DL for financial applications (paper review)

Armine Hayrapetyan

Introduction

Stock market forecasting, algorithmic trading, credit risk assessment, portfolio allocation, asset pricing and derivatives market are among the areas where ML researchers focused on developing models that can provide real-time working solutions for the financial industry.

Our focus in this paper is to present different implementations of the developed financial DL models in such a way that the researchers and practitioners that are interested in the topic can decide which path they should take.

Introduction

In this paper, we tried to provide answers to the following research questions:

- What financial application areas are of interest to DL community?
- How mature is the existing research in each of these application areas?
- What are the areas that have promising potentials from an academic/industrial research perspective?
- Which DL models are preferred (and more successful) in different applications?
- How do DL models pare against traditional soft computing / ML techniques?
- What is the future direction for DL research in Finance?

Machine learning in finance

Finance has always been one of the most studied application areas for ML, starting as early as 40 years ago. So far, thousands of research papers were published in various fields within finance, and the overall interest does not seem to diminish anytime soon. some insights about previous ML studies by citing the related surveys within the last 20 years.

Deep learning

Deep Learning is a particular type of ML that consists of multiple ANN layers. It provides high-level abstraction for data modelling [21]. In the literature, different DL models exist: Deep Multilayer Perceptron (DMLP), CNN, RNN, LSTM, Restricted Boltzmann Machines (RBMs), Deep Belief Networks (DBNs), and Autoencoders (AEs).Each neuron in every layer has input (x), weight (w) and bias (b) terms. An output of a neuron in the neural network is illustrated in Equation

$$y_i = \sigma(\sum_i W_i x_i + b_i)$$

Deep Multi Layer Perceptron (DMLP)

With multi-layer deep ANNs, more efficient classification and regression performances are achieved when compared against shallow nets. DMLPs' learning process is implemented through backpropagation. The amount of the output error in the output layer neurons is also reflected back to the neurons in the previous layers. In DMLP, Stochastic Gradient Descent (SGD) method is (mostly) used for the optimization of learning (to update the weights of the connections between the layers)

