LSTM-BASED AUTOMATED LEARNING WITH SMART DATA TO IMPROVE MARKETING FRAUD DETECTION AND FINANCIAL FORECASTING

Zair Bouzidi [1]
Abdelmalek Boudries [2]
Mourad Amad [3]

DOI: https://doi.org/10.31410/EMAN.2021.191

Abstract: *This proposed model is based on a deep recurrent neural network trained with Long Short-Term Memory Network (LSTM), used because of its ability to learn long term dependencies, taking the concatenated function and Financial data as input, while integrating encapsulations, using Smart Data and retrieving information by combining multiple search results (all the Web). It combines representation training with financial data while integrating encapsulations from multiple sources and retrieving information by combining multiple search results. It provides some good ideas that we have extended to improve Corporate Marketing and Business Strategies. We show that the proposed model learns to localize and recognize different aspects of Corporate Marketing and Business Strategies. We evaluate it on the challenging task of detecting Fraud in Financial Services and Financial Time Series Forecasting and show that it is more accurate than the state-of-the-art of other neural networks and that it uses fewer parameters and less computation.*

Keywords: *Business, Marketing, Forecasting of financial times series, Fraud detection, LSTM, Smart Data.*

1. INTRODUCTION

The prime goal of a financial time series model is to provide reliable future forecasts which are crucial for investment planning, fiscal risk hedging, governmental policy making, etc. These time series often exhibit notoriously haphazard movements which make the task of modeling and forecasting extremely difficult. As per the research evidence, the random walk (RW) (Fama, 1995) is so far the best linear model for forecasting financial data. Artificial neural network (ANN) is another promising alternative with the unique capability of nonlinear self-adaptive modeling. Numerous comparisons of the performances of RW and ANN models have also been carried out in the literature with mixed conclusions (Adhikari, 2014).

We propose a new real-time automated learning model based on a recurrent neural network trained with Long Short-Term Memory Network (LSTM) that integrates encapsulations using Smart Data and thus retrieves information by combining multiple search results from multiple sources (all the Web). Thus, we provide not only a solution to this challenge, but also, propose better performances.

[1] LIMPAF Laboratory, Computer Science Dept, Faculty of Science and Applied Science, Bouira University, Algeria

[2] Laboratory LMA, Commercial Science Dept, Faculty of Economics, Business and Management, Bejaia University, Algeria

[3] LIMPAF Laboratory, Computer Science Dept, Faculty of Science and Applied Science, Bouira University, Algeria

In this paper, we try to identify relevant content dealing with financial time series. Once this information is retrieved (distinguished, of course, from large amounts of other content and also distinguished from abusive information), it can be used to improve Corporate Marketing, Business Strategies, Fraud Detection in Financial Services and Financial Time Series Forecasting. Our main contributions are listed below.

1. We develop a recurrent neural network-based model that uses low-level content learning capabilities to automatically separate relevant information from redundant or abusive.
2. We develop a recurrent neural network-based model that uses content learning capabilities of multiple sources (all online channels from social media to websites) to automatically and efficiently capture real-time dynamics of financial data. Using a set of knowledge related to the financial market, this model collects, using Smart data, according to their lexical similarity, accurate forecasting of volatility from financial time series.
3. We have adapted some algorithms to streaming to get Smart Data. Smart Data, a different concept of Big Data, even in opposition to it, is based primarily on real-time data analysis. This term refers to an approach to data analysis that directly analyzes the data at the source, without the need to transmit it to a centralized system. Big Data is the mass of information circulating via the web, connected objects or smartphones, while Smart Data can be defined as the intelligent and relevant way of processing data.
4. Keeping in mind the limitations of the previous work, we develop an event-independent model that can be used directly to filter content on multiple sources at a time in future events. Experiments on multiple financial market-related content flow with diverse characteristics show that our proposed model outperforms forecasting-based approaches. Our approach filters content issued from all online channels from social media to websites.
5. Once we have developed this real-time recurrent neural network-based model, and annotated manually the first information deducted by the recurrent neural network, using multi-source content learning capabilities to automatically and efficiently capture real-time accurate forecasting of volatility from financial time series, using a set of knowledge related to the financial market and a set of tagged contents we collected reliable future forecasts which are crucial for investment planning, fiscal risk hedging, governmental policy making, etc.

The rest of this paper is organized as follows: Section 2 presents the background and related works. In section 3, we describe our proposed model, we provide details on accurate forecasting of volatility from financial time series, and preliminary results, followed by a discussion of the results obtained. Finally, we conclude and give some future works.

2. BACKGROUND & RELATED WORKS

2.1. Corporate Marketing, Fraud Detection in Financial Time Series Forecasting

With a wide variety of products and buying behaviors, the shelves on which products are presented are one of the most important resources in the retail environment. Retailers can not only increase profits but also reduce costs by properly managing shelf allocation and product display (Aloysius, 2013). Using learning models in the organization of shelves in supermarkets by grouping products that are usually bought together, we can extract the following relation: *customers who buy the product X at the end of the week, during the summer, generally also buy the product Y.*

Also, the credit institution, that permits to decide whether or not to grant credit based on the profile of the credit applicant, his / her request, and past loan experiences, is used in data mining. There are also the overbooking (optimization of the number of seats in planes, hotels, etc.), the targeting of offers (organization of advertising campaigns, promotions) and the analysis of business practices, strategies and their impact on sales in Data Mining. This knowledge, unknown at first, may be correlations, patterns, or general trends in that data. Experimental data are necessary to verify the correction of the system or the estimation of some difficult parameters to mathematical modeling. Data Mining is a field that has emerged with the explosion of the amount of information stored, with significant progress notably in processing speeds and storage media. The purpose of data mining is to discover, in large amounts of data, valuable information that can help understand the data or predict the behavior of future data. Since its inception, data mining has used several tools for statistics and artificial intelligence to achieve its objectives. It is an essential component of Big Data technologies, large data analysis techniques and recently data smart streaming. It is often defined as the process of discovering new knowledge by examining large amounts of data (stored in warehouses or streaming) using pattern recognition technologies as well as statistical and mathematical techniques.

Table 1. Comparative table of all economic tasks used

Economic Tasks	**References**	**Our New Approach (ONA)**
Detect novel frauds	(Pumsirirat, 2018), (Schreyer, 2017), (Wang, 2018), (Zheng, 2018), (Dong, 2018), (Gomez, 2018), (Ryman-Tubb, 2018), (Fiore, 2019)	ONA
Trading performance	(Sermpinis, 2019)	
Exchange rate prediction	(Calvez, 2018)	
Stock prediction	(Kodogiannis, 2002)	
Trade on the stock market	(Fischer, 2018)	
Company stock prices	(Pinheiro, 2017)	
Forecasting of financial time series	(Bodyanskiy, 2006), (Lai, 2006), (Ghazali, 2009), (Pradeepkumar, 2017), (Tk, 2016), (Lasfer, 2013), (Gudelek, 2017), (Mohammad, 2018)	

Source: Authors' research

Extraction of element set is the model where some sequence exploration problems lend themselves to the discovery of frequent item sets and their order. For example, we look for rules of the type "if a customer buys a car, it is likely to take out insurance within one week"}, or in the context of stock prices, "if A up and B up, it is likely that C up and D up within one week". Traditionally, the extraction of element sets is used in marketing applications to detect patterns among competing elements in large transactions, for example, by analyzing customer shopping basket transactions in a supermarket, according to Han et al. (Han, 2007).

The extraction of sequential models has many real applications because the data is encoded as sequences in many areas such as bio-informatics, genomics and Proteomics (Fournier-Viger, 2017). We also see the development of market basket analysis which consists of studying sales (Sales receipt analysis) (Masseglia, 2005).

Web Mining is the next model. Thanks to the huge amount of information available online, the World Wide Web is a fertile ground for research in data mining. Web mining research is at the crossroads of research conducted by several research communities, such as databases, information retrieval and within Artificial Intelligence (AI), particularly the sub-domains of learning and natural language processing (Kosala, 2000).

Text Mining is a branch of Data mining that specializes in the processing of text corpora to analyze the content and extract knowledge. The main tasks to be accomplished are the recognition of the information presented in the document and its interpretation. It refers to linguistic technologies that make it possible to switch from text (full text) to a digital vector (presence-absence or frequency).

String Mining is the next model that channel exploration usually deals with a limited alphabet of elements that appear in a sequence, but the sequence itself can usually be very long. Examples of alphabets may be those of the ASCII character set used in the natural language text, nucleotide bases “A”, “G”, “C” and “T” in DNA sequences or amino acids for protein sequences and examining gene and protein sequences to determine their properties (Abouelhoda, 2009).

The last model Sequential Pattern Mining consists of discovering unexpected and useful models in data sets. It consists of discovering interesting sub-sequences in a set of sequences, where the interest of a sub-sequence can be measured according to various criteria such as its frequency of appearance, its length and its profit.

