REFERENCES

[1]   A. Ritter, S. Clark, Mausam and O. Etzioni, "Named Entity Recognition in Tweets: An Experimental Study".

[2]   Y. B. P. JungHyen An, "Methodology for Automatic Ontology Generation Using Database Schema Information," *Mobile Information Systems,* vol. 2018, p. 13, 2018.

[3]   X. Tao, Z. Tang, C. Xu and Y. Wang, "Logical Labeling of Fixed Layout PDF Documents Using Multiple Contexts," in *11th IAPR International Workshop on Document Analysis Systems*, 2014.

[4]   J. Chung and T. Delteil, "A Computationally Efficient Pipeline Approach to Full Page Offline Handwritten Text Recognition," in *International Conference on Document Analysis and Recognition Workshops (ICDARW)*, 2019.

[5]   W. Li, W. Li and Y. Wu, "A Unified Model for Document-Based Question Answering," in *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*.

[6]   T. Saba, A. S. Almazyad and A. Rehman, "Language Independent Rule Based Classifcation of Printed & Handwritten Text (Classication ofPrinted & Handwritten Text)," in *2015 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, Douai, France, 2015.

[7]   R. Karpinski and A. Bela¨ıd, "Combination of Structural and Factual Descriptors for Document Stream Segmentation," in *12th IAPR Workshop on Document Analysis Systems*, 2016.

[8]   T. T. H. Nguyen, A. Doucet and M. Custaty, "Enhancing Table of Contents Extraction by System Aggregation," in *14th IAPR International Conference on Document Analysis and Recognition*, 2017.

[9]   X. Yang, E. Yumer, P. Asente, M. Kraley, D. Kifer and C. L. Giles, "Learning to Extract Semantic Structure from Documents Using Multimodal Fully Convolutional Neural Networks," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[10]   M. K. Elhadad, K. M. Badran and G. I. Salama, "A Novel Approach for Ontology-based Dimensionality 5eduction for Web Text Document Classification," in *IEEE ICIS 2017*, Wuhan, China, 2017.

[11]   M. Marrero, S. Sánchez-Cuadrado, J. M. Lara and G. Andreadakis, "Evaluation of Named Entity Extraction Systems," in *Advances in Computational Linguistics Research in Computing Science*, 2009, pp. 47-58.

[12]   M. Marrero, J. Urbano, S. Sánchez-Cuadrado, J. Morato and M. Gómez-Berbís, "Named Entity Recognition: Fallacies, Challenges and Opportunities".

[13]   B. W. Locke, "Named entity recognition: adapting to microblogging," in *Computer Science Undergraduate Contributions*, 2009, p. 29.

[14]   A. Toral and R. Munoz, "A proposal to automatically build and maintain gazetteers for Named Entity Recognition by using Wikipedia".

[15]   S. Jayalakshmi and D. Sheshasaayee, "Automated Question Answering System Using Ontology and Semantic Role," in *International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2017)*, 2017.

[16]   V. G., D. Gupta, A. S. and S. Shaji, "A Graph-Based Relation Extraction Method for Question Answering System," 2017.

[17] Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, Toutanova and Kristina, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv preprint arXiv:1810.04805,* 2018.

[18] Peters, Matthew E., Neumann, Mark, Iyyer, Mohit and Gardner, Matt and Clark, Christopher and Lee, Kenton and Zettlemoyer and Luke, "Deep contextualized word representations," in *Proc. of NAACL*, 2018.

[19] "Kaggle," [Online]. Available: https://www.kaggle.com/docs/tpu. [Accessed 20 03 2020].

[20] "TensorFlow Features | Why TensorFlow Is So Popular," [Online]. Available: https://data-flair.training/blogs/tensorflow-features/. [Accessed 20 03 2020].

[21] "Various Uses of TensorFlow," [Online]. Available: https://dzone.com/articles/learn-various-uses-of-tensorflow#:~:text=Other%20major%20TensorFlow%20Applications%20include,Self%20Driving%20Cars. [Accessed 20 04 2020].

[22] "TensorFlow 2.0 is now available!," [Online]. Available: https://blog.tensorflow.org/2019/09/tensorflow-20-is-now-available.html#:~:text=Today%2C%20we're%20delighted%20to,TensorFlow%202.0%20is%20now%20available!&text=TensorFlow%202.0%20provides%20a%20comprehensive,build%20scalable%20ML%2Dpowered%20applications.. [Accessed 01 05 2020].

