

Embracing Textual Data Analytics in Auditing with Deep Learning

Ting Sun. Rutgers Business School. NJ, USA tings@scarletmail.rutgers.edu

Miklos A. Vasarhelyi. Rutgers Business School. NJ, USA miklosv@business.rutgers.edu

Abstract. While the massive volume of text documents from multiple sources inside and outside of the company provides more information for auditors, the lack of efficient and effective technology solutions hampers the full use of text data. Powered by the emerging data analytics technology of deep learning, the value of the text can be better explored to deliver a higher quality of audit evidence and more relevant business insights. This research analyzes the usefulness of the information provided by various textual data in auditing and introduces deep learning, an evolving Artificial Intelligence approach. Furthermore, it provides a guide for auditors to implement deep learning techniques with pre-developed tools and open-source libraries.

Keywords: Text Analysis, Deep Learning, Artificial Intelligence, Big Data, Auditing

1. INTRODUCTION

Deep learning is an artificial intelligence technology employing deep hierarchical neural network to analyze data to extract complex and abstract features underlying the raw data. It is a powerful big data analytics tool (Najafabadi, et al., 2016) and has been successfully implemented in a variety of industries such as automotive,

health, financial and advertising. Textual data is the most common type of big data in auditing. As compared to other types of big data like video or audio, text can be obtained from readily available sources within or outside of the company. Nevertheless, the application of deep learning in text analysis in auditing is just evolving. Pioneers in this area like the Big Four accounting firms have started employing deep learning to harness the power of textual big data to surface deeper insights and highlight risky areas. For example, Deloitte combines its business insights in cognitive technologies with Kira Systems¹, advances in deep learning to conduct document analysis tasks (e.g., investigations, mergers, contract management, and leasing arrangements). Furthermore, Deloitte-customized instances of Kira platform have been applied to its audit business (under the name Argus) and consulting services (under the name D-ICE) (Deloitte, 2016). In May 2017, Deloitte, together with Mainspring Consulting Group as well as legal services providers Axiom, Cogna Law, and Elevate, joined a Certified Service Provider Program launched by Kira. This is a certification program designed to provide extensive training to service providers to enable them to effectively use Kira's AI technology and enhance their work in areas like contract data extraction, regulatory compliance and transactional review (Artificial Lawyer, 2017). However, the application of deep learning to examine textual data in auditing remains largely unexplored. In the academic area, while it has been well recognized that big data can be used as additional audit evidence (Brown-Liburd and Vasarhelyi, 2015; Yoon, et al., 2015; Alles and Gray, 2016), few studies provide solutions for the implementation of big data analytical tools in the context of auditing.

This research introduces a state-of-the-art big data analytics technique, deep learning. It aims to guide accounting professionals and researchers in implementing deep learning by employing publicly available deep learning software; this research study also provides future directions for audit-specific text analysis. Second, this study explores the value of text data as well as the underlying information that can be extracted and used as additional audit evidence. More importantly, it proposes basic approaches to adopt deep learning in the context of auditing and analyzes their advantages and disadvantages. Finally, this paper contributes to the literature by providing the direction for future audit data analysis.

¹ Kira is an Canada-based legal AI company. <https://kirasystems.com/>

The remaining parts of this paper are organized as follows. Section 2 highlights the importance of big data analytics in auditing and provides examples of the textual big data and the possible metadata that is useful for audit decision-making. Section 3 reviews traditional textual analysis techniques, including bag-of-words and machine learning. Section 4 introduces deep learning and points out its superiority as opposed to traditional ones. Section 5 discusses two basic approaches to implement deep learning in audit-related text analysis. Specifically, it summarizes some popular pre-developed deep learning applications like IBM Watson's Natural Language Understanding and open-source deep learning frameworks such as Tensor Flow and Theano. The advantages and disadvantages of these two approaches are analyzed as well. This section also discusses the future directions for the implementation of deep learning to improve the performance of textual big data analytics in auditing. Finally, section 6 concludes the paper and lists several ideas for future research.

2. OBTAINING AUDIT INSIGHTS FROM THE TEXTUAL BIG DATA

Big data permeates almost all aspects of business decision making (Griffin and Wright 2015). Comments from social media, news articles, emails, recordings of phone calls, financial and non-financial data extracted from web server logs, and various machine-generated and sensor-detected data, may be analyzed to assist key business decision-making (Chui, et al., 2010; Bertolucci 2013; O'Leary 2013). Accounting professionals wish to use intelligent analytics of the potentially vast source of evidence to deliver a higher quality of audit and more relevant business insights (EY, 2015; Rose, et al., 2017). Deloitte's Audit of the Future Survey (2016) shows that, among the 250 financial statement preparers, audit committee members, and financial statement users who participated in the survey, more than 70% agree that auditors should more extensively use advanced data analytic technologies and consider information beyond traditional financial statements. Auditors can acquire and collect textual data easily from a variety of sources. Therefore, text analysis is a good starting point for big data analysis in auditing. Table 1 lists some examples of textual data that can be used by the auditor as well as the metadata that can be extracted from the text and used as additional audit evidence.

