = ax + b$

The regression model takes input x and generates a model in the form of a fitted line of the form $y = f(x)$. Here, x is the independent variable that may be one or more attributes and y is the dependent variable. In Figure 1.8, linear regression takes the training set and tries to fit it with a line – product sales = 0.66 × Week + 0.54. Here, 0.66 and 0.54 are all regression coefficients that are learnt from data. The advantage of this model is that prediction for product sales (y) can be made for unknown week data (x). For example, the prediction for unknown eighth week can be made by substituting x as 8 in that regression formula to get y.

One of the most important regression algorithms is linear regression that is explained in the next section.

Both regression and classification models are supervised algorithms. Both have a supervisor and the concepts of training and testing are applicable to both. What is the difference between classification and regression models? The main difference is that regression models predict continuous variables such as product price, while classification concentrates on assigning labels such as class.

1.4.2 Unsupervised Learning

The second kind of learning is by self-instruction. As the name suggests, there are no supervisor or teacher components. In the absence of a supervisor or teacher, self-instruction is the most common kind of learning process. This process of self-instruction is based on the concept of trial and error.

Here, the program is supplied with objects, but no labels are defined. The algorithm itself observes the examples and recognizes patterns based on the principles of grouping. Grouping is done in ways that similar objects form the same group.

Cluster analysis and Dimensional reduction algorithms are examples of unsupervised algorithms.

Cluster Analysis

Cluster analysis is an example of unsupervised learning. It aims to group objects into disjoint clusters or groups. Cluster analysis clusters objects based on its attributes. All the data objects of the partitions are similar in some aspect and vary from the data objects in the other partitions significantly.

Some of the examples of clustering processes are — segmentation of a region of interest in an image, detection of abnormal growth in a medical image, and determining clusters of signatures in a gene database.

An example of clustering scheme is shown in Figure 1.9 where the clustering algorithm takes a set of dogs and cats images and groups it as two clusters-dogs and cats. It can be observed that the samples belonging to a cluster are similar and samples are different radically across clusters.

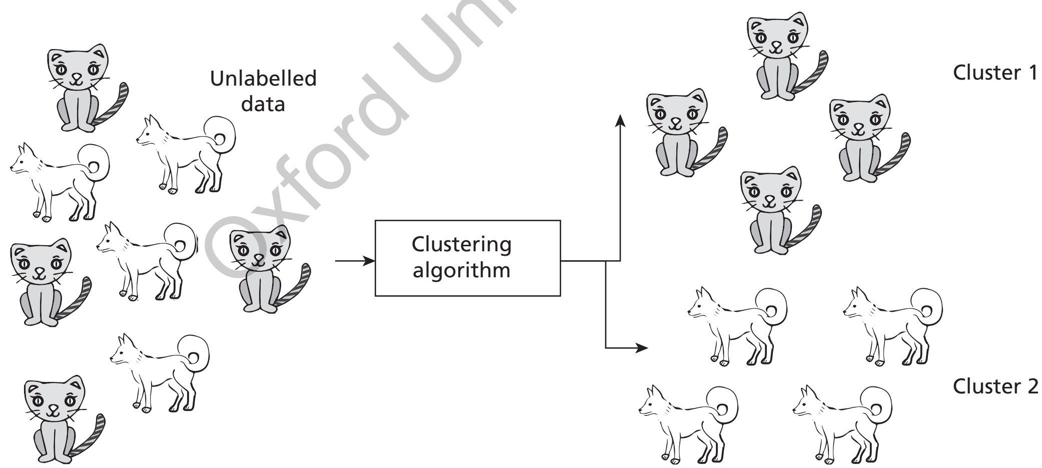


Figure 1.9: An Example Clustering Scheme

Some of the key clustering algorithms are:

- k-means algorithm
- Hierarchical algorithms

Dimensionality Reduction

Dimensionality reduction algorithms are examples of unsupervised algorithms. It takes a higher dimension data as input and outputs the data in lower dimension by taking advantage of the variance of the data. It is a task of reducing the dataset with few features without losing the generality.

The differences between supervised and unsupervised learning are listed in the following Table 1.2.