Fraud detection is an interesting problem in that it can be formulated in unsupervised and/or supervised classification. In unsupervised learning category, class labels are either unknown or assumed to be unknown, and clustering techniques are employed to figure out (i) distinct clusters containing fraudulent samples or (ii) far off fraudulent samples that do not belong to any cluster, where all clusters contained genuine samples, in which case, it is treated as an outlier detection problem. In the supervised learning category, class labels are known and a binary classifier is built in order to classify fraudulent samples. Fraud (including cyber fraud) detection is increasingly becoming menacing and fraudsters always appear to be few notches ahead of organizations in terms of finding new loopholes in the system and circumventing them effortlessly. On the other hand, organizations make huge investments in money, time and resources to predict fraud in near real-time, if not real time and try to mitigate the consequences of fraud. Financial fraud manifests itself in various areas such as banking, insurance and investments (stock markets). It can be both offline as well as online. Online fraud includes credit/debit card fraud, transaction fraud, cyber fraud involving security, while offline fraud includes accounting fraud, forgeries, etc.

Advances in technology and breakthrough in deep learning models have seen an increase in intelligent automated trading and decision support systems in financial markets, especially in the stock and foreign exchange (FOREX) markets.

However, time series problems are difficult to predict especially financial time series (Cavalcante, 2016). On the other hand, Neural and Deep learning models have shown great success in forecasting financial time series (Li, 2009) despite the contradictory report by efficient market hypothesis (EMH) (Fama, 1995), that the FOREX and stock market follows a Random Walk (RW) and any profit made is by chance. This can be attributed to the ability of Neural Networks to self-adapt to any nonlinear data set without any static assumption and prior knowledge of the data set (Lu, 2009). Deep learning used both fundamental and technical analysis data, which is the two, most commonly, used techniques for financial time series forecasting, to train and build deep learning models (Cavalcante, 2016).

Fundamental analysis is the use or mining of textual information like financial news, company financial reports and other economic factors like government policies, to predict price movement.

2.2. Automated Learning

Learning is a set of mechanisms leading to the acquisition of know-how and knowledge. While Automated learning is a branch of Artificial Intelligence (AI) that deals with the development of algorithms that make capable of accomplishing complex tasks without having been explicitly programmed for that purpose, making extensive use of tools and concepts of AI, mathematics, other cognitive sciences and so on. It can rely on statistical approaches to give the ability to "learn" from data using two phases. The first one, namely Model Design Phase (Training) consists of estimating a model from data, called observations. The second one is the Production Phase where the model is being determined; new data can then be submitted to obtain the result corresponding to the desired task. Handwriting recognition is a good and complex example because two similar characters are never exactly equal. An automatic learning system can be designed to learn to recognize characters by observing "examples", that is, known characters.

Table 2. Comparative table of all AI Concepts used

AI Concepts	References	Our New Approach (ONA)
Elman ANN (EANN)	(Adhikari, 2014), (Bouzidi, 2018), (Bouzidi, 2019)	ONA
Multi-Layer Feed-forward (MFF)	(Adhikari, 2014), (Tk, 2016), (Pandey, 2018), (Bouzidi, 2020)	
ConvNets/Autoencoder	(Pinheiro, 2017), (Lai, 2006), (Lasfer, 2013), (Gudelek, 2017), (Bao, 2017)	
RNN	(Bouzidi, 2020b), (Ghazali, 2009), Mohammad2018)	
LSTM	(Adhikari, 2014), (Cavalcante, 2016), (Sermpinis, 2019), (Calvez, 2018), (Fischer, 2018), (Pradeepkumar, 2017), (Tk, 2016), (Gudelek, 2017), (Pandey, 2018)	
Memory Networks	(Adhikari, 2014), (Cavalcante, 2016), (Sermpinis, 2019), (Calvez, 2018), (Fischer, 2018), (Pradeepkumar, 2017), (Tk, 2016), (Gudelek, 2017), (Pandey, 2018)	
Social Media	(Adhikari, 2014), (Cavalcante, 2016), (Sermpinis, 2019), (Calvez, 2018), Fischer, 2018), (Ghazali, 2009), (Pradeepkumar, 2017), (Tk, 2016), (Gudelek, 2017), (Pandey, 2018)	
Analysis	(Sermpinis, 2019), (Pradeepkumar, 2017), (Tk, 2016), (Lasfer, 2013), (Mohammad, 2018), (Pandey, 2018)	

Source: Authors' research

According to the information available during the learning phase, it is qualified in different ways. If the data is tagged, it is supervised learning. We are talking about "classification" if the labels are discrete, or "regression" if they are continuous. If the model is learned incrementally based on a reward, it is called "reinforcement learning". When data (or "tags") are missing, the model must use untagged examples that can still provide information. In medicine, for example, it can be an aid to diagnosis. It is said that learning is "semi-supervised". While the labeling of data is partial when a model states that a data does not belong to a class A, but perhaps to a class B or C (A, B and C is 3 diseases for example evoked as part of a differential diagnosis). This is called *partially supervised learning*. In the most general, unlabeled case, we try to determine the underlying structure of the data: this is *unsupervised learning* (Abouelhoda, 2009).

Automated learning is used for a wide range of applications, such as diagnostic aid (Abouelhoda, 2009), outlier detection, missing data detection, relevant information retrieval from multiple sources (social media) (Bouzidi, 2018), (Bouzidi, 2019), (Bouzidi, 2020), & (Bouzidi, 2020b), fraud detection, financial market analysis (Bouzidi, 2020b) & (Masseglia, 2005) and so on.

It is not just about a set of algorithms, but a list of steps to take into account and execute in order to reach an optimal result. Data Acquisition is the first step of this list where the algorithm feeds on input data and where the success of the project is collecting relevant data and in sufficient quantity. Preparation and cleaning of the data is the second step. The third is the creation of the model. The fourth step is Evaluation which consists of evaluating the trained model on the other (second) set of data. The fifth is Deployment where the model will be deployed in production to make predictions, and potentially use new input data to re-train and improve its model.

However, care must be taken to use an adequate number of neurons and hidden layers, to detect and thus avoid over-learning. Thus, the data is divided into two subsets (LeCun, 2015): the learning set which allows changing the weight of the neural network. The validation set allows verifying the relevance of the network while avoiding over-learning.

2.3. Social Networks

Social networking forms an important part of the online activities of Web users. There are two types of social networks: Centralized social networks and Decentralized social networks. Current Online social networks (OSN) are Web services run on logically centralized infrastructure (Datta, 2010). They use content distribution networks and thus distribute some of the load by caching for performance reasons, while keeping a central repository for user and application data. This centralized nature of Online Social Networks has several drawbacks including scalability, privacy, dependence on a provider, and the need for being online for every transaction (Yeung, 2009). Web sites such as Facebook, MySpace and Orkut have millions of users using them every day. A decentralized online social network (DOSN) is a distributed system for social networking with no or limited dependency on any dedicated central infrastructure (Datta, 2010), while being a solution to the violation of privacy, especially thanks to p2p architecture (Qamar, 2016).

Recent trends in the use of social networking highlight the fact that there is not only an increasing number of users of social networking applications but also a significant increase in the number of such applications. In a short time, social networks have invaded the daily lives of Internet users and Web professionals. The social media giants Facebook and Twitter were seen establishing, growing and evolving. They have been followed by a multitude of other more specific networks: Instagram, LinkedIn, etc. The list is long. Among the existing research studies, a group of studies identifies useful social networking information, using machine learning, to successfully extract structured information from unstructured textual social media contents.

People use social networking to post situational updates in various forms (Imran, 2015) such as text messages, images and videos. Numerous studies (Imran, 2014) have shown that this online information is useful for a quick response to a particular situation. Communication via social networking is direct, easy and instant and can simplify quick responses. Custom sites like Facebook, Twitter, Instagram, YouTube and Xing can subjectively offload the first contact of authorities and service providers. These analyzes of the use of social networking in events have identified a distinct role for users, who are more likely to generate useful information to improve situational awareness. Social networking can be considered as a practical and reliable emergency communication tool. While the predominant function of social networking remains social interaction, social networking sites are also considered the fourth most popular source of information. Different social networks have different characteristics and are therefore more or less suitable for use during a given situation. Social networking can support the exchange of information before, during and after an eventual event. With the proliferation of social media, knowledge is transformed from expert knowledge to everyday knowledge co-produced by various stakeholders thanks to Web 2.0. In recent years, a growing number of studies have examined the use of social networking data to gain knowledge of areas of human activities that are as diverse as detecting diseases such as epidemics or stock market forecasts. However, understanding these voluminous and high velocity data is a difficult task.

Table 3. Comparative table of all techniques and methods used in Models including our approach

References	Identification Methods	Used OSN
(Zaini, 2020)	Flood Disaster Game-based Learning	Twitter
(Vivakaran, 2018)	Educational Purposes among the Faculty of Higher Education with Special Reference	
(He, 2016)	Summarization with social-temporal context	
(Dussart, 2020)	Capitalizing on a TREC Track to Build a Tweet Summarization Dataset	
(Lamsal, 2020)	Semi-automated artificial intelligence-based classifier for Disaster Response	
(Rudra, 2019)	Summarizing situational tweets in crisis scenarios: An extractive-abstractive approach	
(Bouzidi, 2018), (Bouzidi, 2019)	Based on Artificial Neural Network (ANN)	Twitter & Facebook
(Bouzidi, 2020)	Based on FeedForward Neural Network (FFNN)	All the Web
(Bouzidi, 2020b)	Based on Recurrent Neural Network (RNN)	
Our New approach	Deep Learning from Social Media and Big Data to improve Marketing, Business Strategies, Fraud Detection and Financial Time Series Forecasting	

Source: Authors' research

This section presents the most relevant related works, namely information retrieval models, in general, and those from several sources, in particular.