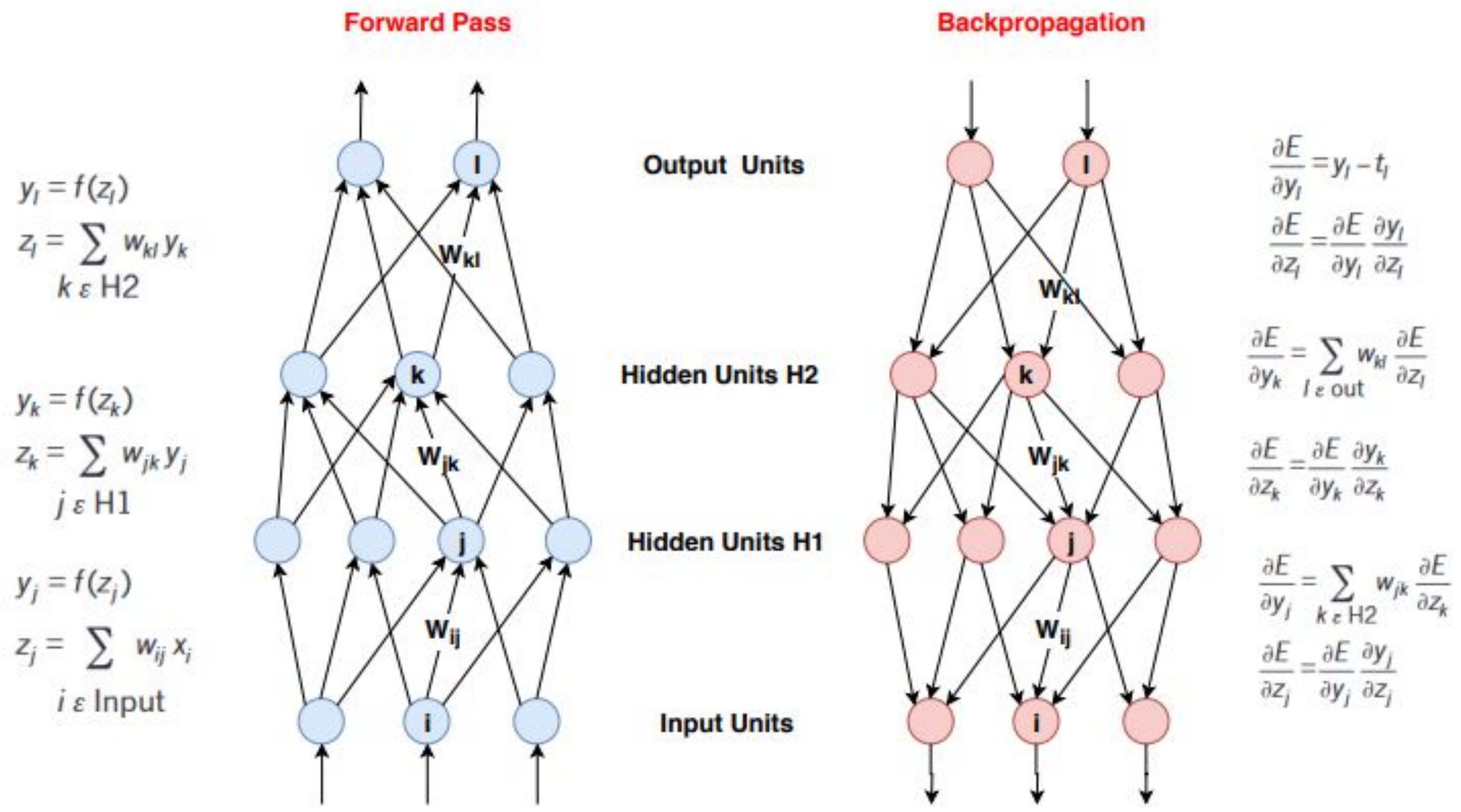

Figure 1: Deep Multi Layer Neural Network Forward Pass and Backpropagation [21]

Convolutional Neural Networks (CNNs)

CNN is a type of Deep Neural Network (DNN) that is mostly used for image classification, image recognition problems. In its methodology, the whole image is scanned with filters. In the literature, 1x1, 3x3 and 5x5 filter sizes are mostly used. In most of the CNN architectures, there are different types of layers: convolutional, pooling (average or maximum), fully connected layers. CNN consists of convolutional layers based on the convolutional operation.