Examples of textual data in auditing	contracts, confirmations, transcripts of conference call, earnings announcements, emails, trading publications, posts on social media, customer reviews, news articles, press release, analyst reviews, emails, MD&A sections of 10-Ks, audit working papers, transcripts of inquiries and interviews, background documents about the senior managers, code of conduct, internal control reports, management letter, invoices, shipping documents, minutes of board of directors meetings
Possible metadata that can be extracted from the text	topics, keywords, entities, sentiments, emotions, date, time, events, anomalies
Applicable audit phases	planning (including pre-engagement planning), internal control evaluation, substantive testing, completion

Table 1. Textual data and the metadata in auditing

For instance, content analysis of social media postings and news articles could inform auditors of potential litigation risk, business risk, internal control risk, or risk of management fraud. For example, auditors may identify troublesome products or services by analyzing customers' reviews on the product or service of the company. Measures of sentiment and emotion of MD&A sections of 10-Ks can be merged with accounting data to develop models to predict financial frauds (e.g., Sun, et al., 2018). Similar sentiment scores of the Q&A section of earnings conference calls can help the auditor predict internal control material weakness (Sun, 2017). As these examples highlight, auditors can engage in analysis of textual data to monitor the risk of the company in all audit phases (Richins, et al., 2017).

The accounting profession has recognized the impact of textual big data analysis on enhancing the quality and relevance of the audit (EY, 2015). The Big Four public accounting firms have launched a variety of projects to adapt to the advent of Big Data. Deloitte has cooperated with Kira to use the text-mining software for machine learning to quickly "read" thousands of contracts and other complex documents, extracting and structuring textual information to support their audit and consulting practices (Deloitte, 2016). Specifically, Deloitte's application can be taught to look for particular terms or entities such as the contract parties or the governing law (Legal Insider, 2016).

3. TRADITIONAL TEXT ANALYSIS TECHNIQUES

The main goal of text analysis is to extract insightful information, such as sentiment and keywords, from textual resources. Text analysis is challenging as it

is difficult to parse the semi-structured text. Data analysts and researchers typically use the following two approaches to analyze text.

3.1. Bag-of-words

As a prevalent text mining approach, “bag-of-words” represents a text as the bag of its words. The features of the text are obtained by counting the frequency of particular words based on the predefined dictionary (“bag-of-words”). For instance, General Inquirer (GI) is one of the most frequently used generic dictionaries. GI was developed at Harvard in 1961 for content analysis in behavioral research and then merged with the Hunt Concept Learner for automatic theme analysis in 1962 (Stone and Hunt, 1963). GI system maps each text file with counts on dictionary-supplied categories, and each category is a list of words. Currently, it has 182 categories which combine the "Harvard IV-4" dictionary content-analysis categories, the "Lasswell" dictionary content-analysis categories, as well as 5 categories based on prior social cognition work. The category "negative" has 2291 entries, which is the largest category². Based on the negative and weak words listed in GI, Tetlock (2007) measures the investor’s pessimism sentiment in the Wall Street Journal. Likewise, the negative and positive word lists of GI are employed by Kothari, Li, and Short (2009) to analyze the effect of disclosures by management, analysts, and business press on the cost of capital, return volatility, and analyst forecasts. Since the dictionaries in GI are typically for general language, researchers argue that it is not appropriate in business domain and call for finance-specific word lists (Henry and Leone, 2009; Loughran and McDonald, 2011). Loughran and McDonald (2011) develop a financial dictionary (also known as L&M List) that is composed of all words occurred in at least 5% of the 10-Ks during 1994 -2008. The L&M list has been widely used by subsequent studies (e.g., Garcia, 2013; Jegadeesh and Wu, 2012; Loughran and McDonald, 2013) because it is a finance-specific dictionary that is designed for 10-K text analysis. A limitation of bag-of-words approach is that it cannot understand the context or semantics of words. This approach encodes every word in the vocabulary as one-hot-encoded vector for the vocabulary of size $|V|$. In other words, each word is represented by a $|V|$ dimensional sparse vector, with 1 at index corresponding to the word and 0 at every other index. As a result, the vectors corresponding to words with similar meaning could be orthogonal. For

² For more information please visit: <http://www.wjh.harvard.edu/~inquirer/3JMoreInfo.html>

example, “buy used cars” and “purchase old automobiles” are represented by orthogonal vectors in Bag-of-words model (Bhardwaj, 2017). Another drawback of this technique is that it ignores the ordering of the words. For example, it is not able to tell the difference between “This is good” and “Is this good” (Salton and McGill, 1983).

