REINFORCEMENT AND SYSTEMIC MACHINE LEARNING FOR DECISION MAKING

Parag Kulkarni
REINFORCEMENT AND SYSTEMIC MACHINE LEARNING FOR DECISION MAKING
Dedicated to the late D.B. Joshi and the late Savitri Joshi, who inspired me to think differently
# CONTENTS

## Preface xv

## Acknowledgments xix

## About the Author xxi

### 1 Introduction to Reinforcement and Systemic Machine Learning 1

1.1. Introduction 1

1.2. Supervised, Unsupervised, and Semisupervised Machine Learning 2

1.3. Traditional Learning Methods and History of Machine Learning 4


1.5. Machine-Learning Problem 8

1.5.1. Goals of Learning 8

1.6. Learning Paradigms 9

1.7. Machine-Learning Techniques and Paradigms 12

1.8. What Is Reinforcement Learning? 14

1.9. Reinforcement Function and Environment Function 16

1.10. Need of Reinforcement Learning 17

1.11. Reinforcement Learning and Machine Intelligence 17


1.15. Reinforcement Machine Learning and Systemic Machine Learning 19

1.16. Case Study Problem Detection in a Vehicle 20

1.17. Summary 20

### Reference 21

### 2 Fundamentals of Whole-System, Systemic, and Multiperspective Machine Learning 23

2.1. Introduction 23


2.1.2. History 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2. What Is Systemic Machine Learning?</td>
<td>27</td>
</tr>
<tr>
<td>2.2.1. Event-Based Learning</td>
<td>29</td>
</tr>
<tr>
<td>2.3.1. System Definition</td>
<td>31</td>
</tr>
<tr>
<td>2.4. Multiperspective Decision Making and Multiperspective Learning</td>
<td>33</td>
</tr>
<tr>
<td>2.4.1. Representation Based on Complete Information</td>
<td>40</td>
</tr>
<tr>
<td>2.4.2. Representation Based on Partial Information</td>
<td>41</td>
</tr>
<tr>
<td>2.4.3. Uni-Perspective Decision Scenario Diagram</td>
<td>41</td>
</tr>
<tr>
<td>2.4.4. Dual-Perspective Decision Scenario Diagrams</td>
<td>41</td>
</tr>
<tr>
<td>2.4.5. Multiperspective Representative Decision Scenario Diagrams</td>
<td>42</td>
</tr>
<tr>
<td>2.4.6. Qualitative Belief Network and ID</td>
<td>42</td>
</tr>
<tr>
<td>2.5. Dynamic and Interactive Decision Making</td>
<td>43</td>
</tr>
<tr>
<td>2.5.1. Interactive Decision Diagrams</td>
<td>43</td>
</tr>
<tr>
<td>2.5.2. Role of Time in Decision Diagrams and Influence Diagrams</td>
<td>43</td>
</tr>
<tr>
<td>2.5.3. Systemic View Building</td>
<td>44</td>
</tr>
<tr>
<td>2.5.4. Integration of Information</td>
<td>45</td>
</tr>
<tr>
<td>2.5.5. Building Representative DSD</td>
<td>45</td>
</tr>
<tr>
<td>2.5.6. Limited Information</td>
<td>45</td>
</tr>
<tr>
<td>2.5.7. Role of Multiagent System in Systemic Learning</td>
<td>46</td>
</tr>
<tr>
<td>2.6. The Systemic Learning Framework</td>
<td>47</td>
</tr>
<tr>
<td>2.6.1. Mathematical Model</td>
<td>50</td>
</tr>
<tr>
<td>2.6.2. Methods for Systemic Learning</td>
<td>50</td>
</tr>
<tr>
<td>2.6.3. Adaptive Systemic Learning</td>
<td>51</td>
</tr>
<tr>
<td>2.6.4. Systemic Learning Framework</td>
<td>52</td>
</tr>
<tr>
<td>2.7. System Analysis</td>
<td>52</td>
</tr>
<tr>
<td>2.8. Case Study: Need of Systemic Learning in the Hospitality Industry</td>
<td>54</td>
</tr>
<tr>
<td>2.9. Summary</td>
<td>55</td>
</tr>
<tr>
<td>References</td>
<td>56</td>
</tr>
<tr>
<td>3. Reinforcement Learning</td>
<td>57</td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>57</td>
</tr>
<tr>
<td>3.2. Learning Agents</td>
<td>60</td>
</tr>
<tr>
<td>3.3. Returns and Reward Calculations</td>
<td>62</td>
</tr>
<tr>
<td>3.3.1. Episodic and Continuing Task</td>
<td>63</td>
</tr>
<tr>
<td>3.4. Reinforcement Learning and Adaptive Control</td>
<td>63</td>
</tr>
<tr>
<td>3.5. Dynamic Systems</td>
<td>66</td>
</tr>
<tr>
<td>3.5.1. Discrete Event Dynamic System</td>
<td>67</td>
</tr>
<tr>
<td>3.6. Reinforcement Learning and Control</td>
<td>68</td>
</tr>
</tbody>
</table>
5.6. Bayesian Paradigm and Inference 113
   5.6.1. Bayes’ Theorem 113
5.7. Time-Based Inference 114
5.8. Inference to Build a System View 114
   5.8.1. Information Integration 115
5.9. Summary 118
References 118

