Machine Learning stream evolved from the broad area of Artificial Intelligence, which adopts the intelligent capacity of humans. The main aim of machine learning is to use previous data or experience for solving a particular problem and to make machine “learn” in an intelligent way. The components of machine learning are used for last few years to solve real world problems. It includes the development of new learning algorithms and its applications in various disciplines. Machine Learning deals with the design and development of new algorithms based on various types of data and has wide range of applications in today’s world.

This book mainly focuses on improvement in machine learning algorithms and its use for real world applications. The book addresses state-of-the art solutions for many real-world problems in business, science and the engineering disciplines. Some of the applications include: image processing/retrieval, security, computer vision, handwritten characters recognition, bioinformatics, health, law, text analysis.

The book is intended for academics, researchers, scientists and professionals who are engaged in research and development in the area of machine learning. It provides the development of new algorithms and their implementations as well as recent advances in the field.

The book includes 21 chapters that are broadly organized into three sections. Section 1, titled Machine Learning Applications, contains nine book chapters and details new algorithms for solving real world problems in the field of image retrieval, computer vision, handwritten character recognition, medicine, bioinformatics, text analysis, and legal databases. These chapters cover detail information about methodology, experiments, and analysis of their results. Section 2, titled Computational Intelligence Techniques and Applications, contains six chapters and deals with novel algorithms and hybrid techniques in machine learning. Computational intelligent techniques such as artificial neural networks and fuzzy logic are proposed for image segmentation, biomedical image processing, floodplain filtration system, noisy data systems, classifier ensembles, and photocatalytic reactor. Section 3, titled Miscellaneous Techniques in Machine Learning, contains six chapters and details in the area of pattern recognition, nature-inspired toolbox to design and optimize systems, intelligent systems for recognition of cancerous cervical cells, ontology-based clustering of the web search results, decision tree induction algorithm, and network security.

Section 1: Machine Learning Applications

Chapter 1 is titled Machine Learning Approach for Content Based Image Retrieval. In this chapter, Kulkarni investigates the problem of efficient and fast searching of images based on their low level features. Colour is prominent feature for most of the images. This chapter reviews various image retrieval techniques based on computational intelligence such as fuzzy logic and neural networks and
proposes these techniques for posing a natural language query, classification, and fusion of multiple queries. Proposed technique was developed for a single query and later extended for multiple queries. It works effectively for searching specific images on large image database. Number of experiments was conducted, and promising results were obtained. Machine learning approach for content based image retrieval would be used for retrieving images in health, security, arts and many other areas.

Chapter 2 is titled Machine Learning Techniques in Handwriting Recognition: Problems and Solutions. In this chapter, Lee, Verma, Li, and Rahman focus on offline handwritten character recognition. Handwritten character recognition is an active research area for more than four decades, but some of the major problems are still unsolved. The proposed system gets a scanned handwritten page as input and provides an output in the form of editable recognized text. Problems and solutions for handwritten recognition techniques using machine learning have been identified in the chapter. The state of the art techniques and useful applications of handwriting recognition have been reviewed and presented. The binary segmentation with neural validation processes were discussed for improvement of the segmentation accuracy with expectation of better word recognition rate. In addition, very advanced architectural models, such as multiple experts systems and ensemble systems, have been discussed to achieve better performance and accuracy.

Chapter 3 is titled Semi Blind Source Separation for Application in Machine Learning. In this chapter, Naik and Kumar discuss semi Blind Source Separation (BSS) machine learning applications on audio and bio signal processing. BSS is one form of unsupervised learning and is a class of computational data analysis techniques for revealing hidden factors that underlie sets of measurements or signals. BSS assumes a statistical model whereby the observed multivariate data, typically given as a large database of samples, are assumed to be linear or nonlinear mixtures of some unknown latent variables. This chapter also establishes the applicability of BSS for biomedical applications and identifies the shortcomings related to order and magnitude ambiguity.

Chapter 4 is titled Machine Learning Algorithms for Analysis of DNA Data Sets. In this chapter, Yearwood, Bagirov, and Kelarev investigate that the applications of machine learning algorithms to analysis of data sets of DNA sequences are very important. This chapter discusses applications of several machine learning algorithms for the analysis of a JLA data set consisting of DNA sequences derived from non-coding segments in the junction of the large single copy region and inverted repeat A of the chloroplast genome in Eucalyptus. Machine learning algorithms based on local alignment scores achieve very good agreement with known biological classes for this data set. A new machine learning algorithm based on graph partitioning performed best for clustering of the JLA data set. Authors also claims that a new machine learning algorithm based on graph partitioning performed best for clustering of the JLA data set.

Chapter 5 is titled Machine Learning Applications in Radiation Therapy. In this chapter, Zhang, Meyer, Shi and D’Souza propose machine learning technique for health informatics. Cancer is one of major cause of death, and most patients go through radiation therapy during their treatment. Authors discuss the proposed use of computational techniques such as machine learning in radiation therapy for better diagnostic for obtaining tumor location for more accurate treatment. In addition, the chapter also reveals the review of recent advancement of machine learning applications in radiation therapy.

