CAPCO

THE CAPCO INSTITUTE JOURNAL OF FINANCIAL TRANSFORMATION

DATA INTELLIGENCE

Natural language understanding: Reshaping financial institutions’ daily reality

BERTRAND K. HASSANI

DATA ANALYTICS

50TH EDITION | NOVEMBER 2019

THE CAPCO INSTITUTE

JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

CONTENTS

DATA MANAGEMENT

DATA ANALYTICS

DATA INTELLIGENCE

DEAR READER,

Welcome to the milestone 50th edition of the Capco Institute Journal of Financial Transformation.

Launched in 2001, the Journal has covered topics which have charted the evolution of the financial services sector and recorded the fundamental transformation of the industry. Its pages have been filled with invaluable insights covering everything from risk, wealth, and pricing, to digitization, design thinking, automation, and much more.

The Journal has also been privileged to include contributions from some of the world's foremost thinkers from academia and the industry, including 20 Nobel Laureates, and over 200 senior financial executives and regulators, and has been co-published with some of the most prestigious business schools from around the world.

I am proud to celebrate reaching 50 editions of the Journal, and today, the underlying principle of the Journal remains unchanged: to deliver thinking to advance the field of applied finance, looking forward to how we can meet the important challenges of the future.

Data is playing a crucial role in informing decision-making to drive financial institutions forward, and organizations are unlocking hidden value through harvesting, analyzing and managing their data. The papers in this edition demonstrate a growing emphasis on this field, examining such topics as machine learning and AI, regulatory compliance, program implementation, and strategy.

As ever, you can expect the highest caliber of research and practical guidance from our distinguished contributors, and I trust that this will prove useful to your own thinking and decision making. I look forward to sharing future editions of the Journal with you.

Lance Levy, **Capco CEO**

FOREWORD

Since the launch of the Journal of Financial Transformation nearly 20 years ago, we have witnessed a global financial crisis, the re-emergence of regulation as a dominant engine of change, a monumental increase in computer processing power, the emergence of the cloud and other disruptive technologies, and a significant shift in consumer habits and expectations.

Throughout, there has been one constant: the immense volume of data that financial services institutions accumulate through their interactions with their clients and risk management activities. Today, the scale, processing power and opportunities to gather, analyze and deploy that data has grown beyond all recognition.

That is why we are dedicating the 50th issue of the Journal of Financial Transformation to the topic of data, which has the power to change the financial industry just as profoundly over the coming 20 years and 50 issues. The articles gathered in this issue cover a broad spectrum of data-related topics, ranging from the opportunities presented by data analytics to enhance business performance to the challenges inherent in wrestling with legacy information architectures. In many cases, achieving the former is held back by shortcomings around the quality of, and access to, data arising from the latter.

It is these twin pillars of opportunity and challenge that inform the current inflection point at which the financial industry now stands. Whilst there is opportunity to improve user experiences through better customer segmentation or artificial intelligence, for example, there are also fundamental challenges around how organizations achieve this – and if they can, whether they should.

The expanding field of data ethics will consume a great deal of senior executive time as organizations find their feet as they slowly progress forward into this new territory. In my view, it is critical that organizations use this time wisely, and do not just focus on short-term opportunities but rather ground themselves in the practical challenges they face. Financial institutions must invest in the core building blocks of data architecture and management, so that as they innovate, they are not held back, but set up for long-term success.

I hope that you enjoy reading this edition of the Journal and that it helps you in your endeavours to tackle the challenges of today's data environment.