6 Adaptive Learning 119
   6.1. Introduction 119
   6.2. Adaptive Learning and Adaptive Systems 119
   6.4. Adaptation and Learning Method Selection Based on Scenario 124
      6.4.1. Dynamic Adaptation and Context-Aware Learning 125
   6.5. Systemic Learning and Adaptive Learning 127
      6.5.1. Use of Multiple Learners 129
      6.5.2. Systemic Adaptive Machine Learning 132
      6.5.3. Designing an Adaptive Application 135
      6.5.4. Need of Adaptive Learning and Reasons for Adaptation 135
      6.5.5. Adaptation Types 136
      6.5.6. Adaptation Framework 139
   6.6. Competitive Learning and Adaptive Learning 140
      6.6.1. Adaptation Function 142
      6.6.2. Decision Network 144
      6.6.3. Representation of Adaptive Learning Scenario 145
   6.7. Examples 146
      6.7.1. Case Study: Text-Based Adaptive Learning 147
      6.7.2. Adaptive Learning for Document Mining 148
   6.8. Summary 149
References 149

7 Multiperspective and Whole-System Learning 151
   7.1. Introduction 151
   7.2. Multiperspective Context Building 152
   7.3. Multiperspective Decision Making and Multiperspective Learning 154
      7.3.1. Combining Perspectives 155
      7.3.2. Influence Diagram and Partial Decision Scenario Representation Diagram 156
7.3.3. Representative Decision Scenario Diagram (RDSD) 160
7.3.4. Example: PDSRD Representations for City Information Captured from Different Perspectives 160

7.4. Whole-System Learning and Multiperspective Approaches 164
7.4.1. Integrating Fragmented Information 165
7.4.2. Multiperspective and Whole-System Knowledge Representation 165
7.4.3. What Are Multiperspective Scenarios? 165
7.4.4. Context in Particular 166

7.5. Case Study Based on Multiperspective Approach 167
7.5.1. Traffic Controller Based on Multiperspective Approach 167
7.5.2. Multiperspective Approach Model for Emotion Detection 169

7.6. Limitations to a Multiperspective Approach 174
7.7. Summary 174
References 175

8 Incremental Learning and Knowledge Representation 177
8.1. Introduction 177
8.2. Why Incremental Learning? 178
8.3. Learning from What Is Already Learned. . . 180
8.3.1. Absolute Incremental Learning 181
8.3.2. Selective Incremental Learning 182
8.4. Supervised Incremental Learning 191
8.5. Incremental Unsupervised Learning and Incremental Clustering 191
8.5.1. Incremental Clustering: Tasks 193
8.5.2. Incremental Clustering: Methods 195
8.5.3. Threshold Value 196
8.6. Semisupervised Incremental Learning 196
8.7. Incremental and Systemic Learning 199
8.8. Incremental Closeness Value and Learning Method 200
8.8.1. Approach 1 for Incremental Learning 201
8.8.2. Approach 2 202
8.8.3. Calculating C Values Incrementally 202
8.9. Learning and Decision-Making Model 205
8.10. Incremental Classification Techniques 206
8.11. Case Study: Incremental Document Classification 207
8.12. Summary 208
9 Knowledge Augmentation: A Machine Learning Perspective 209

