Abstract

The thesis is based on a series of papers reporting the studies on the application of selected Artificial Intelligence (AI) methods to time series analysis. It consists of 5 published papers and one submitted manuscript.

Over recent years Artificial Intelligence has attracted a great deal of interest from both academia and industry. AI in a way connects the beauty of mathematics and the utility of computer science and this, probably, is a key factor of its recent success. Also, AI is fed by data and nowadays, we are collecting a tremendous amount of data every minute. With the increase of variety of data types, there is a need to develop specialized algorithms. One of such data types are time series and they will be the main topic of this dissertation.

The thesis covers several topics related to time series. Much experimental effort has been made in the area of distance measures for time series. The problem of the choice of the proper distance measure is present in many different time series tasks as classification, clustering, anomaly detection or motif discovery to name only a few. To the best of my knowledge, there exist over 50 distance measures proposed so far. As a result of joint work, 2 papers on this topic were published and 1 was submitted, in which we created the biggest available overview, comparing 55 distance measures on 128 benchmark data sets.

The next part of the thesis deals with Functional Data Analysis (FDA) which considers each sample element to be a function. This is an interesting approach, that lets us work with time series represented in a novel way and thus benefit from it. Moving from a universe of discrete data points to a universe of function, the need of adapting existing methods arises. Within this area, I worked on relationships between multivariate functional data, defining the functional Procrustes coefficient.

The third part of the dissertation is an application study in chemometrics. Modern appliances as spectrophotometer can provide us data describing the spectral specification of substances, which can be seen as time series. My research work showed the usage of data analysis tools in evaluating the quality

of apple juice.

The last part of my research activity was devoted to a general classification problem. I observed, inspired with Random Forest classifier, that we can propose a relatively simple but powerful extension of Neural Network classifier – RAndom Neural Networks. (RANNs). The initial results show that RANNs outperform other popular classifiers (such as Random Forest, SVC or XGBoost) on a selected benchmark data sets.

Keywords: artificial intelligence, time series, distance measures, functional data, chemometrics

Paweł Piasecki