[23] J. Alammar, "The Illustrated Transformer," [Online]. Available: http://jalammar.github.io/illustrated-transformer/. [Accessed 30 05 2020].

[24] "Transformers (State-of-the-art Natural Language Processing)," [Online]. Available: https://towardsdatascience.com/transformers-state-of-the-art-natural-language-processing-1d84c4c7462b. [Accessed 30 09 2020].

[25] J. Taylor, "ELMo: Contextual language embedding," [Online]. Available: https://towardsdatascience.com/elmo-contextual-language-embedding-335de2268604. [Accessed 04 01 2020].

[26] E. Ma, "ELMo helps to further improve your sentence embeddings," 30 10 2018. [Online]. Available: https://towardsdatascience.com/elmo-helps-to-further-improve-your-word-embeddings-c6ed2c9df95f.

[27] S. Vajpayee, "Understanding BERT — (Bidirectional Encoder Representations from Transformers)," [Online]. Available: https://towardsdatascience.com/understanding-bert-bidirectional-encoder-representations-from-transformers-45ee6cd51eef. [Accessed 23 08 2020].

[28] J. Alammar, "The Illustrated BERT, ELMo, and co. (How NLP Cracked Transfer Learning)," [Online]. Available: http://jalammar.github.io/illustrated-bert/. [Accessed 30 08 2020].

[29] R. Horev, "BERT Explained: State of the art language model for NLP," [Online]. Available: https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270. [Accessed 03 03 2020].

[30] Yiheng Xu, Minghao Li, Lei Cui, Shaohan Huang, Furu Wei and Ming Zhou, *LayoutLM: Pre-training of Text and Layout for Document Image Understanding,* 2019.

[31] J. Walter, *receipts,* kaggle.com.

[32] Guillaume Jaume, Hazim Kemal Ekenel and Jean-Philippe Thiran, "FUNSD: A Dataset for Form Understanding in Noisy Scanned Documents," 2019.

[33] guillaumejaume, "FUNSD: Form Understanding in Noisy Scanned Documents," [Online]. Available: https://guillaumejaume.github.io/FUNSD/. [Accessed 01 07 2020].

[34] Bowman, Samuel R. and Angeli, Gabor and Potts, Christopher, Manning and Christopher D., "A large annotated corpus for learning natural language inference," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2015.

[35] vishakha, *med_train.*

[36] A. Walia, *entity-annotated-corpus.*

[37] Jawahar, Minesh Mathew, Dimosthenis Karatzas and R. Manmatha, *{DocVQA}: A Dataset for VQA on Document Images,* 2020.

[38] "Overview - Document Visual Question Answering," [Online]. Available: https://rrc.cvc.uab.es/?ch=17.

[39] P. Rajpurkar, J. Zhang and K. Lopyrev, " SQuAD: 100,000+ questions for machine comprehension of text," in *EMNLP*, 2016.

[40] "CVPR2020 Workshop on Text and Documents in the Deep Learning Era," [Online]. Available: https://cvpr2020text.wordpress.com/.

[41] "google-research/bert," [Online]. Available: https://github.com/google-research/bert.

[42] t. f. e. Wikipedia, "F-score," [Online]. Available: https://en.wikipedia.org/wiki/F-score.

[43] "Research Group on Computer Vision and Artificial Intelligence," [Online]. Available: http://www.fki.inf.unibe.ch/databases.

[44] "neuromorphic computing," [Online]. Available: https://searchenterpriseai.techtarget.com/definition/neuromorphic-computing. [Accessed 06 09 2020].

[45] N. K. M. &. S. M. Kumar, "Automated ontology generation from a plain text using statistical and NLP techniques," *Int J Syst Assur Eng Manag,* vol. 7, p. 282–293 , 2016.

[46] D. S. Jeong, "Neuromorphic spiking neural networks for temporal learning," Hanyang University, Seoul.

[47] J. Devlin, M.-w. Chang, K. Lee and K. Toutanova, *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* 2019.

[48] A. I. F. AI, *COVID-19 Open Research Dataset Challenge (CORD-19).*

[49] L.L. Wang, K. Lo, Y. Chandrasekhar, R. Reas, J. Yang, D. Eide, K. Funk, R.M. Kinney, Z. Liu, W. Merrill, P. Mooney, D. Murdick, D. Rishi, J. Sheehan, Z. Shen, B. Stilson, A. Wade, K. Wang, C. Wilhelm, B. Xie, D. Raymond and D. S. Weld, "CORD-19: The Covid-19 Open Research Dataset," *ArXiv,* 2020.