9.1. Introduction 209
9.2. Brief History and Related Work 211
9.3. Knowledge Augmentation and Knowledge Elicitation 215
  9.3.1. Knowledge Elicitation by Strategy Used 215
  9.3.2. Knowledge Elicitation Based on Goals 216
  9.3.3. Knowledge Elicitation Based on Process 216
9.4. Life Cycle of Knowledge 217
  9.4.1. Knowledge Levels 219
  9.4.2. Direct Knowledge 219
  9.4.3. Indirect Knowledge 219
  9.4.4. Procedural Knowledge 219
  9.4.5. Questions 220
  9.4.6. Decisions 220
  9.4.7. Knowledge Life Cycle 220
9.5. Incremental Knowledge Representation 222
9.6. Case-Based Learning and Learning with Reference to Knowledge Loss 224
9.7. Knowledge Augmentation: Techniques and Methods 224
  9.7.1. Knowledge Augmentation Techniques 225
  9.7.2. Knowledge Augmentation Methods 226
  9.7.3. Mechanisms for Extracting Knowledge 227
9.8. Heuristic Learning 228
9.9. Systemic Machine Learning and Knowledge Augmentation 229
  9.9.1. Systemic Aspects of Knowledge Augmentation 230
  9.9.2. Systemic Knowledge Management and Advanced Machine Learning 231
9.10. Knowledge Augmentation in Complex Learning Scenarios 232
9.11. Case Studies 232
  9.11.1. Case Study Banking 232
  9.11.2. Software Development Firm 233
  9.11.3. Grocery Bazaar/Retail Bazaar 234
9.12. Summary 235
References 235

10 Building a Learning System 237

10.1. Introduction 237
10.2. Systemic Learning System 237
  10.2.1. Learning Element 240
  10.2.2. Knowledge Base 240
  10.2.3. Performance Element 240
10.2.4. Feedback Element 240
10.2.5. System to Allow Measurement 241
10.3. Algorithm Selection 242
  10.3.1. $k$-Nearest-Neighbor ($k$-NN) 242
  10.3.2. Support Vector Machine (SVM) 243
  10.3.3. Centroid Method 243
10.4. Knowledge Representation 244
  10.4.1. Practical Scenarios and Case Study 244
10.5. Designing a Learning System 245
10.6. Making System to Behave Intelligently 246
10.7. Example-Based Learning 246
10.8. Holistic Knowledge Framework and Use of Reinforcement Learning 246
  10.8.1. Intelligent Algorithms Selection 249
10.9. Intelligent Agents—Deployment and Knowledge Acquisition and Reuse 250
10.10. Case-Based Learning: Human Emotion-Detection System 251
10.11. Holistic View in Complex Decision Problem 253
10.12. Knowledge Representation and Data Discovery 255
10.13. Components 258
  10.13.1. Example 258
10.15. Summary 259

Appendix A: Statistical Learning Methods 261

Appendix B: Markov Processes 271

Index 281
There has been movement for years to make machines intelligent. This movement began long ago, even long before the computer era. Event-based intelligence in those days was incorporated in appliances or the ensemble of appliances. This intelligence was very much guided, and human intervention was mandatory. Even feedback control systems are a rudimentary form of intelligent system. Later adaptive control systems and hybrid control systems added flair of intelligence in these systems. This movement has received more attention with the advent of computers. Simple event-based learning with computers became a part of many intelligent systems very quickly. The expectation from intelligent systems kept on increasing. This led to one of the very well-received paradigms of learning, which is pattern-based learning. This allowed the systems to exhibit intelligence in many practical scenarios. It included patterns of weather, patterns of occupancy, and different patterns where patterns could help to make decisions. This paradigm evolved into a paradigm of behavioral pattern-based learning. This was more a behavioral pattern than a simple pattern of a particular measurement parameter. Behavioral patterns attempted to give a better picture and insight. This helped to learn and make decisions in case of networks and business scenarios. This took the intelligent systems to the next level. Learning is a manifestation of intelligence. Making machines to learn is a major part of the movement to make machines intelligent.

The complexities in decision scenarios and making machines to learn in complex scenarios raised many questions on the intelligence of a machine. Learning in isolation is never complete. Human beings learn in groups, develop colonies, and interact to build intelligence. The collective and cooperative learning of humans allows them to achieve supremacy. Furthermore, humans learn in association with the environment. They interact with the environment and receive feedback in the form of a reward or penalty. Their learning in association gives them power for exploration-based learning. Exploitation of already learned facts and exploration with reference to actions takes place. The paradigm of reinforcement learning added a new dimension to learning and could cover many new aspects of learning required for dynamic scenarios.

