

Deep Learning and the Future of Auditing

How an Evolving Technology Could Transform Analysis and Improve Judgment

By Ting Sun and Miklos A. Vasarhelyi

IN BRIEF

This article introduces deep learning technology—an emerging form of artificial intelligence that can be trained to recognize patterns in vast volumes of data that would be impossible for humans to process. This still evolving technology represents a way to utilize big data to create supplementary audit evidence that improves the effectiveness and efficiency of audit automation and decision making. The authors also discuss the application of these techniques to audit procedures.

In the current business environment, the development of data-intensive technologies (e.g., ERP systems, sensors, cloud storage, remote communication tools) facilitates the production and maintenance of large amounts of data, which necessitates a new data environment and serves as a motivator for audit automation. Leading accounting firms have leveraged deep learning, a cutting-edge use of artificial intelligence, to conduct audit tasks. For example, KPMG applies IBM Watson’s deep learning–powered systems to analyze banks’ credit files for commercial mortgage loan portfolios, and Deloitte has allied with Kira Systems to review contracts, leases, invoices, and tweets. The adoption of deep learning within the accounting profession is still, admittedly, at an early stage. To accelerate the wider use of this technology, it is necessary to create economies of scale by integrating its cognitive capabilities in the areas of textual analysis, voice recognition, image and video parsing, and judgment support into the audit process. This article discusses how the cognitive capabilities of deep learning could be applied to various audit procedures to enable audit automation and improve decision making.

Introduction

Some tasks that individuals accomplish effortlessly are actually extremely complex computational problems. Imagine how easily a human can distinguish cats from dogs, read traffic signs, or identify handwriting. In order for a machine to perform these tasks, the object must be converted to a machine-readable format and analyzed pixel-by-pixel. Due to variations in the object’s position, viewpoint, pose, lighting, and background, the complexities of the image—and thus, the object itself—can be represented in a variety of ways (Nicolas Pinto, David D. Cox, and James J. DiCarlo, “Why is Real-World Visual Object Recondition Hard?” *PLOS Computational Biology*, Jan. 25, 2008, <http://bit.ly/2qMUuG4>). Tasks like object recognition require the construction of deep artificial neural networks that simulate the multiple layers of neurons in a human brain and the ways that sensory data streams are processed through those layers (Robert D. Hof, “10 Breakthrough Technologies—Deep Learning,” *MIT Technology Review*, 2013, <http://bit.ly/2rIzj5Y>).

Human brains contain about 100 billion interconnected neurons. Each neuron receives input signals (e.g., the visual signal of a cat) from other connected neurons. If the combination of these signals exceeds a certain threshold (or activation level), the neuron transmits an output signal to other neurons. Humans’ deep layers of neurons and their complex interconnections result in a “thinking machine,” the human brain (“Neuralyst Users Guide,” Cheshire Engineering Corporation, 1994, <http://bit.ly/2qOM3bk>).

Although the idea of artificial neural networks dates back to the 1950s, such networks could not be called real artificial intelligence until recent advances in computational power and data storage enabled the development of deep neural networks that model the structure and thinking process of the brain. The hidden layers of a deep neural network automatically “learn” from



massive amounts of data (especially semi-structured or unstructured data) received by the input layer (e.g., millions of images, years' worth of speeches, terabytes of text files), recognize data patterns in more and more abstract representations as the data is processed and transmitted from one hidden layer to the next, and classify the data into predefined categories in the output layer.

The Effectiveness and Efficiency of Deep Learning

In order to ensure predictive effectiveness, a deep neural network is trained with a relatively large amount of data (e.g., millions of images of human faces) to allow the system to fine tune the parameters to minimize prediction errors. First, the input layer receives and identifies the most

essential data element of the input images, the intensity of the pixels. Second, the input neuron transforms the data with a nonlinear algorithm. During this process, the output of each neuron in the input layer is assigned a random weight, then combined and sent to the neuron of the next layer (the first hidden layer); now the first hidden layer can extract more abstract data representation, the edge (or the simple shape) comprising the pixels. Third, the neurons in the first hidden layer perform another round of complex nonlinear transformations to the received data and assign them new weights. The successor hidden layers then receive, transform, and send the data to the next layer in their turn; the system detects more and more complex data features (e.g., parts of the face) with each successive layer. Finally, the output

layer identifies the face. The output will be compared with the actual observation to check for errors, and the system will adjust the output weights and other parameters in the next round of training. The entire "learning" process will be repeated millions, or even billions, of times until error is minimized. The more examples the model is trained with, the higher accuracy the system can achieve. The performance of such a deep learning-based model (e.g., for financial misstatement detection) improves the sample size. More importantly, out-of-sample tests can be conducted with new samples to validate the effectiveness of the model using a set of metrics (e.g., F score, precision, recall, AUC).