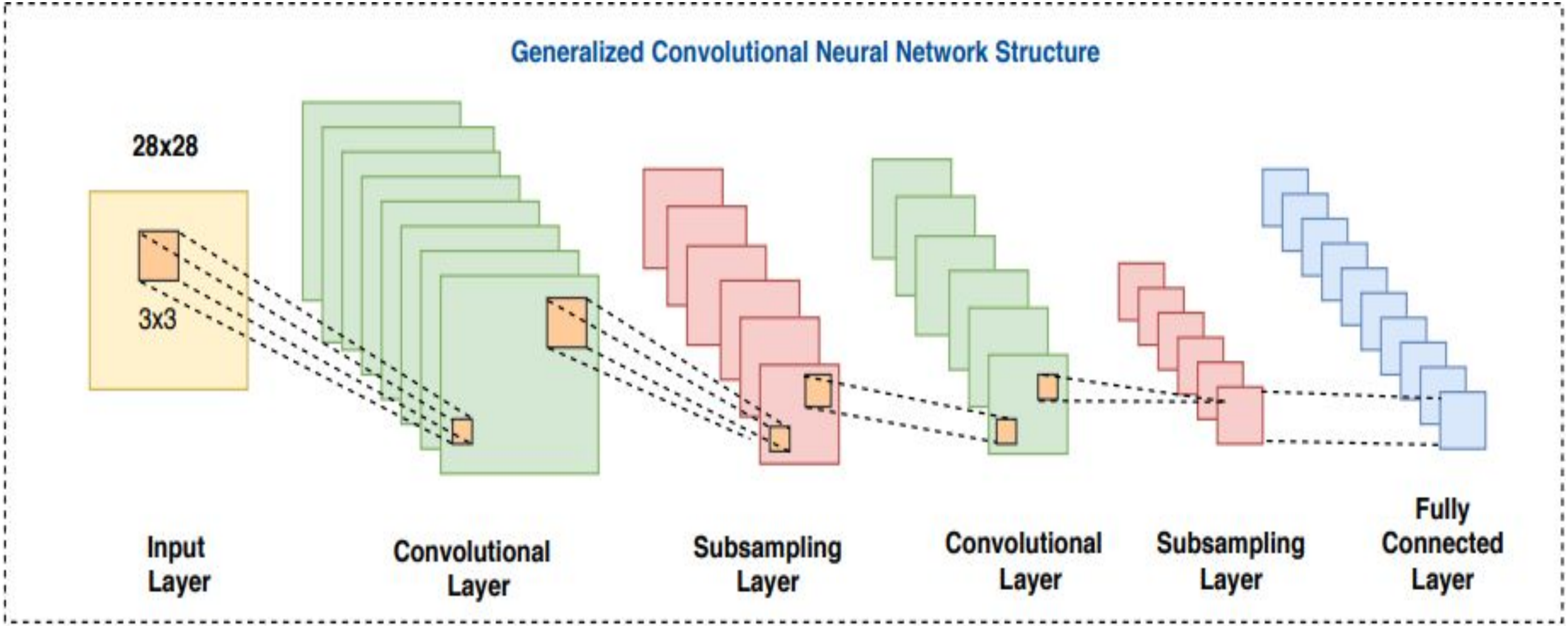

Figure 2: Generalized Convolutional Neural Network Architecture

Recurrent Neural Network (RNN)

In the literature, RNN has been mostly used on sequential data such as time-series data, audio and speech data, language. It consists of RNN units that are structured consecutively. Unlike feed-forward networks, RNNs use internal memory to process the incoming inputs. RNNs are used in the analysis of the time series data in various fields (handwriting recognition, speech recognition, etc).

There are different types of RNN structures: one to many, many to one, many to many. Generally, RNN processes the input sequence series one by one at a time, during its operation. Units in the hidden layer hold information about the history of the input in the "state vector" [21]. RNNs can be trained using the Backpropagation Through Time (BPTT) method. Using BPTT, the differentiation of the loss at any time t has reflected the weights of the network at the previous time. Training of RNNs are more difficult than Feedforward Neural Networks (FFNNs) and the training period of RNNs takes longer.