Table 4. Comparative table of all techniques and methods used in Models including our approach

Models	Identification Usage of Big Data
(Ofli, 2016)	Human Computing and Machine Learning to Make Sense of Big Data
(Horita, 2017)	Decision-Making and emerging Big Data
(Immonen, 2015)	Quality of Social Media Data in Big Data
(Smith, 2012)	Big Data privacy in public Social Media
Our Previous approach (Bouzidi, 2020b)	Smart Data privacy in public Social Media
Our New approach	Automated Learning using Big Data to improve Marketing, Business Strategies, Fraud Detection and Financial Time Series Forecasting

Source: Authors' research

Contents were collected from all online channels tracked automatically by the Online Listening Tool, namely Radian6 (Ruggiero, 2014) & (Young, 2014) or any of its competitors, such as Awario, Brand24.com, Brandwatch, Mention, Keyhole, Socialert.net, SocialPilot.co, Simplify360, etc. from websites to social media, such as Twitter, Facebook, LinkedIn, Instagram, Google+, YouTube and so on. Many networking platforms allow access to their data via the Application Programming Interface (API) (Imran, 2020). Online listening tools provide the model, which reasonably represents the essentials, namely, as shown in figure 1:

- Harvesting contents: (such as conversations at the social media, news or any information on the Web);
- Cleaning the data of duplication and replication content: eliminating, from the content, any dubbed information like a retweet, and any information harmful or redundant;
- Enabling relevance: thanks to neural learning, obtaining relevant information by using machine learning with the learning corpus obtained thanks to the tagged messages. These tagged messages are realized by volunteers;
- Analyzing the results: during this stage, the verification and analysis of the results are carried out in order to ensure adequacy so that it is obtained to build disaster information such as situational awareness, damage assessment and/or disaster education.

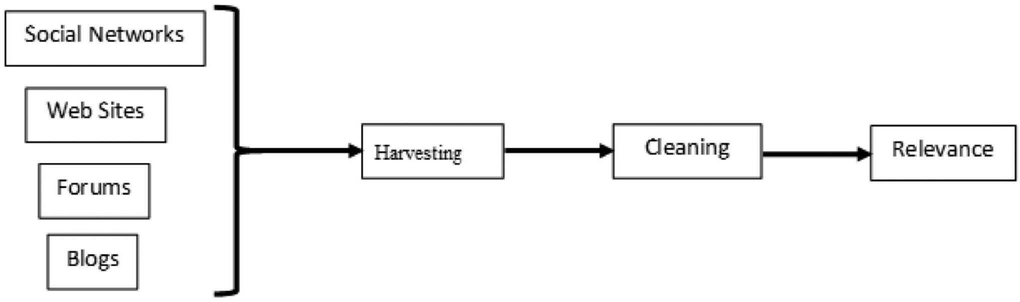

Figure 1. Online Methodology Reflecting our Listening and Monitoring Approach
Source: Authors' research

Some of the benefits of using the Twitter and Facebook APIs include:

- using the development space of the two social media;
- encouraging development environments;
- scientific recognition of development environments;
- encouraging other social media to involve themselves in research development;
- helping social media to feel imbued with this desire for development and scientific research in parallel with their commercial activity;
- contributing to the development of further improvements in networking services.

3. AUTOMATED LEARNING ENVIRONMENT FOR FRAUD & FINANCIAL TIME SERIES FORECASTING

3.1 Automated Learning Environment

Artificial intelligence (AI) is a combination of reinforcement learning (RL) and deep learning (DL) (Abiodun, 2018), represented mathematically, as:

$$\mathbf{AI = RL + DL} \tag{1}$$

where:

AI represents Artificial Intelligence, RL represents Reinforcement Learning, and
DL represents Deep Learning.

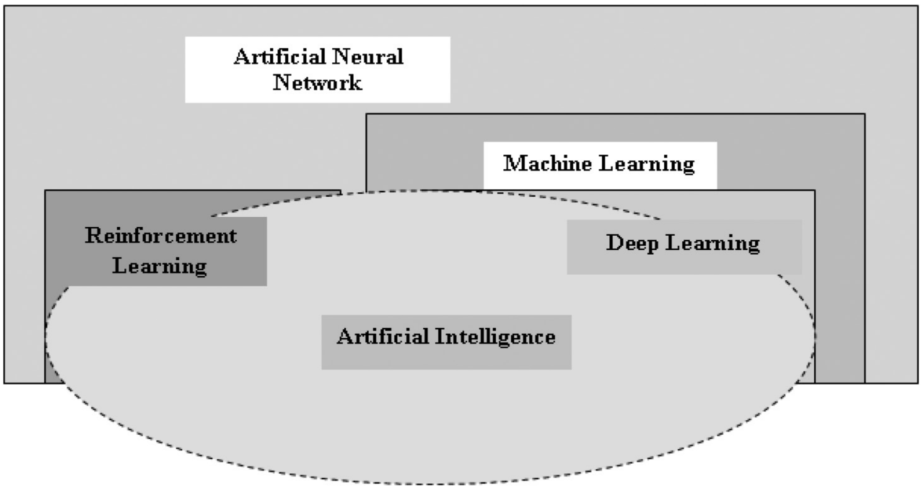

Figure 2. Fields of Artificial Intelligence
Source: Authors' research

A feedforward neural network (FFNN) is an automated learning classification algorithm that is made up of organized layers, similar to human neuron processing units. In FFNN, each unit in a layer (known as a node) relates to all other units in these layers. These layers' connections can have a different weight. These weights measure the potential amount of knowledge of the network. The information processing in the network involves data entry from the input units and passes through the network, flowing from one layer to another until the output units. When a neural network operates normally, that is when it acted as a classifier, then there will be no feedback between layers (Abiodun, 2018). FFNN can logically handle tasks according to "first come first serve" bases of inputs.

Unlike FFNN, the feed-backward neural network (FBNN) can use internal state "memory" to process a sequence of data inputs, such as Recurrent Neural Network (RNN).

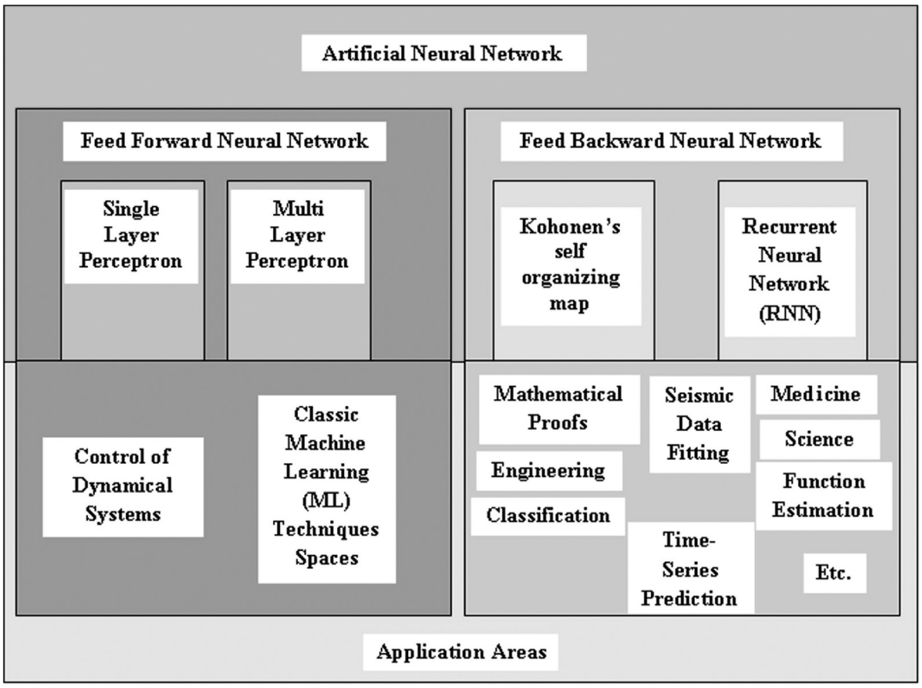

Figure 3. Artificial Neural Network Classification with their Application Areas
Source: Authors' research

3.2. Recurrent Neural Network (RNN)

As a class of feedforward neural networks, recurrent neural networks (RNNs) are augmented by the inclusion of recurrent edges connecting adjacent time steps. Figure 1 shows the well-known Elman, (1990) recurrent neural network (Elman, 1990) & (Wu, 2018).

We can use two equations to describe this type of RNN; all the calculations necessary for computation at each time step on the forward pass are:

$$h_t = \alpha\,(\sigma_{hx} \cdot x_t + \sigma_{hh} \cdot h_{t-1} + b_h) \tag{2}$$

$$y_t = \beta\,(\sigma_{yh} \cdot h_t + b_y) \tag{3}$$

where:

The σ terms denote weight matrices (e.g. σ_{hx} is the matrix of weights between the input and hidden layers);
The b terms denote bias vectors (e.g. b_h is hidden bias vector) which allows each node to learn an offset;
α denotes the hidden layer function h at the level t.

In general, α is usually an element-wise application of a sigmoid function and β is the output layer function.