Guest Editor
Chris Probert, **Partner, Capco**

NATURAL LANGUAGE UNDERSTANDING: RESHAPING FINANCIAL INSTITUTIONS' DAILY REALITY

BERTRAND K. HASSANI | Université Paris 1 Panthéon-Sorbonne, University College London, and Partner, AI and Analytics, Deloitte[1]

ABSTRACT

Though in the past, data captured by financial institutions and used to understand customers, processes, risks, and, more generally, the environment of financial institutions was mainly structured, i.e., sorted in "rigid" databases, today, that is no longer the case. Indeed, the so-called structured data is representing no more than a drop in an ocean of information. The objective of this paper is to present and discuss opportunities offered by natural language processing and understanding (NLP, NLU) to analyze the unstructured data, and automate its treatment. Indeed, NLP and NLU are essential to understanding and analyzing banks' internal way of functioning and customer needs in order to bring as much value as possible to the firm and the clients it serves. Consequently, while we will briefly describe some algorithms and explain how to implement them, we will focus on the opportunities offered as well as the drawbacks and pitfalls to avoid in order to make the most out of these methodologies.

1. INTRODUCTION AND MOTIVATION

Financial institutions' current objective or fantasy – as it is a matter of opinion – of full automation requires several things to become reality, such as a complete capture of information (structured and unstructured data) pertaining to the customers and the bank itself, i.e., behaviors and needs evolution over time, dynamic risk exposure, perception of bank activity, and so on and so forth. Though techniques that rely on structured data to score customers, to understand their needs, and the products that might suit them have been used for years, and is nowadays quite advanced, the solutions using unstructured data are still far from being fully deployable at an industrial level, in particular when it comes to natural language processing, natural language understanding, and even more when it comes to natural language generation.

In order to make sure that the terms introduced above are clear, the concepts behind are now introduced. The first one, natural language processing (NLP) [Collobert et al. (2011)], simultaneously belongs to the subfields of linguistics, computer science, information engineering, and artificial intelligence (AI). NLP deals with the interactions between computers and human languages, and as such how computers are processing and analyzing large quantities of natural language data. Natural language understanding (NLU) [Liu et al. (2019)] is itself a subtopic of NLP that deals with machine accurate comprehension of languages, i.e., tone, nuances, etc. NLU is considered an AI-hard problem, i.e., it implies that the complexity of these computational problems is equivalent to that of solving the central AI problem; in other words, making computers as intelligent as people. This would require advanced approaches as the problem would not be solved by a simple algorithm. NLU is usually used on top of

[1] The opinions, ideas and approaches expressed or presented are those of the author(s) and do not necessarily reflect any past or future positions of Deloitte. As a result, Deloitte cannot be held responsible for them.

NLP algorithms utilizing context from recognition devices (automatic speech recognition (ASR): Qin et al. (2019), personalized profiles, etc.), in all of its forms, to decipher the meaning of sentences to execute the implied intent. NLU aims at informally assessing the probability of that intent. Finally, natural language generation (NLG) [Tran and Nguyen (2019)] consists of a program able to answer queries as if a human being was talking or writing.

There is considerable commercial interest in the field of NLP-NLU. Its application to automated reasoning, machine translation, question answering, news-gathering, text categorization, voice activation, archiving, and large-scale content analysis generates a genuine interest from financial institutions. Swedbank's famous application of NLP for customer service illustrates the efforts made by financial institutions. At the very least, using their customer base, and the data pertaining to it, financial institutions are able to develop tools to handle simple and common requests, and accurately pass along the most complex to human beings for dedicated processing as unfortunately AI systems are often unable to deal with complex customer requests.

Fundamental improvements in AI and ML methodologies are required to cover this gap, and it will be years before a customer service can be fully automated, if it ever happens; as it might not be the case considering that we are social animals. Fortunately, handling simple requests and routing complex requests is still highly valuable for banks with huge customer service costs, allowing them to reallocate human resources to tasks of higher added value.

Furthermore, it is noteworthy to mention that U.S. Bank[2] and ING[3] already allow using Siri or Alexa to interact with them. In these cases, methodologies implemented belong to ASR and, therefore, implies signal processing before converting the speech to text for further analysis. However, Voice ID, as rolled out by Santander in the U.K.,[4] does not necessarily imply NLP as signal matching (after or before encryption) is the only required thing. Banks are not the only ones looking at NLP techniques. Fintechs are increasingly relying on these approaches, as illustrated by the three following examples. For instance, B2B-oriented fintech venture, Clinc,[5] offers a conversational AI platform to banks for personal finance, wealth management, and customer services. Cleo[6] provides B2C solutions helping customers to gain insight in, and to improve the management of their daily spending. Invyo[7] helps financial institutions to identify opportunities in fintech using NLP-NLU.

