Application of Deep Neural Network for Stock Value Prediction and Analysis fromText Data

[1]N.Jagannadha Rao , [2]Kosna Durga Srinivasu,[3]Pothumudi Manasa,[4]Yennam Ramesh Kumar

Department of CSE, PRAGATI Engineering College (Autonomous),Surampalem, A.P, India.

ABSTRACT:

The recent advance of deep learning has enabled trading algorithms to predict stock price movements more accurately.Unfortunately, there is a significant gap in the real-world deployment of this breakthrough. For example, professionaltraders in theirlong-term careers have accumulated numerous trading rules, the mythof which they can understand quite well. On the other hand, deep learning models have been hardlyinterpretable. This paper presents Deep Clue, a system built to bridge text-based deep learning modelsand end users through visually interpreting the key factors learned in the stock price prediction model. We make three contributions inDeepClue. First, by designing the deep neural network architecture for interpretation and applying an algorithm to extract relevantpredictive factors, we provide a useful case on what can be interpreted out of the prediction model for end users. Second, by exploringhierarchies over the extracted factors and displaying these factors in an interactive, hierarchical visualization interface, we shed light onhow to effectively communicate the interpreted model to end users. Specially, the interpretation separates the predictable fromthe unpredictables for stock prediction through the use of intercept model parameters and a risk visualization design. Third, we evaluate theintegrated visualization system through two case studies in predicting the stock price with online financial news and company-relatedtweets from social media. Quantitative experiments comparing the proposed neural network architecture with state-of-the-art models andthe human baseline are conducted and reported. Feedbacks from an informal user study with domain experts are summarized and discussed in details. All the study results demonstrate the effectiveness of DeepClue in helping to complete stock market investment andanalysis tasks.

1. INTRODUCTION

DEEP learning techniques [1] are reshaping the landscape ofpredictive analysis in the big data research area and havemade major breakthroughs in image and speech recognition [2],question answering [3], machine translation [4] and many otherapplication domains. In thispaper, we focus on the financialanalytics domain. It has been shown that texts such as financialnews and tweets on stock markets are useful in predicting stockprice movements. For example, financial news such as “Amazonprofit beats forecasts” was accompanied with a surge of Amazon’sstock price, while “Oil price hits a record high” triggered worrieson the auto industry and weakened their performance in the stockmarket. Previous work has demonstrated an over 60% accuracyin predicting the daily stock price movement using deep neural. networks over a large collection of financial news. Nevertheless, end users can hardly benefit from these successfuldeep learning models in their primitive form. We consider two classes of users in this work: stock traders from public/privatefunds (or independent investors) who manage the stock tradingoperations; and stock market analysts, who provide the stockprediction models for traders. First, the everyday job of traders isto make trading decisions, i.e., to buy/sell which stock at which particular time. Such a decision is typically based on multiplesources of information known as trading signals, coming out ofa large number of trading rules accumulated in the long term.To cope with the trader’s job, there should be a method to helptraders detect signals from the prediction model, so that traders cancombine these signals with their traditional source of informationto finalize the decision. The automatic stock trading based only onthe prediction model can be an option, but it will require a muchhigher accuracy than that of the latest model. In some cases, aclose to 60% accuracy can even lead to losses (Section 6.2). On theother hand, analysts’ workis to fine-tune the stock price predictionmodel for particular stocks and market trends, in order to optimize the prediction accuracy. This will require analysts to have a deepunderstanding of failure cases of the prediction model.To this end, both classes of end users will benefit from deeplearning technology only if they can interpret the predictionmodel on where, when and why it works or does not work. Thisknowledge can then be assembled with the domain expertise to improve the investment in the stock market. Unfortunately, oninterpretability deep learning models suffer from a well-knowndrawback in contrast to traditional machine learning methods suchas linear regression and support vector machines (SVM). In someareas such as image recognition, the mechanism of deep learninghas been partially known, e.g., working as level-of-detail featureselectors, from the basic visual feature up to motifs and finally to objects [8] [9]. For most other domains, there is still little clue onhow deep learning models work. In our scenario, the use of textinput introduces an additional work embedding stageto map textcollections onto the feature space, which makes it more difficult tointerpret the prediction model.In this paper, we target the research problem of how tointerpret text-based deep stock prediction model for end users,so that they can make up their stock tradingdecisions as wellas improve the prediction model based on the interpretation. Inparticular, we investigate research questions includingwhat kind ofinformation can be efficiently extracted from the prediction

