On Bias-Variance Analysis for Probabilistic Logic Models

Huma Lodhi, Imperial College London, UK

ABSTRACT

The article introduces bias-variance decomposition in probabilistic logic learning. We use Stochastic Logic Programs for probabilistic logic representation. In order to learn probabilistic logic models we use Failure Adjusted Maximization (FAM) that is an instance of the Expectation Maximization (EM) algorithm for first order logic. Experiments are carried out by concentrating on one kind of application: quantitative modelling of metabolic pathways that is a complex and challenging task in computational systems biology. We apply bias-variance definitions to analyze quantitative modelling of amino acid pathways of Saccharomyces cerevisiae (yeast). The results show the phenomenon of bias-variance trade-off in probabilistic logic learning.

Keywords: algorithms; artificial intelligence; data modeling; information accuracy; probabilistic systems; relational model; systems evaluation

INTRODUCTION

During the past few years, many methods have been proposed that are at the intersection of logic and probability. The distinguishing characteristic of these techniques is the provision of efficient representation for complex real world problem. For example, Stochastic Logic Programs (SLPs) (Muggleton, 1996) combine probabilistic models such as Stochastic Context Free Grammars (SCFG) and Hidden Markov Models (HMM) with first order logic. SLPs provide an efficient representation for complex biological systems such as metabolic pathways. Another competitive approach is termed Bayesian Logic Programs (BLPs) (Kersting &
De Raedt) that are considered as generalization of Bayesian nets and logic programs. Newly proposed Markov Logic networks (MLNs) (Richardson & Domingos, 2006) are also a useful example of combining probabilistic graphical models and logic.

More recently, there has been a growing interest in learning probabilistic logic representations. For instance, Failure Adjusted Maximization (FAM) (Cussens, 2001) is a useful tool to learn parameters in SLPs and Balios (Kersting & Dick, 2004) is a system that performs inference and learning in BLPs. While systems and techniques have been proposed for probabilistic logic learning, research has not been conducted in the important direction of analyzing the performance of probabilistic logic learners. Tools and methods to study the performance of probabilistic logical learning algorithms have not been investigated. In this article we focus on the unexplored research direction. We propose bias-variance (BV) decomposition to analyze and investigate the prediction (estimation) performance of the probabilistic logic learning algorithms for parameter estimation task.

In order to empirically analyze the BV definitions we focus on a challenging and fundamental task in computational systems biology, namely quantitative modelling of metabolic pathways. Recent research has shown critical importance of quantitative aspects of biological information stored in complex networks and pathways for the system level understanding of biology (Kitano 2002a, 2002b). We have conducted experiments using metabolic pathways in Saccharomyces cerevisiae. We have applied the proposed bias-variance definitions for estimating the rates of reactions catalyzed by enzymes in pathways using FAM.

The article is organized as follows. Section 2 briefly explains metabolic pathways. In Section 3 we describe logic programming concepts, SLPs and FAM. We present BV decomposition for parameter estimation task in Section 4. Experimental results are described in Section 5 and Section 6 concludes the article.

**METABOLIC PATHWAYS**

Metabolic pathways, an important class of biological systems, represent chemical reactions within the confines of a cell. They comprise metabolites and enzymes and may be viewed as series of enzyme-catalyzed reactions in which product of one reaction becomes substrate for the next reaction. Dynamics of biological system and behaviour of enzymes in metabolic pathways can be studied by applying the Michaelis-Menten (MM) enzyme kinetic function,
Information Technology Industry Dynamics: Impact of Disruptive Innovation Strategy
www.igi-global.com/chapter/information-technology-industry-dynamics/5520?camid=4v1a

Best Practices for Effective Virtual Teams
www.igi-global.com/chapter/best-practices-effective-virtual-teams/14247?camid=4v1a