Table 1.2: Differences between Supervised and Unsupervised Learning

S.No.	Supervised Learning	Unsupervised Learning
1.	There is a supervisor component	No supervisor component
2.	Uses Labelled data	Uses Unlabelled data
3.	Assigns categories or labels	Performs grouping process such that similar objects will be in one cluster

1.4.3 Semi-supervised Learning

There are circumstances where the dataset has a huge collection of unlabelled data and some labelled data. Labelling is a costly process and difficult to perform by the humans. Semi-supervised algorithms use unlabelled data by assigning a pseudo-label. Then, the labelled and pseudo-labelled dataset can be combined.

1.4.4 Reinforcement Learning

Reinforcement learning mimics human beings. Like human beings use ears and eyes to perceive the world and take actions, reinforcement learning allows the agent to interact with the environment to get rewards. The agent can be human, animal, robot, or any independent program. The rewards enable the agent to gain experience. The agent aims to maximize the reward.

The reward can be positive or negative (Punishment). When the rewards are more, the behavior gets reinforced and learning becomes possible.

Consider the following example of a Grid game as shown in Figure 1.10.

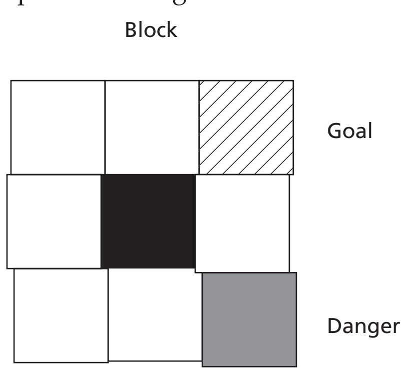


Figure 1.10: A Grid game

In this grid game, the gray tile indicates the danger, black is a block, and the tile with diagonal lines is the goal. The aim is to start, say from bottom-left grid, using the actions left, right, top and bottom to reach the goal state.

To solve this sort of problem, there is no data. The agent interacts with the environment to get experience. In the above case, the agent tries to create a model by simulating many paths and finding rewarding paths. This experience helps in constructing a model.

It can be said in summary, compared to supervised learning, there is no supervisor or labelled dataset. Many sequential decisions need to be taken to reach the final decision. Therefore, reinforcement algorithms are reward-based, goal-oriented algorithms.

Scan for information on '*Important Machine Learning Algorithms*'

1.5 CHALLENGES OF MACHINE LEARNING

What are the challenges of machine learning? Let us discuss about them now.

Problems that can be Dealt with Machine Learning

Computers are better than humans in performing tasks like computation. For example, while calculating the square root of large numbers, an average human may blink but computers can display the result in seconds. Computers can play games like chess, GO, and even beat professional players of that game.

However, humans are better than computers in many aspects like recognition. But, deep learning systems challenge human beings in this aspect as well. Machines can recognize human faces in a second. Still, there are tasks where humans are better as machine learning systems still require quality data for model construction. The quality of a learning system depends on the quality of data. This is a challenge. Some of the challenges are listed below:

1. Problems – Machine learning can deal with the 'well-posed' problems where specifications are complete and available. Computers cannot solve 'ill-posed' problems.

   Consider one simple example (shown in Table 1.3):

   Table 1.3: An Example

Input (x_1, x_2)	Output (y)
1, 1	1
2, 1	2
3, 1	3
4, 1	4
5, 1	5

   Can a model for this test data be multiplication? That is, $y = x_1 \times x_2$. Well! It is true! But, this is equally true that y may be $y = x_1 \div x_2$, or $y = x_1^{x_2}$. So, there are three functions that fit the data. This means that the problem is ill-posed. To solve this problem, one needs more example to check the model. Puzzles and games that do not have sufficient specification may become an ill-posed problem and scientific computation has many ill-posed problems.

2. Huge data – This is a primary requirement of machine learning. Availability of a quality data is a challenge. A quality data means it should be large and should not have data problems such as missing data or incorrect data.
3. High computation power – With the availability of Big Data, the computational resource requirement has also increased. Systems with *Graphics Processing Unit* (GPU) or even *Tensor Processing Unit* (TPU) are required to execute machine learning algorithms. Also, machine learning tasks have become complex and hence time complexity has increased, and that can be solved only with high computing power.
4. Complexity of the algorithms – The selection of algorithms, describing the algorithms, application of algorithms to solve machine learning task, and comparison of algorithms have become necessary for machine learning or data scientists now. Algorithms have become a big topic of discussion and it is a challenge for machine learning professionals to design, select, and evaluate optimal algorithms.
5. Bias/Variance – Variance is the error of the model. This leads to a problem called bias/variance tradeoff. A model that fits the training data correctly but fails for test data, in general lacks generalization, is called overfitting. The reverse problem is called underfitting where the model fails for training data but has good generalization. Overfitting and underfitting are great challenges for machine learning algorithms.

