Modeling Affective State using Learning Vector Quantization

de Vries, Jan

Hoofdstuk 1

GENERAL INTRODUCTION

Each research field has interesting challenges, but most fascinating challenges arise where multiple research fields meet. This thesis cross-fertilizes various research fields from life-sciences, including psychology and physiology, with computing science and artificial intelligence. By studying and embedding machine learning techniques in affective sciences, i.e., those that study human emotion, this thesis outlines advances in the recognition of human affect using various measurement modalities.

From a technical perspective, we employ and study learning dynamics of machine learning techniques. These methods can be subdivided into supervised and unsupervised learning. The former type aims at separating data from different classes by learning from labelled samples, the latter aims at finding optimal separation of unlabelled data samples into clusters that represent the overall data structure well. Unsupervised learning is beyond the scope of this thesis. Classification methods can be further subdivided into black-box and white-box methods. Black-box methods can be observed through input and output, but do not have an insightful internal structure. White-box methods, on the contrary, allow for more natural inspection of their internals and thereby provide insights into what data properties separate each class from the others. In this thesis, we primarily focus on Learning Vector Quantization (LVQ) which is a typical white-box method. Amongst the black-box methods are many of the popular classification techniques such as Support Vector Machine (SVM), Artificial Neural Network (ANN), k-Nearest Neighbors (kNN) as well as various ensemble methods such as Random Forest (RF). There have been developed techniques that can extract information from such classifiers (Tickle et al. 1998, Martens et al. 2007, Cortez and Embrechts 2013), however they do require significant additional analysis. For that reason white-box methods could be preferable over black-box methods, hence our interest to study the learning dynamics of LVQ. We will include several black-box methods in our experiments for comparison, given their popularity in other studies.

From the perspective of the application domain, we study the recognition of

various affective states from a variety of measurable quantities. The research on affect has a long history in psychology. The book by Charles Darwin titled *On the Expression of the Emotions in Man and Animals* is considered the pioneering work (Darwin 1872). Since then the topic has been studied widely in controlled (laboratory) conditions. With the recent technological developments that enabled miniaturization of measurement equipment also studies of emotion in daily life became feasible (Picard and Scheirer 1999, Westerink et al. 2009). Traditionally, psychological theories were developed and tested in experiments. In recent years the field of emotions was also discovered by computing science as an area with challenging problems and data driven studies, such as studied in this thesis became possible. The term *Affective Computing* for this multidisciplinary field was first introduced by Rosalind Picard in 1995 (Picard 1995), but the notion of monitoring one's emotional state with the help of computers originates in the 1970's (Kuechenmeister et al. 1970).

1.1 Scope of this study

Whereas a substantial part of affective research focusses on the fundamental research of developing theories on underlying mechanisms involved in emotion expression or detection, this study focusses on applying such knowledge in systems for affect recognition. Although this work is not the first to address affect recognition, it does comprise pioneering work in the application of white-box methods, in particular the prototype based LVQ methods to affective computing. The application of such methods allows for a deeper study of the aspects that most influence the automated affect recognition process and could potentially lead to a better understanding of the processes involved in human affect. In order to reflect the diversity in approaches to study affective phenomena, three of the following chapters will approach affective computing from different angles. In addition, we perform a theoretical study of various LVQ methods to better understand the learning dynamics involved.

1.2 Outline

In the following chapter we will further define and describe the technical and domain concepts that are studied in this thesis. After that, Chapter 3 will present a theoretical study of the performance and other properties of various LVQ variants

during the training phase of these classifiers. The three chapters following that will describe different affective classifiers using a variety of input modalities, thereby approaching affective computing from three different perspectives: Stress detection from physiological measurements reflecting the bodily perspective (Chapter 4), facial expression recognition using facial images reflecting the facial perspective (Chapter 5), and emotion detection using mental appraisals reflecting the cognitive perspective (Chapter 6). Chapter 7 will describe potential applications that make use of the various affective classifiers. Finally, a summary will be given in Chapter 8.