Deep Learning Methods Application in Finance: A Review of State of Art

Dovilė Kuizinienė[a], Tomas Krilavičius[a]

[a]*Department of Applied Informatics Vytautas Magnus University Kaunas, Lithuania*

Abstract
Artificial intelligence uses in financial markets or business units forms financial innovations. These innovations are the key indicator for economic grow and intelligent finance system formation. Recants years scientist and most innovation driving companies, such as Google, IBM, Microsoft and other, are focusing in deep learning methods. These methods have achieved significant performances in diverse areas: image recognition, natural language processing, speech recognition, video processing, etc. Therefore, it is necessary to understand the variety of deep learning methods and only then their applicability in the financial field. Accordingly, in this paper firstly is presented differences in science community already settled deep learning method's architectures. Secondly, is shown a big picture of developing scientific articles of deep learning uses in finance field, where the most used deep learning methods were identified. Finally, the conclusion, limitations and future work have been shown.

Keywords
Artificial intelligence, Machine Learning, Deep Learning, Convolution Neural Network, Deep Belief Network, Deep Boltzmann Machine, Deep neural network, Deep Q-Learning, Deep reinforcement learning, The extreme learning machine, Generative adversarial network, Recurrent Neural Learning, Long short-term memory, Gated Recurrent Unit, Finance, Financial innovations

1. Introduction

The global financial industry is quietly changing under the catalysis of artificial intelligence (AI) [1]. AI represents a clear opportunity to advance the transformation of the finance industry by providing users with greater value and increasing firms' revenues [2]. The goal of AI is to invent a machine which can sense, remember, learn, and recognize like a real human being [3]. The deep integration of AI technology and finance is the inevitable result of deepening development and Exploring Innovation in these fields [1]. These innovations have the potential to directly influence both the production and the characteristics of a wide range of products and services, with important implications for productivity, employment, and competition [4]. AI also improve work efficiency at the business and create a whole process of intelligent finance [1]. Applications of AI systems are generally viewed as positive for economic growth and productivity [2]. Deep learning is a recently-developed field belonging to Artificial Intelligence [3]. It attempts to learn hierarchical representations from raw data and is capable of learning simple concepts first and then successfully building up more complex concepts by merging the simpler ones [5, 6, 7]. Companies such as Google, Facebook, IBM, Microsoft and others use this algorithm for developing next-generation intelligent applications [8]. In finances there are two major problems:1) to predict future returns (i.e., stock prices, currencies, indices, product demand); or 2) to make categorical classification (i.e. credit scoring ("good, "bad"), bankruptcy ("True", "False")). While the issues in finances remain almost the same over the last several decades, novel methods, and growing amount of data are changing the field, especially Machine Learning and Artificial Intelligence techniques [9]. Furthermore, exploitation of additional data sources allows to achieve better results, e.g. satellite images can be used for predicting economic activity, voice information provides information about emotions, textual information, extracted from news and comments gives sentiments of writers and audience, etc [10]. However, extraction of useful knowledge out of such data heap is not trivial, it requires considerable effort [11, 12]. Portfolio management tasks have more challenges, because there are two main issues with portfolio formation: (1) selection of assets with highest revenue, and (2) determination other value composition of assets in the portfolio to achieve the goal of maximal potential returns with minimal risk [13]. Therefore, this paper is divided in two parts: 1) different deep learning architectures are discussed; 2) application of the aforementioned methods in finances is discussed.

IVUS 2020: Information Society and University Studies, 23 April 2020, KTU Santaka Valley, Kaunas, Lithuania
dovile.kuiziniene@vdu.lt (D. Kuizinienė);
tomas.krilavicius@vdu.lt (T. Krilavičius)

2. Literature review

The term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving" [14]. In other words, it tries to mimic the human brain, which is capable of processing the complex input data, learning different knowledges intellectually and fast, and solving different kinds of complicated tasks well [3].

AI has been part of human thoughts and slowly evolving in academic research labs [14]. Machine learning is the subset of AI. Machine learning is the study of computer algorithms that can be improved automatically through experience [1]. Machine learning algorithms overcome following strictly static program instructions by making data-driven predictions or decisions, through building a model from sample inputs [14]. In machine learning, artificial neural networks are a family of models that mimic the structural elegance of the neural system and learn patterns inherent in observations [15], see Fig. 1. The term “deep” refers to the number of layers in the network—the more layers, the deeper the network [16]. Traditional neural networks contain only 2 or 3 layers, while deep networks can have hundreds [16]. Deep learning has been explosively developed today. Compared with shallow learning, deep learning reaches the state of arts in many researches [17].

In contrast to the shallow architectures like kernel machines which only contain a fixed feature layer (or base function) and a weight-combination layer (usually linear), deep architectures refers to the multi-layer network where each two adjacent layers are connected to each other in some way [3]. This introduces the unprecedented flexibility to model even highly complex, non-linear relationships between predictor and outcome variables, a quality that has allowed deep neural networks to outperform models from traditional machine learning in a variety of tasks [18]. Deep learning methods have only now become so powerful, due to technical reasons of: computational power (hardware), availability of large datasets and optimization algorithms [18],[19].

2.1. Convolution Neural Network

Convolution neural network (CNN) algorithm is separated into two main parts: feature detection and classification (see Fig. 2).