1.6 MACHINE LEARNING PROCESS

The emerging process model for the data mining solutions for business organizations is CRISP-DM. Since machine learning is like data mining, except for the aim, this process can be used for machine learning. CRISP-DM stands for Cross Industry Standard Process – Data Mining. This process involves six steps. The steps are listed below in Figure 1.11.

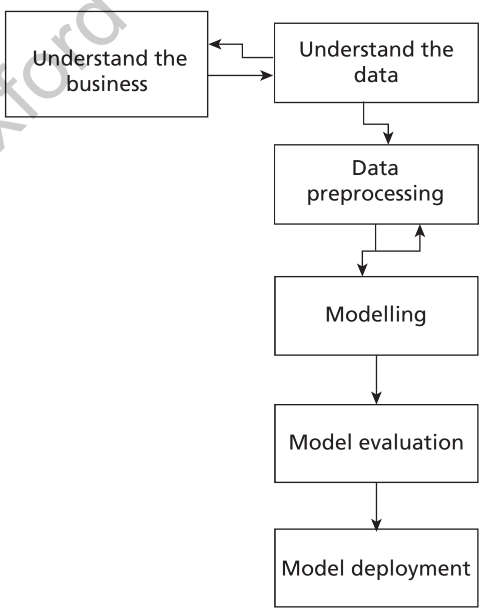


Figure 1.11: A Machine Learning/Data Mining Process

1. Understanding the business – This step involves understanding the objectives and requirements of the business organization. Generally, a single data mining algorithm is enough for giving the solution. This step also involves the formulation of the problem statement for the data mining process.
2. Understanding the data – It involves the steps like data collection, study of the characteristics of the data, formulation of hypothesis, and matching of patterns to the selected hypothesis.
3. Preparation of data – This step involves producing the final dataset by cleaning the raw data and preparation of data for the data mining process. The missing values may cause problems during both training and testing phases. Missing data forces classifiers to produce inaccurate results. This is a perennial problem for the classification models. Hence, suitable strategies should be adopted to handle the missing data.
4. Modelling – This step plays a role in the application of data mining algorithm for the data to obtain a model or pattern.
5. Evaluate – This step involves the evaluation of the data mining results using statistical analysis and visualization methods. The performance of the classifier is determined by evaluating the accuracy of the classifier. The process of classification is a fuzzy issue. For example, classification of emails requires extensive domain knowledge and requires domain experts. Hence, performance of the classifier is very crucial.
6. Deployment – This step involves the deployment of results of the data mining algorithm to improve the existing process or for a new situation.

1.7 MACHINE LEARNING APPLICATIONS

Machine Learning technologies are used widely now in different domains. Machine learning applications are everywhere! One encounters many machine learning applications in the day-to-day life. Some applications are listed below:

1. Sentiment analysis – This is an application of natural language processing (NLP) where the words of documents are converted to sentiments like happy, sad, and angry which are captured by emoticons effectively. For movie reviews or product reviews, five stars or one star are automatically attached using sentiment analysis programs.
2. Recommendation systems – These are systems that make personalized purchases possible. For example, Amazon recommends users to find related books or books bought by people who have the same taste like you, and Netflix suggests shows or related movies of your taste. The recommendation systems are based on machine learning.
3. Voice assistants – Products like Amazon Alexa, Microsoft Cortana, Apple Siri, and Google Assistant are all examples of voice assistants. They take speech commands and perform tasks. These chatbots are the result of machine learning technologies.
4. Technologies like Google Maps and those used by Uber are all examples of machine learning which offer to locate and navigate shortest paths to reduce time.

The machine learning applications are enormous. The following Table 1.4 summarizes some of the machine learning applications.

