Chapter 8
A Study on Various Applications of Data Mining and Supervised Learning Techniques in Business Fraud Detection

Amit Majumder

JIS College of Engineering, India

Ira Nath

JIS College of Engineering, India

ABSTRACT

Data mining technique helps us to extract useful data from a large dataset of any raw data. It is used to analyse and identify data patterns and to find anomalies and correlations within dataset to predict outcomes. Using a broad range of techniques, we can use this information to improve customer relationships and reduce risks. Data mining and supervised learning have applications in multiple fields of science and research. Machine learning looks at patterns of data and helps to predict future behaviour by learning from the patterns. Data mining is normally used as a source of information on which machine learning can be applied to solve some of problems in our daily life. Supervised learning is one type of machine learning method which uses labelled data consisting of input along with the label of inputs and generates one learned model (or classifier for classification type work) which can be used to label unknown data. Financial accounting fraud detection has become an emerging topic in the field of academic, research and industries.

1. INTRODUCTION

Fraud is measured as an intentional thing that searches for guarantying its writer or a third party for giving an unlawful advantage, to the disadvantage of an individual. At the business level, it can be specified through changed economic data, misuse of properties, unsuitable expenditures and revenue, amongst others. The significances of a business fraud are financial losses, image and disbelief of clients and stockholders.

1.1 Introduction on Business Fraud

Corporate fraud indicates to happenings commenced by a personal or company that are executed in an untruthful or unlawful way and are calculated to give a benefit to the perpetrating individual or company. Corporate fraud schemes go beyond the scope of an employee's stated position and are marked by their complexity and economic impact on the business, other staffs, and outside parties.

Some examples of business frauds that have happened in the world have been: Enron, the energy company that tried to hide its true level of indebtedness through complex transactions with its subsidiary companies; WorldCom, a telecommunications firm that recorded expenditures of close to 4,000 million dollars as delayed overheads, in order to hide its actual losses; Global Crossing, a fiber optic company that recorded revenue as revenue that should be deferred; Vivendi, European communications entity that manipulated its Ebitda in the consolidation of its data with the business group, in order to acquire superior debt capacity; and Adelphia Communications, cable television society that was undercapitalized by granting enormous individual loans to its members with the agreement of its directors. These, among many others, are some of the most tarnished cases in the world of business fraud.

1.2 Introduction on Data Mining

Data Mining **is** a set of technique that applies to huge and composite databases. This is to eradicate the arbitrariness and determine the unseen outline. These *data mining techniques* are almost always computationally rigorous. We usage data mining tools, practices, and concepts for revealing designs in information. There are too many driving militaries existing. And, this is the cause why data mining has developed such a vital zone of education.

As data mining is having spacious applications. Thus, it is the young and promising field for the present generation. It has attracted a great deal of attention in the information industry and in society. Due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge. Thus, we use information and knowledge for applications ranging from market study. This is the cause why data mining, recognized as knowledge detection from information.

1.3 Introduction on Machine Learning Using Supervised Approach

Machine Learning is indisputably one of the most powerful and influential skills in today's world. More significantly, we are distant from watching its complete possibilities. There is no uncertainty that it will continue to be making headlines for the foreseeable future. This article is designed as an introduction to the Machine Learning concepts, covering all the fundamental ideas without being too high level.

Related Content

Analysis and Implications of Adopting AI and Machine Learning in Marketing, Servicing, and Communications Technology

Priyal J. Borole (2024). *International Journal of Artificial Intelligence and Machine Learning (pp. 1-11).*
www.irma-international.org/article/analysis-and-implications-of-adopting-ai-and-machine-learning-in-marketing-servicing-and-communications-technology/338379

Augmented Reality and Its Significance in Healthcare Systems

Ashish Tripathi, Nikita Chauhan, Arjun Choudharyand Rajnesh Singh (2023). *Meta-Learning Frameworks for Imaging Applications (pp. 103-118).*
www.irma-international.org/chapter/augmented-reality-and-its-significance-in-healthcare-systems/331479

A Method Based on a New Word Embedding Approach for Process Model Matching

Mostefai Abdelkaderand Mekour Mansour (2021). *International Journal of Artificial Intelligence and Machine Learning (pp. 1-14).*
www.irma-international.org/article/a-method-based-on-a-new-word-embedding-approach-for-process-model-matching/266492

Breast Cancer Disease Exploitation to Recure a Healthy Lifestyle

Mudrakola Swapnaand Nagaratna P. Hegde (2023). *Encyclopedia of Data Science and Machine Learning (pp. 2617-2635).*
www.irma-international.org/chapter/breast-cancer-disease-exploitation-to-recure-a-healthy-lifestyle/317700

MHLM Majority Voting Based Hybrid Learning Model for Multi-Document Summarization

Suneetha S.and Venugopal Reddy A. (2019). *International Journal of Artificial Intelligence and Machine Learning (pp. 67-81).*
www.irma-international.org/article/mhlm-majority-voting-based-hybrid-learning-model-for-multi-document-summarization/233890