Feature detection phase consist from convolution, pooling and rectified linear unit (ReLU) layers. Convolutional filters activates certain features from data set unit (image, video, time series). This layer produces huge amount of features that makes overfitting problems and expensive computation [8]. Pooling leayers reducses this problem by aggregating multiple feature values into a single value. Max-pooling is mostly used pooling operation, in Keras instead of this operation could be used Average-pooling, Global-max-pooling or Global-average-pooling operations [20]. Rectified linear unit (ReLU) is an activation function meant to zero out negative values, whereas a sigmoid “squashes” arbitrary values into the interval [0, 1] producing something that can be interpreted as a probability [19].

These three operations are repeated over tens or hundreds of layers, with each layer learning to detect different features [16]. The classification phase consists from two layers dropout and fully connected. Dropout consists of randomly dropping out (setting to zero) a number of output features of the layer during training [19].

The fully connected layer that produces a vector of K dimensions where K is the number of classes that the network will be able to predict. This vector contains the probabilities for each class of any image being classified [16, 21]. The quality of model is evaluated by the cost function in fully connected layer (sigmoid, softmax or other).

2.2. Deep Belief Network

The power of Deep Belief Network (DBN) (Fig. 3 and Fig. 4) lies in their ability to reconstruct both the input vector and the learning feature vectors, which is implemented using a layer-by-layer learning strategy [22]. Each layer of a DBN consists of a Restricted Boltzmann Machines (RBM). RBMs follow the principle of the probability distribution to complete its learning cycle [23]. Each RBM is concluded from a visible layer (v) and a hidden layer (h). Number of neurons is set up in each layer. The neurons between different layers are fully connected, and the neurons in the same layer are not connected [23]. When an RBM has learned, its feature activations are used as the “data” for training the next RBM in the DBNs [24]. RBMs is as an unsupervised network which considers the visible layer to the hidden layer as a subnetwork. Then, this hidden layer is considered as a visible layer to the next layer and so on [24]. The higher-level features are learned from the previous layers and the higher-level features are believed to be more complicated and better reflects the information contained in the input data’s structures [3]. DBN training is divided into two steps: forward pre-training process and reverse fine-tuning process [25]. During the pre-training phase, the RBMs are

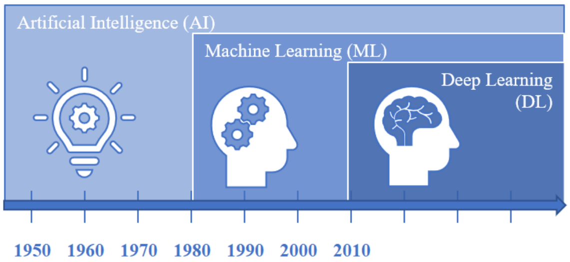

Figure 1: The connection of AI, ML and DL

trained one-by-one until the hidden layer of the last RBM. During this phase, the parameter of each RBM can be obtained [23]. The back-propagation network (BP) is set in the last hidden layer of the DBN [25]. BP is applied to fine tune the parameter using the output labels of the sample data [23].

2.3. Deep Boltzmann Machine

Deep Boltzmann Machine (DBM) have only one undirected network [24]. DBM as DBN is comprised of a Restricted Boltzmann Machines (RBM). The main difference is related to the interaction among layers of RBMs [25]. For the computation of the conditional probability of the hidden units h1, both the lower visible layer v and the upper hidden layer h2 are incorporated, that makes DBM differentiated from DBN and also more robust to noisy observation [15]. There are no direct connections between the units in the same layers. DBM parameters of all layers can be optimized jointly by following the approximate gradient of a variational lower-bound on the likelihood objective [26].

Different from the DBN, the DBM can incorporate top-down feedback, which can better propagate uncertainty and hence deal with ambiguous input more robustly [27].

2.4. Deep Neural Network

Due to the novelty of the concept Deep neural network (DNN) (Fig. 5)in the scientific literature can be identified for all the algorithms analyzed in this paper. However, in recent years the concept of DNN has become known as Artificial Neural Network (ANN) with hidden layers [9] [28]. DNN typically is feedforward network so it can be understood as the Multilayer Perceptron (MLP or MP). MLP consists of an input layer, several hidden layers and one output layer ant it's widely used for pattern classification, recognition and prediction [29].

2.5. Deep Q-Learning or Deep Reinforcement Learning

Deep Q-Learning (DQL) or Deep reinforcement learning (DRL) concept is replaceable in scientific literature (6). In DQL is always used reinforcement learning algorithm, or in DRL is often used Q-learning function, because of it is dealing with high-dimensional state space inputs [30], [31]. A reinforcement learning (RL) process involves an agent learning from interactions with its environment in discrete time steps in order to update its mapping between the perceived state and a probability of selecting possible actions (policy) [32]. In other words, RL is commonly used to solve an sequential decision making problem [30]. The RL problem is normally formalized using the Markov decision process (MDP) and includes a set of states S, set of actions A, transition function T as action distributions, reward function R and discount factor γ [33]. The solution to the MDP is a policy $\pi : S \longrightarrow A$ and the policy should maximize the expected discounted cumulative reward [30]. Q-learning, as a typical reinforcement learning approach, mimics human behaviors to take actions to the environment, in order to obtain the maximum long-term rewards [34]. The DQL process can be viewed as iteratively optimizing network parameters process according to gradient direction of the loss function at each stage [35]. Therefore, the inexact approximate gradient estimation with a large variance can largely deteriorate the representation performance of deep Q network by driving the network parameter deviated from the optimal setting, causing the large variability of DQL performance [35]. The advantages of deep Q-learning is good results and ease of use (code can be modified easy for different physical problems) [36].

