Anomaly detection using Long Short Term Memory Networks and Support Vector Data Description: application to consumer demand in supply chain management

May 13, 2019

Abstract

Anomaly detection refers to the problem of finding patterns in data that do not comply with an expected behavior. We can consider different kinds of applications like intrusion detection, fraud detection, fault detection, system health monitoring, image processing and sensor networks. For example, an anomalous traffic pattern in a computer network could mean that a hacked computer is sending out sensitive data to an unauthorized destination. An extensive survey covering both the anomaly detection techniques and its applications has been conducted in Chandola et al. Recently, Mehrotra et al. provided a comprehensive research on anomaly detection principles and algorithms.

In this project, we will focus on techniques that detect anomalies in Supply Chain Management (SCM) and more specifically in consumer demands. Supply chains are characterized by huge amounts of data corresponding among others to a lot of different order cycles (products, locations...). Monitoring this information is a key factor for companies evolving nowadays in a globally competitive market.

As discussed in Bontemps et al., most of the current studies on anomaly detection do not consider recent/past events to detect possible new incoming outlier. However, complex time dependence structure of consumer demands is most of the time observed in supply chains, see Bousqaoui et al.. Using Long Short Term Memory (LSTM) networks is then proposed to deal with challenges associated with time dependent anomaly detection problems. In deep learning, a LSTM network is a special type of recurrent artificial neural network that accounts for the long-term dependencies between data nodes. In the literature, most applications of LSTM networks are in predictive analytics. For example, a stacked LSTM network is proposed for anomaly prediction of space shuttle engine in Malhotra et al.. The bi-directional LSTM is applied to capture long-term dependence for machining tool wear prediction in Zhao et al.. Wu et al. suggest to use a vanilla LSTM to estimate the remaining useful
life of an aircraft turbofan engine under complex operating conditions and strong background noise. In the field of SCM, Bousqaoui et al. propose a LSTM-based prediction model for demand forecasting. A brief review of the studies related to applying LSTM to time series and anomaly detection issues is presented in Bontemps et al. In this context, Malhotra et al. suggest to use stacked LSTM networks; however, they assumed a multivariate Gaussian distribution for the error vectors, which is not often realistic.

In this master thesis, we will thus consider a LSTM-based prediction model for consumer demands. Anomalies will then result from too large prediction errors. Several tools will be considered to determine thresholds defining detection limits in this context. Some of them are outlined in the sequel. The aim of Support Vector Data Description (SVDD) is to describe a realistic domain for the data, excluding superfluous space. The resulting boundary can then be used to detect outliers. The method is inspired by support vector classifiers: kernel functions allow for flexible boundary as in support vector machines. The Hotelling’s $T^2$ control chart is also a tool that helps to detect outliers, see Montgomery; it however assumes normal data. Other nonparametric methods can also be applied: in particular, quantiles derived from a kernel density estimator or the smoothing technique developed by Sheather and Marron can be considered. Comparison between the different techniques and the possible interests of using SVDD, which has not been studied in the above context so far, should be discussed.

**Keywords:** Long short term memory networks, Anomaly detection, Supply chain management, Support vector data description.

**References**