The efficiency of deep learning can be demonstrated by the fact that it is able

to “learn” patterns in data without human intervention and requires fewer data preprocessing steps compared to traditional data mining approaches. For example, in the case of textual analysis, a person is usually needed to remove HTML tags, translate HTML characters into text characters, drop stop words and non-linguistic marks, and develop word lists. With a deep learning model, however, this time-consuming process is unnecessary. IBM Watson develops deep learning-based textual analysis that can directly read text files or even URLs; it automatically removes advertisements, navigation links, and other irrelevant content

and creates a list of data features such as authors, keywords, concepts, relationships among those concepts, and embedded sentiments or emotions (“Overview of the IBM Watson Natural Language Understanding Service,” IBM, <https://ibm.co/2rIsSQh>). The tool provides a high level of accuracy, and it is continuously improving as more and more data is used to train the model; as of February 2014, it claims to have processed over 3 billion documents a month for 40,000 users across 36 countries (Janet Wagner, “Deep Learning: 6 Real World Use Cases,” AlchemyAPI, 2014, <http://bit.ly/2qczSTI>). A recent study (Ting Sun, Yue Liu, and Miklos A. Vasarhelyi, “The Performance of Sentiment Feature of MD&As for Financial Misstatements Prediction,” working paper, 2017) found that this technique, without text preprocessing work, generally outperforms the tradi-

The Challenge of Big Data Analysis

tional text mining approach for predicting financial misreporting using the sentiment of management’s discussion and analysis (MD&A).
Semi-structured or unstructured big data contains a variety of information that allows auditors to freely explore the status of their clients’ products, services, and operations and reduces auditors’ dependence on their clients for data (Kyunghee Yun, Lucas Hoogduin, and Li Zhang, “Big Data as Complementary Audit Evidence,” *Accounting Horizons*, June 2015, <http://bit.ly/2rIFR10>). As a

reports that are issued quarterly or annually. Thus, the use of big data analysis in auditing has been hampered due to the lack of effective and efficient technologies to solve the issues in data extraction, transformation, and validation. While the challenges of big data analysis require a willingness to adopt more advanced data analytical technologies, such as deep learning, the availability of massive amounts of financial data facilitates the implementation and improvement of this technology in auditing.

The Need for Automated Audit Procedures

The automation of some tedious and repetitive audit processes would significantly enhance audit effectiveness and efficiency (Jon Raphael, “How Artificial Intelligence Can Boost Audit Quality,” *CFO*, June 2015, <http://bit.ly/2q301ZM>). Given a sufficiently large sample of how auditors make decisions under various circumstances, a deep learning system allows auditors to automate many tasks that have traditionally been conducted manually, such as checking inventories, processing paperwork, reviewing contracts, and even drafting audit reports.

Possible Applications

The application of deep learning to improve audit efficiency and effectiveness is especially relevant to facilitating repetitive audit procedures and supporting audit judgments. Deep learning can add value for routine tasks that involve massive amounts of data and require significant effort for auditors to solve, such as text analysis, speech recognition, and parsing images and videos. It can also be used to reduce manual work by automating some substantive procedures, such as confirmation and examination. Furthermore, these competencies will allow auditors to perform tasks, such as examining all corporate contracts, that are currently cost prohibitive or too complex for unaided human minds.

Given a sufficiently large sample of how auditors make decisions under various circumstances, a deep learning system allows auditors to automate many tasks that have traditionally been conducted manually.

result, mining and extracting meaningful patterns from big data has great value for audit decision-making, especially in the area of risk assessment. Analyzing big data is not easy, however; big data analysis was cited as one of the top challenges of the future by 25% of respondents to a 2014 AICPA survey (“2014 AICPA Survey on International Trends in Forensic and Valuation Services,” <http://bit.ly/2q3avIT>).

The vast majority of big data is semi-structured or unstructured, and requires labeling and categorizing; however, auditors cannot do this manually, since the data contains a variety of types and sources and is too voluminous for humans to process it all. In addition, portions of big data are usually generated on a real-time basis, so it requires timely responses. Furthermore, automated trading, which accounts for the majority of stock trades, cannot work well with

Text analysis. A large amount of textual data is generated and disseminated during a company’s operating processes, such as regulatory filings, transcripts of conference calls, press releases, earnings announcements, MD&As, business contracts, news articles, and social media messages. Textual data provides information on multiple aspects of a business from various perspectives. For example, MD&A contains management’s perspective on the company’s current financial situation and future prospects, analysts’ reports include retrospective analysis of past events and forecasts of future earnings and cash flows, and social media posts can include advertisements, product reviews, and news announcements.