Table 1.4: Applications' Survey Table

S.No.	Problem Domain	Applications
1.	Business	Predicting the bankruptcy of a business firm
2.	Banking	Prediction of bank loan defaulters and detecting credit card frauds
3.	Image Processing	Image search engines, object identification, image classification, and generating synthetic images
4.	Audio/Voice	Chatbots like Alexa, Microsoft Cortana. Developing chatbots for customer support, speech to text, and text to voice
5.	Telecommunication	Trend analysis and identification of bogus calls, fraudulent calls and its callers, churn analysis
6.	Marketing	Retail sales analysis, market basket analysis, product performance analysis, market segmentation analysis, and study of travel patterns of customers for marketing tours
7.	Games	Game programs for Chess, GO, and Atari video games
8.	Natural Language Translation	Google Translate, Text summarization, and sentiment analysis
9.	Web Analysis and Services	Identification of access patterns, detection of e-mail spams, viruses, personalized web services, search engines like Google, detection of promotion of user websites, and finding loyalty of users after web page layout modification
10.	Medicine	Prediction of diseases, given disease symptoms as cancer or diabetes. Prediction of effectiveness of the treatment using patient history and Chatbots to interact with patients like IBM Watson uses machine learning technologies.
11.	Multimedia and Security	Face recognition/identification, biometric projects like identification of a person from a large image or video database, and applications involving multimedia retrieval
12.	Scientific Domain	Discovery of new galaxies, identification of groups of houses based on house type/geographical location, identification of earthquake epicenters, and identification of similar land use

Summary

1. Machine learning can enable top management of an organization to extract the knowledge from the data stored in various archives to facilitate decision making.
2. Machine learning is an important subbranch of Artificial Intelligence (AI).
3. A model is an explicit description of patterns within the data.
4. A model can be a formula, procedure or representation that can generate data decisions.
5. Humans predict by remembering the past, then formulate the strategy and make a prediction. In the same manner, the computers can predict by following the process.
6. Machine learning is an important branch of AI. AI is a much broader subject. The aim of AI is to develop intelligent agents. An agent can be a robot, humans, or other autonomous systems.

7. Deep learning is a branch of machine learning. The difference between machine learning and deep learning is that models are constructed using neural network technology in deep learning. Neural networks are models constructed based on the human neuron models.
8. Data science deals with gathering of data for analysis. It is a broad field that includes other fields.
9. Data analytics aims to extract useful knowledge from crude data. There are many types of analytics. Predictive data analytics is an area that is dedicated for making predictions. Machine learning is closely related to this branch of analytics and shares almost all algorithms.
10. One can say thus there are two types of data – labelled data and unlabelled data. The data with a label is called labelled data and those without a label are called unlabelled data.
11. Supervised algorithms use labelled dataset. As the name suggests, there is a supervisor or teacher component in supervised learning. A supervisor provides the labelled data so that the model is constructed and gives test data for checking the model.
12. Classification is a supervised learning method. The input attributes of the classification algorithms are called independent variables. The target attribute is called label or dependent variable. The relationship between the input and target variables is represented in the form of a structure which is called a classification model.
13. Cluster analysis is an example of unsupervised learning. It aims to assemble objects into disjoint clusters or groups.
14. Semi-supervised algorithms assign a pseudo-label for unlabelled data.
15. Reinforcement learning allows the agent to interact with the environment to get rewards. The agent can be human, animal, robot, or any independent program. The rewards enable the agent to gain experience.
16. The emerging process model for the data mining solutions for business organizations is CRISP-DM. This model stands for Cross Industry Standard Process – Data Mining.
17. Machine Learning technologies are used widely now in different domains.

Key Terms

- **Machine Learning** – A branch of AI that concerns about machines to learn automatically without being explicitly programmed.
- **Data** – A raw fact.
- **Model** – An explicit description of patterns in a data.
- **Experience** – A collection of knowledge and heuristics in humans and historical training data in case of machines.
- **Predictive Modelling** – A technique of developing models and making a prediction of unseen data.
- **Deep Learning** – A branch of machine learning that deals with constructing models using neural networks.
- **Data Science** – A field of study that encompasses capturing of data to its analysis covering all stages of data management.
- **Data Analytics** – A field of study that deals with analysis of data.
- **Big Data** – A study of data that has characteristics of volume, variety, and velocity.
- **Pattern Recognition** – A field of study that analyses a pattern using machine learning algorithms.