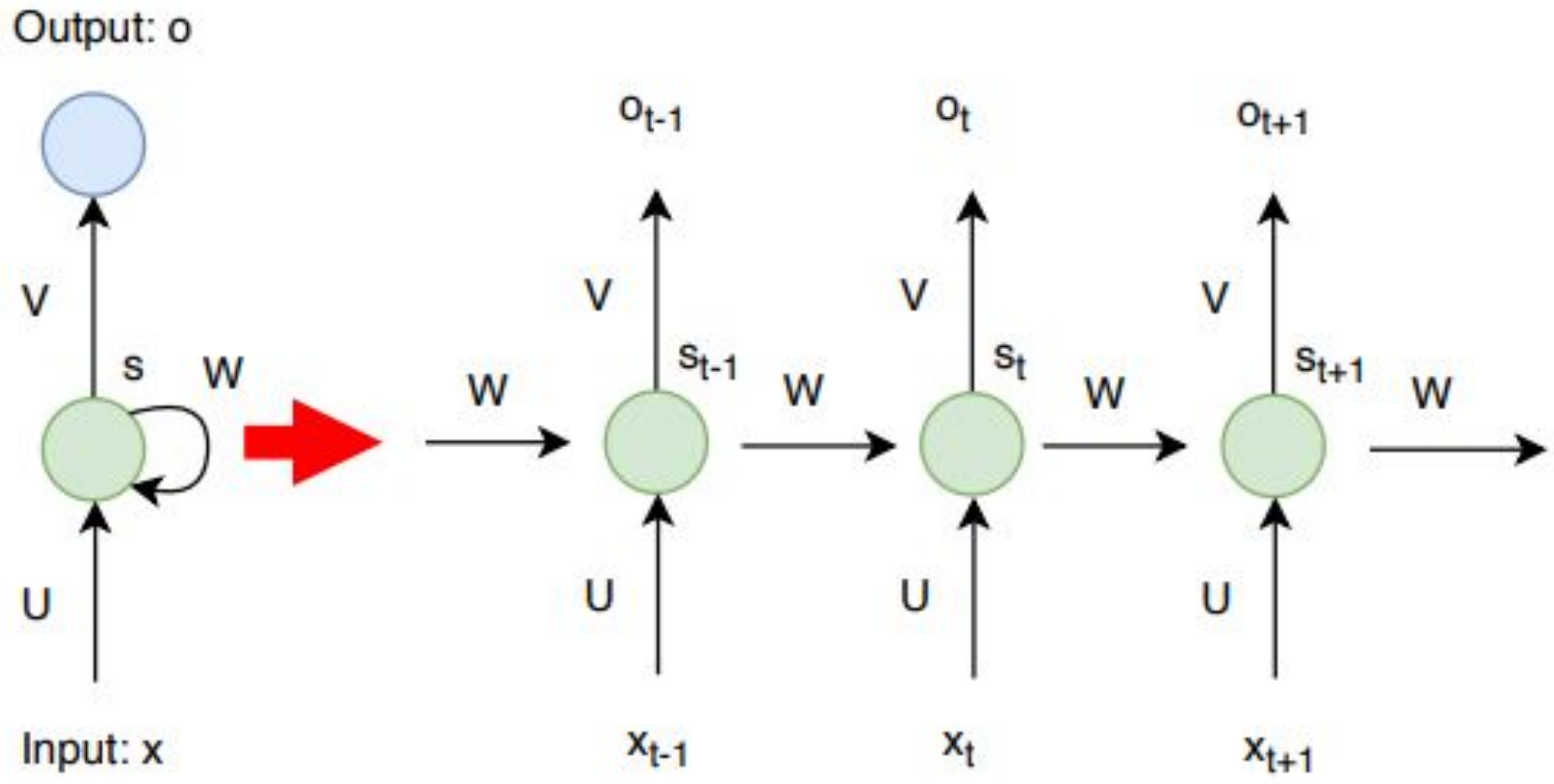

Figure 3: RNN cell through time[21]

Long Short Term Memory (LSTM)

LSTM network [28] is a different type of DL network specifically intended for sequential data analysis. The advantage of LSTM networks lies in the fact that both short term and long term values in the network can be remembered. Therefore, LSTM networks are mostly used for sequential data analysis (automatic speech recognition, language translation, handwritten character recognition, time-series data forecasting, etc.) by DL researchers. LSTM networks consist of LSTM units. LSTM unit is composed of cells having input, output and forget gates. These three gates regulate the information flow. With these features, each cell remembers the desired values over arbitrary time intervals. LSTM cells combine to form layers of neural networks.