In this paper we will discuss the state-of-the-art, the latest trends, the use-cases, the opportunities, and the challenges that implementing these methodologies will engender. We will also present the path from NLP to NLG, considering that NLP itself offers a large scope of opportunities and possibilities, and already allows addressing several fundamental issues faced by financial institutions.

2. METHODOLOGY: STATE-OF-THE-ART

Natural language treatment-related concepts have been previously introduced, hence in the following subsections we will present some of the most widely used methodologies to achieve automated treatment of textual data. To facilitate the understanding of the underlying methodologies, a distinction will be made in what follows between approaches requiring tremendous computational power (referred to as "heavy methodologies") and methodologies allowing a local implementation (referred to as "light methodologies" in this paper).

2.1 Light methodologies

Light methodologies, as mentioned previously, can be developed at a local level and does not necessarily require the consideration of an alternative IT infrastructure.

2.1.1 RETRIEVING INFORMATION

The methodology traditionally used for retrieving information is usually referred to as TF–IDF (term frequency–inverse document frequency), which is a numerical statistic reflecting how important a word is to a document in a collection or corpus [Luhn (1957), Jones (1972)]. The value increases proportionally to the number of times a word appears in the target document and is offset by the number of documents in the corpus that contain the word, which helps adjusting for the fact that generally some words appear more frequently. This approach is a first step towards document classification, as by retrieving specific words it is possible to sort documents by topics if in the set considered the words have the same meanings.

[2] https://bit.ly/351nrja
[3] https://bit.ly/2Dq2KBQ
[4] https://bit.ly/2WhtQBW
[5] https://bit.ly/358LWLB
[6] https://bit.ly/1UhoArW
[7] https://bit.ly/2Vc07e7

2.1.2 CAPTURING THE CONTEXT

The first methodology of interest is called "word2vec". This approach consists of a set of related models used to produce word embeddings. Word embedding consists of numerically capturing the context of a word in a document, semantic, and syntactic similarity, as well as relations with other words. These models are not very deep as they consist of two-layer neural networks trained for the reconstruction of linguistic contexts of words. Word2vec takes as input a large corpus of text and produces a very large vector space (usually hundreds of dimensions), in which each unique word in the corpus is assigned a corresponding vector in the space. Word vectors are positioned in the vector space such that words that share common contexts in the corpus are located close to one another in the space [Mikolov et al. (2013), Goldberg and Levy (2014), Rong (2014)].

However, though very powerful in a homogeneous context, the methodology is rather limited in an open environment, as only one word embedding can be obtained per word, i.e., word embeddings can only store one vector for each word. Consequently, with respect to the methodology, the word "bank" has only one meaning for "I withdrew some money from my bank account" and "I went walking on the river bank." This issue can be highly misleading. Furthermore, one main drawback of being easy to implement on a small infrastructure is that this approach is difficult to train on large datasets, and it is very challenging to fine tune them and tailor them to a particular domain.

An alternative approach is called GloVe (Global Vectors for Word Representation) [Pennington et al. (2014)]. GloVe is an unsupervised learning algorithm for obtaining vector representations for words. Training is performed on aggregated global words co-occurrence statistics from a corpus, and the resulting representations showcase interesting linear substructures of the word vector space.

Both word2vec and GloVe learn vectors of words from how often they co-occur in a textual corpus. However, contrary to word2vec, which is a predictive model, GloVe is a count-based approach [Almasian et al. (2019)].

2.1.3 SUMMARIZING A TEXT

While the previous approaches were interesting to capture contextual information or identify similarities between words, here we introduce TextRank, a graph-based ranking model designed for text processing, and which can be used for finding the most relevant sentences in text as well as the keywords [Mihalcea and Tarau (2004)].