A recurrent neural network (RNN) is referred to as a standard kind of neural network which extended over time, with edges, that feed into the next time step, rather than feeding into the next layer. RNN is constructed to sequences recognition, for instance, a text or a speech signal. It has cycles within that indicate the presence of short-memory in the net. RNN is like a hierarchical network in which the input needs processing hierarchically in the form of a tree because there is no time to the input sequence. Recurrent neural networks (RNNs) can be adapted to powerful sequence learning tasks. RNNs have proven to be an excellent pattern for recognition and prediction engines, especially in a task involving machine learning of sequences such as text or speech recognition. RNNs have feedback loops in their recurring layer. This helps them to keep information in “memory” for an extended period.

Although RNN is not deep in space, it is inherently deep in the time since each hidden state is a function of all previous hidden states (Wu, 2018). The problems of vanishing and exploding gradients will occur when propagating errors across many time steps. However, it can be difficult to train a standard RNN to solve problems that require learning over a long period of dependency. This is because the gradient of the loss function decays exponentially over time (this is called the gradient vanishing problem).

RNN is Discriminative, Supervised, Gradient Descent, Backpropagation through Time, Natural Language Processing and Language Translation. However, it is difficult time series inference and unsupervised in negative time (Berglund, 2015).

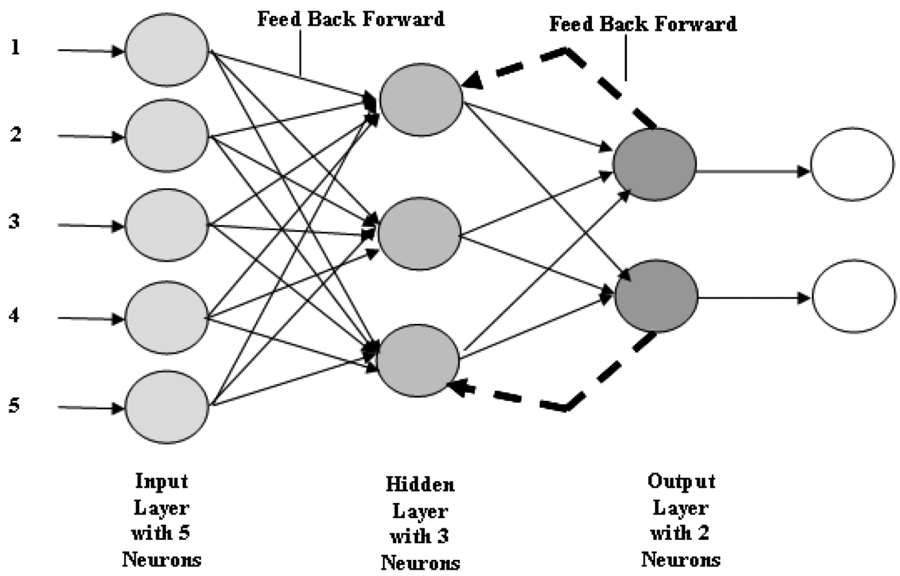

Figure 4. Recurrent Neural Network Architecture (Abiodun, 2019)
Source: Authors' research

3.3. Long Short-Term Memory (LSTM)

To overcome the deficiency of RNN, Hochreiter and Schmidhuber (1997) (Hochreiter, 1997) proposed Long Short-Term Memory, one of the most successful RNN architectures for sequence learning. Compared with the Elman RNN, LSTM introduces the memory cell, a computation unit replacing conventional artificial neurons in the hidden layer.

Long Term Memory Networks (LSTMs) is a kind of RNN that uses specific units with standard units. A "memory cell" is a component of LSTM units that can hold information in memory for a long time. LSTMs are often referred to as Sophisticated RNNs.

The mathematical definition of the computation of the LSTM model can be described as follows:

$$i_t = \gamma\ (\omega_{ix} \cdot x_t + \omega_{ih} \cdot h_{t-1} + b_i) \tag{4}$$

$$f_t = \gamma\ (\omega_{fx} \cdot x_t + \omega_{fh} \cdot h_{t-1} + b_f) \tag{5}$$

$$c_t = f_t\ \mathit{O}\ c_{t-1} + i_t\ \mathit{O}\ \tanh\ \omega_{cx} \cdot x_t + \omega_{ch} \cdot h_{t-1} + b_c) \tag{6}$$

$$\beta_t = \gamma\ (\omega_{\beta x} \cdot x_t + \omega_{\beta h} \cdot h_{t-1} + b_\beta \tag{7}$$

$$h_t = \beta_t\ \mathit{O}\ \tanh(c_t) \tag{8}$$

where:

O denotes element-wise multiplication.
ω is the logistic sigmoid function.
i, f, β, and c are respectively the input gate, forget gate, output gate, and cell activation vectors, all of which are in the same size as the hidden vector h at the level t.

3.4. Recurrent Neural Network-based Automated Learning Environment

Artificial Neural Networks (ANN) are inspired by biological neural networks as the human brain. Such systems learn to perform tasks by considering examples, usually without being programmed, with task-specific rules. The neural network is based on connected units'/nodes' collection called neurons, which model biological brain neurons. Each connection, like the synapses of a biological brain, can transmit a signal from an artificial neuron to another.

It is not just about a set of algorithms, but the following steps list to take into account and execute in order to reach an optimal result:

- Data Acquisition where the algorithm feeds on input data and where the success of the project is collecting relevant data and in sufficient quantity;
- Preparation and cleaning of the data;
- Model Creation;
- Model evaluation: it consists of evaluating the trained model on the other (second) set of data.
- Validation/ Deployment: where the model will be deployed in production to make predictions, and potentially use new input data to re-train and improve its model.

However, care must be taken to use an adequate number of neurons and hidden layers and to detect over-learning. Thus, the data is divided into two subsets (Lagmay, 2017):

- Learning set: which allows changing the weight of the neural network, and
- Validation set: This allows verifying the relevance of the network.

We can reduce the size of the network and start learning again. To avoid over-learning, a form of regularization is used. The weight decay method is a regularization technique used to limit over-learning in a neural network.

As for online methodology reflecting, our listening and monitoring approach consists of Harvesting all the contents issued from conversations at the social media, news or any information on the Web; Cleaning the data of duplication and replication content; Enabling relevance thanks to the neural learning, with the learning corpus obtained thanks to the tagged messages. The verification and analysis of the relevant content are carried out by the emergency management model, in order to ensure adequacy and to alert the disaster managers.

The benefits are huge: such as finding some details of an ultimate discussion or predicting and preventing the outbreak of a disaster. Making certain decisions that could save lives, has no commercial value but a value of great morality.

Figure 5 shows the functional architecture of our proposed model. So, the disaster management model based on an artificial neural network works by extracting, from Social Media, messages that contain predefined keywords; once the message is properly cleaned: it is not redundant and is not replication (insult). The message cleaned up from duplicated and replication contents, retained according to keywords, must be checked by the content manually annotated by volunteers in our laboratory. The message, issued from these checks, is considered relevant. It will be rapidly addressed to disaster managers to make quick and efficient decisions that can save lives if not relieve them.

3.4.1. Recurrent Neural Network: Modeling

The entire ANN modeling procedure has been studied to introduce systematic methods leading to always efficient ANN models, namely the collection of learning data, preprocessing and post-processing of data, different types of activation functions, initialization of weights, learning algorithms and error functions. Although all of these factors affect ANN's performance, increased attention has been focused on finding the best architecture. There is no theoretical background on how this architecture will be found or on its appearance (Thawakar, 2019). The most typical method followed is a repetitive trial and error process, in which a large number of different architectures are examined and compared to each other. Therefore, this process is time-consuming and relies mainly on the experience and intuition of the human expert, which implies a high degree of uncertainty. Nevertheless, we cite different approaches, namely: the empirical or statistical methods used to study the effect of internal parameters of an ANN and choose appropriate values depending on the performance of the model, the hybrid methods such as fuzzy inference, the constructive and/or pruning algorithms, and finally the evolutionary strategies. The training data is created from the November 10th, 2001 (Algiers Floods) and the May 21st, 2003 (Boumerdes earthquake disasters). This information, easily obtained using the neural network, is manually annotated by volunteers.

A. Foundation of neural learning

We use a neural network with a hidden layer that takes, as input to the network, a content e, as:

$$e = (w_1, \dots , w_i, \dots , w_n) \tag{9}$$

containing words w, each coming from a finite vocabulary Υ. C^n is the set of contents issued from social media.

Let:

$$e_i \in C_n = \mathrm{E}\ \forall\ i \in [1, \mathbf{N}] \text{ with } e_i = (w_{i1}, w_{i2}, \dots , w_{in}) \tag{10}$$

containing words each coming from the set of words \^{**W**} where each word comes from a finite vocabulary Υ, the incorporation of the content of the source message is relevant for, at least, a keyword or a hashtag such as:

$$\exists\ j \in [1, \mathbf{M}]\ /\ h_j \in \mathbf{H} \tag{11}$$

We want the learning of a generic space with the neural network, as:

$$\mathbf{E_K} = \max_{1 < k < K} \{e_k\} = \max_{1 < k < K} \{e_k\} \tag{12}$$

which normalizes the differences:

$$\mathbf{E_K} = [\mathbf{E} - \mathbf{RDF}] \text{ where } \mathbf{RDF} = [\mathbf{R} + \mathbf{D} + \mathbf{F}] \tag{13}$$

Thanks to the neural network, the transformation of $\{e_i\}$ into $\{e_k\}$ can be explained by:

$$\exists\, j \in [1, \mathbf{M}] \;/\; h_j \in \mathbf{H} \;\&\; \exists\, l \in [1, \mathbf{L}] \;/\; w_l \in \mathbf{W} \;/\; \{e_i \rightarrow \mathrm{e}_k = \{e_i \;/\; e_i \text{ is relevant for } h_j \text{ and } w_l\} \text{ with } i \in [1, \mathbf{N}]\} \;\&\; e_i \in [\mathbf{R} + \mathbf{D} + \mathbf{F}] \tag{14}$$

where

R, **D** and **F** denote respectively the set of duplicate re-tweets, duplicate contents and false alerts.