3.2. Machine Learning

In the context of text analysis, we can leverage unsupervised and supervised machine learning algorithms to differentiate between documents. Supervised machine learning requires a training dataset with labels, which can be seen as the prior knowledge of the data, for features such as parts of speech and entities. This training dataset is used to build sophisticated models; the trained model is then applied to new dataset to evaluate its predictive performance. Through the process of model validation, the parameters of the model are configured until the best predictive accuracy is achieved (Kotsiantis, et al., 2007). In this way, supervised machine learning groups documents into predetermined categories based on the labels (the prior knowledge) in the training set. Examples of supervised machine learning algorithms used frequently include Support Vector Machines, Bayesian Networks, Maximum Entropy, Conditional Random Field. Unsupervised machine learning can extract data features from a collection of text without labels. In other words, this technique attempts to differentiate between documents without any prior knowledge (Kloo, 2015). For example, Clustering divides a set of objects into clusters so that objects in the same cluster are similar to each other, and objects in different clusters are dissimilar. Other examples of prevalent unsupervised machine learning algorithms are latent Semantic Indexing and Matrix Factorization.

4. DEEP LEARNING

Deep Learning (also called deep neural network) is an emerging machine learning/Artificial Intelligence technique, which enables machines to analyze complex and abstract data features through hierarchical deep neural networks. While the basic idea of artificial neural networks dates back to as early as the 1950s, deep neural network as a real artificial intelligence technique gained momentum in the recent decade due to the advance in computational power and data storage (Issa, et al., 2016; Sun and Vasarhelyi, 2017). Figure 1 shows a basic

deep neural network: Multilayer Perceptrons (MLP)³. It consists of one input layer, one output layer, and multiple hidden layers between the input and output layer. Unlike traditional neural networks which usually do not have more than two hidden layers, deep neural networks have three or more interconnected hidden layers, which facilitate the processing of complex and unstructured data, like text, video, and audio. Each layer in a deep neural network consists of artificial neurons which are modified with weights and combined. The combination of the neurons in a predecessor layer is the input of its successive layer. The input layer receives raw input data, transforms, and passes the modified version of data to the next layer. As the neural network goes further, the data is transmitted from one layer to another, and the network can extract more abstract data features (Nielsen, 2015). In other words, in a deep neural network, the abstract data representation in a layer is built upon other, simpler data representation in the last layer (Goodfellow, et al., 2016; Sun and Vasarhelyi, 2017).

Besides weights, deep neural networks use nonlinear activation function such as rectified linear unit (ReLU) or Hyperbolic Tangent (Tanh) (Sharma, 2017) to transform the input data and compute the losses, which is the difference between the true output and computed output, measured by Mean Square Error(MSE), Logloss, and etc. The loss will be backward propagated to adjust the weights and other hyper parameters of the deep neural network to improve the prediction accuracy. Learning from a large amount of training data, the process of layer-wise feature extraction of deep neural networks is automatic without any human intervention⁴, which makes it work more effectively (with better accuracy performance) and efficiently (with less time) than traditional machine learning techniques for text and other big data analysis.

³ MLP is generally the starting phase of building other more sophisticated deep network and can be used for any supervised regression or classification problems. In fact, the deep learning model typically used for sentiment analysis of text data is Recurrent Neural Network (Goodfellow, Bengio, and Courville, 2016).

⁴ This is called unsupervised feature extraction.

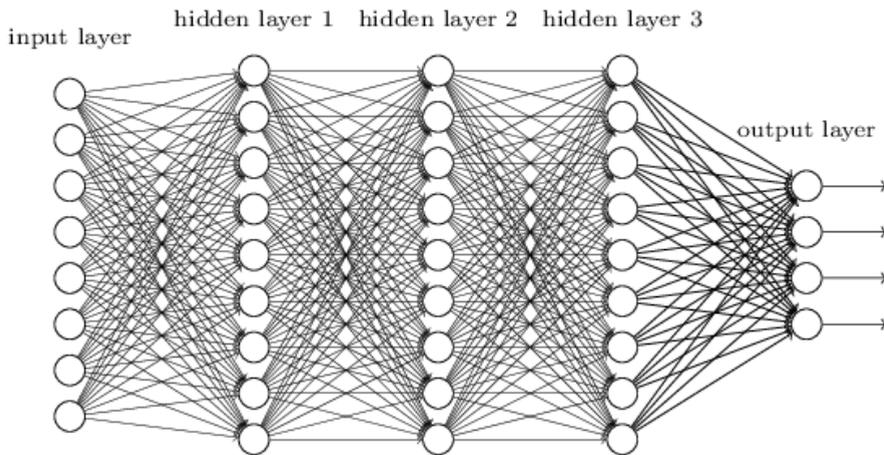


Figure 1. The structure of a basic deep neural network (Adapter from Nielsen, 2015)