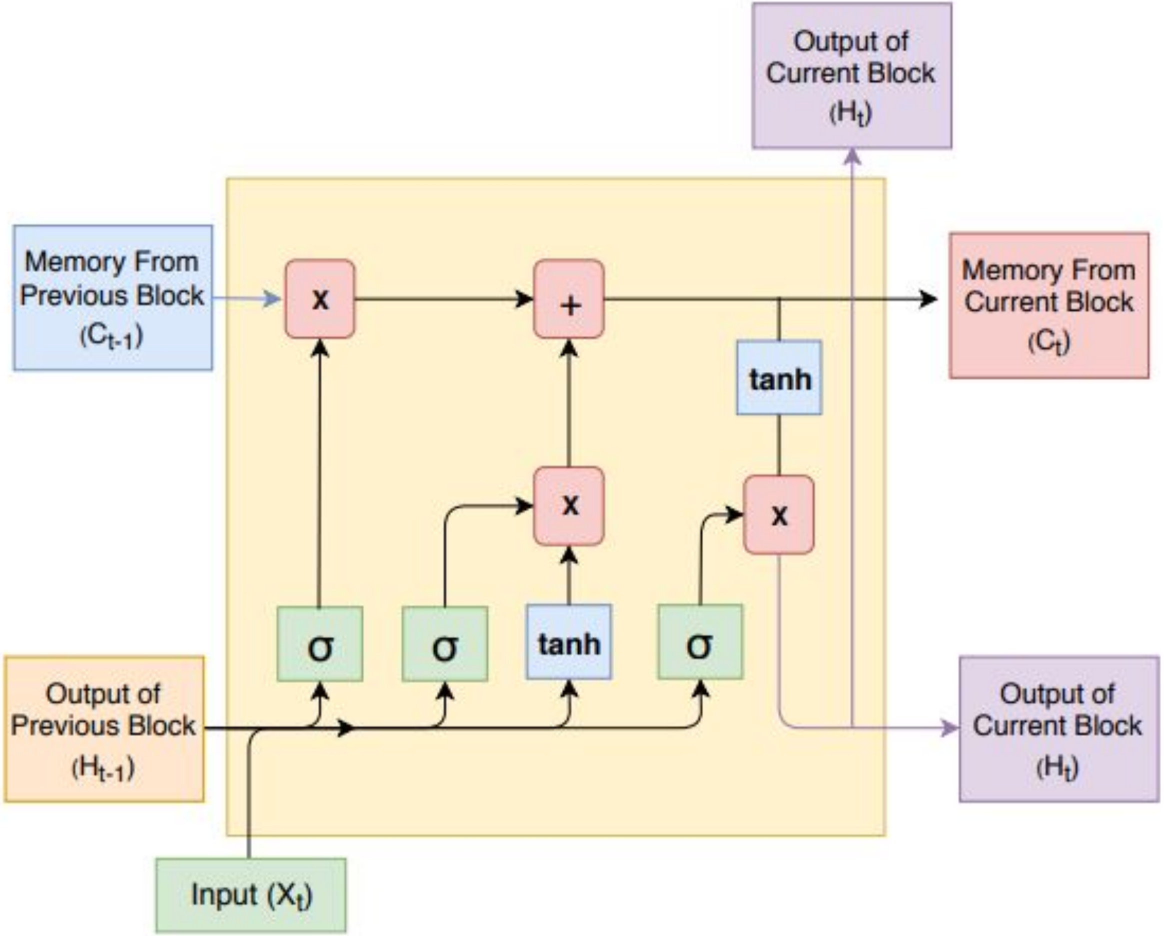

Figure 4: Basic LSTM Unit [28]

Financial Applications

Algorithmic trading

Algorithmic trading (or Algo-trading) is defined as buy-sell decisions made solely by algorithmic models. These decisions can be based on some simple rules, mathematical models, optimized processes, or as in the case of machine/deep learning, highly complex function approximation techniques. With the introduction of electronic online trading platforms and frameworks, algorithmic trading took over the finance industry in the last two decades. As a result, Algo-trading models based on DL also started getting attention.

Most of the Algo-trading studies were concentrated on the prediction of stock or index prices. Meanwhile, LSTM was the most preferred DL model in these implementations. In [35], market microstructures based trade indicators were used as the input into RNN with Graves LSTM to perform the price prediction for algorithmic stock trading.

Risk Assessment

Another study area that has been of interest to DL researchers is Risk Assessment which identifies the “riskiness" of any given asset, firm, person, product, bank, etc. Several different versions of this general problem exist, such as bankruptcy prediction, credit scoring, credit evaluation, loan/insurance underwriting, bond rating, loan application, consumer credit determination, corporate credit rating, mortgage choice decision, financial distress prediction, business failure prediction. Correctly identifying the risk status in such cases is crucial, since asset pricing is highly dependent on these risk assessment measures.

Fraud Detection

Financial fraud is one of the areas where the governments and authorities are desperately trying to find a permanent solution. Several different financial fraud cases exist such as credit card fraud, money laundering, consumer credit fraud, tax evasion, bank fraud, insurance claim fraud. This is one of the most extensively studied areas of finance for ML research and several survey papers were published accordingly

Portfolio Management

Portfolio Management is the process of choosing various assets within the portfolio for a predetermined period. As seen in other financial applications, slightly different versions of this problem exist, even though the underlying motivation is the same. In general, Portfolio Management covers the following closely related areas: Portfolio Optimization, Portfolio Selection, Portfolio Allocation. Sometimes, these terms are used interchangeably.

Asset Pricing and Derivatives Market (options, futures, forward contracts)

Accurate pricing or valuation of an asset is a fundamental study area in finance. There are a vast number of ML models developed for banks, corporates, real estate, derivative products, etc. However, DL has not been applied to this particular field and there are some possible implementation areas that DL models can assist the asset pricing researchers or valuation experts. There were only a handful of studies that we were able to pinpoint within the DL and finance community. There are vast opportunities in this field for future studies and publications.

Cryptocurrency and Blockchain Studies

In the last few years, cryptocurrencies have been the talk of the town due to their incredible price gain and loss within short periods. Even though price forecasting dominates the area of interest, some other studies also exist, such as cryptocurrency Algo-trading models. Meanwhile, Blockchain is a new technology that provides a distributed decentralized ledger system that fits well with the cryptocurrency world. As a matter of fact, cryptocurrency and blockchain are highly coupled, even though blockchain technology has a much wider span for various implementation possibilities that need to be studied. It is still in its early development phase, hence there is a lot of hype in its potentials.

