

# Differential Diagnosis of Erythemato-Squamous Diseases Using Ensemble of Decision Trees

Mohamed El Bachir Menai and Nuha Altayash

Department of Computer Science,  
College of Computer and Information Sciences,  
King Saud University  
P. O. Box 51178, Riyadh 11543, Saudi Arabia  
{menai,nstayash}@ksu.edu.sa

**Abstract.** The differential diagnosis of erythemato-squamous diseases (ESD) in dermatology is a difficult task because of the overlapping of their signs and symptoms. Automatic detection of ESD can be useful to support physicians in making decisions if the model gives comprehensible explanations and conclusions. Several approaches have been proposed to automatically diagnosis ESD, including artificial neural networks (ANN) and support vector machines (SVM). Although, these methods achieve high performance accuracy, they are not attractive for dermatologists because their models are not directly usable. Decision trees can be converted into a set of if-then rules, which makes them particularly suitable for rule-based systems. They have been already used for the diagnosis of ESD. In this paper, we investigate the performance of boosting decision trees as an ensemble strategy for the diagnosis of ESD. We consider two decision tree models, namely unpruned decision tree and pruned decision tree. The experimental results obtained on UCI dermatology data set show that boosting decision trees leads to a relative increase in accuracy that attains 5.35%. Comparison results with other related methods demonstrate the competitiveness of the ensemble of unpruned decision trees. It performs 96.72% accuracy, which is better than those of some methods, such as genetic algorithms and  $K$ -means clustering.

## 1 Introduction

Psoriasis is a common dermatology disease that affects 2% to 3% people from the total population around the world [6]. Erythemato-squamous diseases (ESD) include six groups of skin diseases that share some psoriasis signs and symptoms with redness (erythema) causing loses in cells (squamous). ESD are psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis and pityriasis rubra pilaris. They all share the clinical features of erythema and scaling, and many histopathological features as well. The exact causes of these diseases remain in general unknown but they could be a combination of hereditary and environmental factors that can affect people in any ages.

The differential diagnosis of ESD is a difficult task in dermatology, which depends on the analysis of features obtained from the evaluation of both clinical

and histopathological features [7], and thus requires a detailed observation and experience to be correctly diagnosed. For example, pityriasis rubra pilaris is a rare disease that is usually mistakenly diagnosed in its first stages as psoriasis.

Medical decision support systems model the diagnosis task as a classification problem to assist physicians in making reliable and accurate decisions. As a result, several methods have been proposed for the classification of ESD, including genetic algorithms (GAs), support vector machines (SVM), artificial neural networks (ANN),  $K$ -means clustering, instance-based learning, decision tree learning, and other hybrid methods. Comprehensibility and interpretability are key factors of the usefulness of those methods. Indeed, methods that provide physicians with high diagnosis accuracy, but without explanations and conclusions, are not practically useful to physicians.

Decision tree learning is among the most practical methods for inductive inference. Moreover, decision trees can be converted to a set of if-then rules to improve their readability [11], which makes them suitable for rule-based systems, particularly in the medical domain. The purpose of this study is to investigate the performance of diagnosis models for ESD based on single decision tree and ensemble of decision trees (or forests) using boosting [5] as an ensemble strategy. The decision tree models considered are unpruned decision tree and pruned decision tree. The rest of this paper is organized as follows. The next Section describes the differential diagnosis of ESD. Section 3 outlines some related work. Section 4 describes the diagnosis models to be examined. Section 5 presents and discusses some experimental results. Section 6 concludes this work and outlines some future research directions.

## 2 Differential Diagnosis of ESD

A differential diagnosis of ESD is a systematic diagnostic procedure used by physicians to evaluate the differences between the diseases in order to identify the specific class of ESD in a patient. This procedure involves both a clinical and histopathological diagnostics for the evaluation of 12 clinical features and 22 histopathological features [6] described in Table 1.

The clinical diagnostic can distinguish between some ESD. For instance, the erythema and scaling are high in psoriasis and low in chronic dermatitis. Psoriasis, lichen planus and pityriasis rosea have the koebner phenomenon. Lichen planus has itching and polygonal papules where follicular papules are clinical feature for pityriasis rubra pilaris. Oral mucosa is usually seen in lichen planus, where the knee, elbow and scalp are usually preferred location of psoriasis. Family history is important in psoriasis. Pityriasis rubra pilaris is mostly seen during childhood [6]. The histopathological diagnostic consists in analyzing tissue samples under microscope to distinguish more accurately ESD. For instance, melanin incontinence is a diagnostic feature for lichen planus, fibrosis of the papillary dermis is for chronic dermatitis, exocytosis may be seen in lichen planus, pityriasis rosea and seborrheic dermatitis. Acanthosis and parakeratosis can be seen in all the diseases in different degrees. Follicular horn plug and perifollicular