5. THE IMPLEMENTATION OF DEEP LEARNING FOR TEXT ANALYSIS IN AUDITING

The biggest challenge of analyzing text is to parse the semi-structured text to identify trends and extract insights. It is especially difficult for text from sources like social media posts, articles, and blogs, which constitutes a large mass of unstructured data with a lot of irrelevant content, i.e., advertisements and links (Daffron, 2017).

5.1. Pre-developed deep learning tools

The first approach of conducting text analysis in auditing with deep learning is to employ pre-developed deep learning algorithms. Examples are discussed next.

Natural Language Understanding

Natural Language Understanding, a cloud-based service provided by IBM Watson, offers a suite of natural language processing (NLP) capabilities that allow the user to extract key metadata from the text, including entities, relations, concepts, sentiment, and emotion. Watson Natural Language Understanding builds on the legacy of its predecessor Alchemy Language, whose deep learning algorithm is trained with text from billions of webpages, covering a wide range of topics. With a few simple API calls, auditors can use Natural Language Understanding to automate NLP and machine learning capabilities to analyze text data that is relevant to the audit. Specifically, it analyzes the sentiment and the

emotion for a document and returns corresponding scores such as MD&A's and conference calls. In addition, it measures targeted sentiment and emotion towards keywords in the text. For instance, while a CFO's response to an analyst's question during a conference call may have an overall positive sentiment, particular keywords in the response may have a negative tone.

Sun (2017) applies *Alchemy Language* to generate the overall sentiment score and the score of joy for transcripts of quarterly conference calls from 2004 to 2014 and utilizes those scores as additional indicators to predict internal control material weakness. The result shows that, with the incorporation of the sentiment features, especially the score of joy, the explanatory ability of the model improves significantly compared to the baseline model that merely utilizes the major internal control weakness determinants suggested by prior literature (i.e., Doyle, et al., 2007; Ashbaugh-Skaife, et al., 2007).

Watson Analytics for Social Media

Based on the topics, keywords, time periods, and other specifications that the user inputs, IBM Watson Analytics for Social Media is able to automatically collect relevant social media content from millions of interactions on Twitter, Facebook, YouTube, Amazon, Google Plus, WordPress, and so forth, display data patterns, and create data visualizations on sentiment, geography, authors and demographics for insightful discoveries⁵. For example, it can provide the result of analysis regarding the sentiment and the popularity of the discussion about a company, such as the number of tweets with positive sentiment and the maximum number of retweets mentioned the company during a certain period. Sun and Vasarhelyi (2018) employ similar content data on Twitter to predict audit fees.

Bottlenose's Nerve Center® Platform

A similar tool of Watson Analytics is Bottlenose's Nerve Center® Platform of cognitive computing, which employs NLP, trend discovery, supervised and unsupervised machine learning, and live interactive visualization to detect anomalies and trends, and anticipate change underlining the data spanning across more than two million different sources including social, news, blogs, TV, radio, and print media⁶.

⁵ For more information please visit: <https://www.ibm.com/watson-analytics/social-media> and <https://www.ibm.com/us-en/marketplace/social-media-data-analysis>

⁶ For more information please visit: <http://bottlenose.com/product>

Advantages

The pre-developed tools require a minimum level of programming skill as they usually have point-and-click user interfaces. While conventional text analysis approaches require laborious data preprocessing (i.e., removing HTML tags, transferring HTML characters to text characters, converting words from upper case to lower case, and deleting punctuation and stop words), the pre-developed tools can directly use the input without data pre-processing (Sun, et al., 2018). For example, the user of Natural Language Understanding only needs to input the text/html raw file or even the URL. Natural Language Understanding automatically removes irrelevant content such as links and advertisements and returns results. Similarly, Watson Analytics for Social Media is fast and easy to use. The user only needs to provide topic or keywords and period. The tool immediately displays the visualizations of the result. It even provides cognitive themes and topics that help the user tailor searches. A recent study (Sun, et al., 2018) demonstrates the superiority of Alchemy Language regarding data processing efficiency and prediction effectiveness as opposed to the bag-of-words approach for the task of predicting financial frauds with the sentiment and emotion of Management Discussion and Analysis of 10-K filings.