Financial Sentiment Analysis and Behavioral Finance

One of the most important components of behavioral finance is emotion or investor sentiment. Lately, advancements in text mining techniques opened up the possibilities for successful sentiment extraction through social media feeds. There is a growing interest

in Financial Sentiment Analysis, especially for trend forecasting and Algo-trading model development. Kearney et al. [149] surveyed ML-based financial sentiment analysis studies that use textual data.

Nowadays there is broad interest in the sentiment analysis for financial forecasting research using DL models. Table 10 provides information about the sentiment analysis studies that are focused on financial forecasting and based on text mining.

Financial Text Mining

With the rapid spreading of social media and real-time streaming news/tweets, instant text-based information retrieval became available for financial model development. As a result, financial text mining studies became very popular in recent years. Even though some of these studies are directly interested in the sentiment analysis through crowdsourcing, there are a lot of implementations that are interested in the content retrieval of news, financial statements, disclosures, etc. through analyzing the text context. There are a few ML surveys focused on text mining and news analytics. Among the noteworthy studies of such, Mitra et al. [159] edited a book on news analytics in finance, whereas Li et al. [160], Loughran et al. [161], Kumar et al. [162] surveyed the studies of textual analysis of financial documents, news and corporate disclosures. It is worth to mention that there are also some studies [163, 164] of text mining for financial prediction models.

Current Snapshot of DL research for Financial Applications

For the survey, we reviewed 144 papers from various financial application areas. Each paper is analyzed according to its topic, publication type, problem type, method, dataset, feature set and performance criteria. Due to space limitations, we will only provide the general summary statistics indicating the current state of the DL for finance research.