In order to find the most relevant sentences in a text, a graph is constructed where the nodes of the graph represent each sentence in a document and the edges between sentences are reflecting content overlap, usually obtained by calculating the number of common words contained in two sentences.[8]

Following the creation of this network of sentences, these ones are entered the PageRank algorithm [Page (2001)], which aims to identify the most important – and – theoretically – the most relevant sentences. To create a summary of the text, we only have to gather and combine the most important sentences.

Alternatively, if we are interested in finding relevant keywords, the TextRank algorithm can build a network of words. This network is built by looking at which words follow one another. A link is created between two words if they follow one another, and this link is attributed a higher weight if these two words frequently materialize next to each other in the text considered.

As for the creation of a summary, the PageRank algorithm is laid on the resulting network to obtain each word's importance. The words ranked at the top are kept, as these are considered relevant (according to the algorithm). Then, a keyword table is obtained by gathering the relevant words together if they come forth following one another in the considered text.

An important aspect of TextRank is that the algorithm does not need deep linguistic knowledge, nor language- or domain-specific annotated corpus. Consequently, this aspect makes it highly portable and generalizable [Barrios et al. (2016)].

2.2 Heavy methodologies

In the following sections, we briefly introduce the latest methodologies developed by institutions such as OpenAI or various branches of Google, which usually require high specifications in terms of required IT infrastructure and computing power. It is noteworthy to mention that it is not always necessary to retrain the full models on the corpus of interest, as fine-tuning them on specific tasks might be sufficient. This possibility drastically reduces the required computing power.

[8] https://bit.ly/2LI0SZk

2.2.1 BERT

BERT stands short for Bidirectional Encoder Representations from Transformers [Devlin et al. (2018)] and is closely related to GPT [Radford et al. (2018)]. This large language model is trained on free text and then fine-tuned on specific tasks without customized network architectures. BERT improves on the GPT approach by making the training bidirectional. Consequently, the model learns to predict context on each side of the target. BERT allowed obtaining "best-in-class" results in multiple NLP tasks, such as "question answering" or "natural language inference".

A previously mentioned, BERT's key technical innovation is to apply the bidirectional training of Transformer, a popular attention model [Vaswani (2017)], to language modeling. This is in contrast to previous efforts, which looked at a text sequence either from left to right or combined left-to-right and right-to-left training. The paper's results show that a language model that is bidirectionally trained can have a deeper sense of language context and flow than single-direction language models. It is not surprising that a representation able to learn the context around a word rather than just after the word is able to better capture its meaning, both syntactically and semantically. The main achievement of this approach is that it predicts the missing words without any information regarding which words have been replaced or which words should be predicted.[9]

2.2.2 OPENAI GPT-2

As with Google's BERT, GPT-2 [Radford et al. (2018)] is Open AI's successor to GPT. It was originally trained to predict the next word in 40GB of Internet text. GPT-2 is a large transformer-based language model with 1.5 billion parameters, trained on a dataset of 8 million web pages (i.e., ten times the number of parameters of GPT and ten times the size of the initial training set), supporting our choice to put it in the Heavy Methodology section. GPT-2 is trained with a simple objective, which is to predict the next word, given all of the previous words within some text. As claimed by the authors, GPT-2 seems to have the ability to generate conditional "synthetic text samples of unprecedented quality." Furthermore, GPT-2 outperforms other language models trained on specific domains without requiring the use of domain-specific training sets.[10]

2.2.3 XLNET

XLNET is a generalized autoregressive model. An item is dependent on the previous ones. XLNET captures bi-directional context through "permutation language modeling". It combines auto-regressive modeling with a bi-directional context approach. It outperforms BERT on tasks such as question answering, natural language inference, sentiment analysis, and document ranking.[11]

Permutation language modeling allows capturing context in both directions by training an autoregressive model on all rearrangements of words possible in a sentence [Yang et al. (2019)].