Disadvantages

The pre-developed tools are not free, and the deep learning models are not exclusively trained with financial text. The words or phrases used in the finance-related document may have a different meaning when it is used in general context. Certain positive phrases in general language are negative in accounting domain, such as “qualified opinion”. “Expand” and “expansion” can have any polarity in general language but are always positive in financial topics. On the other hand, some generally negative words such as “liability” are neutral in the financial context (Henry and Leone, 2009; Loughran and McDonald, 2011). As a result, directly applying algorithms trained with general text to financial text may lead to biased prediction.

5.2. Open-source deep learning programming libraries

The second way is to use open-source deep learning libraries to train finance-specific models. One of the most prevalent deep learning software is Tensor Flow.

Tensor Flow

Tensor Flow⁷ is the deep learning framework developed by Google's machine intelligence research team – Brain Team. With its pre-built functions, it can easily train different neural networks (MLP, CNN, RNN, etc.) with our own data. Tensor Flow provides both Python API and C++ API. With over 6000 open-source online repositories online, Tensor Flow, in its first year, has helped users in a variety of industries make progress in a wide range of tasks from language translation to early detection of skin cancer and blindness in people with diabetes.

Theano

Theano⁸ is a Python library for Deep Learning developed in 2007 by the LISA (now MILA) group at the University of Montreal, Quebec, Canada. It is a compiler for mathematical expressions in Python, which can be considered an industry standard for Deep Learning research and development. Similar to Tensor Flow, Theano allows the user to directly create deep learning models or wrapper libraries that greatly simplify the process (Brownlee, 2016).

Other deep learning libraries include Apache MX Net and Gluon, Microsoft Cognitive Toolkit, Caffe, Caffe2, Torch, Pytorch, and Keras.

Advantages

Unlike the pre-developed deep learning tools which only provide certain built-in functions (i.e., returning the overall sentiment score, scores of certain types of emotions like anger, joy, sadness, disgust), the open-source deep learning software is more flexible. Auditors can train sophisticated, custom AI models, or experiment with any new algorithms.

Disadvantages

Meanwhile, it requires a higher level of programming skill, such as python and C++. Also, the user needs to understand the basics of math (especially calculus, probability, and linear algebra), Statistics, Machine Learning (Shaikh, 2016). The user should also have good knowledge of neural networks and deep learning. Therefore, the mainstream use of deep learning could be hampered due to the learning curve.

⁷ For more information please visit: <https://www.tensorflow.org/>

⁸ For more information please visit: <http://deeplearning.net/software/theano/#>

6. THE MIDDLE GROUND APPROACH

Another tool called H2O flow is provided by H2O.ai⁹, a leading AI company in Silicon Valley. H2O flow is a web-based interactive computational environment that combines the concept of a command-line, text-based shell with that of a modern graphical purpose-built point-and-click user interfaces¹⁰. As a result, it has the advantage of the pre-developed tool, user-friendly, but allows users to build deep learning models based on their data. However, Flow only has one type of deep neural network, the MLP. Table 2 summarizes the advantages and disadvantages of the deep learning approaches.

Categories	Advantages	Disadvantages	Examples
Pre-developed deep learning algorithms	User-friendly Does not require programming skills	Not free Lacks flexibility The training data is not finance-specific	Watson Natural Language Understanding, Watson Analytics for Social Media, Bottlenose's Nerve Center®
Open-source deep learning programming libraries	Free, Flexible, Allows users to train their own model, Can be trained with finance-specific data	Requires high level of programming skill and other data analytics knowledge	TensorFlow, Theano, Apache MXNet and Gluon, Microsoft Cognitive Toolkit, Caffe, Caffe2, Torch, Pytorch, and Keras
The middle ground tool	Free, User-friendly, Does not require programming skills, Allows users to train their own model, Can be trained with finance-specific data	Only has one type of deep neural network, MLP	H2O Flow

Table 2. A summary of deep learning approaches for audit textual data analytics

7. NEXT STEP: DEVELOPING FINANCE-SPECIFIC TEXT DATASET

While pre-developed deep learning models and open-source deep learning libraries facilitate the development of deep learning applications in auditing, they

⁹ For more information please visit: <https://www.h2o.ai/>

¹⁰ For more information please visit: <https://blog.h2o.ai/2014/11/introducing-flow/>

still lack sufficient audit-specific text datasets with labels of syntactic annotation due to issues with data capture and concerns about data privacy and security. The Penn tree bank dataset is based on 2,499 stories from a three-year Wall Street Journal collection of 98,732 stories as well as IBM computer manuals, nursing notes, and transcribed telephone conversations for syntactic annotation¹¹. It produces a large number of words of part-of-speech tagged text, skeletally parsed text, text parsed for predicate-argument structure, and transcribed spoken text annotated for speech disfluencies (Taylor, et al., 2003). Although Wall Street Journal is a business-focused magazine, its coverage spans art, fashion, entertainment, design, food, architecture, travel and more, so Penn tree bank is not a strictly finance-specific dataset. Future work can be conducted to prepare more finance-specific and audit-specific text data (for instance, using documents like audit working paper, 10-K, 10-Q, transcripts of conference calls, MD&As, and press release).