As mentioned by Rutherford D. Roger: “We are drowning in information and starving for knowledge.” More and more information becomes available for our disposal. This information is in heterogeneous forms. There are many information sources and numerous learning opportunities. The practical assumptions while
learning can make learning restrictive. Actually there are relationships among
different parts of the system, and one of the basic principles of system thinking
states is that the cause and effect are separated in time and space. The impact of the
decision or any action can be felt beyond visible boundaries. Failing to consider
this systemic aspect and relationship will lead to many limitations while learning,
and hence the traditional learning paradigms suffer in highly dynamic and complex
real-life problems. The holistic view and understanding of the interdependencies
and intradependencies can help us to learn many new aspects and understand,
analyze, and interpret the information in a more realistic way. The aspect of
learning based on available information, building new information, mapping it to
knowledge, and understanding different perspectives while learning can really
help to make learning more effective. Learning is not just getting more data and
arranging that data. It is not even building more information. Basically, the purpose
of learning is to empower individuals to make better decisions and to improve their
ability to create value. In machine learning, there is a need to expand the ability of
machines with reference to different information sources and learning opportu-
nities. In machine learning, it is also about empowering machines to make better
decisions and improving their ability to create value.

This book is an attempt to put forth a new paradigm of systemic machine
learning and research opportunities in machine learning with reference to different
aspects of machine learning. The book tries to build the foundation for systemic
machine learning with elaborate case studies. Machine learning and artificial
intelligence are interdisciplinary in nature. Right from statistics, mathematics,
psychology, and computer engineering, many researchers contributed to this field
to make it rich and achieve better results. Based on these numerous contributions
and our research in machine learning field, this book tries to explore the concept of
systemic machine learning. Systemic machine learning is holistic, multiperspec-
tive, incremental, and systemic. While learning we can learn different things from
the same data sets, we can also learn from already learned facts, and there can be
number of representations of knowledge. This book is an attempt to build a
framework to make the best use of all information sources and build knowledge
with reference to the complete system.

In many cases, the problem is not static. It changes with time and depends on
environment, and the solution even depends on the decision context. Context
may not be just limited to a few parameters, but the overall information about a
problem builds the context. A general-purpose system without context may not
be able to handle context-specific decision. This book discusses different facets
of learning as well as the need of a new paradigm with reference to complex
decision problems. The book can be used as a reference book for specialized
research and can help readers and researchers to appreciate new paradigms of
machine learning.

This book is organized as depicted in the following figure:
Chapter 1 introduces concepts of systemic and reinforcement machine learning. It builds a platform for the paradigm of systemic machine learning while highlighting the need of the same. Chapter 2 throws more light on the fundamentals of systemic machine learning, whole system learning, and multiperspective learning. Chapter 3 is about reinforcement learning while Chapter 4 deals with systemic machine learning and model building. The important aspects of decision making such as inference are covered in Chapter 5. Chapter 6 discusses adaptive machine learning and various aspects of adaptive machine learning. Chapter 7 discusses the paradigm of multiperspective machine learning and whole system learning. Chapter 8 addresses the need for incremental machine learning. Chapters 8 and 9 deal with knowledge representation and knowledge augmentation. Chapter 10 discusses the building learning system.

This book tries to include different facets of learning while introducing a new paradigm of machine learning. It deals with building knowledge through machine learning. This book is for those individuals who are planning to contribute to make a machine more intelligent by making it learn through new experiments, are ready to try new ways, and are open for a new paradigm for the same.

Parag Kulkarni
For the past two decades I have been working with various decision-making and AI-based IT product companies. During this period I worked on different Machine Learning algorithms and applied them for different applications. This work made me realize the need for a new paradigm for machine learning and the need for change in thinking. This built the foundation for this book and started the thought process for systemic machine learning. I am thankful to different organizations I worked with, including Siemens and IDEaS, and to my colleagues in those organizations. I would also like to acknowledge the support of my friends and coworkers.

I would like to thank my Ph.D. and M.Tech. students—Prachi, Yashodhara, Vinod, Sunita, Pramod, Nitin, Deepak, Preeti, Anagha, Shankar, Shweta, Basawraj, Shashikanth, and others—for their direct and indirect contribution that came through technical brainstorming. They are always ready to work on new ideas and contributed through collective learning. Special thanks to Prachi for her help in drawing diagrams and formatting the text.