2.6. The Extreme Learning Machine

The extreme learning machine (ELM) is a single-hidden layer feedforward network, proposed by Huang in 2012. In the traditional feed-forward ANN, the training of the network is iterative, while the process is transformed into an analytical equation in the ELM [37]. In ELM the weights between input and hidden layer are assigned randomly following a normal distribution and the weights between hidden and output layers are learnt in single step by a pseudo-inverse technique. During the training, the hidden layer is not learned but the weight matrix of output layer is obtained by solving the optimization problem formulated by some learning criteria and regularizations [38], as showed in the theory the output weights solved from regular-

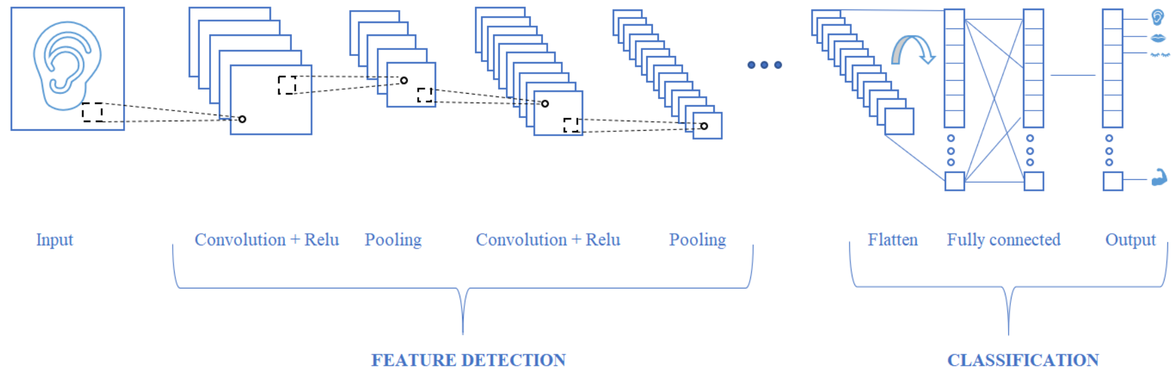

Figure 2: Convolution neural network architecture.

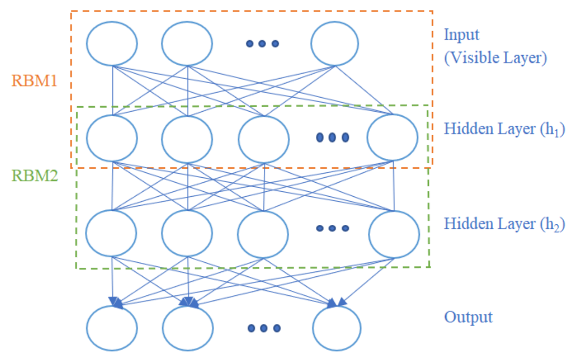

Figure 3: Deep Belief Network architecture.

ized least squares problem [39]. Therefore, ELM offers benefits such as fast learning speed, ease of implementation, and less human intervention when compared to the standard neural networks [40]

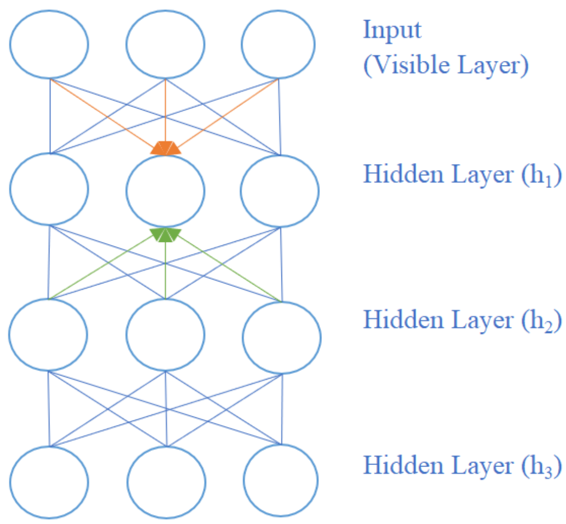

Figure 4: Deep Belief Network architecture

2.7. Generative Adversarial Network

The general idea of Generative adversarial network (GAN) is that it aims to train a generator to reconstruct high-resolution images for fooling a discriminator that is trained to distinguish generative images from real ones [41] (Fig. 8). This idea involves two competing neural network models: one of them takes noise as input and produces some samples (generator) and the other model (discriminator) accepts both the data outputted by the generator and the real data, meanwhile, separates their sources [42]. The Discriminator trains itself to discriminate real data and generated data better while the Generator trains itself to fit the real data distribution so as to fool Discriminator [43]. These two neural networks are trained at the same time, and finally the output is almost the same as the real data [44].

2.8. Recurrent Neural Learning

Recurrent Neural Learning (RNN) (Fig. 9) is different from the traditional feedforward neural networks, because have feedback connections, which can be between hidden units or from the output to the hidden units [44, 45]. This connections address the temporal relationship of inputs by maintaining internal states that have memory . An RNN is able to process the sequential inputs by having a recurrent hidden state whose activation at each step depends on that of the previous step [5, 46]. In other words, RNN not only processes the current element in the sequence, but also

