



Chapter VII

Bayesian Learning

Paula Macrossan
University of New England, Australia

Kerrie Mengersen
University of Newcastle, Australia

Learning from the Bayesian perspective can be described simply as the modification of opinion based on experience. This is in contrast to the Classical or “frequentist” approach that begins with no prior opinion, and inferences are based strictly on information obtained from a random sample selected from the population. An Internet search will quickly provide evidence of the growing popularity of Bayesian methods for data mining in a plethora of subject areas, from agriculture to genetics, engineering, and finance, to name a few. However, despite acknowledged advantages of the Bayesian approach, it is not yet routinely used as a tool for knowledge development. This is, in part, due to a lack of awareness of the language, mechanisms and interpretation inherent in Bayesian modeling, particularly for those trained under a foreign paradigm. The aim of this chapter is to provide a gentle introduction to the topic from the KDD perspective. The concepts involved in Bayes’ Theorem are introduced and reinforced through the application of the Bayesian framework to three traditional statistical and/or machine learning examples: a simple probability experiment involving coin tossing, Bayesian linear regression and Bayesian neural network learning. Some of the problems associated with the practical aspects of the implementation of Bayesian learning are then detailed, and various software freely available on the Internet is introduced. The advantages of the Bayesian approach to learning and inference, its impact on diverse scientific fields and its present applications are identified.

INTRODUCTION

Learning from the Bayesian perspective can be described very simply as the modification of opinion based on experience. The Bayesian approach to learning combines a prior subjective opinion with new information, in the form of data, to develop a revised opinion. This process can be repeated any number of times and occurs in a cyclical fashion, with the revised opinion becoming the prior opinion with the arrival of new data. This is in contrast to the classical, traditional or “frequentist” approach that begins with no prior opinion, and inferences are based strictly on information obtained from the data.

This chapter is intended to provide a gentle introduction of Bayesian methods by beginning with simple examples and explanations and later reviewing more complex applications in data mining. The concepts involved in Bayes’ Theorem are established with the example case of a simple coin-tossing experiment, after which the typical hierarchical methods used in Bayesian modeling are described. Graphical models, in which the essential structures implicit in a model are defined, are introduced as a convenient method of breaking down a complex model into simple components that provide the basis for Bayesian computation. Bayesian linear regression is examined to illustrate the Bayesian approach to a well-known statistical technique. The application of Bayesian inference to hybrid methods of machine learning is illustrated with the discussion of Bayesian neural network learning. Some of the problems and pitfalls associated with the practical implementation of the Bayesian framework are discussed, after which the reader is then introduced to powerful software for Bayesian inference and diagnostics freely available on the Internet. In conclusion, the present and predicted future impact of the Bayesian approach to learning and inference on diverse scientific fields is discussed. Some texts are recommended for further reading.

BACKGROUND

The essential characteristic of Bayesian methods is their explicit use of probability for quantifying uncertainty of parameter estimates of interest in scientific analysis. Bayesian statistics are used to analyse the plausibility of alternative hypotheses that are represented by probabilities, and inference is performed by evaluating these probabilities and making consequent decisions or estimates.

There has been a long-standing debate between Bayesian and classical statisticians regarding the approach to data analysis and knowledge acquisition. The interested reader can pursue this largely philosophical discussion by referring to Press and Tanur (2001). Recently there has been a revival of interest, fuelled by pragmatic rather than philosophical concerns, in the Bayesian approach in all areas of scientific research. It is no coincidence that this revival has occurred in parallel with the burgeoning of computer technology, allowing computationally intensive algorithms to be implemented easily and inexpensively. The most far-reaching of these algorithms is the Markov Chain Monte Carlo (MCMC) approach for simulat-

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/bayesian-learning/22152

Related Content

Hierarchical Clustering Using Evolutionary Algorithms

Monica Chis (2008). *Mathematical Methods for Knowledge Discovery and Data Mining* (pp. 146-156).

www.irma-international.org/chapter/hierarchical-clustering-using-evolutionary-algorithms/26138

Optimizing and Managing Digital Telecommunication Systems Using Data Mining and Knowledge Discovery Approaches

Adnan I. Al Rabeaand Ibrahim M. M. El Emary (2011). *Knowledge Discovery Practices and Emerging Applications of Data Mining: Trends and New Domains* (pp. 66-84).

www.irma-international.org/chapter/optimizing-managing-digital-telecommunication-systems/46891

Drivers of Innovation

Neeta Baporikar (2015). *Knowledge Management for Competitive Advantage During Economic Crisis* (pp. 250-270).

www.irma-international.org/chapter/drivers-of-innovation/117852

Formation of Faden Quartz Druses in Mid-Carboniferous Sandstones of the Donetsk Basin

Oleg Krisakand Vyacheslav Bezrukov (2018). *Dynamic Knowledge Representation in Scientific Domains* (pp. 155-162).

www.irma-international.org/chapter/formation-of-faden-quartz-druses-in-mid-carboniferous-sandstones-of-the-donetsk-basin/200175

The LBF R-Tree: Scalable Indexing and Storage for Data Warehousing Systems

Todd Eavis (2010). *Complex Data Warehousing and Knowledge Discovery for Advanced Retrieval Development: Innovative Methods and Applications* (pp. 1-27).

www.irma-international.org/chapter/lbf-tree-scalable-indexing-storage/39585