- **Statistics** – A branch of mathematics that deals with learning from data using statistical methods.
- **Hypothesis** – An initial assumption of an experiment.
- **Learning** – Adapting to the environment that happens because of interaction of an agent with the environment.
- **Label** – A target attribute.
- **Labelled Data** – A data that is associated with a label.
- **Unlabelled Data** – A data without labels.
- **Supervised Learning** – A type of machine learning that uses labelled data and learns with the help of a supervisor or teacher component.
- **Classification Program** – A supervisory learning method that takes an unknown input and assigns a label for it. In simple words, finds the category of class of the input attributes.
- **Regression Analysis** – A supervisory method that predicts the continuous variables based on the input variables.
- **Unsupervised Learning** – A type of machine leaning that uses unlabelled data and groups the attributes to clusters using a trial and error approach.
- **Cluster Analysis** – A type of unsupervised approach that groups the objects based on attributes so that similar objects or data points form a cluster.
- **Semi-supervised Learning** – A type of machine learning that uses limited labelled and large unlabelled data. It first labels unlabelled data using labelled data and combines it for learning purposes.
- **Reinforcement Learning** – A type of machine learning that uses agents and environment interaction for creating labelled data for learning.
- **Well-posed Problem** – A problem that has well-defined specifications. Otherwise, the problem is called ill-posed.
- **Bias/Variance** – The inability of the machine learning algorithm to predict correctly due to lack of generalization is called bias. Variance is the error of the model for training data. This leads to problems called overfitting and underfitting.
- **Model Deployment** – A method of deploying machine learning algorithms to improve the existing business processes for a new situation.

Short Questions

1. Why is machine learning needed for business organizations?
2. List out the factors that drive the popularity of machine learning.
3. What is a model?
4. Distinguish between the terms: Data, Information, Knowledge, and Intelligence.
5. How is machine learning linked to AI, Data Science, and Statistics?
6. List out the types of machine learning.
7. List out the differences between a model and pattern. Patterns are local and model is global for entire dataset – Justify.
8. Are classification and clustering are same or different? Justify.

9. List out the differences between labelled and unlabelled data.
10. Point out the differences between supervised and unsupervised learning.
11. What are the differences between classification and regression?
12. What is a semi-supervised learning?
13. List out the differences between reinforced learning and supervised learning.
14. List out important classification and clustering algorithms.
15. List out at least five major applications of machine learning.

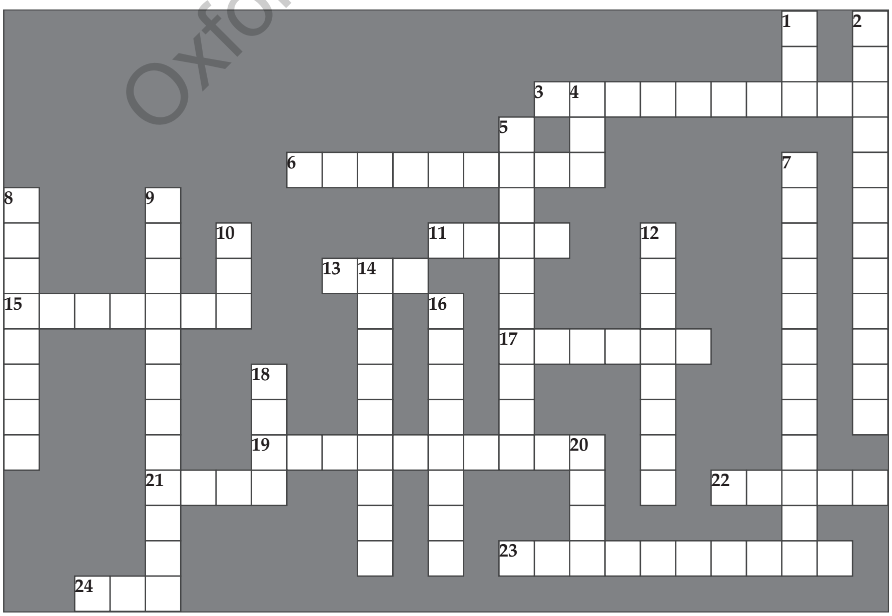
Long Questions

1. Explain in detail the machine learning process model.
2. List out and briefly explain the classification algorithms.
3. List out and briefly explain the unsupervised algorithms.