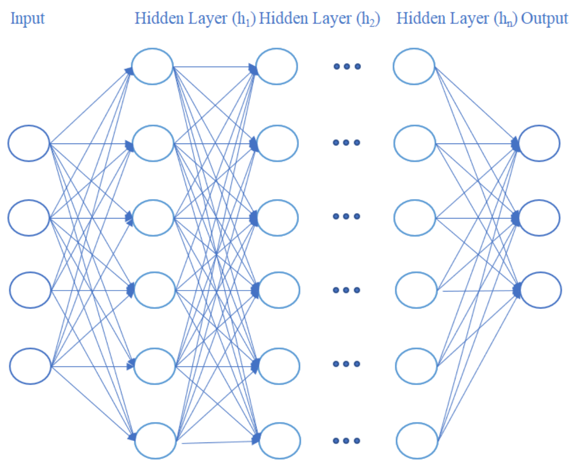

Figure 5: Deep neural network architecture

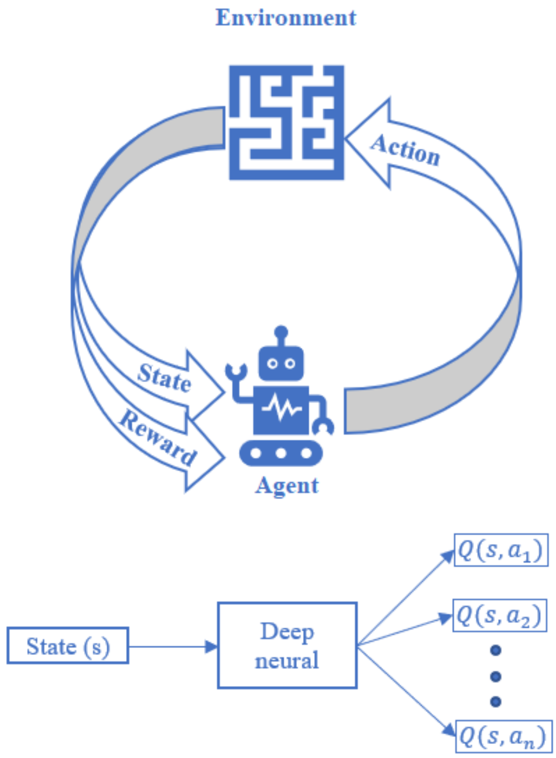

Figure 6: Deep Q-Learning architecture

draws upon the hidden layer of the previous element in the sequence [18]. For example, the states produced by an RNN at time t-1 will have some impacts on the states produced by the RNN at time t [17]. Hidden units can be regarded as the storage of the whole network, which remember the end-to-end information [47].

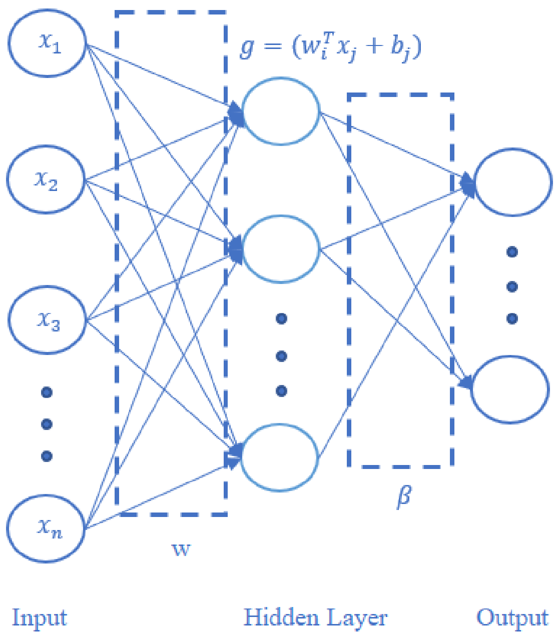

Figure 7: The extreme learning machine architecture

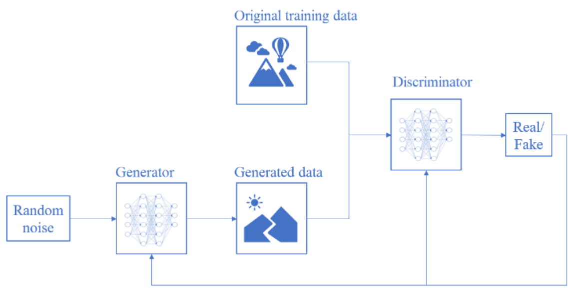

Figure 8: Generative adversarial network architecture

However, it has been observed that it is difficult to train the RNNs to deal with long-term sequential data, as the gradients tend to vanish [5].

2.9. Long short-term memory

Long short-term memory (LSTM) (Fig. 10) in literature is called one of the classes [13], advanced or extension [48] of RNN. The main advantage of LSTM is capability to learn longer dependencies in data [49] compared with RNN. Information sequentially is processed in LSTM, but there is a memory cell, which remembers and forgets information [48]. In each memory cell is three multiplication units: input gate, output gate and forget gate, which controls the flow of information [50]. The input gate determines how much cur-