modelas interpretations, and how to communicate such information inan effective way to end users.Throughout this work, we dependon an interactive visualization interface to bridge the predictionmodel and end users, which turns outa natural and straightforwardchoice. Yet, designing and prototyping such a visualizationsystemcan be quite

challenging. First,traditional patterns discovered fromdata can be presented by visually distinct channels in the same dataview, while in this case, theinformation extracted from the modellies in a higher order than the data pattern. Multiple coordinatedviews should be

designed elaborately to illustrate the relationshipamong data, model, and interpretation. Second, the deep learningmodel is designed in a bottom-up structure to take advantage ofthe machine's capability in processing huge amount of data, whilethe visual information-seeking mantra is"overview first, details ondemand" [10]. Third,it is commonly accepted that the stock market is information efficient [11], but not all stock price movements are predictable or reflected intext

information. Ingenuity is

required toseparate predictable and unpredictable price changes.In the literature, there is a recent surgeon the topic

ofvisualizing deep neural networks (DNN) for model interpretation.A large portion of these methods focused on the display of neuralnetwork architecture to help users understand the functionality ofindividual neurons and features [8] [12], interpret the mechanism of both small-scale neural networks [13] and large-scale multi-layerDNNs [14]. Another thread of research proposed to visualizethe model output (e.g. the image class model [15]) or theircorrespondence to the input data through algorithms similar toback propagation [9]. While our study aligns with these successfulmethods on DNN model interpretation, thegoalis fundamentallydifferent. Instead of visually illustrating DNN structures, we targetat extracting useful information from the prediction model, andincorporating this interpretation with domain expertise to improvethe performance of stock trading and modeling. In addition, existingliterature mostly studied model interpretation for image recognitionand object detection tasks, while to our knowledge, we are thefirst to visually interpret the hidden linkage between public textcollections and stock prices through deep learning models.In summary, we make the following contributions.

1. Based on a customized DNN architecture for stock priceprediction (Section 3), we apply a model interpretationalgorithm, i.e., the pixel-based layer-wise relevance propagation[16], to extract the textual factors relevant to thedaily prediction result (Section 4.1). Notably, the extractedfactors (i.e., keywords, bigrams, titles) are analyzed to forma factor hierarchy for effective visual interpretation by endusers (Section 4.2).
2. An integrated visualization system called DeepClue isdesigned and applied to the stock price prediction scenario,which visually correlates algorithm-extracted textual factorswith stock price movements and the risks associated withthe text-based prediction. Flexibilities are granted to endusers in model configuration, factor analysis, and detailedreasoning. (Section 5)
3. We evaluate the proposed system through real-life casesin analyzing the text-based deep stock prediction modelbuilt from financial

news (Section 6.1) and social mediacollections (Section 6.2) on US stock markets. Quantitative 2experiments are conducted to compare the proposed neuralnetwork architecture with state-of-the-art models and thehuman baseline (Section 6.3). Informal user studies are thencarried out with private-fund stock traders and deep learningmodel builders, which demonstrate the value of DeepCluein optimizing stock trading operations and improving theprediction model of stock price movement (Section 6.4).