The objective is then to maximize the size **K** of the set $\mathbf{E_K}$. Figure 4 shows the Recurrent Neural Network-based Emergency Management Architecture.

Algorithm 1 determines, using a set of keywords, information that will be annotated manually to enrich the neural network in its possible learning. Algorithm 2 shows the functioning, during the following passages, of this emergency management model to learn the relevant information that will be used to inform managers so they can take quick and effective decisions.

Figure 5 shows the functioning of the Recurrent Neural Network-based Automated Learning Environment to retrieve relevant information.

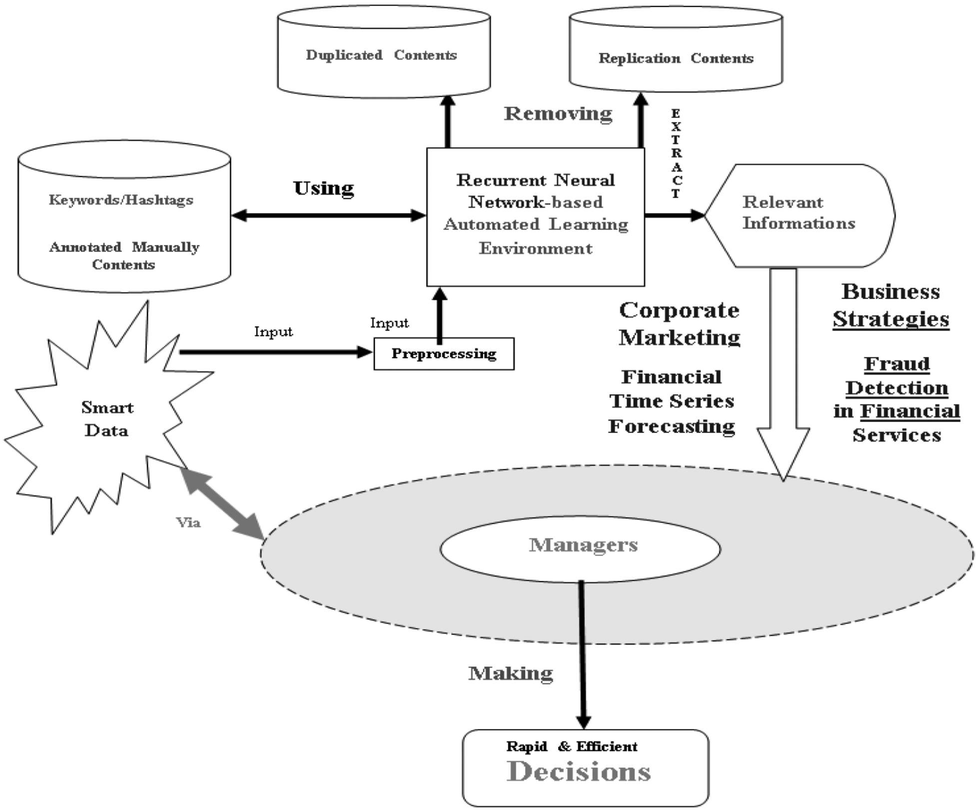

Figure 5. Recurrent Neural Network-based Automated Learning Environment to retrieve relevant information

Source: Authors' research

3.4.2. Sequential Pattern Mining: Algorithms Implementation

A pattern of a formal Content Set (O; P; R) is a subset of P. The set of all the motives of a base is, therefore, the set of parts of P, denoted 2^P. We say that an object x ϵ O has a pattern m if, ∀ p ϵ m; x **Re** p. For the example of the content set, we have:

Size pattern 0 = ø; (C_5^0 = 0 reasons).

Size reasons 1 = {a}; {b}; {c}; {d} and {e}, which will be noted, for simplicity, a; b; c; d and e. (C_5^1 = 5 reasons)

Grounds of size 2 = ab; ac; ad; ae; bc; bd; be; cd; this; of (C_5^2 = 10 patterns)

Size reasons 3 = abc; abd; abe; acd; ace; ade; bcd; ecb; bde; cde (\$$C_5^3$ = 10 units)

Size reasons 4 = abcd; abce; abde; acde; bcde (C_5^4 = 5 patterns).

Size 5 = abcde (C_5^5 = 1 patterns). In the previous formal set, x1 has the patterns: at; c; d; ac; ad; cd and acd. Among the global set of 2^p patterns, it will look for those that appear frequently. Let m_1 and m_2 be two patterns. An association rule is an implication of the form:

$$m_1 \rightarrow m_2 \text{ where } m_1, m_2 \in 2^p; \ \& \ m_1 \cup m_2 = \emptyset \qquad (15)$$

The rule $m_1 \rightarrow m_2$ is verified in data set E with a support s, where s is the percentage of objects in E containing $m_1 \cup m_2$. Let $m \in 2^p$ be a pattern. The support of m is the proportion of objects in O that have the pattern:

Support: $2^p \rightarrow [0;\ 1]$, $m \rightarrow$ Support (m) = (f (m)) / O)

For example, in the previous database, we have:

Support (a) = 3/6;

Support (b) = 5/6;

Support (ab) = 2/6;

Support (ø) = 1 and

Support (P) = 0. The support is decreasing by (2^p; subset) in ([0; 1]; ≤). In other words, if m is a sub-pattern of m' (m ⊂ m') then Support (m) ≤ Support (m'). The medium measures the frequency of a pattern: the higher it is, the more frequent the pattern. Frequent patterns of non-frequent patterns can be distinguished using a threshold ω (Masseglia, 2005). Streaming is a way of broadcasting and reading streaming content, which is widely used on the Internet. It is opposed to the {\textit file download} which requires recovering all the data of a file. However, playing streaming content involves being connected to an Internet server. Smart Data (https://www.lebigdata.fr/smart-data-definition-differences-big-data), a different concept of Big Data, is based primarily on real-time data analysis. This term refers to an approach to data analysis that directly analyzes the data at the source, without the need to transmit it to a centralized system.

3.4.3 Dynamic Counting Algorithm

The DIC algorithm, proposed by Brin et al. (1997) (Brin, 1997), to reduce the number of runs of the database, is suitable for streaming. Thus, DIC, which partitions the database into blocks of M transactions, has adapted to work with blocks of M contents. During the computation of the k-item sets supports, after the search of a partition of size M of D, we verify the k-item sets candidates who have already reached the minimum support; DIC then uses them to generate candidates of size (k+1) and starts counting their supports. Thus, the supports of candidates of different sizes are calculated during the same course of D. As the number of scans in the database decreases, there is only one passage of the content, whereas DIC considers candidate item sets of different sizes simultaneously. This poses the problem of storing the candidate item sets processed simultaneously and the cost of calculating the candidates' media which is greater than for the Apriori algorithm (Agrawal, 1996).

3.4.4. The AprioriTlD Algorithm

The Apriori algorithm (Agrawal, 1996), requiring N passages on the database, a possible optimization consists in generating memory, during the first pass, the identifiers (TID) of the transactions for each 1-item set (together pattern size 1) frequent. This algorithm is also suitable for streaming. The TID lists corresponding to each k-item set of the K contents are kept. The calculation of a k-item set is always done from the two (k-1)-item sets containing one less element, but the counting is done simply by intersection of the two TID lists of the two (k-1)-item sets source. We build the list of TIDs after determining the frequent 1-item sets, which is more efficient in streaming (Smart data). When reading the first content, this eliminates the infrequent products and thus reduces the lists of TID in memory. TID lists in parallel memory should be generated as soon as the first content is more efficient. Here is the Apriori reference algorithm adapted to streaming (Smart data).

3.4.5 Apriori Algorithm

Algorithm 1 shows the AprioriTID algorithm (Agrawal, 1996) which is the algorithm adapted to streaming.

Algorithm 1. AprioriTlD algorithm

begin
  Require: Content of Social Networks D, Minimum Support Threshold σ;
  Ensure: Set of frequent items;
  input (a content $V(w_1, w_2, ..., w_i, ... , w_n)$) and σ
  initialization $i \leftarrow 1$;
  initialization $C_i \leftarrow 1$ set of size 1 patterns (one item);
  while ($C_i \neq \emptyset$) **do**
    Calculate the Support of each pattern $m \in C_i$ in the content set:
    $F_i \leftarrow \{m \in C_i \mid \text{support } (m) \geq \sigma\}$;
    $C_{i+1} \leftarrow$ {all possible combinations of F_i patterns of size i + 1};
    incrementation $i \leftarrow i + 1$;
  endwhile
  output $\bigcup_{i \geq 1} \mathbf{Fi}$
end

1. The application of the algorithm on the basis of the given contents read in the streaming given in the example with $\omega = 0.25$ happens as follows. Supports: 3/6, 5/6, 5/6, 1/6 and 5/6. Hence, F_1 = {a; b; c; e} (no frequent reason will contain d).