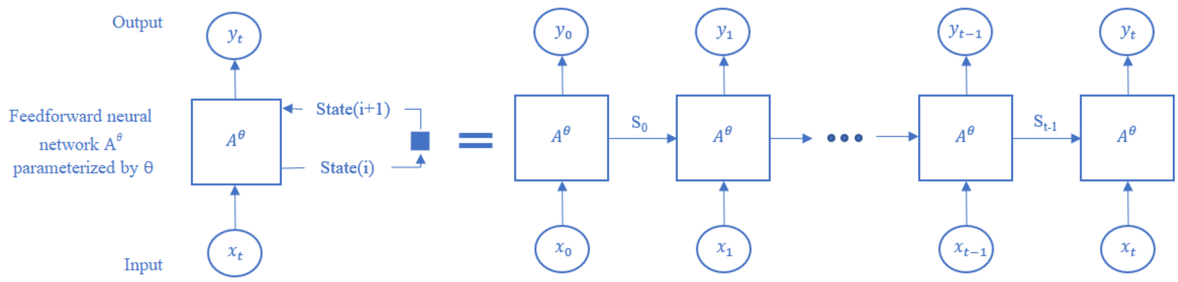

Figure 9: Recurrent Neural Learning architecture

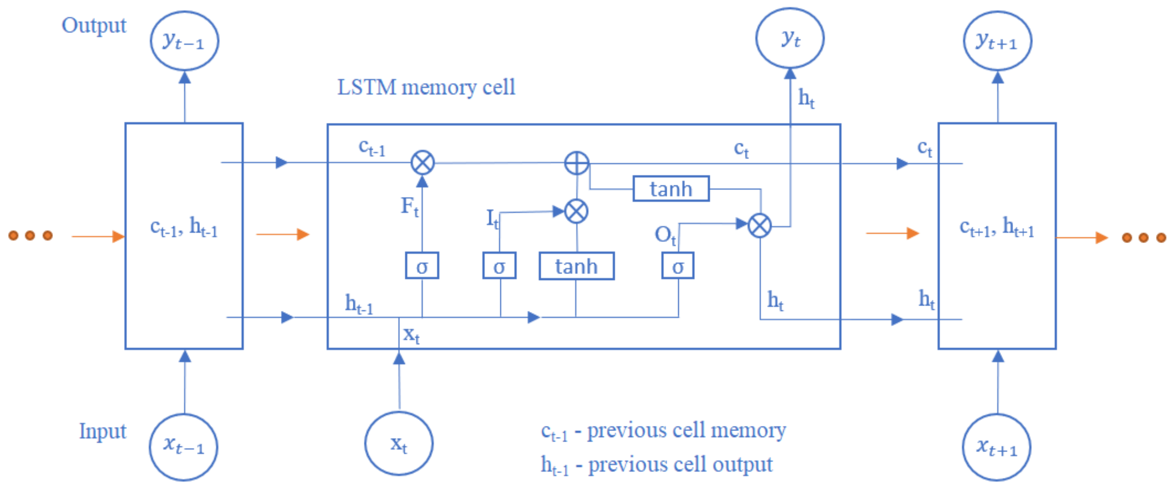

Figure 10: Long short-term memory architecture.

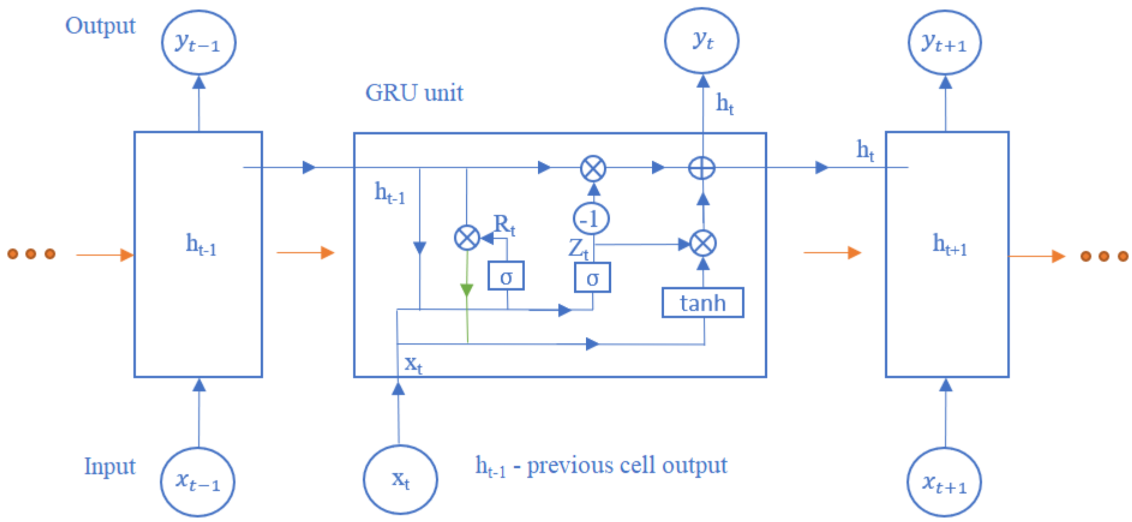

Figure 11: The extreme learning machine architecture

rent information should be treated as input in order to generate the current state [51], whilst the forget gate determine which information to be forgotten from the memory state [52]. Finally, the output gate filters the information that can be actually treated as significant and produces the output [52]. The "gate" structure is implemented using the sigmoid function, which denotes how much information can be allowed to pass. For one hidden layer in LSTM, activation function is used in forward propagation, and gradient is used in backward propagation [38].

2.10. Gated recurrent unit

Gated Recurrent Unit (GRU) is aimed to solve the vanishing gradient problem which comes with a standard RNN [53]. GRU consists of two gates: update gate (zt) and reset gate (rt). Update gate decides how much the unit updates its activation, or content, and reset gate allows forget the previously computed state [54]. GRU is a less complex compared with LSTM, it does not possess any internal memory and output gate like LSTM [49].

3. Application of Deep Learning Methods

Articles were included from electronic libraries: Science Direct, IEEE, Scopus, ACM, Emerald, SpringerLink, JSTOR, EBSCO and others.

Analyzed period started from 2017 till 2020. The review was conducted in January 2020. Keywords "Deep learning" and "Finance" were used for the article's selection. All methods presented in this review matches a term "Deep learning", wherefore individually search by each method was not developed. The same with a term "Finance", which includes accounting, financial markets, risks and etc. Therefore, this paper presents a big picture of developing scientific articles in Deep learning in Finance category. 33 papers were selected and analyzed. The analyzed articles can be categorized by the problematic of given task: to predict future returns or two make classification of results. Sometimes, for better results are used natural language processing algorithms (Fig. 12 and Tab 1).