2 RELATED WORKS

DNN Interpretation and Visualization

Early research related to the DNNinterpretation can be foundin Erhan et al. [8], who introduced the concept of understandinga particular unit of DNN by visualizing inputsthat maximizethe unit's response. This activation maximization method wascompared with other alternatives including sampling the unitand linear combination of previous filters. Experiment results onimage data sets showedthat the activation maximizationmethodsproduced more interesting interpretation results. Zeiler et al. [9]proposed to map feature activations in neural nets back to inputsby deconvolution layers. Simonyan etal. [15] developed a class model visualization that generates a representative image for eachclass of interest, and a class saliency map for a single inputimage based on gradients with respect to the input pixel. Bachet al. [16] [17] introduceda class of algorithms named layer wiserelevance propagation (LRP), which decompose a neural netprediction layer by layer into scores for each neural unit, andapplied it to state of the art deep networks in image classification.These scores, when computed for the inputs, explain the amountof contribution of a pixel or region to the prediction value for a given class. Dudovskiy et al. [18] trained neural networks

Journal of Management & Entrepreneurship
ISSN 2229-5348
UGC Care Group I Journal
Vol-07 Issue-02 Nov 2018

to reconstruct inputs from feature representations. Zintgraf etal. [19] developed an elaborate conditional sampling algorithmto analyze how deep neural networks respond to perturbed inputs.Yosinski et al. [12] introduced tools to visualize the activationson neural network layers, and the features extracted at eachlayer through regularization. Liu et al. proposed CNNVis [14], avisual analytics approach that employs layer and neuron clustering.CNNVis introduced several novel visualization algorithms suchas hierarchical rectangle packing and matrix reordering to displayfeatures on clustered neurons.Visualizations have made their way into deep learning toolboxes.Besides the well-known deep dream [20], TensorFlowPlayground by Google [13] provided an online visualization toolfor non-experts to understand deep learning architecture and theirtraining process through a direct manipulation design. Overall, previous literature on deep learning model visualization concentrateson the scenario of image classification with CNNs. DeepClue, incontrast to these systems, is dedicated to stock market investorsfor better understanding the association between text streams andstock price time series. Moreover, rather than opening the Blackboxstructure of neural networks and interpreting the functionalityof each individual unit, our method focuses on extracting input levelinterpretable information from the DNN model and visuallyincorporating such information with domain expertise to improve the performance of stock trading and modeling.

Text-based Stock Prediction and Visualization

It has been pointed out by Kearney and Liu [21] that the complexand time-varying relationship between textual information and

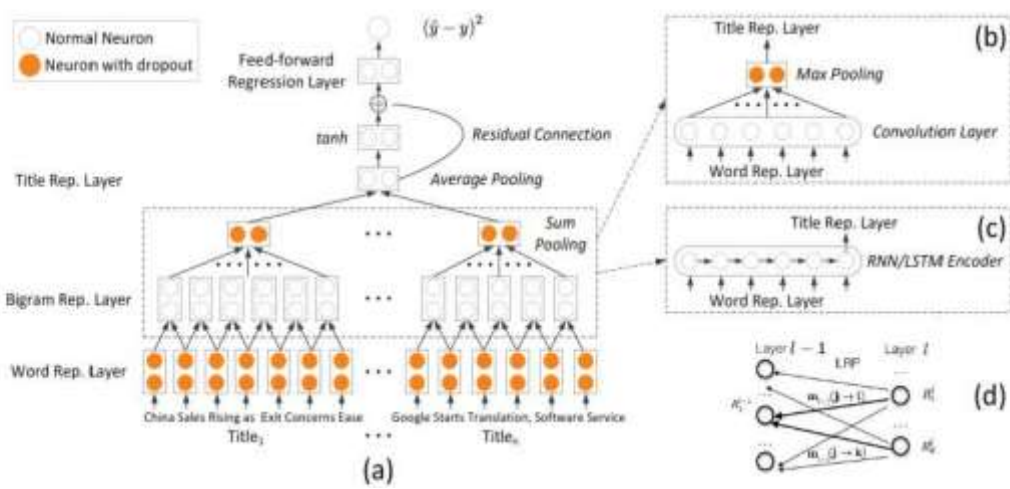

Fig. 1: Neural network architecture in this work: (a) the overall hierarchical structure featuring multiple representation layers fromsingle word to bigram and news/tweet title; (b) an alternative convolution layer that can replace the bigram representation and the sumpooling on each title; (c) another design by a recurrent neural network layer with LSTM cells to represent each title as time series; (d)the mechanism of LRP algorithm that propagates relevance scores back to the input features over the neural network.