2. Generation of size 2 candidates. Combine 2 to 2 F_1 size 1 candidate. So, C_2 = {ab; ac; ae; bc; be} and Supports are 3/6, 2/6, 4/6, 5/6 and 4/6. $F_2 = C_2$: all C_2 patterns are common.

3. Generation of size 3 candidates: Combine 2 to 2 F_2 size candidates (and consider only those who give size 3 patterns). So, C_3 = {abc; abe; ace; bc} and Supports are 2/6, 2/6, 2/6 and 4/6. $F_3 = C_3$: all C_3 motifs are common.

4. Generation of size 4 candidates. So, C_4 = {abce} and Support is 2/6.

5. Generation of size 5 candidates: $C_5 = ø; So, $F_5 = ø;

6. The algorithm then returns all of the common patterns, namely: $F_1 U F_2 U F_3 U F_4$.

Among the goals of the optimizations, this algorithm is to facilitate streaming playback with a single read with sufficient computation and storage. The threshold σ is set by the analyst. This can follow an iterative approach by setting a threshold at the start and, depending on the result, change the threshold value. The algorithm proposed by Savasere (Zaki, 1997) solves the memory space problem of the previous algorithm. The advantage of this algorithm is that it requires only one reading at most.

3.5. Discussion About Sequential Pattern Mining And Dynamic Counting Algorithm

Figures 6 and 8 show Examples of Content Occurrences Number in the Automated Learning Environment.

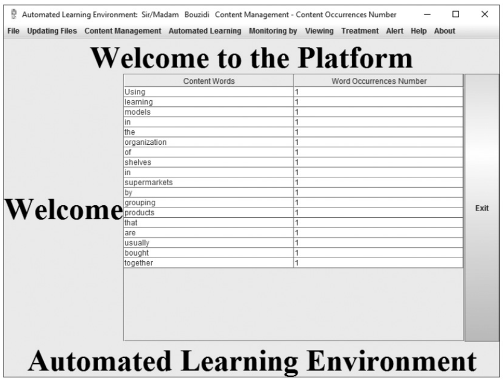

Figure 6. Example of Content Occurrences Number in the Automated Learning Environment
Source: Authors' research

Figure 9 shows the flow chart showing the functioning, during the first passages, of the Recurrent Neural Network-based Automated Learning Environment model to learn the first information that will be manually annotated by volunteers.

Figure 10 shows the flow chart showing the functioning, during the following passages, of this Recurrent Neural Network-based Automated Learning Environment model to learn relevant information that will be used to alert public opinion, and in particular disaster managers, so they can take quick and effective decisions that can save lives.

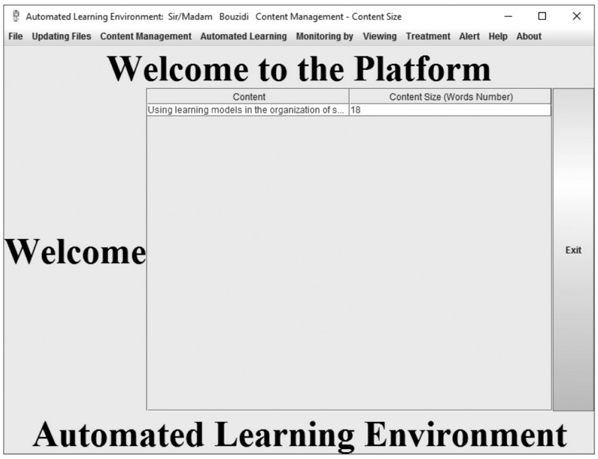

Figure 7. Example of Content Size in the Automated Learning Environment

Source: Authors' research

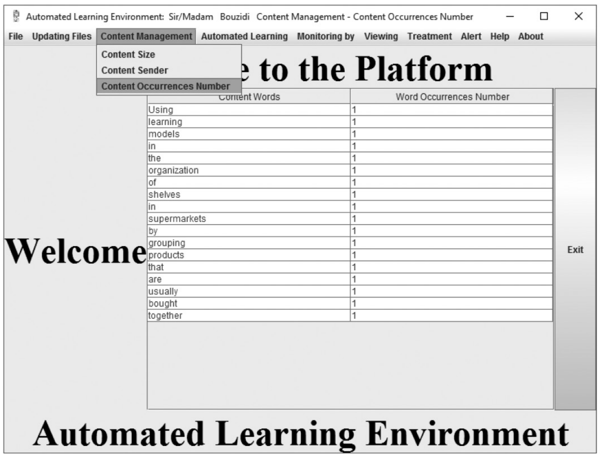

Figure 8. Example of Content Occurrences Number in the Automated Learning Environment

Source: Authors' research

Algorithm 2 Pseudo codes for one-step performance prediction.

Require: Input feature series $S = S^t$, $t = 1, 2, ..., T_s$
function OneStepPrediction(S)
  Initialize $h_0 = 0$, $f = 0$, $(m_1, m_2) = (1, 0)$
  Initialize $S^0 = (f, m_1, m_2)$, $t = 0$
  while $t \leq T_s$ **do**
    Generate h_{t+1} and y_{t+1} by h_t & S^t
    Sample S^{t+1} using y_{t+1}
    $t = t + 1$
  end while
  return $S = \{S^t\}$, $t = 1, 2, ..., T_s$
end function

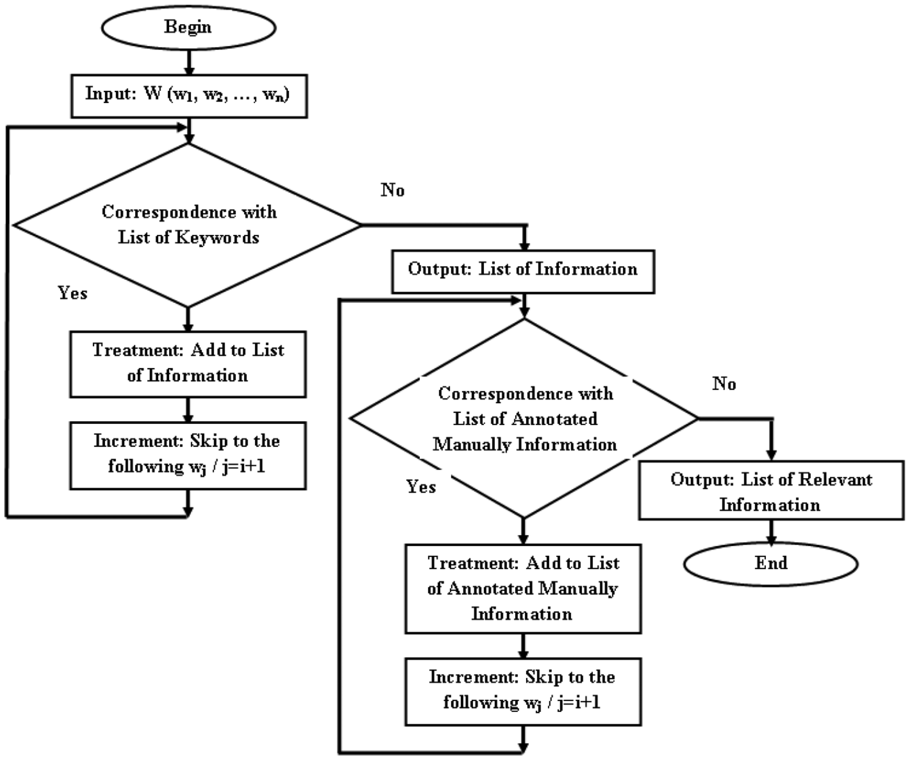

Figure 9. Flow chart for determining, using a set of keywords, information that will be annotated manually to enrich the neural network in its possible learning

Source: Authors' research

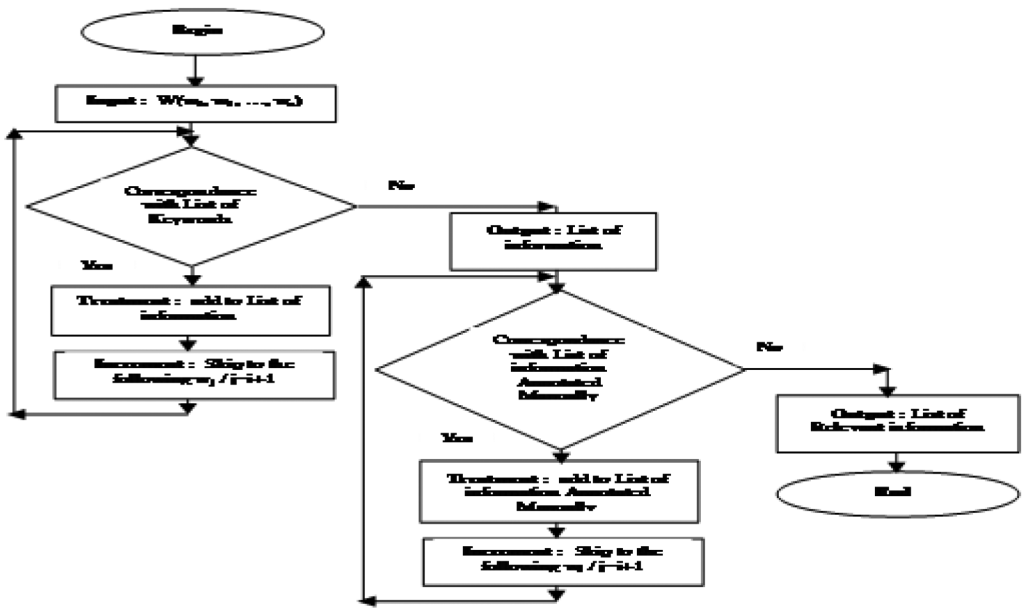

Figure 10. Flow chart for determining relevant information using a set of manually annotated information

Source: Authors' research

4. CONCLUSION

A new ad hoc real-time automated learning environment with accurate forecasting of volatility from financial time series is presented here, based on a new multi-view recovery model from multiple sources using Smart data. Such an approach is really useful for paramount in financial decision making, but also for help to make strategic appropriate decisions. This work has some limitations as follows.