3. USE-CASES

Considering that most data is unstructured (photo, text, audio…) and that textual data represent the largest part, the combination of NLP-NLU algorithms such as those described above, and the availability of enhanced computational power permitting processing this data, is reshaping financial institutions' internal management on the one hand and the way they interact with their customers on the other. In this section, we will present use-cases detailing how relying on methodologies presented in the previous sections we can improve financial institutions processes, starting with what in our opinion is the most important: customer experience.

Indeed, customer experience [McColl-Kennedy et al. (2019)] is arguably the most important thing in a commercial relationship, as that is what defines clients' perception of a brand, a shop, or a professional. Customer experience is what makes people buy and what makes consumer come back. The key to offering a good customer experience is to understand their needs precisely, to answer them in the most customized manner possible, and following up proactively to make the customer feel as if they were experiencing a "valet" service.

To achieve such a performance, data must be analyzed (i.e., customer information, interactions with the financial institutions, products already available, social medias, etc.) in a holistic fashion. To analyze this data, techniques described above to structure the interactions in an actionable manner (i.e., such that we would be able to push the right product at the right time, or to demonstrate empathy during interactions between an agent and a customer) are very helpful.

[9] https://bit.ly/2S8w6Jt: This link contains the code and the documentation explaining to fine-tune the model using TPUs on the Cloud.
[10] https://bit.ly/2M9B9b3: This link contains the code to run GPT-2 on the Cloud
[11] https://bit.ly/31LxDdt

For instance, claims have to be properly dealt with. Indeed, claims are part of the customer journey, and if not satisfied with the bank's services, the customer needs to be able to express their grief or disappointment using the multiple channels usually available: a web portal, an email address, a telephone line, or directly in a branch. Multiple channels imply multiple data formats; for instance, audio or texts. Note that the text might be pure as it has not been modified by anyone, or might be reported by an agent and consequently might suffer from perception bias, and may, therefore, require additional treatment. Besides, audio format requires a first transformation, implementing a speech-to-text strategy [Bansal et al. (2018)]. Once the data is pre-processed, the methodologies presented above might be implemented for various purposes. Indeed, claims or, more generally, bank-customer interactions could be classified by type using keyword extractions. Besides, some claims might be properly dealt with using a bot. Advanced chatbots (using NLP and NLU) may allow a precise and customized treatment of a particular matter.

From an internal point of view, an appropriate use of resources is the essence of appropriate management. Considering that we are in an era of specialization and professionalization, a precise understanding of resume, skills, and evolution are critical for success. However, analyzing, qualifying, and routing resumes towards the most appropriate recipient is gradually becoming more complex. Consequently, for keywords extraction, to understand how people sell their skills, or to capture the confidence transpiring from resumes, the algorithms presented earlier can be used as these have literally been designed to tackle these specific tasks.

NLU can also be used for risk management purposes. For instance, named-entity recognition ([Khalifa and Shaalan (2019)] can be used to screen each and every contract in a folder to check that none of them has been signed by a blacklisted third-party, and can, therefore, be used for anti-money laundering purposes. Named-entity recognition or NER consists of extracting and classifying relevant information from unstructured text into predefined categories, such as the person names, organizations, locations, or monetary values. This approach combined with or lauded in a graph database, would allow for making connections between customer, investments, location, etc.; mechanically enhancing compliance assurance.

Related to non-compliance issues is the so-called reputational risk. Indeed, reputational risk is today one of the most damaging exposures bank face. Tackling the issue requires gathering a large quantity of information coming from traditional medias, podcast, or social medias for the matter at hand, as well as the implementation of a sentiment analysis [Cambria et al. (2019] approach relying once again on NLU and somehow on one of the methodology presented in the second section of this paper. As a matter of fact, the Heavy Methodologies are very interesting to obtain a precise classification of articles, whether these are positive, negative, or neutral with all the nuances that can be captured in between. A score might even be built.