Chapter 6 is titled Insights from Jurisprudence for Machine Learning in Law. In this chapter, Strenieri and Zeleznikow focus on use of machine learning for legal database. The chapter provides basic information about approach of machine learning based on previous court cases and their outcomes. In addition, the chapter provides detailed literature for overview of machine learning techniques in law, jurisprudence concepts, and related case studies. It also introduces the novel concept of machine learning
in family law-split up. Limitations of knowledge discovery are extracted and proposed solutions based on various parameters are detailed in this chapter.

Chapter 7 is titled Machine Learning Applications in Computer Vision. In this chapter, Harandi, Taheri, and Lovell review recent advances in the area of automated object recognition mainly considering various learning frameworks for solving object recognition problems. Intelligent techniques such as reinforcement learning and machine learning are discussed, as well as their use for solving sequential decision problems. In addition, this chapter discusses local and global appearance models for object recognition and how similarities between objects can be learnt and evaluated.

Chapter 8 is titled Applications of Machine Learning for Linguistic Analysis of Texts. In this chapter, Torney, Yearwood, Vamplew, and Kelarev describe a novel multistage method for linguistic clustering of large collection of text available on the Internet. This method addresses the practicalities of applying clustering operations to a very large set of text documents by using a combination of unsupervised clustering and supervised classification. The method relies on creating a multitude of independent clustering of a randomized sample selected from the International Corpus of Learner English. Experimental results compare the performance of several classification algorithms incorporated in the multistage scheme and demonstrate that several of these classification algorithms achieve very high precision and recall and can be used in practical applications.

Chapter 9 is titled An Automatic Machine Learning Method for the Study of Keyword Suggestion. In this chapter, Chen describes two semantic analysis models to build a keyword suggestion system. The suggested keywords returned from the system have certain semantic relationship with similarity measure. The proposed system overcomes the problem of synonymy and polysemy over the information retrieval field by using vector space model. In addition, the benefit of semantic graph is to find the terms easily with semantic relationships by using a graph search. This system solves the problem of finding any two objects with a semantic relationship and calculating the similarity degree between these two objects.

Section 2: Computational Intelligence Techniques and Applications

Chapter 10 is titled Emergence Phenomenon and Fuzzy Logic in Meaningful Image Segmentation and Retrieval. In this chapter, Deb and Kulkarni discuss the problem of extracting meaningful image segmentation using emergence phenomenon. The chapter provides useful information about basics of emergence index, definitions, and its performance on various real world images. Extracting meaningful objects from images is a challenging task in content based image retrieval and helps to overcome the problem of semantic gap. Computational Intelligence techniques such as artificial neural networks and fuzzy logic are proposed for identifying the problem of image classification and similarity measure. Artificial neural networks have been proposed to classify images based on their structure, features, and objects. Fuzzy logic has been used to calculate the similarity between segmented objects and shapes and their colours.

Chapter 11 is titled Predicting Adsorption Behavior In Engineered Floodplain Filtration System Using Backpropagation Neural Networks. In this chapter, Rene, Behera, and Park propose artificial neural networks to approximate and interpret the complex input/output relationships, essentially to understand the breakthrough times in EFF (EFF). Authors claim that engineered floodplain filtration system is an eco-friendly low-cost water treatment process wherein water contaminants can be removed, by adsorption and-or degraded by microorganisms, as the infiltrating water moves from the wastewater treatment plants to the rivers. The chapter discusses various related areas such as wastewater treatment processes,
flood plain filtration for wastewater treatment, and mechanism and usefulness of flood plain filtration. The chapter is detailed with various experiments and their analysis.

Chapter 12 is titled *Computational Intelligence Techniques for Pattern Recognition in Biomedical Image Processing Applications*. In this chapter, Hemanth and Anitha first describe medical image classification and challenges in abnormality detection in the anatomy of the human body. The application of AI techniques for pattern recognition is explored in the context of abnormal Magnetic Resonance (MR) brain image classification. This chapter also illustrates the theory behind the AI techniques and their effectiveness for practical application in medical image classification. Algorithmic approach of neural networks, fuzzy theory, and genetic algorithm are also discussed in this chapter.

Chapter 13 is titled *A PSO-Based Framework for Designing Fuzzy Systems from Noisy Data Set*. In this chapter, Singh, Siani, and Khosla present a framework to design fuzzy logic systems using Particle Swarm Optimization (PSO). The chapter describes basics of fuzzy sets and systems along with lower and upper membership functions and fuzzy operators. PSO based framework is described using Nature-Inspired Toolbox in detailed steps and its application for designing fuzzy logic systems from noisy data.