stock price poses an important area of study for financial analysis.The textual data for stock price prediction comes primarily fromthree data sources: public corporate filings, news articles, and theemerging social media content. This work focuses on the latter twosources. On using news articles for the prediction, Engelberg [22]and Tetlock et al. [23] employed firm-specific news from multiplesources to predict firms' fundamentals, which inherently influencetheir stock prices. In addition, social media content, especiallytextual sentiments, has shown implicit effects on the stock market.

The work by Chen et al. revealed that the views expressed on apopular social media site for investors have strong associationswith the related firm's stock returns, thus helps to predict theirstock price changes [24].On stock market visualization, most state-of-the-art literaturefocused on the display of stock price time series. Similarity andcluster analysis have been employed to group the stock pricetime series into trajectories to optimize the visualization. Forexample, Ziegler et al. visually analyzed the distribution of timeseries trajectories among different market sectors [25]. Keimet al. presented the Growth Matrix visualization [26] for thesimultaneous display of growth rates of all possible subintervalsin a time series. Beyond the time series visualization, many otherdesigns incorporate the related news and events to the displayof stock price time series. Contexture [27] produced annotatedstock price visualizations given news articles as the informationsource. Sorenson and Brath proposed a system to visualize a largecollection of stock-related events in a single view [28]. The eventdisplay can be visually correlated with the stock price time seriesfor reasoning. Compared to the DeepClue visualization, most aboveworks feature a direct visualization of raw stock price time seriesand the temporally correlated news/events. There has been littleresearch on visualizing the predictive linkage between the stockprice and the textual information, which is extracted from a state of-the-artdeep learning model.

3 TEXT-BASED STOCK PRICE PREDICTION

Data Collection

We consider S&P 500 stocks in the US stock market from 2006to 2015. Their historical prices are acquired from Yahoo Finance.We crawled financial news from Reuters and Bloomberg, obtainingin total 341,310 news articles. For each news, we extracted thetitle, textual content, and timestamp from their raw HTML file. Tomap each news to the corresponding stocks, we maintained a listof keywords for each firm (e.g., Apple: AAPL, AAPL.O, APPLE,AAPL.N, Apple Inc, etc.).

The stock-related tweets were collected through Twitter APIin a period from April 2015 to November 2015, by matching thefirm's cashtags in the tweet content. Cashtag [29] is a new way ofsharing financial information on social media developed by Twitterand other providers. The firm's stock ticker symbols are prefixedwith a dollar sign to compose the cashtag, e.g., Apple=$AAPL,Google=$GOOG. In total, we obtained 6,869,771 stock-relatedtweets. For each tweet, we recorded the create time, textual content,source, user, location and related firms.