The classification algorithms in finance most often have been applied for credit scoring, which divides loans into "good" and "bad". For this problem solving author's used DBN [29], modified LSTM [52] and CNN [55] networks. The results cannot be compared due to different classifier evaluation methods used and data source differences. In credit scoring topic is a big

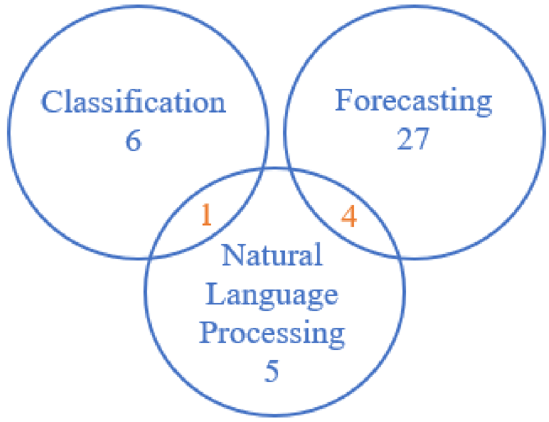

Figure 12: Categorical classification of analyzed articles

problem with unbalanced data set, i.e. authors [55] used data set, where credit worthy instances were 91,55 proc. and CNN accuracy rate was 91,64 proc. In the bankruptcy and investment market structure was used CNN network or in tax evasion DQL network. Articles in financial field is interested to obtain knowledge from words and used it as some indicators. Therefore, is seen a trend to use natural language processing techniques. The goal of natural language processing (NLP) is to process text using the computational linguistics, text analysis, machine learning, statistical and linguistics knowledge in order to analyze and extract significant information [56]. Researches in financial field are using sentiment analysis for better stock price prediction or bankruptcy classification. Sentiment analysis is the essential task for NLP, which can be divided into three categories: lexicon-based sentiment analysis, machine learning-based sentiment analysis and the hybrid approach [56].

Lexicon-based sentiment analysis was used only in one article [11], due to the need of opinion lexicon in this field. Machine learning-based sentiment analysis uses in bag-of-words method [48], [57] and word embeddings [48, 58, 57, 10] with CNN [58, 57], LSTM [48, 10] methods.

In [4] research was used bag of words and word embeddings methods for LSTM, results showed that LSTM models can outperform all traditional machine learning models based on the bag-of-words approach, especially when further pre-train word embeddings with transfer learning. The main financial article's focus is in future returns prediction, especially in stock prices or stock indexes. The main reason is data source availability for scientific research. In this field very often, scientific researches combine different methods

Table 1
Detailed topics from Finance perspective

Classification	Forecasting
Bankruptcy	Currencies, cryptocurrencies
Credit scoring	Demand, X prices
Investment market structure	Investment strategies, risk behavior
Tax evasion	Option, Derivatives
Natural Language Processing	Portfolio Management (forecasting stock's prices in the portfolio and portfolio optimization)
Bankruptcy	Stock Index
Stock Price	Stock Price

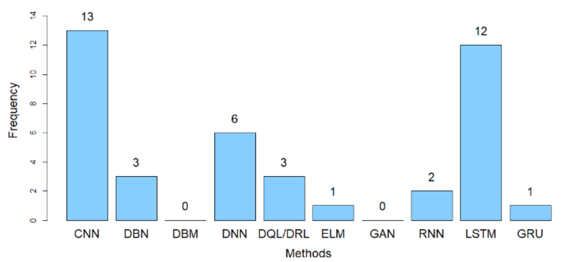

Figure 13: Use of deep learning methods in financial context.

together [49, 59, 60] or make some model modification's [13, 50, 61, 62] for better prediction results. Some authors [48, 63] analyze several different deep learning models results for the deeper future model development, see Fig. 13.

Most popular methods are CNN and LSTM. However, DBM and GAN method's was not found any adjustment in finance field.

In some papers data is not normalized, i.e. cryptocurrency prices [51] or demand [18]. Therefore, predictive accuracy measurements, such as RMSE, MPE and others, can be comparable with different other authors works or sometimes even in the same paper, i.e. RMSE for Bitcoin is 2.75×103 or for Ripple 0.0499 [51].

4. Conclusions

learning machine, generative adversarial network, recurrent neural learning, long short-term memory, gated recurrent unit; and it's applicability in finance field. This review reveals that financial article's:

1. mainly focus for the forecasting task than classification;
2. starts using natural language processing techniques, mostly sentiment analysis, for better results prediction;
3. uses not 'basic' the deep learning methods, i.e. they are often combined with several different models or merged to voting classifier.

Furthermore, this analysis has shown the importance of balanced data set and normalization of the data, which is submitted to deep learning networks.

The main limitation of this work is representation only a big picture of developing scientific articles in Deep learning in Finance category. Therefore, in future research is needed to extend search keyword's in electronic libraries, i.e. search by each method

References

[1] C. Huang, X. Wang, Financial innovation based on artificial intelligence technologies, in: Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science, 2019, pp. 750–754.

[2] P. Yeoh, Artificial intelligence: accelerator or panacea for financial crime?, Journal of Financial Crime (2019).

[3] D. Mo, A survey on deep learning: one small step toward ai, Dept. Computer Science, Univ. of New Mexico, USA (2012).

[4] I. M. Cockburn, R. Henderson, S. Stern, The impact of artificial intelligence on innovation, Technical Report, National bureau of economic research, 2018.

[5] L. Mou, P. Ghamisi, X. X. Zhu, Deep recurrent neural networks for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 55 (2017) 3639–3655.

