CS 410:\ Search Engines & Text Mining

BERT For Sentiment Analysis Tech Review

What in-house small business applications are there & how accessible are they technically & financially?

The author with Sebastian Thrun & David Silver, graduating from a different kind of degree

Introduction

This “tech review” will cover the BERT unsupervised representation language model and some possible use cases for small (and mid-sized) businesses to utilize said BERT model for sentiment analysis by employing in-house talent which is not otherwise focused on technical work in their day-to-day job-related activities.

Body

[Extremely] Brief [Partial] History of NLP

Several decades ago [2], Natural Language Processing relied on hard coded rule based systems. For example, if you see the word "good" [2] assume the sentiment is positive would be a prototypical example of one such rule that could conceivably exist in such a system. Some time after that, *probabilistic systems [2]* were used, which were a lot more effective but still required substantial amounts of "hand engineering". Finally, as Andrew Ng says in [2], "to now where NLP relies much more on machine learning and deep learning".

Around 2014 the state of the art in NMT (neural machine translation) were sequence-to-sequence [8] models which are "deep learning models … that have achieved a lot of success… that take a sequence of items … and outputs another sequence of items", and "are explained in the two pioneering papers (Sutskever et al., 2014 [9], Cho et al., 2014 [10])."

The Google brain paper *Attention is All You Need* [6] introduced the *Transformer*. At that point in time (circa June 2017) attention was already a cutting edge methodology for connecting encoders to decoders in NLP deep neural networks for e.g. machine translation (from one language to another) [5], but the major innovation [5] of the Attention [6] paper was that it completely cut out the RNN's and only retained the Transformers.

Finally, at the end of our short elided history of NLP we have *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* by Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova [7]. BERT & friends marked a sea-change in the world of NLP in 2018 "For many, the introduction of deep pre-trained language models in 2018 (ELMO, BERT, ULMFIT, Open-GPT, etc.) signals the same shift to transfer learning in NLP that computer vision saw." [11].

How Small Business [SMBs, Small & Mid-sized Businesses] Can Utilize Sentiment Analysis & Deep Learning

There are not many examples of very small non-tech businesses doing much in the way of high technology as they typically would not have the personnel or know-how to implement anything themselves. There are entire constellations of companies that use whatever proprietary methodologies they are using behind the scenes (perhaps BERT for instance if they are doing sentiment analysis, but how they are leveraging that becomes proprietary as

they do whatever they are doing behind the scenes to make their sentiment analysis [for example] services available to business customers) to service up such things to other businesses including for example the Fortune 500. Sometimes such companies will have tiers of service that are entirely within the range of financial feasibility even for very small businesses (e.g. $49 USD / month). However, in this review, we are more interested in the feasibility of non-tech small businesses “rolling their own” in terms of using someone in-house who is not employed in a technical position to utilize in this case BERT for sentiment analysis.

Sentiment Analysis Applications for Small Business

There are many ways that SMBs can utilize sentiment analysis. For example in [4] we see that *sentiment analysis* or *social sentiment analysis* or *opinion mining* or *emotion AI* [12] finds how people are perceiving a business online (or in general, classifies user sentiment about a subject but our application is sentiment expressed anywhere online about a particular business and the ability of that business to find and use that expressed sentiment en masse to improve its products and services, marketing, and user-sentiment, i.e. its public image). Because we can do sentiment analysis given user sentiment which will be expressed anywhere that we have user generated content [4], it makes sense to enable many channels for user-feedback and user generated content, for example in blogs, on the business’s website, and in its various social media channels; as well as leveling up social media marketing campaigns that allow for user interaction. By interacting with customers, crises can be diverted and re-routed to become a source of business service, product and processes improvement. By interacting with customers, we can find out their wants and needs and find out how we are failing them and how we can improve and work to satisfy them. By leveraging technologies like sentiment analysis we can find out when many people are sharing opinions that may not be exactly what we want, or where we are doing very well. Every area of business that already exists can be improved by finding out what our customers want and need and their opinions on how we are serving and failing those desires and focusing our energies on satisfying them better. By leveraging sentiment analysis technologies our feedback cycle gets much larger, faster, and more informative.

Deep learning applications for small businesses

There are a huge variety of applications of AI (the main driver of which these days is deep learning) for small businesses [14, 15, 16, 17, 18]. However, most such applications are either baked in to software services being utilized by SMBs such as CRM software which can now already have a variety of AI capabilities built into it, or purchased perhaps on a subscription model from a company which mostly focuses on provided AI related services typically to larger companies such as the S&P 500, but happens to have a lower-priced tier that is attainable or even quite reasonable for many small and mid-sized, even very small businesses. However, that is not our focus here but rather we are interested in the feasibility of small businesses using in-house talent that is otherwise preoccupied with non-tech related activities in their day to day job description related activities, as this is a huge swath of the workforce in the USA, which can serve as a proxy but by extension to the First World. Googling this yields nothing. “It looks like there aren't many great matches for your search Tip: Try using words that might appear on the page you’re looking for. For example, ‘cake recipes’ instead of ‘how to make a cake.’ Need help? Check out other tips for searching on Google.” [19].