From an operational risk management [Hassani (2016)] point of view, the use of external data to either feed internal databases, scenario analysis, or risk control self-assessment procedures are already consuming large quantities of textual data. These tasks require both manual classification and light sentiment analysis to reduce perception bias. These tasks could be fully automated using methodologies presented above.

Last but not least, both credit scoring and segment hunting based on credit scoring strategies are increasingly reliant on external unstructured data capture and analysis, cobbling the way towards reduced risk taken by financial institutions. The better understanding of customer profiles from a credit worthiness point of view is mechanically improving the accuracy of the pricing of the loans [Wang et al. (2018), Crouspeyre et al. (2019)].

4. LIMITATIONS

After presenting some of the opportunities offered by NLP-NLU techniques, this section will address their limitations. Though, some genuine value can be obtained from NLP-NLU approaches, the methodologies suffer from limitations worth bearing in mind.

The first limitation that comes to mind is the quantity of data required to achieve good results. On case specific tasks, this data might not be available, making the validity of models questionable.

The second aspect is related to the tremendous computational power required. If your company is not cloud computing oriented, though not impossible, the tasks might be extremely complicated to fulfill.

Another main issue associated with NLP-NLU strategies as deployed in banks is relevance. As the devil is in the detail, it is possible that actual understanding of real meanings is not appropriate and as a result the provided response not accurate. One may wonder what is the impact of an inappropriate action triggered by the model?

Besides, the number of parameters associated with the models, in particular with the Heavy Methodologies, may potentially generate a model risk in the future, as its interpretability is questionable.

Furthermore, from a business point of view, people's fickle natures have not been taken into account and as such the responses are assumed stationary, implying that past data is informative of the future. As of today, state-of-the-art results lead us to think that AI algorithms allow for understanding humans better than humans themselves; as if processes were always linear. We believe that this will not always be the case, due to the fact that people may not be completely honest in what they say, they might use nuances due to their education, they might not be direct, they might be biased, or they might be fed up with something they used to enjoy. NLP-NLU is more about having it right all the time than the underlying algorithms.

CONCLUSION

Today, as discussed in this paper NLP-NLU is reshaping the way financial institutions are working, impacting every aspect of the value chain. Considering the number of tasks involved in the three domains of NLP, NLU, and NLG, it is better to state the objective to achieve or the problem to solve clearly and carefully. It is from having clear objectives that we will be able to derive the type of data to be gathered, the most appropriate IT architecture, or the most effective modeling strategy. We need to select the models solving the problems, not the problems being fine-tuned by the models.

In this paper, after presenting the current market appetite for NLP-NLU methodologies and use-cases, we discussed some of the pertaining limitations, and though applications are very attractive, to ensure the durability of the solutions these limitations have to be borne in mind at all time. In any case, though NLP-NLU might sometimes fails, the recent evolutions of the research in the field support the statement that it should attract major investments from financial institutions in the coming years.

REFERENCES

Almasian, S., A. Spitz, and M. Gertz, 2019, "Word embeddings for entity-annotated texts," European Conference on Information Retrieval, 307–322. Springer

Bansal, S., H. Kamper, K. Livescu, A. Lopez, and S. Goldwater, 2018, "Pre-training on high-resource speech recognition improves low-resource speech-to-text translation," arXiv preprint arXiv:1809.01431

Barrios, F., F. López, L. Argerich, and R. Wachenchauzer, 2016, "Variations of the similarity function of textrank for automated summarization," arXiv preprint arXiv:1602.03606

Cambria, E., S. Poria, and A. Hussain, 2019, "Speaker-independent multimodal sentiment analysis for big data," in Seng, K. P., L.-m. Ang, A. W.-C. Liew, and J. Gao (eds.), Multimodal analytics for next-generation big data technologies and applications, 13–43, Springer

Collobert, R., J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, 2011, "Natural language processing (almost) from scratch," Journal of Machine Learning Research 12, 2493–2537