Deep Neural Network Architecture

We take news data as an example to introduce the architectureof the neural network model adopted in this work. The model isbuilt for each particular S&P 500 firm. The goal of the model isto predict a stock price ˆy that is close to the real stock price y ofthe firm. The raw input of each model is the set of financial newstitles collected on the target firm. Intuitively, news content canbe useful for further enhancing the prediction accuracy. However,preliminary experiments using both news title and content as inputs(Supplemental Material–Table I) show that our model does notbenefit from the additional content information, compared withonly using the news title (Figure 11(a)). This is consistent withthe observations of Ref. [5], who extract event information fromboth news title and content, showing that it does not substantiallyimprove a model with only new title as the information source.Therefore, we leave it to future work to further exploit theusefulness of news content information.Figure 1(a) shows our proposed deep regression model organizedin a hierarchical neural network structure. The networkconsists of four layers: a word representation layer, a bigramrepresentation layer, a title representation layer, and a feed-forwardregression layer. The word representation layer accepts all the newstitles as input and turns each word in the title into a real-valuedword embedding vector [30]. The bigram representation layerconstructs representation vectors for word bigrams based on therepresentation vector of individual words. The title representationlayer summarizes representations of word bigrams and encodeseach title into a title vector. The feed-forward regression layerreceives the output of the title encoder and maps the output to areal-valued prediction through a feed-forward neural network withresidual connections [31].In addition to the prediction, the proposed model is alsooptimized for the interpretation purpose by three key designs.First, we explicitly extract representation vectors (i.e., features)from the input news titles in a hierarchical, interpretable way (word! bigram ! title), which provides the opportunity to efficientlyvisualize a large amount of contributing factors. Second, we makeuse of a combination of techniques to prevent overfitting, e.g. thedropout mechanism. Third, as the hierarchical method lengthensthe backward propagation path, we introduce residual connectionsto ease the burden of training a deep neural network.Note that the proposed deep stock prediction model can beupgraded by introducing state-of-the-art deep neural networkstructure, such as Convolutional Neural Network (CNN) andRecurrent Neural Network (RNN). In Section 6.3, we will describetwo such alternative designs by replacing the bigram-based titlerepresentation method with convolution layers (Figure 1(b)) andRNN with LSTM cells (Figure 1(c)). The prediction performanceof these alternative designs is also studied in Section 6.3.

4 MODEL INTERPRETATION

4.1 Relevant Keyword Extraction

We introduce a method to identify the importance of textual factorsto the stock price change by analyzing the neural network model. The goal is to compute a relevance score with respect to theprediction result of each trading day, denoted as f (.), for eachword, bigram, and news title. Take the word relevance score asan example, f (v) defines how much contribution a word with thevector representation v has made to the stock price prediction. Apositive (negative) score indicates that this word is evidence forthe rise (fall) of the stock price.

$$f = \sum_i f_i^l \tag{5}$$

$$\Delta f_{i \to j}^{l-1} = \frac{m_{l-1}(j \to i)}{\sum_k m_{l-1}(k \to i)} f_i^l \tag{6}$$

$$f_j^{l-1} = \sum_i \Delta f_{i \to j}^{l-1} \tag{7}$$

$$f(s_k) = f(h) \cdot \frac{w_k s_k}{(\mathbf{w} \times \mathbf{s} + b) + \varepsilon \cdot \mathrm{sgn}(\mathbf{w} \times \mathbf{s} + b)} \tag{8}$$

$$f(b) = f(h) \cdot \frac{b}{(\mathbf{w} \times \mathbf{s} + b) + \varepsilon \cdot \mathrm{sgn}(\mathbf{w} \times \mathbf{s} + b)} \tag{9}$$

Finally, the overall relevance score of each word and bigram isobtained by summing up their relevance scores propagated from allthe titles at all vector dimensions. Note that each relevance score iscomputed once every day according to the daily prediction result.

4.2 Factor Analysis

By the LRP algorithm, a list of words having nonzero relevancescore can be obtained on each day, which is defined as key-words here. These keywords, together with relevant bigrams andnews/tweet titles, compose the potential influencing factors to thechange of stock price. A straightforward method to visualize thesefactors is to juxtapose the top factors and their relevance scoretime series in a list view. This view can be aligned with the timeseries of the actual/predicted stock price changes for multi-factoranalysis. Due to the large number of relevant factors (e.g.,

1801keywords for Apple/APPL), there is an obvious constraint that thelist view can quickly grow beyond the limit of the screen space.These keywords can be shortlisted after sorting by the overallrelevance (i.e., `1 norm of the relevance vector), but an overviewof all relevant factors will be missing.To provide such an overview and allow users to drill-down toeach interested factor, we propose to construct a factor hierarchybased on the list of relevant keywords extracted. As shown in Figure2, the factor hierarchy is composed of four levels: the top level arekeyword clusters that include all extracted relevant keywords; thesecond level are keywords themselves; the third level are bigramphrases stemmed from relevant keywords; and the bottom levelare individual documents (news, tweets, etc.) containing these
keywords/phrases. The lower two levels of the factor hierarchy canhave overlaps in the same level. For example, one bigram phrase can yield two relevant keywords, and one document can have multiplebigram phrases.
This factor hierarchy offers an initial overview ofall factors relevant to stock price changes. The navigation on factorhierarchy through expand/collapse operations allows analyzing thedetails of every factor.