Feasibility of applying BERT to sentiment analysis for small non-tech businesses

Thanks to the incredible ease of use of Python, PyTorch, and Huggingface’s libraries that work with the Python, PyTorch and TensorFlow ecosystem, such as the HuggingFace transformer library, it is very reasonable for an experienced developer, software engineering student, and really a very broad range of lay people who are hobbyists or otherwise strongly interested in software development, with some related experience, to spin up a working sentiment analysis classifier in just a few hours which leverages the hundreds of hours that Google had to use behind the scenes to train BERT, to say nothing of the hundreds-to-thousands of dollars that it takes to train said model to say nothing of the brilliance it takes to come up with something like that [13]. For example, there are numerous tutorials which will enable anyone with some experience in Python (PyTorch, etc.) to come up to speed with creating a BERT sentiment analysis or other NLP library in a few hours [11, 20, 21, 22]. The author perused any number of these and numerous others. Sometimes invoking a BERT model for something like sentiment analysis can feel like finding out what line of code needs to be written. It’s that simple.

Conclusion

While the pervasive belief seems to be that SMBs need Venture Capital if they are going to implement deep learning for natural language processing applications, the author begs to differ. Getting up to speed with creating simple applications that utilize BERT for sentiment analysis can happen in a few hours or a few days by following free or very reasonably priced online blogs, tutorials, courses, videos, YouTube playlists and books on a variety of platforms. Some basic familiarity with Python, Pytorch, etc. will really help. There are many, many laypeople who are hobbyists with strong interest in programming, machine learning, mathematics, and other technical topics. Small businesses often have access to small data sets which could be used by creating an in-house BERT application for sentiment analysis with many purposes. Training only takes a few hours because you are just doing transfer learning. For awhile computer vision was getting eaten alive by deep learning and was at the cutting edge of deep learning applications in the wild. Around 2018, with the introduction of BERT and friends, NLP began to experience the same.

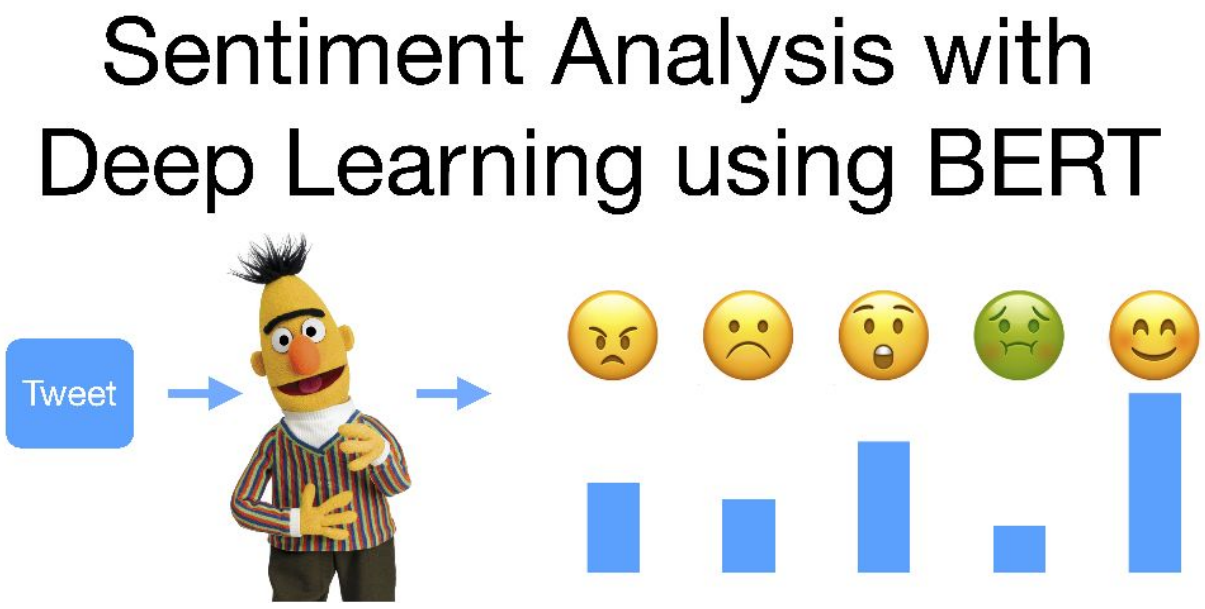

Image taken from [22]

References

[1] **Tatman, R. [Kaggle]. (2019, January 9th).** *Kaggle Reading Group: Bidirectional Encoder Representations from Transformers (aka BERT)* | *Kaggle* [Video]. YouTube. https://www.youtube.com/watch?v=BaPM47hO8p8&list=PLnwa8JsPg8WiUcZVun_XtysTU-flkwm3n&index=1

[2] **Course 1 from [3],** Natural Language Processing with Classification and Vector Spaces, introduction video retrieved from: https://www.coursera.org/learn/classification-vector-spaces-in-nlp/lecture/dDdRc/welcome-to-the-nlp-specialization

[3] deeplearning.ai's Natural Language Processing Specialization **on Coursera** https://www.coursera.org/specializations/natural-language-processing

[4] **Perzynska, K. (2019, September 20). How You Can Use Sentiment Analysis to Improve Your Business**. Retrieved November 14, 2020, from https://www.sotrender.com/blog/2018/10/sentiment-analysis/

[5] **Tatman, R. [Kaggle]. (2018, November 28th).** ***Kaggle Reading Group: Attention is All You Need*** **|** ***Kaggle*** **[Video]. YouTube. https://www.youtube.com/watch?v=54uLU7Nxyv8**

[6] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In Advances in Neural Information Processing Systems, pages 6000–6010. (As cited in [7]). Retrieved from: https://arxiv.org/pdf/1706.03762.pdf on 14 November 2020 at 7:36 p.m. CST.