[6] F. Beritelli, G. Capizzi, G. Lo Sciuto, C. Napoli, M. Woźniak, A novel training method to preserve generalization of rbpnn classifiers applied to ecg signals diagnosis, Neural Networks 108 (2018) 331–338.

[7] F. Beritelli, G. Capizzi, G. Lo Sciuto, C. Napoli, F. Scaglione, Rainfall estimation based on the intensity of the received signal in a lte/4g mobile terminal by using a probabilistic neural network, IEEE Access 6 (2018) 30865–30873.

[8] J. Hearty, Advanced Machine Learning with Python, Packt Publishing Ltd, 2016.

[9] O. Lachiheb, M. S. Gouider, A hierarchical deep neural network design for stock returns prediction, Procedia Computer Science 126 (2018) 264–272.

[10] D. Katayama, Y. Kino, K. Tsuda, A method of sentiment polarity identification in financial news using deep learning, Procedia Computer Science 159 (2019) 1287–1294.

[11] M.-Y. Day, C.-C. Lee, Deep learning for financial sentiment analysis on finance news providers, in: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), IEEE, 2016, pp. 1127–1134.

[12] G. Capizzi, G. Lo Sciuto, C. Napoli, D. Polap, M. Wozniak, Small lung nodules detection based on fuzzy-logic and probabilistic neural network with bioinspired reinforcement learning, IEEE Transactions on Fuzzy Systems 28 (2020).

[13] Wang, Wuyu, Li, Weizi, Zhang, Ning, Liu, Kecheng, Portfolio formation with preselection using deep learning from long-term financial data, Expert Systems with Applications 143 (2020) 11–42.

[14] Ongsulee, Pariwat, Artificial intelligence, machine learning and deep learning, in: 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), IEEE, 2017, pp. 1–6.

[15] Suk, Heung-Il, An introduction to neural networks and deep learning, in: Deep Learning for Medical Image Analysis, Elsevier, 2017, pp. 3–24.

[16] MATHWORKS, Introducing Deep Learning with MATLAB, MATHWORKS, 2020.

[17] Zheng, Chao, Wang, Shaorong, Liu, Yilu, Liuand, Chengxi, A novel rnn based load modelling method with measurement data in active distribution system, Electric Power Systems Research 166 (2019) 112–124.

[18] Kraus, Mathias, Feuerriegel, Stefan, Oztekin, Asil, Deep learning in business analytics and operations research: Models, applications and managerial implications, European Journal of Operational Research 281 (2020) 628–641.

[19] Chollet, Francois, Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek, MITP-Verlags GmbH & Co. KG, 2018.

[20] keras, Guide to the Sequential model - Keras Documentation, keras, 2020.

[21] G. Capizzi, G. L. Sciuto, P. Monforte, C. Napoli, Cascade feed forward neural network-based model for air pollutants evaluation of single monitoring stations in urban areas, International Journal of Electronics and Telecommunications 61 (2015) 327–332.

[22] D. Saif, S. El-Gokhy, E. Sallam, Deep belief networks-based framework for malware detection in android systems, Alexandria engineering journal 57 (2018) 4049–4057.

[23] Balakrishnan, Nagaraj, Rajendran, Arunkumar, Pelusi, Danilo, Ponnusamy, Vijayakumar, Deep belief network enhanced intrusion detection system to prevent security breach in the internet of things, Internet of Things (2019) 100112.

[24] J. Karhunen, T. Raiko, K. Cho, Unsupervised deep learning: A short review, in: Advances in Independent Component Analysis and Learning Machines, Elsevier, 2015, pp. 125–142.

[25] Fan, Chaodong, Ding, Changkun, Zheng, Jinhua, Xiao, Leyi, Ai, Zhaoyang, Empirical mode decomposition based multi-objective deep belief network for short-term power load forecasting, Neurocomputing 388 (2020) 110–123.

[26] N. Srivastava, R. R. Salakhutdinov, Multimodal learning with deep boltzmann machines, in: Advances in neural information processing systems, 2012, pp. 2222–2230.

[27] S. Wang, M. He, Z. Gao, S. He, Q. Ji, Emotion recognition from thermal infrared images using deep boltzmann machine, Frontiers of Computer Science 8 (2014) 609–618.

[28] K. Akyol, Comparing of deep neural networks and extreme learning machines based on growing and pruning approach, Expert Systems with Applications 140 (2020) 112875.

[29] C. Luo, D. Wu, D. Wu, A deep learning approach for credit scoring using credit default swaps, Engineering Applications of Artificial Intelligence 65 (2017) 465–470.

[30] W. Huang, Y. Wang, X. Yi, Deep q-learning to preserve connectivity in multi-robot systems, in: Proceedings of the 9th International Conference on Signal Processing Systems, 2017, pp. 45–50.

[31] Matta, Cardarilli, D. Nunzio, Fazzolari, Giardino, Nannarelli, Re, Spano, A reinforcement learning-based qam/psk symbol synchronizer, Ieee Access 7 (2019) 124147–124157.

[32] M. Ramicic, A. Bonarini, Attention-based experience replay in deep q-learning, in: Proceedings of the 9th International Conference on Machine Learning and Computing, 2017, pp. 476–481.

[33] H. Shen, H. Hashimoto, A. Matsuda, Y. Taniguchi, D. Terada, C. Guo, Automatic collision avoidance of multiple ships based on deep q-learning, Applied Ocean Research 86 (2019) 268–288.