5 VISUALIZATIONS

Design Principle

The DeepClue interface is composed of four coordinated views,as shown in Figure 4(a)(b)(c)(d). We follow two principles in thevisualization design.First, the visualization interface should help users completethree key tasks in the scenario of stock price prediction and analysis.Understanding stock market: The baseline task is to examine theunderlying stock data, including price movements, trading volume,historical rise&fall trends, and the potential temporal patterns.Visualizing prediction result: Over the stock data, users shouldget access to the result produced by the model, i.e., whether acertain stock is predicted to rise or fall on the next day. S/he alsoneeds to navigate the input data to the prediction model, i.e., thenews/tweets collection in our scenario. Interpreting prediction model: Finally, users are expected tounveil the myth of the model by learning why and how eachrise&fall prediction is decided. In DeepClue, this is achieved byvisualizing the key textual factors that jointly make up the decision.Second, we design DeepClue for financial domain users, i.e.,stock traders and investors. These users are mostly accustomed toclassical financial visualization interfaces (e.g., Yahoo Finance),especially for the first two tasks in presenting stock data and
their predictions. The classical visualization depends heavily on

Fig. 4: DeepClue interface: (a) stock timeline view showing the stock price history in an overview+detail design; (b) factor hierarchy view displaying the relevant textual factors to the stock prediction in a hierarchical structure; (c) document list view showing the related documents with the selected factor; (d) keyword map view depicting the relationship of relevant keywords.

statistical charts. Therefore, to reduce the user's learning cost, webuild DeepClue from commodity statistical charts, both in the stock data and prediction visualization (Figure 4(a)(c)) and in illustratingtheir relevant predictive factors (Figure 4(b)(d)).

Visualization Components

In details, the stock timeline view in Figure 4(a) displays the pricemovement of a selected stock over time. This view is organizedin an overview detail design to support flexible navigation of thetimeline. The overview timeline chart on the top row allows usersto specify a focused window on the timeline. The selected timelineis enlarged in the detail chart on the bottom row of Figure 4(a).On the detail timeline chart, four time series can be shown in twotabbed groups, as indicated by the draggable legends in the topleft corner of the detail timeline chart. The first tab on the leftincludes the actual stock price (solid line), the predicted stockprice (dashed line), and the risk of the prediction (shading area).Note that there is little difference between these charts in Figure4(a) because the selected timeline is almost ten years and the dailyprediction introduces little variation. More separation of these timeseries can be observed in Figure 7 and Figure 10(a). In anotherlegend tab on the right, the investment yield curve according tothe prediction model is displayed, which is highly suggested in theexpert study as the key indicator of model's success. As shown inFigure 9, the yield curve always starts from one. The model earnswhen the curve is above one and loses when it is below one.

User Interactions for Model Analysis

The basic user interactions in DeepClue are the customization offour coordinated views according to the analysis task. On the stocktimeline view, the "+" on the left panel allows users to juxtaposetwo bar chart time series for comparison purpose (Figure 4(a)).These bar charts can be dragged vertically to avoid overlaps withthe timeline chart. In the factor hierarchy view, the hierarchicalstructure can be expanded and collapsed. Each factor can be deleted

6 EVALUATIONS

Learning from Financial News

In the first case, we study the use of DeepClue in interpreting thestock price prediction model over financial news. The details of thedata set and model can be found in Section 3. We invited a stocktrader from a private fund, one of our target users, to work withthe DeepClue system. He was interested in investing on Apple Inc.(NASDAQ: AAPL), so the DeepClue configuration is set to displaythe prediction of Apple's stock price from 2006 to 2015.