I am thankful to Prof. Chande, the late Prof. Ramani, Dr. Sinha, Dr. Bhanu Prasad, Prof. Warnekar, and Prof. Navetia for useful comments and reviews. I am also thankful to Institutes such as COEP, PICT, GHRIET, PCCOE, DYP COE, IIM, Masaryk University, and so on, for allowing me to interact and present my thoughts in front of students. I am also thankful to IASTED, IACSIT, and IEEE for giving me the platform to present my research through various technical conferences. I am also thankful to reviewers of my research papers.

I am thankful to my mentor, teacher, and grandfather, the late D.B. Joshi, for motivating me to think differently. I also would like to take the opportunity to thank my mother. Most importantly I would like to thank my wife Mrudula and son Hrishikesh for their support, motivation, and help.

I am also thankful to IEEE/Wiley and the editorial team of IEEE/Wiley for their support and helping me to present my research, thoughts, and experiments in the form of a book.

Parag Kulkarni
Parag Kulkarni, Ph.D. D.Sc., is CEO and Chief Scientist at EKLaT Research, Pune. He has more than two decades of experience in knowledge management, e-business, intelligent systems and machine learning consultation, research and product building. An alumnus of IIT Kharagpur and IIM Kolkata, Dr. Kulkarni has been a visiting professor at IIM Indore, visiting researcher at Masaryk University Czech Republic, and Adjunct Professor at the College of Engineering, Pune. He has headed companies, research labs, and groups at various IT companies including IDeaS, Siemens Information Systems Ltd., and Capilson, Pune, and ReasonEdge, Singapore. He has led many start-up companies to success through strategic innovation and research. The UGSM Monarch Business School, Switzerland, has conferred higher doctorate D.Sc. on Dr. Kulkarni. He is a coinventor of three patents and has coauthored more than 100 research papers and several books.
1.1 INTRODUCTION

The expectations from intelligent systems are increasing day by day. What an intelligent system was supposed to do a decade ago is now expected from an ordinary system. Whether it is a washing machine or a health care system, we expect it to be more and more intelligent and demonstrate that behavior while solving complex as well as day-to-day problems. The applications are not limited to a particular domain and are literally distributed across all domains. Hence domain-specific intelligence is fine but the user has become demanding, and a true intelligent and problem-solving system irrespective of domains has become a necessary goal. We want the systems to drive cars, play games, train players, retrieve information, and help even in complex medical diagnosis. All these applications are beyond the scope of isolated systems and traditional preprogrammed learning. These activities need dynamic intelligence. Dynamic intelligence can be exhibited through learning not only based on available knowledge but also based on the exploration of knowledge through interactions with the environment. The use of existing knowledge, learning based on dynamic facts, and acting in the best way in complex scenarios are some of the expected features of intelligent systems.

The learning has many facets. Right from simple memorization of facts to complex inference are some examples of learning. But at any point of time, learning is a holistic activity and takes place around the objective of better decision-making. Learning results from data storing, sorting, mapping, and classification. Still one of the most important aspects of intelligence is learning. In most of the cases we expect learning to be a more goal-centric activity. Learning results from an inputs from an experienced person, one’s own experience, and inference based on experiences or past learning. So there are three ways of learning:

- Learning based on expert inputs (supervised learning)
• Learning based on own experience
• Learning based on already learned facts

In this chapter, we will discuss the basics of reinforcement learning and its history. We will also look closely at the need of reinforcement learning. This chapter will discuss limitations of reinforcement learning and the concept of systemic learning. The systemic machine-learning paradigm is discussed along with various concepts and techniques. The chapter also covers an introduction to traditional learning methods. The relationship among different learning methods with reference to systemic machine learning is elaborated in this chapter. The chapter builds the background for systemic machine learning.

1.2 SUPERVISED, UNSUPERVISED, AND SEMISUPERVISED MACHINE LEARNING

Learning that takes place based on a class of examples is referred to as supervised learning. It is learning based on labeled data. In short, while learning, the system has knowledge of a set of labeled data. This is one of the most common and frequently used learning methods. Let us begin by considering the simplest machine-learning task: supervised learning for classification. Let us take an example of classification of documents. In this particular case a learner learns based on the available documents and their classes. This is also referred to as labeled data. The program that can map the input documents to appropriate classes is called a classifier, because it assigns a class (i.e., document type) to an object (i.e., a document). The task of supervised learning is to construct a classifier given a set of classified training examples. A typical classification is depicted in Figure 1.1.