1. The contents are known to be written informally, the contents follow no syntax, no logic, are noisy, may contain spelling mistakes, abbreviations, etc.
2. There is only collected English content posted. As a result, domain-specific biases may exist in the dataset.
3. Side by side, content published in other languages may contain different types of reasons in relation to English content.
4. The features of the automated learning environment have been developed based on the analysis of specific content.

As for future works, this study can have many potential applications for the future. The proposed model can be completed with a series of new research questions and perspectives. Pure improvements can start from:

1. An improvement of the validation of the information before launching this update information in order to avoid errors inaccurate forecasting of volatility from financial time series, with abusive information.
2. The extension of the Real-Time automated learning environment to process images and videos of social media with the recurrent neural network trained with Long Term Memory Network.
3. The use of multiple languages, notably local languages.

REFERENCES

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. In *Heliyon*, 4(11), DOI:10.1016/j.heliyon.2018.e00938

Abouelhoda, M. & Ghanem, M. (2009). String Mining in Bioinformatics. *Scientific Data Mining and Knowledge Discovery, 207–247.* https://doi.org/10.1007/978-3-642-02788-8_9

Adhikari, R. & Agrawal, R.K., (2014), A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications, vol. 24, 1441–1449.* https://doi.org/10.1007/s00521-013-1386-y

Agrawal, Rakesh, Mehta, Manish, Shafer, John, Srikant, Ramakrishnan, Arning, Andreas & Bollinger, Toni, (1996), The Quest Data Mining System. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining KDD'96, 244-249.* doi:10.5555/3001460.3001511

Aloysius, G. & Binu D. (2013). An approach to products placement in supermarkets using Prefix Span algorithm. *Journal of King Saud University – Computer and Information Sciences 25(1), 77–87.* https://doi.org/10.1016/j.jksuci.2012.07.001

Bao, W., Yue, J. & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory, *PloS one, 12(7), e0180944.*

Berglund, M., Raiko, T., Honkala, M., Karkkainen, L., Vetek, A. and Karhunen, J., (2015), Bidirectional Recurrent Neural Networks as Generative Models, MIT Press, Cambridge, MA, USA

Bodyanskiy, Yevgeniy & Popov, Sergiy, (2006), Neural network approach to forecasting of quasiperiodic financial time series. *European Journal of Operational Research, 175(3), 1357-1366.* https://doi.org/10.1016/j.ejor.2005.02.012

Bouzidi Z., Boudries A. and Amad M., (2018), A New Efficient Alert Model for Disaster Management, Proceedings of Conference AIAP'2018: Artificial Intelligence and Its Applications, El-Oued, Algeria,

Bouzidi, Z., Amad, M. and Boudries, A., (2019), ``Intelligent and Real-time Alert Model for Disaster Management based on Information retrieval from Multiple Sources``, International Journal of Advanced Media and Communication, Vol. 7, No. 4, pp. 309-330, doi:10.1145/253260.253325

Bouzidi, Z., Boudries, A. & Amad, M. (2020). Towards a Smart Interface-based Automated Learning Environment Through Social Media for Disaster Management and Smart Disaster Education. *Advances in Intelligent Systems and Computing. SAI 2020. Vol 1228. Springer, Cham, 443-468.* https://doi.org/10.1007/978-3-030-52249-0_31

Bouzidi, Z., Amad, M. and Boudries, A. (2020b). Deep Learning-based Automated Learning Environment using Smart Data to improve Corporate Marketing, Business Strategies, Fraud Detection in Financial Services and Financial Time Series Forecasting. In *International Conference on "Managing Business through Web Analytics – (ICMBWA2020)"*, Khemis Miliana University, Algeria

Brin, Sergey, Motwani, Rajeev, Ullman, Jeffrey D. & Tsur Shalom, (1997), Dynamic items and counting and implication rules for market basket data. *Proceedings of the 1997 ACM SIGMOD international conference on Management of data, SIGMOD'97, 255-264.* https://doi.org/10.1145/253260.253325

Calvez, A. (le) and Cliff, D., (2018), Deep learning can replicate adaptive traders in a limit-order-book financial market, 2018 IEEE Symposium Series on Computational Intelligence, (SSCI 2018), pp. 1876-1883

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P. & Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications, vol. 55,194–211.*

Datta, A., Buchegger, S., Vu, LH., Strufe, T. & Rzadca, K., (2010), Decentralized Online Social Networks, Handbook of Social Network Technologies and Applications, Springer, Boston, MA, DOI:org/10.1007/978-1-4419-7142-5_17

Dong, M., Yao, L., Wang, X., Benatallah, B., Huang, C. & Ning, X. (2018). Opinion fraud detection via neural autoencoder decision forest. *Pattern Recognition Letters.* http://www.sciencedirect.com/science/article/pii/S0167865518303039

Dussart, A., Pinel-Sauvagnat, K. & Hubert, G. (2020). Capitalizing on a TREC Track to Build a Tweet Summarization Dataset. In *Text Retrieval Conference*, (TREC'2020)

Elman, J., (1990), Finding structure in time. *Cognitive Science, 14(2), 179–211.* doi:10.1016/0364-0213(90)90002-e

Fama, E. F. (1995). Random walks in stock market prices. *Financial analysts journal, 51 (1), 75–80.*

Fischer J., Jiang W. and Moran S., (2012), AtomicOrchid: A mixed reality game to investigate coordination in disaster response, Proceedings of the 11th International Conference on Entertainment Computing (ICEC 2012), pp. 572-577

Fiore, U., Santis, A. D., Perla, F., Zanetti, P. & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences, vol. 479, 448 – 455.* http://www. sciencedirect.com/science/article/pii/S0020025517311519

Fournier-Viger, P., Lin, J. C.-W., Kiran, R. U., Yun S. K. & Rincy, T. (2017). A Survey of Sequential Pattern Mining. *Data Science and Pattern Recognition, Ubiquitous International 1(1), 54–78.*

Ghazali, Rozaida, Hussain, Abir Jaafar, Nawi, Nazri Mohd & Mohamad, Baharuddin, (2009), Non-stationary and stationary prediction of financial time series using dynamic ridge polynomial neural network. *Neurocomputing, 72(10–12), 2359-2367.* https://doi.org.10.1016/j.neucom.2008.12.005

Gomez, J. A., Arvalo, J., Paredes, R. & Nin, J. (2018). End-to-end neural network architecture for fraud scoring in card payments. *Pattern Recognition Letters, vol. 105, 175 – 181, Machine Learning and Applications in Artificial Intelligence.* http://www.sciencedirect.com/science/article/pii/S016786551730291X

Gudelek, M. U., Boluk, S. A. & Ozbayoglu, A. M. (2017). A deep learning based stock trading model with 2-d cnn trend detection. *Computational Intelligence (SSCI), 2017 IEEE Symposium Series on. IEEE, 1–8.*

Han, J.-H., Kushner, S. A., Yiu, A. P., Cole, C.J., Matynia, A., Brown, R.A. & Josselyn. S.A. (2007). Neuronal Competition and Selection During Memory Formation. *Science 316, 5823, 457–460.* https://doi.org/10.1126/science.1139438

He, R., Liu, Y., Yu, G., Tang, J., Hu, Q. & Dang, J. (2016). Twitter summarization with social-temporal context. In *World Wide Web*, 20(2), pp. 267-290, DOI:10.1007/s11280-016-0386-0

Hochreiter, S. & Schmidhuber, J., (1997), Long short-term memory. *Neural Computation, 9(8), 1735–1780.* DOI=10.1162/neco.1997.9.8.1735

Horita, Flavio E.A. and Albuquerque, Joao Porto (de) and Marchezini, Victor and Mendiondo, Eduardo M., (2017), Bridging the gap between decision-making and emerging big data sources: An application of a model-based framework to disaster management in Brazil, Decision Support Systems, vol. 97, pp. 2-22, doi:10.1016/j.dss.2017.03.001

Immonen, A. and Paakkonen, P. and Ovaska, E., (2015), Evaluating the Quality of Social Media Data in Big Data Architecture, IEEE Access, vol. 3, pp. 2028-2043, doi:10.1109/ACCESS.2015.2490723