8. CONCLUSION

Big data analytics is expected to make substantive contributions to auditing (Earley, 2015; Tang and Karim, 2017). Incorporating big data as additional audit evidence facilitates anomalies detecting and fraud prediction, which enhances the audit quality (Yoon, et al., 2015). This paper focuses on the analysis of textual data because it is the most common data format in auditing and investigates how to apply an innovative big data analytics technique, deep learning, to explore audit-related text documents. It claims that, deep learning is an effective and efficient approach to automatically extract topics, keywords, entities, sentiments, emotions, events, and other important metadata from semi-structured text data such as contracts, earnings announcements, emails, posts on social media, news articles, press release, analyst reviews, and emails. Deep learning supports audit decision-making in all audit phases, including planning, internal control evaluation, substantive test, and completion. Moreover, due to the depth of the hidden layers in deep neural networks and the unsupervised feature extraction, deep learning is superior to traditional text mining techniques. To implement this technology, the auditor can directly use pre-developed deep learning models such as Natural Language Understanding from IBM Watson to obtain data features of the text. Alternatively, the auditor can employ deep learning libraries such as

¹¹ For more information please visit: <https://catalog ldc.upenn.edu/ldc99t42>

Tensor Flow to train audit-specific deep learning model to solve specified audit problems. To facilitate the application of deep learning to auditing, the following ideas can be considered for future research.

(1) Developing finance-specific textual datasets for the development of audit-specific deep neural networks. The most difficult part of building a supervised machine learning model with deep neural network is to source or create large volumes of labelled training data. For instance, for the sentiment analysis of MD&As, we need sufficient amount of MD&As with predefined classes of positive, neutral, negative, and etc. Unfortunately, the most common mechanism for AI is human-based labeling. The task of labeling is laborious, time-consuming, and expensive because experts are rare. Currently, several approaches have been proposed to solve this problem, for example, Generative Adversarial Networks (Kingma, et al., 2014) and Pseudo-Labeling (Lee, 2013). More efforts in this research line are expected.

(2) Providing an illustrative framework to guide the application of deep learning to a variety of audit tasks and procedures in audit phases of planning, field work, reporting, and follow-up.

(3) Exploring more data sources such as news or magazine articles and analyst reports to extract insights to enhance audit quality.

(4) Opening the “black box” of deep neural networks to enable auditors to better understand how a deep learning model learns. Deep learning is called a black box as it is non-parameterized. Although the choice of hyper-parameters of deep learning models, such as the number of layers, the activation function, and the learning rate, as well as the predictor importance is known. It is still unclear how machines learn and deduce conclusions. Scholars are seeking answers to this question. Last year, a YouTube video¹² of a conference presentation made by Naftali Tishby offered a possible answer. In the presentation, Tishby proposed the “information bottleneck” theory to explain how deep learning works. The idea is that a network processes “noisy input data of extraneous details as if by squeezing the information through a bottleneck, retaining only the features most relevant to general concepts” (Wolchover, 2017).

(5) Demonstrating the applications of deep learning to other audit types such as compliance audit, information systems audit, operational audit, and tax audit.

¹² <https://www.youtube.com/watch?v=bLqJHjXihK8&feature=youtu.be>

9. REFERENCES

- ALLES, M.; GRAY, G. L. (2016): “Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors”, *International Journal of Accounting Information Systems*, vol. 22: 44–59. <https://doi.org/10.1016/j.accinf.2016.07.004>
- ASHBAUGH-SKAIFE, H.; COLLINS, D. W.; KINNEY, W. R. (2007): “The discovery and reporting of internal control deficiencies prior to SOX-mandated audits”, *Journal of Accounting and Economics*, vol.44, n. 1-2: 166–192. <https://doi.org/10.1016/j.jacceco.2006.10.001>
- ARTIFICIAL LAWYER. (2017): “Kira Starts Legal AI training Program; Deloitte, Axiom + Cognition Join”. <https://www.artificiallawyer.com/2017/06/01/kira-starts-legal-ai-training-program-deloitte-axiom-cognition-join/>
- BERTOLUCCI, J. (2013): “Big Data: A practical definition. InformationWeek (August 26)”. <http://www.informationweek.com/big-data/big-data-analytics/big-data-a-practical-definition/d/d-id/1111290>
- BHARDWAJ, P. (2017): “What are the limitations of the Bag-of-Words model”, Quora. <https://www.quora.com/What-are-the-limitations-of-the-Bag-of-Words-model>
- BROWN-LIBURD, H.; VASARHELYI, M. A. (2015): “Big Data and Audit Evidence”, *Journal of Emerging Technologies in Accounting*, vol. 12, n. 1: 1–16. <https://doi.org/10.2308/jeta-10468>
- BROWNLEE, J. (2016): “Introduction to the Python Deep Learning Library Theano”. <https://machinelearningmastery.com/introduction-python-deep-learning-library-theano/>
- CHUI, M.; LOFFLER, M.; ROBERTS, R. (2010): “The Internet of things”, *McKinsey Quarterly*2, 1–9. <https://www.mckinsey.com/industries/high-tech/our-insights/the-internet-of-things>
- DAFFRON, S.C. (2017): “Watson Natural Language Understanding for advanced text analytics”. <https://www.ibm.com/blogs/watson/2017/04/watson-natural-language-understanding-advanced-text-analytics/>

