Early Warning System Framework Proposal, Based on Big Data Environment

Goran Klepac, Raiffeisenbank Austria, Magazinska, Croatia
Robert Kopal, University College Algebra Zagreb, Zagreb, Croatia
Leo Mrsic, University College Algebra Zagreb, Zagreb, Croatia

ABSTRACT

Early warning systems are made with the purpose to efficiently recognize deviant and potentially dangerous trends related to company business as early as possible. The Big Data environment gives new opportunities and new approaches in analytical processes. There are numerous ways how to set up early warning systems within a company. The Big Data environment forces companies to apply new ways of thinking and use new disposable data sources. This article gives a novel concept for an early warning system design within a company, which is applicable in different industries. The core of the proposed framework is a hybrid fuzzy expert system which can contain a variety of data mining predictive models responsible for some specific areas as addition to traditional rule blocks. It can also include social network analysis metrics based on linguistic variables and incorporated within the rule blocks. As a part of this framework, SNA methods are also explained and introduced as powerful and unique tool to be used in modern early warning systems.

KEYWORDS

Big Data Environment, Early Warning Systems, Hybrid Fuzzy, Social Network Analysis Metrics

1. INTRODUCTION

The Big Data environment gives new opportunities in analytics. From one hand it demands new way of thinking in combination of unstructured data sources. Unstructured data sources bring new approaches and opportunities in better understanding business areas like early warning systems. Traditional approach in business analytic and usage of data mining methods are mostly concentrated on solving single problem like early warning system segmentation, fraud detection, churn detection on data disposable within local databases. Those data could be huge by volume, collected through transactional databases and consolidated through data warehouses. Even it seems complex and sufficient, it often represents narrow set of data for specific purposes like segmentation, fraud detection, churn detection, etc. Reason for that lays in fact that transactional databases are not constructed, or has been rarely constructed with intention to fulfill analytical needs. Especially, it is evident for some specific needs like fraud detection or churn detection. Transactional databases are constructed
with general scope, and it is transactional business needs. It does not mean that such databases are useless for analytical purposes; it simply means that it provides narrow set of information with which some predictive analytical data mining model or other model or reporting for specific business needs could be constructed.

With the value of the existing data, dislocated within different transactional systems could be increased by integration into data warehouse system. It still does not mean that company does not have limited information about some problem space. Other problem in relation with traditional analytical approach is often avoidance of unstructured data source usage for business modeling purposes, even unstructured data exists within systems like call centers data or similar sources. Unstructured data usage is mainly focused on pure text mining analytics in connection with specific problems based on textual data sources, or in web mining analytics. Integration of structured and unstructured data sources as well as internal or external data sources (from web, blogs, social networks) is not the case when we are speaking about traditional business analytics. All that leads us to the question about potential business strategy based on analytical models. Data mining is defined as a discipline, which reveals hidden unexpected patterns from the data. From perspective of business planning, it should be in service of choosing efficient business strategy. Traditional approach offers revealing hidden unexpected patterns from the internal, mostly structured, relational data sets, and result from those models could be plausible and usable for decision-making. Internal data, collected within company data warehouse has (could have) limited scope on some problematic like fraud, because it does not contain enough relevant information for efficient fraud detection modeling and analytic. Developed model (in this case fraud detection model based on internal data warehouse) could show plausible performance, but relevant patterns could stay undetected, because other data sources like blogs, social networks, web sources are not used for analytical purposes. This chapter is not criticism on traditional approach in business analytics especially on data mining analytics. This chapter is concentrated on additional benefits, in way when external data sources as well as unstructured data sources are also used in analytical process with complex analytical techniques, and how all that could improve strategic business decision making. All this fact leads us to conclusion, that spreading views on analytical side as consequence has wider horizons on business decision side. Managers gets more opportunities for better planning, as well as making requests to analysts for making much more complex and much more reliable models. Big data phenomena are not area only based on data volume; it also covers problematic of using adequate methodology which links structured and unstructured data sources. It also covers problematic how to make adequate connection between different data sources, where data like primary and secondary key does not exist, and it is impossible to make solid connection between internal and external data sources. Most important thing from the perspective of business decision making is to design conceptual solution based on business needs, which will provide valuable information and signals relevant for solving business problems.

Modern global environment includes company’s constant exposure to market influences of different kinds. While in distant past company’s analytic value was measured by its ability to find information, recent past introduce value as ability to collect data. Modern decade fights with problem how to interact and efficiently use large available data collections. Different industries up until today developed large scale analytic methods used for market, customer and profitability analysis and forecasting. Ability to “understand” data still remains one of top analytic priorities. However, due to development of industry specific expert knowledge and due to fact that many industries face significant changes in past years followed by globalization and technology evolutions, concept of developing black box indicator models is more and more used every day. Those models combine industry and expert knowledge but are easy to use. Described as Early Warning Systems (EWS) those models are packed with state-of-the-art knowledge and KPI’s which helps business people to deal with numerous influences, large data sets and market trends. Early Warning Systems are also in close connection with risk assessment and modeling. Risk management methods often precede development phase and are used as starting point for early warning systems development.
SGO A New Approach for Energy Efficient Clustering in WSN
www.igi-global.com/article/sgo-a-new-approach-for-energy-efficient-clustering-in-wsn/214868?camid=4v1a

Guiding Self-Organization in Systems of Cooperative Mobile Agents
www.igi-global.com/chapter/guiding-self-organization-systems-cooperative/4983?camid=4v1a