[34] C. Qiu, F. R. Yu, F. Xu, H. Yao, C. Zhao, Blockchain-based distributed software-defined vehicular networks via deep q-learning, in: Proceedings of the 8th ACM Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications, 2018, pp. 8–14.

[35] W.-Y. Zhao, X.-Y. Guan, Y. Liu, X. Zhao, J. Peng, Stochastic variance reduction for deep q-learning, arXiv preprint arXiv:1905.08152 (2019).

[36] I. Sajedian, H. Lee, J. Rho, Design of high transmission color filters for solar cells directed by deep q-learning, Solar Energy 195 (2020) 670–676.

[37] B. Çil, H. Ayyıldız, T. Tuncer, Discrimination of β-thalassemia and iron deficiency anemia through extreme learning machine and regularized extreme learning machine based decision support system, Medical Hypotheses 138 (2020) 109611.

[38] J. Chen, Y. Zeng, Y. Li, G.-B. Huang, Unsupervised feature selection based extreme learning machine for clustering, Neurocomputing 386 (2020) 198–207.

[39] H. Li, X. Yang, Y. Li, L.-Y. Hao, T.-L. Zhang, Evolutionary extreme learning machine with sparse cost matrix for imbalanced learning, ISA transactions 100 (2020) 198–209.

[40] S. Shukla, B. S. Raghuwanshi, Online sequential class-specific extreme learning machine for binary imbalanced learning, Neural Networks 119 (2019) 235–248.

[41] R. Jiang, X. Li, A. Gao, L. Li, H. Meng, S. Yue, L. Zhang, Learning spectral and spatial features based on generative adversarial network for hyperspectral image super-resolution, in: IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2019, pp. 3161–3164.

[42] Y. Cui, W. Wang, Colorless video rendering system via generative adversarial networks, in: 2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), IEEE, 2019, pp. 464–467.

[43] Z. Zhai, J. Zhai, Identity-preserving conditional generative adversarial network, in: 2018 International Joint Conference on Neural Networks (IJCNN), IEEE, 2018, pp. 1–5.

[44] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep learning, volume 1, MIT press Cambridge, 2016.

[45] F. Bonanno, G. Capizzi, C. Napoli, Some remarks on the application of rnn and prnn for the charge-discharge simulation of advanced lithium-ions battery energy storage, in: International Symposium on Power Electronics Power Electronics, Electrical Drives, Automation and Motion, IEEE, 2012, pp. 941–945.

[46] M. Miljanovic, Comparative analysis of recurrent and finite impulse response neural networks in time series prediction, Indian Journal of Computer Science and Engineering 3 (2012) 180–191.

[47] C. Yin, Y. Zhu, J. Fei, X. He, A deep learning approach for intrusion detection using recurrent neural networks, Ieee Access 5 (2017) 21954–21961.

[48] M. Kraus, S. Feuerriegel, Decision support from financial disclosures with deep neural networks and transfer learning, Decision Support Systems 104 (2017) 38–48.

[49] A. J. Balaji, D. H. Ram, B. B. Nair, Applicability of deep learning models for stock price forecasting an empirical study on bankex data, Procedia computer science 143 (2018) 947–953.

[50] Y. Chen, K. He, G. K. Tso, Forecasting crude oil prices: a deep learning based model, Procedia computer science 122 (2017) 300–307.

[51] S. Lahmiri, S. Bekiros, Cryptocurrency forecasting with deep learning chaotic neural networks, Chaos, Solitons & Fractals 118 (2019) 35–40.

[52] C. Wang, D. Han, Q. Liu, S. Luo, A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism lstm, IEEE Access

7 (2018) 2161–2168.
[53] Y. Santur, Sentiment analysis based on gated recurrent unit, in: 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), IEEE, 2019, pp. 1–5.
[54] J. Kim, H. Kim, et al., Classification performance using gated recurrent unit recurrent neural network on energy disaggregation, in: 2016 international conference on machine learning and cybernetics (ICMLC), volume 1, IEEE, 2016, pp. 105–110.
[55] B. Zhu, W. Yang, H. Wang, Y. Yuan, A hybrid deep learning model for consumer credit scoring, in: 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), IEEE, 2018, pp. 205–208.
[56] A. Abdi, S. M. Shamsuddin, S. Hasan, J. Piran, Deep learning-based sentiment classification of evaluative text based on multi-feature fusion, Information Processing & Management 56 (2019) 1245–1259.
[57] H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, K. Muhammad, A local and global event sentiment based efficient stock exchange forecasting using deep learning, International Journal of Information Management 50 (2020) 432–451.
[58] F. Mai, S. Tian, C. Lee, L. Ma, Deep learning models for bankruptcy prediction using textual disclosures, European journal of operational research 274 (2019) 743–758.
[59] L. Ni, Y. Li, X. Wang, J. Zhang, J. Yu, C. Qi, Forecasting of forex time series data based on deep learning, Procedia computer science 147 (2019) 647–652.
[60] H. Yun, M. Lee, Y. S. Kang, J. Seok, Portfolio management via two-stage deep learning with a joint cost, Expert Systems with Applications 143 (2020) 113041.
[61] S. Huang, D. Wang, X. Wu, A. Tang, Dsanet: Dual self-attention network for multivariate time series forecasting, in: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019, pp. 2129–2132.
[62] A. M. Aboussalah, C.-G. Lee, Continuous control with stacked deep dynamic recurrent reinforcement learning for portfolio optimization, Expert Systems with Applications 140 (2020) 112891.
[63] C. Chen, P. Zhang, Y. Liu, J. Liu, Financial quantitative investment using convolutional neural network and deep learning technology, Neurocomputing 390 (2020) 384–390.