- DELOITTE. (2016): “Audit of the future: Deloitte’s latest survey results”. <https://www2.deloitte.com/us/en/pages/audit/articles/the-future-of-audit-survey-and-innovation-report.html>
- DELOITTE. (2016): “Press releases: Deloitte forms alliance with Kira Systems to drive the adoption of artificial intelligence in the workplace”. <https://www2.deloitte.com/us/en/pages/about-deloitte/articles/press-releases/deloitte-forms-alliance-with-kira-systems-to-drive-the-adoption-of-artificial-intelligence-in-the-workplace.html>
- DOYLE, J.; GE, W.; MCVAY, S. (2007): “Determinants of weaknesses in internal control over financial reporting”, *Journal of accounting and Economics*, vol. 44, n. 1: 193-223. <https://doi.org/10.1016/j.jacceco.2006.10.003>
- EARLEY, C.E. (2015): “Data analytics in auditing: Opportunities and challenges. *Business Horizons*”, vol. 58, n. 5: 493-500. <https://doi.org/10.1016/j.bushor.2015.05.002>
- EY. (2015): “How big data and analytics are transforming the audit”, *EY Reporting*, April (9). <http://www.ey.com/gl/en/services/assurance/ey-reporting-issue-9-how-big-data-and-analytics-are-transforming-the-audit#item1>
- GARCIA, D. (2013): “Sentiment during recessions”, *The Journal of Finance*, vol. 68, n. 3: 1267-1300. <https://doi.org/10.1111/jofi.12027>
- GOODFELLOW I.; BENGIO, Y.; COURVILLE, A. (2016): “Deep Learning”, MIT Press. <http://www.deeplearningbook.org>
- GRIFFIN, P.A.; WRIGHT, A.M. (2015): “Commentaries on Big Data’s importance for accounting and auditing”, *Accounting Horizons*, vol. 29, n. 2: 377-379. <https://doi.org/10.2308/acch-51066>
- HENRY, E.; LEONE, A.J. (2009): “Measuring qualitative information in capital markets research”. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1470807
- ISSA, H.; SUN, T.; VASARHELYI, M. A.(2016): “Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation”, *Journal of Emerging Technologies in Accounting*, vol. 13, n. 2. <https://doi.org/10.2308/jeta-10511>
- JEGADEESH, N.; WU, A.D. (2012): “Word power: A new approach for content analysis”, *Journal of Financial Economics*, vol.110,n.3: 712-729

KINGMA, D.P.; MOHAMED, S.; REZENDE, D.J.; WELLING, M. (2014): "Semi-supervised learning with deep generative models. In Advances in Neural Information Processing Systems", Neural Information Processing Systems 2014 Proceedings: 3581-3589. <https://papers.nips.cc/paper/5352-semi-supervised-learning-with-deep-generative-models>

KLOO, I. (2015): "Text Mining: Clustering, Topic Modeling, and Classification".http://data-analytics.net/cep/Schedule_files/Textmining%20%20Clustering,%20Topic%20Modeling,%20and%20Classification.htm

KOTHARI, S.P.; LI, X.; SHORT, J.E. (2009): "The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis", The Accounting Review, vol. 84, n. 5: 1639-1670. <https://doi.org/10.2308/accr.2009.84.5.1639>

KOTSIANTIS, S. B.; ZAHARAKIS, I.; PINTELAS, P. (2007): "Supervised machine learning: A review of classification techniques", Emerging Artificial Intelligence Applications in Computer Engineering, vol. 160: 3-24. <https://dl.acm.org/citation.cfm?id=1566773>

LEE, D.H. (2013): "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In Workshop on Challenges in Representation Learning", ICML, vol. 3. http://deeplearning.net/wp-content/uploads/2013/03/pseudo_label_final.pdf