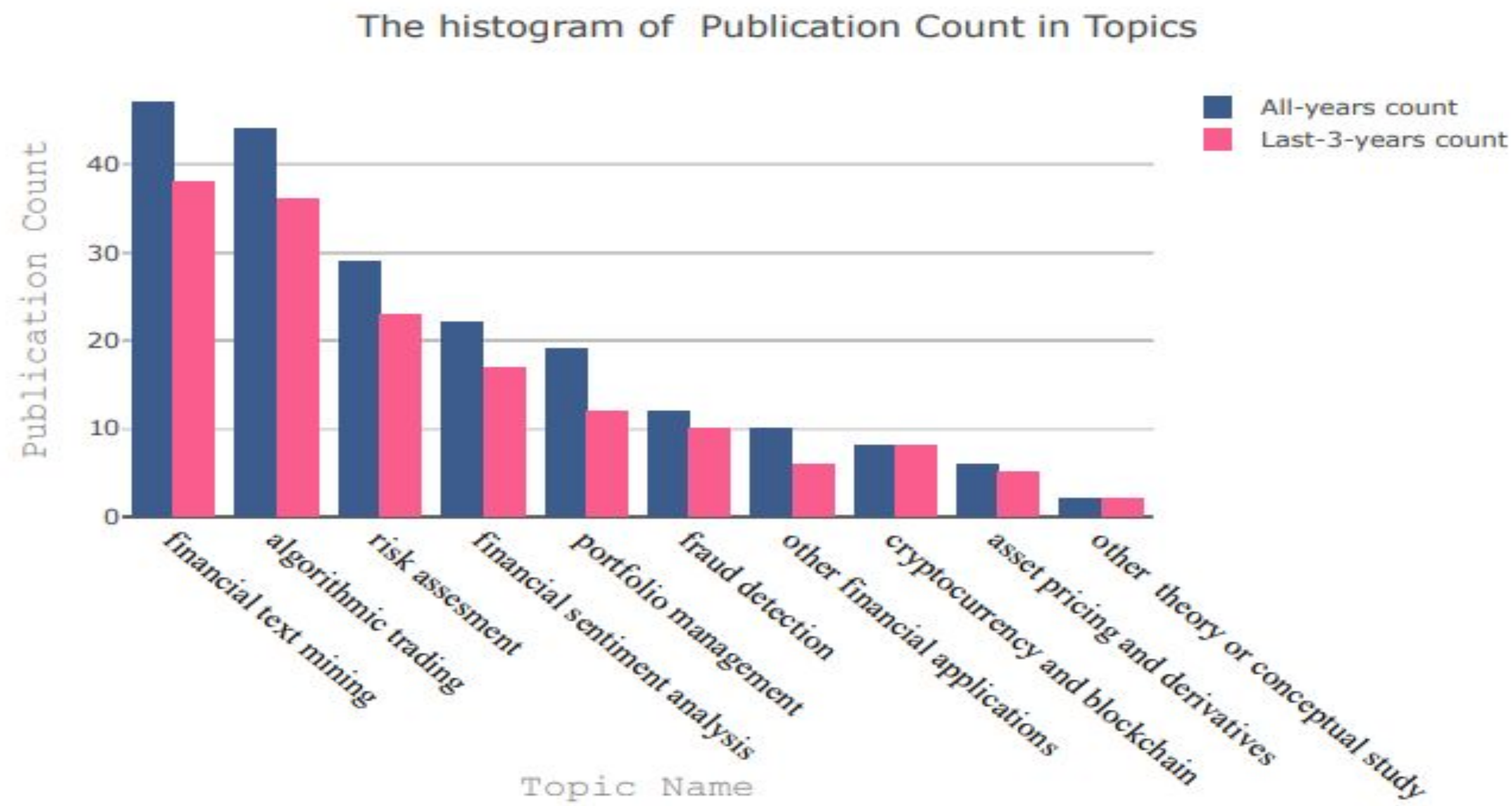

Figure 8: The histogram of Publication Count in Topics

Current Snapshot of DL research for Financial Applications

A quick glance at the figure shows us financial text mining and algorithmic trading are the top two fields that the researchers most worked on followed by risk assessment, sentiment analysis, portfolio management and fraud detection, respectively. The results indicate most of the papers were published within the last 3 years implying the domain is very hot and actively studied.

