

## Preface

In today's digital age, there is huge amount of data generated everyday. Deriving meaningful information from this data is a huge problem for humans. Therefore, techniques such as data mining whose primary objective is to unearth hitherto unknown relationship from data becomes important. The application of such techniques varies from business areas (Stock Market Prediction, Content Based Image Retrieval), Proteomics (Motif Discovery) to Internet (XML Data Mining, Web Personalization). The traditional computational techniques find it difficult to accomplish this task of Knowledge Discovery in Databases (KDD). Soft computing techniques like Genetic Algorithms, Artificial Neural Networks, Fuzzy Logic, Rough Sets and Support Vector Machines when used in combination is found to be more effective. Therefore, soft computing algorithms are used to accomplish data mining across different applications.

Chapter one presents introduction to the book. Chapter two gives details of self adaptive genetic algorithms. An iterative merge based genetic algorithms for data mining applications is given in chapter three. Dynamic association rule mining using genetic algorithms is described in chapter four. An evolutionary approach for XML data mining is presented in chapter five. Chapter six, gives a neural network based relevance feedback algorithm for content based image retrieval. An hybrid algorithm for predicting share values is addressed in chapter seven. The usage of rough sets and genetic algorithms for data mining based query processing is discussed in chapter eight. An effective web access sequencing algorithm using hashing techniques for better web reorganization is presented in chapter nine. An efficient data structure for personalizing the Google search results is mentioned in chapter ten. Classification based clustering algorithms using naive Bayesian probabilistic models are discussed in chapter eleven. The effective usage of simulated annealing and genetic algorithms for mining top- $k$  ranked webpages from Google is presented in chapter twelve. The concept of mining bioXML databases is introduced in chapter thirteen. Chapter fourteen and fifteen discusses algorithms for DNA compression. An efficient algorithm for motif discovery in protein

sequences is presented in chapter sixteen. Finally, matching techniques for genome sequences and genetic algorithms for motif discovery are given in chapter seventeen and eighteen respectively.

The authors appreciate the suggestions from the readers and users of this book. Kindly communicate the errors, if any, to the following email address: [venugopalkr@gmail.com](mailto:venugopalkr@gmail.com).

Bangalore,  
November 2008

K.R. Venugopal  
K.G. Srinivasa  
L.M. Patnaik

Soft Computing for Data Mining Applications  
Venugopal, K.R.; Srinivasa, K.G.; Patnaik, L.M.  
2009, XXII, 341 p., Hardcover  
ISBN: 978-3-642-00192-5