LEGAL INSIDER (2016): "Deloitte partners with Kira Systems to bring AI into the workplace". <https://www.legaltechnology.com/latest-news/deloitte-partners-with-kira-systems-to-bring-ai-into-the-workplace/>

LOUGHRAN, T.; MCDONALD, B. (2011): "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks", The Journal of Finance, vol. 66, n. 1: 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>

LOUGHRAN, T.; MCDONALD, B. (2013): "Measuring readability in financial disclosures", SSRN. <http://ssrn.com/abstract=1920411>

NAJAFABADI, M.M., et al. (2016): "Deep Learning Techniques in Big Data Analytics. Big Data Technologies and Applications", HPCC Systems Big Data book: 133-156. https://doi.org/10.1007/978-3-319-44550-2_5

- NIELSEN, M.A. (2015): “Neural networks and Deep learning”, Determination Press. <http://neuralnetworksanddeeplearning.com/>
- O’LEARY, D. E. (2013): “Big Data, the Internet of things and the Internet of signs”, *Intelligent Systems in Accounting, Finance and Management*, vol. 20, n. 1: 53–65. <https://doi.org/10.1002/isaf.1336>
- RICHINS, G.; STAPLETON, A.; STRATOPOULOS, T. C.; WONG, C. (2017): “Big Data Analytics: Opportunity or Threat for the Accounting Profession?”, *Journal of Information Systems*, vol. 31, n. 3: 63-79. <https://doi.org/10.2308/isys-51805>
- ROSE, A.M.; ROSE, J.M.; SANDERSON, K.; THIBODEAU, J.C. (2017): “When Should Audit Firms Introduce Analyses of Big Data into the Audit Process?”, *Journal of Information Systems*, vol. 31, n. 3: 81-99. <https://doi.org/10.2308/isys-51837>
- SALTON, G.; MCGILL, M. J. (1983): “Introduction to Modern Information Retrieval”, McGraw-Hill. <https://dl.acm.org/citation.cfm?id=576628>
- SHAIKH, F. (2016): “A Complete Guide on Getting Started with Deep Learning in Python”. <https://www.analyticsvidhya.com/blog/2016/08/deep-learning-path/>
- SHARMA, S. (2017): “Activation Functions: Neural Networks”. <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>
- STONE, P.J.; HUNT, E.B. (1963): “A computer approach to content analysis: studies using the general inquirer system”, *ACM spring joint computer conference Proceedings*:241-256. <https://pdfs.semanticscholar.org/2d57/a794a8ae2711285065e84fbe1744fd9e3284.pdf>
- SUN, T. (2017): “The Incremental Informativeness of Management Sentiment in Conference Calls for the Prediction of Internal Control Material Weaknesses”, *Journal of Emerging Technologies in Accounting*. <https://doi.org/10.2308/jeta-51969>
- SUN, T.; LIU, Y.; VASARHELYI, M.A. (2018): “The Performance of Sentiment Features of MD&As for Financial Misstatements Prediction: A Comparison of Deep Learning and Bag of Words Approach”, Working paper.

SUN, T.; VASARHELYI, M.A. (2017): “Deep Learning and the Future of Auditing: How an Evolving Technology Could Transform Analysis and Improve Judgment”, CPA Journal, Jun: 24-29.<https://www.cpajournal.com/2017/06/19/deep-learning-future-auditing/>

SUN, T; VASARHELYI, M.A. (2018): “Predicting Audit Fee with Twitter: Do the 140 Characters reveal the company’s material misstatement risk?”, Working paper.

TANG, J.; KARIM, K.E. (2017): “Big Data in Business Analytics: Implications for the Audit Profession”, The CPA Journal. <https://www.cpajournal.com/2017/06/26/big-data-business-analytics-implications-audit-profession/>

TAYLOR, A.; MARCUS, M.; SANTORINI, B. (2003): “The Penn treebank: an overview”, Abeillé A. (eds) Tree banks. Text, Speech and Language Technology, vol. 20: 5-22. Springer, Dordrecht. https://doi.org/10.1007/978-94-010-0201-1_1

TETLOCK, P.C. (2007): “Giving content to investor sentiment: The role of media in the stock market”, The Journal of Finance, vol. 62, n. 3:1139-1168. <https://doi.org/10.1111/j.1540-6261.2007.01232.x>

WOLCHOVER, N. (2017): “New Theory Cracks Open the Black Box of Deep Learning”, Quanta Magazine. <https://www.quantamagazine.org/new-theory-cracks-open-the-black-box-of-deep-learning-20170921/>

YOON, K.; HOOGDUIN, L., ZHANG, L. (2015): “Big Data as Complementary Audit Evidence”, Accounting Horizons, vol. 29, n. 2: 431-438. <https://doi.org/10.2308/acch-51076>