Textual analysis can be automated by deep learning; specifically, text data can be classified based on features of interest. Furthermore, a deep learning model can be trained using transcripts of Q&A sections of conference calls—labeled as positive, negative, or neutral in terms of sentiment—to predict future calls. The entire procedure can be performed automatically, and the result is machine-readable. In this way, a deep learning model transforms qualitative information that used to require great human effort to analyze into quantitative data that can easily be integrated with other data for further audit analysis.

Deep learning algorithms further enrich audit evidence by identifying related concepts or topics, recognizing entities (e.g., people, place, events, companies), extracting emotions (e.g., anger, joy, sadness, disgust), and understanding subject-action-object relationships. In addition, they can link concepts to a document and tag them accordingly.

This textual analysis is better suited to unprepared content (the Q&A section of conference calls, Facebook status updates, blog posts) than prepared content (e.g., press releases, the presentation section of conference calls). Compared to prepared content, unprepared content

may be filled with slang, idioms, and other linguistic clues reflecting the cognitive process of the speaker (Marina Druz, Alexander F. Wagner, and Richard J. Zeckhauser, “Tips and Tells from Managers: How Analysts and the Market Read between the Lines of Conference

calls.” Although public accounting firms offer deception detection training to help their auditors identify verbal red flags, the information processing of interviews is a daunting task because interviewees exhibit myriad verbal behaviors. It is difficult and inefficient for auditors to analyze all oral

Exhibit 1
Input, Output, and Applicable Audit Procedures for Deep Learning–Based Textual Analysis

	Examples
Input Data	Regulatory filings, press releases, earnings announcements, MD&As, business contracts, comment letters from SEC, news articles, analysts’ reports, e-mails, disclosures on company’s website, social media messages
Output Features	Sentiment, emotion, entity, topic, concept, keywords, authors
Applicable Audit Procedures	Inspection, analytical procedure, confirmation

MD&A=management’s discussion and analysis

Calls,” National Bureau of Economic Research, 2015, <http://bit.ly/2rKXIAH>), which may provide clues to potential risks. *Exhibit 1* provides examples of textual documents, possible output features, and audit tasks for which deep learning is applicable.

Speech recognition. To acquire background information about a client’s business and industry environment and to collect audit evidence, auditors interview management, internal auditors, employees, predecessor auditors, bankers, legal counsel, underwriters, analysts, or other stakeholders. The language that subjects use and how they respond to questions over the course of the interview can be just as important as the answers themselves, because they may indicate deception. For example, the use of terms that suggest uncertainty, such as “kind of,” “maybe,” or “sort of,” as well as response latency, could be signs of concealment or falsifi-

responses or transcribe them into text manually, and even transcribed text documents can still be a chore for auditors to analyze.

The speech recognition function of deep learning can now transcribe and translate speech on a real-time basis regardless of noise or the various accents of speakers, enabling analysis of the text and extraction of emotion, risk factors, and other insights directly. Researchers are now considering using embodied conversational agents (ECA), which are “autonomous computer interfaces capable of human-like interactions” (Matthew D. Pickard, Mary B. Burns, and Kevin C. Moffit, “A Theoretical Justification for Using Embodied Conversational Agents to Augment Accounting-Related Interviews,” *Journal of Information Systems*, Fall 2013, <http://bit.ly/2qcBf4Z>), to conduct interviews automatically. Due to technical limitations, current ECAs ask

Exhibit 2
Input, Output, and Applicable Audit Procedures for Deep Learning–Based Speech Analysis

	Examples
Input Data	Interviews, conference calls, phone calls, project meetings, presentations
Output Features	Risk of deception, sentiment, emotion, entity, topic, concept, keyword, speaker, suggested follow-up actions
Applicable Audit Procedures	Inquiry, inspection

only pre-designed questions and record the responses (Matthew D. Pickard, Ryan Schuetzler, Joseph Valacich, and David A. Wood, “Next Generation Accounting Interviewing: A Comparison of Human and Embodied Conversational Agents as Interviewers,” working paper, 2017). Once integrated with deep learning, ECAs may be able to read verbal signs of deception and ask follow-up questions based on the interviewee’s responses. Similarly, other audio documents such as conference calls, phone calls, and project meetings could be automatically processed by deep learning technology. *Exhibit 2* depicts possible input and output data, as well as applicable areas for deep learning–based speech recognition.