Numerical Problems and Activities

1. Let us assume a regression algorithm generates a model $y = 0.54 + 0.66\,x$ for data pertaining to week sales data of a product. Here, x is the week and y is the product sales. Find the prediction for the 5^{th} and 8^{th} week.
2. Give two examples of patterns and models.
3. Survey and find out atleast five latest applications of machine learning.
4. Survey and list out atleast five products that use machine learning.

Crossword



Across

3. The initial assumption of the experiment is called a ___________.
6. A study that deals with the analysis of data is called ___________.
11. A domain of study that covers all the aspects of data management is called ___________ science.
13. Data is a ___________ fact.
15. CRISP-DM is a ___________ model.
17. Pattern recognition is used for identifying patterns in ___________ and videos.
19. Unsupervised learning uses ___________ data.
21. Reinforcement learning uses feedback from environment for learning —. (True/False)
22. Amazon Alexa is a ___________ assistant.
23. Classification is an example of ___________ learning.
24. ___________ data has the characteristics such as volume, variety and velocity.

Down

1. A problem that has well-posed specification can be solved using machine learning algorithms —. (Yes/No)
2. Cluster analysis is an example of ___________ learning.
4. Learning from data is the aim of statistics —. (Yes/No)
5. Predictive models can predict based on ___________ data.
7. Regression can predict ___________ variables.
8. Learning is ___________ to the environment.
9. Bias and variance cause overfitting and ___________ of model.
10. Lack of generalization in machine learning happens because of Bias —. (Yes/No)
12. Supervised learning uses ___________ data.
14. Machine learning is ___________ learning without being explicity programmed.
16. Model is a description of ___________.
18. A semi-supervised algorithm assigns a pseudo-label for unlabelled data —. (True/False)
20. Machine learning using neural networks from a domain called artificial neural network and ___________ learning.

Word Search

Find and mark the words listed below.

R	C	X	L	M	S	S	H	I	S	T	O	R	I	C	A	L	Q	U	B	N	O	H	X	B
E	H	Y	P	O	T	H	E	S	I	S	M	Y	E	S	D	Y	T	Y	V	P	Z	U	K	K
A	H	D	D	T	H	X	P	I	U	B	J	Q	T	X	I	W	T	G	I	L	C	R	G	L
D	R	B	S	U	P	E	R	V	I	S	E	D	W	P	R	O	C	E	S	S	K	W	R	F
A	Y	U	N	L	A	B	E	L	L	E	D	K	S	U	P	A	T	T	E	R	N	S	R	N
P	Y	N	L	Y	T	D	V	K	U	X	U	Z	W	F	Y	E	S	E	R	C	D	X	I	I
T	E	S	D	B	M	Z	L	A	B	E	L	L	E	D	K	H	W	L	W	B	Y	D	K	T
I	S	U	O	P	N	D	T	S	W	M	U	N	D	E	R	F	I	T	T	I	N	G	J	C
N	N	P	B	Z	B	P	R	T	R	A	W	J	M	V	O	P	R	P	H	F	C	W	K	L
G	F	E	I	Q	U	F	U	V	Q	B	O	B	W	E	Y	C	L	Y	V	W	X	Q	D	E
T	P	R	G	L	X	T	E	Y	J	X	Y	O	Q	U	A	N	T	I	T	A	T	I	V	E
B	X	V	U	V	L	H	X	U	C	P	Z	Q	K	W	I	S	Y	I	V	Q	Z	T	K	V
G	W	I	V	F	U	T	W	S	O	D	R	B	I	M	A	G	E	S	B	O	O	Z	M	A
G	D	S	N	M	Z	V	G	C	N	L	O	K	Y	S	Q	O	M	O	A	F	I	T	B	T
J	Q	E	Q	A	Z	L	A	U	T	O	M	A	T	I	C	F	E	I	J	F	F	C	M	R
G	A	D	K	X	Z	P	V	N	K	W	M	H	L	O	D	E	E	P	Z	B	B	O	E	U
C	X	F	N	D	A	T	A	D	X	K	A	N	A	L	Y	T	I	C	S	C	K	E	K	E

Automatic	Raw	Patterns	Historical	Deep	Data
Analytics	Big	Images	Yes	Hypothesis	Adapting
Labelled	Supervised	Quantitative	Unlabelled	Unsupervised	True
True	Yes	Yes	Underfitting	Process	Voice