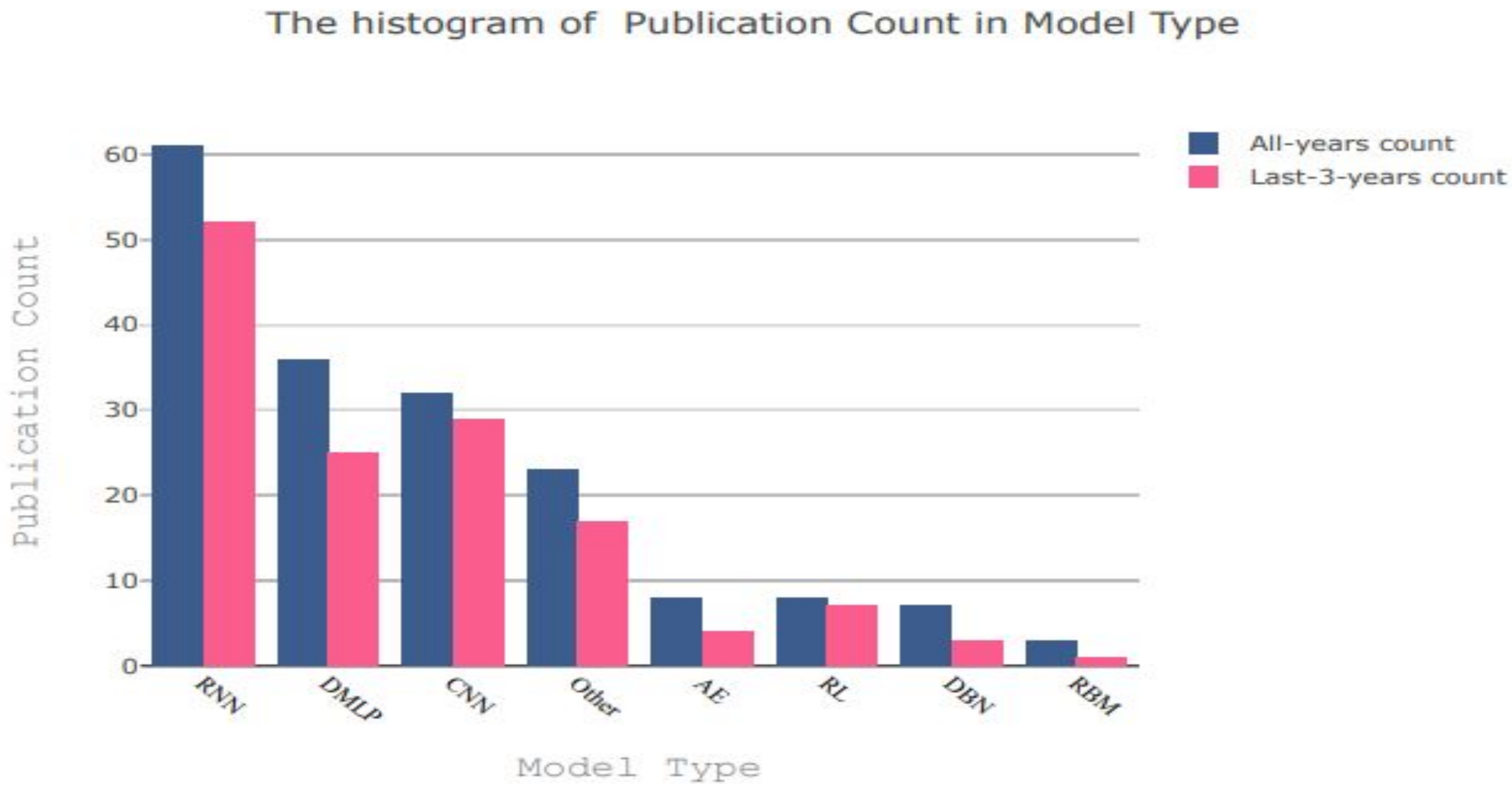

Figure 10: The histogram of Publication Count in Model Types

Popular NN architectures

When the papers were clustered by the DL model type as presented in Figure 10, we observe the dominance of RNN, DMLP and CNN over the remaining models, which might be expected, since these models are the most commonly preferred ones in general DL mplementations. Meanwhile, RNN is a general umbrella model which has several versions including LSTM, GRU, etc.

RNN/DMLP/CNN

Within the RNN choice, most of the models actually belonged to LSTM, which is very popular in time series forecasting or regression problems. It is also used quite often in algorithmic trading. More than 70% of the RNN papers consisted of LSTM models.

Meanwhile, DMLP generally fits well for classification problems; hence it is a common choice for most of the financial application areas. However, since it is a natural extension of its shallow counterpart MLP, it has a longer history than the other DL models.

CNN started getting more attention lately since most of the implementations appeared within the last 3 years. Careful analysis of CNN papers indicates that a recent trend of representing financial data with a 2-D image view in order to utilize CNN is growing. Hence CNN based models might overpass the other models in the future. It actually passed DMLP for the last 3 years.

Responses to our Initial Research Questions

- **What financial application areas are of interest to DL community?**
  - Response: Financial text mining, Algo-trading, risk assessments, sentiment analysis, portfolio management and fraud detection are among the most studied areas of finance research.
- **How mature is the existing research in each of these application areas?**
  - Response: Even though DL models already had better achievements compared to traditional counterparts in almost all areas, the overall interest is still on the rise in all research areas.

Responses to our Initial Research Questions

- **What are the areas that have promising potentials from an academic/industrial research perspective?**
  - Response: Cryptocurrencies, blockchain, behavioral finance, HFT and derivatives market have promising potentials for research.
- **Which DL models are preferred (and more successful) in different applications?**
  - Response: RNN based models (in particular LSTM), CNN and DMLP have been used extensively in implementations. From what we have encountered, LSTM is more successful and preferred in time-series forecasting, whereas DMLP and CNN are better suited to applications requiring classification.

Responses to our Initial Research Questions

- **How do DL models pare against traditional soft computing / ML techniques?**
  - Response: In most of the studies, DL models performed better than their ML counterparts. There were a few occasions where ML had comparable or even better solutions, however the general tendency is the outperformance of the DL methods.
- **What is the future direction for DL research in Finance?**
  - Response: Hybrid models based on Spatio-temporal data representations, NLP, semantics and text mining-based models might become more important in the near future.

Conclusions

The financial industry and academia have started realizing the potentials of DL in various application areas. The number of research work keeps on increasing every year with an accelerated fashion. However, we are just in the early years of this new era, more studies will be implemented and new models will keep pouring in. In this survey, we wanted to highlight the state-of-the-art DL research for the financial applications. We not only provided a snapshot of the existing research status but also tried to identify the future roadway for intended researchers. Our findings indicate there are incredible opportunities within the field and it looks like they will not disappear anytime soon. So, we encourage the researchers that are interested in the area to start exploring.

References

- https://www.quora.com/What-is-the-difference-between-deep-and-shallow-neural-networks
- https://github.com/hayrapetyan-armine/Time-Series-Forecasting/tree/master/papers