Image and video parsing. Certain routine, manual audit procedures can now be automated with the visual recognition function of deep learning. For example, deep learning algorithms can identify the content (e.g., model, quantity, condition of inventory) of an image from a video filmed by a drone in a company’s warehouse. Deep learning systems are able to extract a series of predefined numerical attributes describing the content of the video, attach searchable tags accordingly, and save both the attributes and the images to the auditor’s data warehouse. Furthermore, deep learning video analyt-

ics are capable of recognizing human faces, detecting objects, and identifying concepts and types of scenes on a nearly real-time basis, and the video processing speed is extraordinarily high; Clarifai, for instance, offers deep learning software that can analyze a 3.5-minute video clip within 10 seconds. The internal control observation procedure could eventually be automated using this technology to analyze videos recorded by drones in offices or at large worksites. *Exhibit 3* shows the input video/image data, output features, and audit procedures to which the image and video analysis function of deep learning could apply.

Judgment support. In addition to conducting repetitive and mechanical tasks, deep learning provides a new way to support audit judgement and improve audit quality. Financial statements can be scanned and the financial statement items automatically linked to related supporting evidence such as video clips, press release, news, tweets, and interviews (Raphael 2015), as well as corresponding data attributes (e.g., the overall sentiment and the topic of a news) that have been extracted by deep learning models. For example, auditors can select data attributes in order to predict fraud risk; the selected attributes are then combined with traditional financial or nonfinancial

data fields to develop a new deep learning prediction model. Deep learning performs as an appropriate prediction algorithm in this case because, by introducing the extracted attributes, the number of predictors is much larger than what a traditional machine-learning algorithm could process. For each assertion, the output of the model could be the predicted risk level or suggested follow-up tests, depending upon the nature and the label of the training data.

Implementation

Auditors can leverage deep learning technologies without being experts by using cloud-based services. But to improve the predictive performance of deep learning algorithms for audit automation and judgment support, it is important that auditors and machine learning specialists make a cooperative effort to develop auditing-specific training datasets (e.g., the text dataset of 10-Ks) to establish deep learning models that are designed specifically for audit tasks. It is worth noting that retrieving the auditing-specific data is easier today than it was several years ago. Various data services, such as SeekiNF (<http://www.seekedgar.com>) allow users to search for financial and nonfinancial information in databases containing millions of records in seconds.

The data analytical process must be well planned to balance human resources and technology as well as to ensure that auditors strengthen their professional skills and judgment by concentrating on high-risk items identified through deep learning. In addition, regulators should start considering guidance for the use of data analytics technology. In fact, the Data Analytics Working Group of the IAASB released a Request for Input, “Exploring the Growing Use of Technology in the Audit,” in September 2016, and is seeking advice on determining whether new or revised international standards or guidance are necessary. Furthermore, the PCAOB has set up a research project (Changes in the Use of

Exhibit 3
Input, Output, and Applicable Audit Procedures for Deep Learning–Based Image and Video Analysis

	Examples
Input Data	Inventory counting and other control activity; interviews; video taken in office, warehouse, or store
Output Features	Object, human face, concept and type of scene
Applicable Audit Procedures	Observation, inquiry, inspection

Data and Technology in the Conduct of Audits, <http://bit.ly/2qPdpOF>) on analytics in audits, and Steven Harris, a PCAOB board member, issued a passionate plea

for disruptive change in auditing and the need for analytics (“Technology and the Audit of Today and Tomorrow,” speech at PCAOB/AAA annual meeting, Apr. 20,

2017, <http://bit.ly/2rtLtCD>).

How can smaller firms and sole practitioners acquire this now very experimental and expensive technology? An entire industry is already evolving, with pay-by-usage options available. CPAs, however, will need to beef up their competencies in statistics and IT. □

Ting Sun is a PhD student in the department of accounting and information systems at Rutgers Business School (Rutgers University), Newark, N.J. Miklos A. Vasarhelyi, PhD, is the KPMG Distinguished Professor of Accounting Information Systems and Director of Rutgers Accounting Research Center and Continuous Auditing & Reporting Lab in the department of accounting and information systems at Rutgers Business School, Newark, N.J.



Capstan knows how to accelerate depreciation and maximize tax savings. Expertise in Tangible Property Regulations. A proven track record in Cost Segregation Studies. A true partner to CPAs and their clients.

Choose Capstan for Cost Segregation Services



215-885-7510
capstantax.com

We'll shake on that.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.