Figure 1.1 represents a hyperplane that has been generated after learning, separating two classes—class A and class B in different parts. Each input point presents input–output instance from sample space. In case of document classification, these points are documents. Learning computes a separating line or hyperplane among documents. An unknown document type will be decided by its position with respect to a separator.
There are a number of challenges in supervised classification such as generalization, selection of right data for learning, and dealing with variations. Labeled examples are used for training in case of supervised learning. The set of labeled examples provided to the learning algorithm is called the training set.

The classifier and of course the decision-making engine should minimize false positives and false negatives. Here false positives stand for the result yes—that is, classified in a particular group wrongly. False negative is the case where it should have been accepted as a class but got rejected. For example, apples not classified as apples is false negative, while an orange or some other fruit classified as an apple is false positive in the apple class. Another example of it is when guilty but not convicted is false positive, while innocent but convicted or declared innocent is false negative. Typically, wrongly classified are more harmful than unclassified elements.

If a classifier knew that the data consisted of sets or batches, it could achieve higher accuracy by trying to identify the boundary between two adjacent sets. It is true in the case of sets of documents to be separated from one another. Though it depends on the scenario, typically false negatives are more costly than false positives, so we might want the learning algorithm to prefer classifiers that make fewer false negative errors, even if they make more false positives as a result. This is so because false negative generally takes away the identity of the objects or elements that are classified correctly. It is believed that the false positive can be corrected in next pass, but there is no such scope for false negative.

Supervised learning is not just about classification, but it is the overall process that with guidelines maps to the most appropriate decision.

Unsupervised learning refers to learning from unlabeled data. It is based more on similarity and differences than on anything else. In this type of learning, all similar items are clustered together in a particular class where the label of a class is not known.

It is not possible to learn in a supervised way in the absence of properly labeled data. In these scenarios there is need to learn in an unsupervised way. Here the learning is based more on similarities and differences that are visible. These differences and similarities are mathematically represented in unsupervised learning.

Given a large collection of objects, we often want to be able to understand these objects and visualize their relationships. For an example based on similarities, a kid can separate birds from other animals. It may use some property or similarity while separating, such as the birds have wings. The criterion in initial stages is the most visible aspects of those objects. Linnaeus devoted much of his life to arranging living organisms into a hierarchy of classes, with the goal of arranging similar organisms together at all levels of the hierarchy. Many unsupervised learning algorithms create similar hierarchical arrangements based on similarity-based mappings. The task of hierarchical clustering is to arrange a set of objects into a hierarchy such that similar objects are grouped together. Nonhierarchical clustering seeks to partition the data into some number of disjoint clusters. The process of clustering is depicted in Figure 1.2. A learner is fed with a set of scattered points, and it generates two clusters with representative centroids after learning. Clusters show that points with similar properties and closeness are grouped together.
In practical scenarios there is always need to learn from both labeled and unlabeled data. Even while learning in an unsupervised way, there is the need to make the best use of labeled data available. This is referred to as semisupervised learning. Semisupervised learning is making the best use of two paradigms of learning—that is, learning based on similarity and learning based on inputs from a teacher. Semisupervised learning tries to get the best of both the worlds.

1.3 TRADITIONAL LEARNING METHODS AND HISTORY OF MACHINE LEARNING

Learning is not just knowledge acquisition but rather a combination of knowledge acquisition, knowledge augmentation, and knowledge management. Furthermore, intelligent inference is essential for proper learning. Knowledge deals with significance of information and learning deals with building knowledge. How can a machine can be made to learn? This research question has been posed for more than six decades by researchers. The outcome of this research has built a platform for this chapter. Learning involves every activity. One such example, is the following: While going to the office yesterday, Ram found road repair work in progress on route one, so he followed route two today. It might be possible that route two is worse. Then he may go back to route one or might try route three. Route one is in bad shape due to repair work is knowledge built, and based on that knowledge he has taken action: following route 2, that is, exploration. The complexity of learning increases as the number of parameters and time dimensions start playing a role in decision making.

Ram found that road repair work is in progress on route one.
He hears an announcement that in case of rain, route two will be closed.
He needs to visit a shop X while going to office.
He is running out of petrol.

These new parameters make his decision much more complex as compared to scenario 1 and scenario 2 discussed above.