7. CONCLUSION

We present DeepClue, a system that visually interprets text baseddeep learning models in predicting stock price movements.DeepClue integrates three key designs from the cutting-edgedeep learning technology: a hierarchical neural network modelthat embeds semantics in intermediate processing layers forinterpretation; a backpropagation-like algorithm that effectivelydistributes the decision of prediction back to individual documents,bigrams and words; and an interactive visualization interface thatallows users to navigate and analyze stock price timelines, textualfactors, and their correlations. DeepClue has been deployed topredict S&P 500 stocks using mainstream financial news and firmspecifictweets. Both case studies, quantitative experiments, and theinformal user study with domain experts demonstrate the usefulnessof the proposed system in learning from, evaluating and improvingthe text-based deep stock prediction models.

8.REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521,no. 7553, pp. 436–444, 2015.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classificationwith deep convolutional neural networks," in NIPS'12, 2012, pp. 1097–1105.

[3] A. Bordes, S. Chopra, and J. Weston, "Question answering with subgraphembeddings," in EMNLP'14, 2014, pp. 615–620.

[4] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation byjointly learning to align and translate," in ICLR'15, 2015.

[5] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Using structured events to predictstock price movement: An empirical investigation," in EMNLP'14, 2014,

pp. 1415–1425.

[6] ——, "Deep learning for event-driven stock prediction," in IJCAI'15,2015, pp. 2327–2333.

[7] Y. Peng and H. Jiang, "Leverage financial news to predict stock pricemovements using word embeddings and deep neural networks," inNAACL'16, 2016, pp. 374–379.

[8] D. Erhan, Y. Bengio, A. Courville, and P. Vincent, "Visualizing higherlayerfeatures of a deep network," University of Montreal, vol. 1341,2009.

[9] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutionalnetworks," in ECCV'14, 2014, pp. 818–833.

[10] B. Shneiderman, "The eyes have it: A task by data type taxonomy forinformation

visualizations," in IEEE Symposium on Visual Languages,1996, pp. 336–343.

[11] B. G. Malkiel and E. F. Fama, "Efficient capital markets: A review oftheory and empirical work," The journal of Finance, vol. 25, no. 2, pp.383–417, 1970.

[12] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson, "Understandingneural networks through deep visualization," arXiv:1506.06579, 2015.

[13] "Tensorflow playground," <http://playground.tensorflow.org/>.

[14] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu, "Towards better analysis ofdeep convolutional neural networks," IEEE Transactions on Visualization

and Computer Graphics, vol. 23, no. 1, pp. 91–100, 2017.[12] Rafeeque Pandarachalil Govt. College of Engineering, Kannur "Twitter Sentiment Analysis for Large-Scale Data: An Unsupervised Approach", Springer, 2014

[13] Seven Rill Goethe University Frankfurt, Germany "PoliTwi- Early detection of emerging political topics on Twitter and the impact on concept-level sentiment analysis", Elsevier, 2014

[14] Chirag Sangani Stanford University, USA "Sentiment Analysis of App Store Reviews", 2013

[15] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutionalnetworks: Visualising image classification models and saliency maps,"arXiv:1312.6034, 2013.

[16] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. M¨uller, andW. Samek, "On pixel-wise explanations for non-linear classifier decisionsby layer-wise relevance propagation,"

[17] A. Binder, G. Montavon, S. Lapuschkin, K.-R. M¨uller, and W. Samek,"Layer-wise relevance propagation for neural networks with local renormalizationlayers," ICANN'16, vol.

[18] A. Dosovitskiy and T. Brox, "Inverting visual representations withconvolutional networks," arXiv:1506.02753, 2015.

[19] L. M. Zintgraf, T. S. Cohen, and M. Welling, "A new method to visualizedeep neural networks," arXiv:1603.02518, 2016.