Chapter 14 is titled *Neural Network Based Classifier Ensembles: A Comparative Analysis*. In this chapter, Verma presents the state of the art in classifier ensembles and their comparative performance analysis. Three types of neural classifier ensembles are considered and discussed. The first type is a classifier ensemble that uses a neural network for all its base classifiers. The second type is a classifier ensemble that uses a neural network as one of the classifiers among many of its base classifiers. The third and final type is a classifier ensemble that uses a neural network as a fusion classifier. The chapter reviews recent neural network based ensemble classifiers and compares their performances with other machine learning based classifier ensembles such as bagging, boosting, and rotation forest.

Chapter 15 is titled *Development of an Intelligent Neural Model to Predict and Analyze the VOC Removal Pattern in a Photocatalytic Reactor*. In this chapter, Krishnan, Rene, Lenskiy, and Swaminathan present an application of artificial neural networks to model the removal dynamics of an annular type photoreactor for gas. Volatile Organic Compounds (VOCs) belong to a new class of air pollutant that causes significant effect on human health and environment. Photocatalytic oxidation is an innovative, highly efficient, and promising option to decontaminate air polluted with VOCs, at faster elimination rates. The various steps involved in developing a suitable neural model have been outlined by considering the influence of internal network parameters on the model architecture.

**Section 3: Miscellaneous Techniques in Machine Learning**

Chapter 16 is titled *An Introduction to Pattern Classification*. In this chapter, Nakashima and Schaefer present an overview of pattern classification. The chapter discusses on the mathematical background of pattern classification rather than practical analysis of various pattern classification methods, and presents the derivation of classification rules from a mathematical aspect. Pre-processing and normalisation is often necessary in order to reduce the number of features and use only salient features that are important for classification. In distance-based pattern classification systems, a simple metric such as Euclidean distance between a training pattern and an unseen pattern is used to classify the unseen pattern. On the other hand, probabilistic assumption for patterns and classes are used in statistical pattern classification systems. This chapter also discusses statistical pattern classification from the viewpoint of minimizing classification costs. The property of classification boundaries were also explained for both distance-based and statistical pattern classification.
Chapter 17 is titled *Nature-Inspired Toolbox to Design and Optimize Systems*. In this chapter, Singh, Saini and Khosla discuss the fundamental concepts of Particle Swarm Optimization algorithms in the initial sections, followed by discussions and illustrations of benchmark optimization functions. Various modules of the Graphical User Interface (GUI) of Nature-Inspired Toolbox are explained with necessary figures and snapshots. In addition, the chapter also describes simulations results and comparative performance of various Particle Swarm Optimization models.

Chapter 18 is titled *Adaptive Intelligent Systems for Recognition of Cancerous Cervical Cells Based on 2D Cervical Cytological Digital Images*. In this chapter, Kwintiana and Roller claim that cancer of uterine cervix is still a leading cause of cancer-related deaths in women in the world. Due to scarce number of skilled and experienced cytologists, the screening procedure becomes time consuming and highly prone to human errors that leads to inaccurate and inconsistent diagnosis. This chapter provides detailed information about related studies and machine learning techniques, backpropagation learning algorithm. Experiments are conducted on various sets of images, and their detailed analysis is provided.

Chapter 19 is titled *Ontology-based Clustering of the Web Meta-Search Results*. In this chapter, Boidea, Lipai, and Dascalu present a meta-search tool developed in order to deliver search results structured according to the specific interests of users. Meta-search means that for a specific query, several search mechanisms could be simultaneously applied. Using the clustering process, thematically homogenous groups are built up from the initial list provided by the standard search mechanisms. The results are more user oriented, as a result of the ontological approach of the clustering process. After the initial search made on multiple search engines, the results are pre-processed and transformed into vectors of words. These vectors are mapped into vectors of concepts, by calling an educational ontology and using the WordNet lexical database. The vectors of concepts are refined through concept space graphs and projection mechanisms, before applying the clustering procedure. Implementation details and early experimentation results are also provided.

Chapter 20 is titled *A Beam Search Based Decision Tree Induction Algorithm*. In this chapter, Basgalupp, Barros, Carvalho, and Freitas claim that decision tree induction algorithms are highly used in a variety of domains for knowledge discovery and pattern recognition. They have the advantage of producing a comprehensible classification model and satisfactory accuracy levels in several application domains. This chapter presents a new algorithm that seeks to avoid being trapped in local-optima by doing a beam search during the decision tree growth. In addition, proposed strategy keeps the comprehensibility of the traditional methods and is much less time-consuming than evolutionary algorithms.

Chapter 21 is titled *Learning with Querying and its Application in Network Security*. In this chapter, Lai, Lin, Chang, and Kouh present a framework of query based learning concepts for different models of supervised and unsupervised learning. The chapter introduces a query based learning concept in which all the minds’ beliefs and actions are revised by observing the outcomes of past mutual interactions over time. Experiments show that the proposed method increases the accuracy detection rate for suspicious activity and significantly improve the efficiency of intrusion detection.

*Siddhivinayak Kulkarni*

*University of Ballarat, Australia*