In this chapter, we will discuss various learning methods along with examples. The data and information used for learning are very important. The data cannot be
used as is for learning. It may contain outliers and information about features that may not be relevant with respect to the problem one is trying to solve. The approaches for the selection of data for learning vary with the problems. In some cases the most frequent patterns are used for learning. Even in some cases, outliers are also used for learning. There can be learning based on exceptions. The learning can take place based on similarities as well as differences. The positive as well as negative examples help in effective learning. Various models are built for learning with the objective of exploiting the knowledge.

Learning is a continuous process. The new scenarios are observed and new situations arise—those need to be used for learning. Learning from observation needs to construct meaningful classification of observed objects and situation. Methods of measuring similarity and proximity are employed for this purpose. Learning from observations is the most commonly used method by human beings. While making decisions we may come across the scenarios and objects that we have not used or came across during a learning phase. The inference allows us to handle these scenarios. Furthermore, we need to learn in different and new scenarios and hence even while making decisions the learning continues.

There are three fundamental continuously active human-like learning mechanisms:

1. **Perceptual Learning**: Learning of new objects, categories, and relations. It is more like constantly seeking to improve and grow. It is similar to the learning professionals use.

2. **Episodic Learning**: It is based on events and information about the event, like what, where, and when. It is the learning or the change in the behavior that occurs due to an event.

3. **Procedural Learning**: Learning based on actions and action sequences to accomplish a task. Implementation of this human cognition can impart intelligence to a machine. Hence, a unified methodology around intelligent behavior is the need of time that will allow machines to learn and behave or respond intelligently in dynamic scenarios.

Traditional machine-learning approaches are susceptible to dynamic continual changes in the environment. However, perceptual learning in human does not have such restrictions. Learning in humans is selectively incremental, so it does not need a large training set and is simultaneously not biased by already learned but outdated facts. Learning and knowledge extraction in human beings is dynamic, and a human brain adapts to changes occurring in the environment continuously.

Interestingly, psychologists have played a major role in the development of machine-learning techniques. It has been a movement taken by computer researchers and psychologists together to make machines intelligent for more than six decades. The application areas are growing, and research done in the last six decades made us believe that it is one of the most interesting areas to make machines learn.
Machine learning is the study of methods for programming computers to learn. It is about making machines to behave intelligently and learn from experiences like human beings. In some tasks the human expert may not be required; this may include automated manufacturing or repetitive tasks with very few dynamic situations but demanding very high level of precision. A machine-learning system can study recorded data and subsequent machine failures and learn prediction rules. Second, there are problems where human experts exist and are required, but the knowledge is present in a tacit form. Speech recognition and language understanding come under this category. Virtually all humans exhibit expert-level abilities on these tasks, but the exact method and steps to perform these tasks are not known. A set of inputs and outputs with mapping is provided in this case, and thus machine-learning algorithms can learn to map the inputs to the outputs.

Third, there are problems where phenomena are changing rapidly. In real life there are many dynamic scenarios. Here the situations and parameters are changing dynamically. These behaviors change frequently, so that even if a programmer could construct a good predictive computer program, it would need to be rewritten frequently. A learning program can relieve the programmer of this burden by constantly modifying and tuning a set of learned prediction rules.

Fourth, there are applications that need to be customized for each computer user separately. A machine-learning system can learn the customer-specific requirements and tune the parameters accordingly to get a customized version for a specific customer.

Machine learning addresses many of the research questions with the aid of statistics, data mining, and psychology. Machine learning is much more than just data mining and statistics. Machine learning (ML) as it stands today is the use of data mining and statistics for inferencing to make decisions or build knowledge to enable better decision making. Statistics is more about understanding data and the pattern between them. Data mining seeks the relevant data based on patterns for decision making and analysis. Psychological studies of human learning aspire to understand the mechanisms underlying the various learning behaviors exhibited by people. At the end of the day, we want machine learning to empower machines with the learning abilities that are demonstrated by humans in complex scenarios. The psychological studies of human nature and the intelligence also contribute to different methods of machine learning. This includes concept learning, skill acquisition, strategy change, analytical inferences, and bias based on scenarios.

Machine learning is primarily concerned with the timely response, accuracy, and effectiveness of the resulting computer system. It many times does not take into account other aspects such as learning abilities and responding to dynamic situations, which are equally important. A machine-learning approach focuses on many complex applications such as building an accurate face recognition and authentication system. Statisticians, psychologists, and computer scientists may work together on this front. A data mining approach might look for patterns and variations in image data.

One of the major aspects of learning is the selection of learning data. All the information available for learning cannot be used as it is. It may contain a lot of data