Crouspeyre, C., E. Alesi, and K. Lespinasse, 2019, "From creditworthiness to trustworthiness with alternative NLP/NLU approaches," in Proceedings of the First Workshop on Financial Technology and Natural Language Processing, 96-98

Devlin, J., M-W. Chang, K. Lee, and K. Toutanova, 2018, "Bert: pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805

Goldberg, Y., and O. Levy, 2014, "Word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method," arXiv preprint arXiv:1402.3722

Hassani, B., 2016, Scenario analysis in risk management. Springer International Publishing

Jones, K. S., 1972, "A statistical interpretation of term specificity and its application in retrieval," Journal of Documentation 28:1, 11-21

Khalifa, M., and K. Shaalan, 2019, "Character convolutions for Arabic named entity recognition with long short-term memory networks," Computer Speech & Language 58, 335–346

Liu, X., P. He, W. Chen, and J. Gao, 2019, "Multi-task deep neural networks for natural language understanding," arXiv preprint arXiv:1901.11504

Luhn, H. P., 1957, "A statistical approach to mechanized encoding and searching of literary information," IBM Journal of Research and Development 1:4, 309–317

McColl-Kennedy, J. R., M. Zaki, K. N. Lemon, F. Urmetzer, and A. Neely, 2019, "Gaining customer experience insights that matter," Journal of Service Research 22:1, 8–26

Mihalcea, R., and P. Tarau, 2004, "Textrank: Bringing order into text," Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, 404–411

Mikolov, T., K. Chen, G. Corrado, J. Dean, L. Sutskever, and G. Zweig, 2013, "Word2vec," https://bit.ly/2esteWf

Page, L., 2001, "Method for node ranking in a linked database," September 4, U.S. Patent 6,285,999

Pennington, J., R. Socher, and C. Manning, 2014, "Glove: Global vectors for word representation," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–1543

Qin, Y., N. Carlini, I. Goodfellow, G. Cottrell, and C. Raffel, 2019, "Imperceptible, robust, and targeted adversarial examples for automatic speech recognition," arXiv preprint arXiv:1903.10346

Radford, A., K. Narasimhan, T. Salimans, and I. Sutskever, 2018, "Improving language understanding by generative pretraining," https://bit.ly/2t9cjyM

Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, 2019, "Language models are unsupervised multitask learners," OpenAI Blog 1:8

Rong, X., 2014, "word2vec parameter learning explained," arXiv preprint arXiv:1411.2738

Tran, V-K., and L-M. Nguyen, 2019, "Gating mechanism based natural language generation for spoken dialogue systems," Neurocomputing 325, 48-58

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, 2017, "Attention is all you need," Advances in Neural Information Processing Systems, 5998–6008

Wang, C., D. Han, Q. Liu, and S. Luo, 2018, "A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism LSTM," IEEE Access, 7, 2161-2168

Yang, Z., Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, 2019, "XLNet: generalized autoregressive pretraining for language understanding," arXiv preprint arXiv:1906.08237

ABOUT CAPCO

Capco is a global technology and management consultancy dedicated to the financial services industry. Our professionals combine innovative thinking with unrivalled industry knowledge to offer our clients consulting expertise, complex technology and package integration, transformation delivery, and managed services, to move their organizations forward.

Through our collaborative and efficient approach, we help our clients successfully innovate, increase revenue, manage risk and regulatory change, reduce costs, and enhance controls. We specialize primarily in banking, capital markets, wealth and asset management and insurance. We also have an energy consulting practice in the US. We serve our clients from offices in leading financial centers across the Americas, Europe, and Asia Pacific.

WORLDWIDE OFFICES

APAC
Bangalore
Bangkok
Hong Kong
Kuala Lumpur
Pune
Singapore

EUROPE
Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
London
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA
Charlotte
Chicago
Dallas
Houston
New York
Orlando
Toronto
Tysons Corner
Washington, DC

SOUTH AMERICA
São Paulo

FSC
www.fsc.org
MIX
Paper from responsible sources
FSC® C001648

WWW.CAPCO.COM

CAPCO