Imran, M., Castillo, C., Lucas, J., Meier, P. & Vieweg, S., (2014), AIDR: Artificial intelligence for disaster response, *Proceedings of the 23rd International Conference on World Wide Web, (ICT-DM), 159-162.* https://doi.org/10.1145/2567948.2577034

Imran M., Castillo C., Diaz F. and Vieweg S., (2015), 'Processing social media messages in mass emergency: A survey, ACM Computing Surveys (CSUR), vol.47, No. 67, pp. 1-38, in: doi:10.1145/2771588

Imran, M., Ofli, F., Caragea, D. & Torralba, A., (2020), Using AI and Social Media Multimodal Content for Disaster Response and Management: Opportunities, Challenges, and Future Directions. *Information Processing & Management, 57(5), 1-9.* http://sci-hub.tw/10.1016/j.ipm.2020.102261

Kodogiannis, V. & Lolis, A., (2002), Forecasting Financial Time Series using Neural Network and Fuzzy System-based Techniques. *Neural Computing and Applications, vol. 11, 90–102.* doi:10.1007/s005210200021

Kosala, R. & Blockeel, H., (2000), Web Mining Research: A Survey, ACM SIGKDD Explorations Newsletter, vol. 2, No. 1, ACM

Lagmay A. M., Mendoza J., Cipriano F., Delmendo P. A., Lacsamana M. N., Moises, M. A., Pellejera, III N., Punay, K. N., Sabio, G., Santos, L., Serrano, J., Taniza, H. J. and Tingin, N. E., (2017), Street floods in Metro Manila and possible solutions, Journal of Environmental Sciences

Lai, K.K., Yu, L., Wang, S. and Huang, W., (2006), Hybridizing Exponential Smoothing and Neural Network for Financial Time Series Predication, Proceedings of Computational Science – ICCS 2006, Lecture Notes in Computer Science, (ICCS 2006), vol. 3994, Springer, Berlin, Heidelberg, doi:10.1007/11758549_69

Lamsal, R. & Kumar, T. V. V. (2020). Classifying Emergency Tweets for Disaster Response. In *International Journal of Disaster Response and Emergency Management (IJDREM)*, 3(1), pp. 14-29, DOI:10.4018/IJDREM.2020010102

Lasfer, A., El-Baz, H. and Zualkernan, I. (2013). Neural network design parameters for forecasting financial time series. *Modeling, Simulation and Applied Optimization (ICMSAO), 5th International Conference on. IEEE, 1–4.*

LeCun, Y., Bengio Y. and Hinton G., (2015), ``Deep Learning: Review``, Nature, Vol 521, May 28st

Li, X., Deng, Z. & Luo, J. (2009). Trading strategy design in financial investment through a turning points prediction scheme. *Expert Systems with Applications, 36(4), 7818 – 7826.* http://www.sciencedirect.com/science/article/pii/S0957417408008622

Lu, C.-J., Lee, T.-S. & Chiu, C.-C. (2009). Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems, 47(2), 115 – 125.* http://www.sciencedirect.com/science/article/pii/S0167923609000323

Masseglia, F., Teisseire, M. & Poncelet, P. (2005). Sequential pattern mining: A survey on issues and approaches. *Encyclopedia of Data Warehousing and Mining.*

Mohammad, A. H., Rezaul, K. & Ruppa, T., Neil, D. B. B. & Yang, W. (2018). Hybrid deep learning model for stock price prediction. *IEEE Symposium Series on Computational Intelligence SSCI. IEEE, 1837–1844.*

Ofli, F. and Meier, P. and Imran, M. and Castillo, C. and Tuia, D. and Rey, N. and Briant, J. and Millet, P. and Reinhard, F. and Parkan, M. and Joost, S., (2016), Combining Human Computing and Machine Learning to Make Sense of Big (Aerial) Data for Disaster Response, Big Data, vol. 4, No. 1

Pandey, T. N., Jagadev, A. K., Dehuri, S. & Cho, S.-B. (2018). A novel committee machine and reviews of neural network and statistical models for currency exchange rate prediction: An experimental analysis. *Journal of King Saud University – Computer and Information Sciences.* http://www.sciencedirect.com/science/article/pii/S1319157817303816

Pinheiro, L. dos Santos & Dras, M. (2017). Stock market prediction with deep learning: A character-based neural language model for event based trading. *Proceedings of the Australasian Language Technology Association Workshop 2017, 6–15.*

Pradeepkumar, Dadabada & Ravi, Vadlamani, (2017), Forecasting financial time series volatility using Particle Swarm Optimization trained Quantile Regression Neural Network. *Applied Soft Computing, vol. 58, 35-52.* https://doi.org/10.1016/j.asoc.2017.04.014

Pumsirirat, A. & Yan, L. (2018). Credit card fraud detection using deep learning based on auto-encoder and restricted Boltzmann machine. *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, 9(1), 18–25.*

Qamar, M., Batool, S., Mehmood, S., Malik, A. W. & Rahman, A., (2016), Centralized to Decentralized Social Networks: Factors that Matter, Managing and Processing Big Data in Cloud Computing, DOI:10.4018/978-1-4666-9767-6.ch003

Rappaport, S. D., (2010), Listening Solutions, A Marketer's Guide to Software and Services. *Journal of Advertising Research, 50(2), 197–213.* Doi:10.2501/S00218491009135X

Rudra, K., Goyal, P., Ganguly, N., Imran, M. & Mitra, P. (2019). Summarizing situational tweets in crisis scenarios: An extractive-abstractive approach. In *IEEE Transactions on Computational Social Systems*, 6(5), pp. 981-993, DOI:10.1109/tcss.2019.2937899

Ruggiero, A. and Vos, M., (2014), Social Media Monitoring for Crisis Communication: Process, Methods and Trends in the Scientific Literature, In Online Journal of Communication and Media Technologies, Vol. 4, No. 1

Ryman-Tubb, N. F., Krause, P. & Garn, W. (2018). How artificial intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. *Engineering Applications of Artificial Intelligence, 76, 130 – 157.* http://www.sciencedirect.com/science/article/pii/S0952197618301520

Schreyer, M., Sattarov, T., Borth, D., Dengel, A. & Reimer, B. (2017). Detection of anomalies in large scale accounting data using deep autoencoder networks. *CoRR, vol. abs/1709.05254*. http://arxiv.org/abs/1709.05254

Sermpinis, G., Karathanasopoulos, A., Rosillo, R. and Fuente, D. (de la), (2019), Neural networks in financial trading, Annals of Operations Research, doi:10.1007/s10479-019-03144-y

Smith, M. and Henne, B. and Szongott, C. and Voigt, G. von, (2012), Big data privacy issues in public social media, 6th IEEE International Conference on Digital Ecosystems Technologies, pp. 1-6, doi:10.1109/DEST.2012.6227909

Thawakar, O. & Gajjewar, P., (2019), Training Optimization of Feed Forward Neural Network for Binary Classification. *Proceedings of the International Conference on Computer Communication and Informatics (ICCCI-2019).* https://doi.org/10.1109/ICCCI.2019.8822184

Tk, M. & Verner, R. (2016). Artificial neural networks in business: Two decades of research. *Applied Soft Computing, 38, 788 – 804.* http://www.sciencedirect.com/science/article/pii/S1568494615006122

Vivakaran, M. V. & Neelamalar, M. (2018). Utilization of Social Media Platforms for Educational Purposes among the Faculty of Higher Education with Special Reference to Tamil Nadu. In *Higher Education for the Future*, 5(1), pp. 4-19, DOI:10.1177/2347631117738638

Wang, Z. & Ye, X. (2018). Social media analytics for natural disaster management. In *International Journal of Geographical Information Science*, 32(1), pp. 49-72, In DOI:10.1080/13658816.2017.1367003

Wu, Y., Mao, H., Yi, Z., (2018), Audio classification using attention-augmented convolutional neural network, Knowledge-Based Systems, vol. 161, pp. 90-100

Yeung, C.-m. A., Liccardi, I., Lu, K., Seneviratne, O. & Berners-lee, T., (2009), Decentralization: The future of online social networking, W3C Workshop on the Future of Social Networking Position Papers, 1-5, Barcelona, Spain

Young, S. D., Rivers, C. and Lewis, B., (2014), Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes, Preventive Medicine, vol. 63, pp. 112-115

Zaini, N.A., Noor, S.F.M. & Zailani, S.Z.M. (2020). Design and Development of Flood Disaster Game-based Learning based on Learning Domain. In *International Journal of Engineering and Advanced Technology (IJEAT)*, 9(4), pp. 679-685, DOI:10.35940/ijeat.C6216.049420

Zaki, M. J., Parthasarathy, S., Ogihara, M. and Li, W., (1997), Parallel Algorithms for Discovery of Association Rules. Data Mining and Knowledge Discovery, vol. 1, No. 4, p. 343-373, doi:10.1023/A:1009773317876

Zheng, Y.-J., Zhou, X.-H., Sheng, W.-G., Xue, Y. & Chen, S.-Y. (2018). Generative adversarial network based telecom fraud detection at the receiving bank. *Neural Networks, vol. 102, 78 – 86.* http://www.sciencedirect.com/science/article/pii/S0893608018300698