[7] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. Bert: Pretraining of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018. As cited in Language Models are Unsupervised Multitask Learners by Radford et. al. retrieved from https://arxiv.org/pdf/1810.04805.pdf on 14 November at 7:39 p.m. CST.

[8] Alammar, J. (2018, May 9). Visualizing A Neural Machine Translation Model (Mechanics of Seq2seq Models With Attention). Retrieved November 15, 2020, from

https://jalammar.github.io/visualizing-neural-machine-translation-mechanics-of-seq2seq-models-with-attention/

[9] Ilya Sutskever, Oriol Vinyals, and Quoc VV Le. Sequence to sequence learning with neural networks. In Advances in Neural Information Processing Systems, pages 3104–3112, 2014. (As cited in [6]). Retrieved November 15th at 5:40 p.m. CST from: https://papers.nips.cc/paper/2014/file/a14ac55a4f27472c5d894ec1c3c743d2-Paper.pdf

[10] Cho, K., van Merrienboer, B., Gulcehre, C., Bougares, F., Schwenk, H., and Bengio, Y. (2014a). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Proceedings of the Empiricial Methods in Natural Language Processing (EMNLP 2014). Retrieved from: https://arxiv.org/pdf/1406.1078.pdf November 16th, 2020.

[11] McCormick, C., & Ryan, N. (2019, July 22). BERT Fine-Tuning Tutorial with PyTorch. Retrieved November 15, 2020, from https://mccormickml.com/2019/07/22/BERT-fine-tuning/

[12] https://en.wikipedia.org/wiki/Sentiment_analysis

[13] Synced. (2019, June 27). The Staggering Cost of Training SOTA AI Models. Retrieved November 16, 2020, from https://syncedreview.com/2019/06/27/the-staggering-cost-of-training-sota-ai-models/ (I don't know but probably written by BERT or translated by BERT).

[14] Countants. (2019, December 18). Can Artificial Intelligence and Machine Learning Transform Small Businesses? Retrieved November 16, 2020, from https://medium.com/gobeyond-ai/can-artificial-intelligence-and-machine-learning-transform-small-businesses-4c9d121d9e7f (I don't know but probably written by BERT or translated by BERT).

[15] Otar, C. (2019, April 9). Four Ways Artificial Intelligence Can Help Your Small Business. Retrieved November 16, 2020, from https://www.forbes.com/sites/forbesfinancecouncil/2019/04/09/four-ways-artificial-intelligence-can-help-your-small-business/?sh=5b0246934a1a

[16] Campbell, A. (2018, November 26). How Machine Learning Can Give Small Business a Big Boost. Retrieved November 16, 2020, from https://www.inc.com/comcast/machine-learning-can-give-small-business-a-big-boost.html

[17] Steinmann, L. (2020, May 7). AI and machine learning for small business. Retrieved November 16, 2020, from https://www.microsoft.com/en-us/microsoft-365/business-insights-ideas/resources/ai-and-machine-learning-for-small-business

[18] Mansfield, M. (2018, January 17). 12 Ways Cloud Based Machine Learning Can Boost Small Business Productivity. Retrieved November 16, 2020, from https://smallbiztrends.com/2018/01/machine-learning-can-boost-small-business-productivity.html

[19] Schroeder, E. (2020, November 16). Google Search Results from searching "Feasibility of applying BERT to sentiment analysis for small non-tech businesses" (sans quotes though). Retrieved November 16, 2020, from https://www.google.com/search?sxsrf=ALeKk02ZyG_mXViwgb221IRk1baWt-nrUA%3A1605496205676%2CFeasibility+of+applying+BERT+to+sentiment+analysis+for+small+non-tech+businesses which can be accessed from this bit.ly link I created for ease of use: https://bit.ly/3noL5i3

[20] McCormick, C. (2020, February 26). BERT Research Series. Retrieved November 16, 2020, from https://www.youtube.com/playlist?list=PLam9sigHPGwOBuH4_4fr-XvDbe5uneaf6

[21] Fernandes, G. (2020, July 27). Sentiment analysis with Yelp and Bert (via Passle). Retrieved November 16, 2020, from https://signal.onepointltd.com/post/102gc83/sentiment-analysis-with-yelp-and-bert

[22] Anastassiou, A. (n.d.). Sentiment Analysis with Deep Learning using BERT. Retrieved November 16, 2020, from https://www.coursera.org/projects/sentiment-analysis-bert