References

[1] 'Data Never Sleeps 9.0 | Domo'. [Online]. Available: https://www.domo.com/learn/infographic/data-never-sleeps-9. [Accessed: 12-Oct-2021]

[2] Steven Bird, Ewan Klein, and Edward Loper, *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. 2009.

[3] Sumant Sharma and Amit Arora, 'Adaptive approach for spam detection', *Int. J. Comput. Sci. Issues IJCSI*, vol. 10, no. 4, p. 23, 2013.

[4] Wei Xue and Tao Li, 'Aspect Based Sentiment Analysis with Gated Convolutional Networks', *ArXiv180507043 Cs*, May 2018 [Online]. Available: http://arxiv.org/abs/1805.07043. [Accessed: 15-Jun-2021]

[5] Martin Malmsten, Love Börjeson, and Chris Haffenden, 'Playing with Words at the National Library of Sweden -- Making a Swedish BERT', *ArXiv200701658 Cs*, Jul. 2020 [Online]. Available: http://arxiv.org/abs/2007.01658. [Accessed: 04-Jun-2021]

[6] Matthias Bruckner, Marcelo LaFleur, and Ingo Pitterle, 'The impact of the technological revolution on labour markets and income distribution', *Front. Issues*, p. 51, Jul. 2017.

[7] Emma Strubell, Ananya Ganesh, and Andrew McCallum, 'Energy and Policy Considerations for Deep Learning in NLP', *ArXiv190602243 Cs*, Jun. 2019 [Online]. Available: http://arxiv.org/abs/1906.02243. [Accessed: 14-Jun-2021]

[8] Alhanoof Althnian, Duaa AlSaeed, Heyam Al-Baity, Amani Samha, Alanoud Bin Dris, Najla Alzakari, Afnan Abou Elwafa, and Heba Kurdi, 'Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain', *Appl. Sci.*, vol. 11, no. 2, p. 796, Jan. 2021. DOI: 10.3390/app11020796

[9] 'Natural language processing', *Wikipedia*. 21-May-2021 [Online]. Available: https://en.wikipedia.org/w/index.php?title=Natural_language_processing&oldid=1024275775. [Accessed: 03-Jun-2021]

[10] 'Stack Overflow Developer Survey 2020', *Stack Overflow*. [Online]. Available: https://insights.stackoverflow.com/survey/2020/?utm_source=social-share&utm_medium=social&utm_campaign=dev-survey-2020. [Accessed: 17-Jun-2021]

[11] 'Hugging Face – The AI community building the future.' [Online]. Available: https://huggingface.co/. [Accessed: 15-Jun-2021]

[12] 'TensorFlow'. [Online]. Available: https://www.tensorflow.org/. [Accessed: 17-Jun-2021]

[13] 'Quick Start - Spark NLP'. [Online]. Available: https://nlp.johnsnowlabs.com/docs/en/quickstart. [Accessed: 03-Jun-2021]

[14] 'Apache Spark™ - Unified Analytics Engine for Big Data'. [Online]. Available: https://spark.apache.org/. [Accessed: 03-Jun-2021]

[15] 'Spark SQL and DataFrames - Spark 3.1.2 Documentation'. [Online]. Available: https://spark.apache.org/docs/latest/sql-programming-guide.html#datasets-and-dataframes. [Accessed: 03-Jun-2021]

[16] 'ML Pipelines - Spark 3.1.2 Documentation'. [Online]. Available: https://spark.apache.org/docs/latest/ml-pipeline.html#transformers. [Accessed: 03-Jun-2021]

[17] Dan Jurafsky and James H. Martin, *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, N.J: Prentice Hall, 2000, Prentice Hall series in artificial intelligence, ISBN: 978-0-13-095069-7.

[18] 'Linguistics', *Wikipedia*. 13-Jun-2021 [Online]. Available: https://en.wikipedia.org/w/index.php?title=Linguistics&oldid=1028432773. [Accessed: 15-Jun-2021]

[19] J.R. Firth, *A Synopsis of Linguistic Theory, 1930-1955*. 1957 [Online]. Available: https://books.google.se/books?id=T8LDtgAACAAJ

[20] Jure Zupan, 'Introduction to Artificial Neural Network (ANN) Methods: What They Are and How to Use Them', *Acta Chim. Slov.*, vol. 41, Jan. 1994.

[21] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding', *ArXiv181004805 Cs*, May 2019 [Online]. Available: http://arxiv.org/abs/1810.04805. [Accessed: 03-Jun-2021]

[22] Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang, 'Language-agnostic BERT Sentence Embedding', *ArXiv200701852 Cs*, Jul. 2020 [Online]. Available: http://arxiv.org/abs/2007.01852. [Accessed: 01-Jun-2021]

[23] Emilio Soria Olivas, Jose David Martin Guerrero, Marcelino Martinez Sober, Jose Rafael Magdalena Benedito, and Antonio Jose Serrano Lopez, *Handbook Of Research On Machine*

Learning Applications and Trends: Algorithms, Methods and Techniques, Illustrated edition. Hershey, PA: Information Science Reference, 2009, ISBN: 978-1-60566-766-9.

[24]'Amazon Machine Learning - Developer Guide', *Amaz. Web Serv.*, vol. 2016, p. 146.

[25]Jason Brownlee, 'Difference Between a Batch and an Epoch in a Neural Network', *Machine Learning Mastery*. 19-Jul-2018 [Online]. Available: https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/. [Accessed: 03-Jun-2021]

[26]Jason Brownlee, 'How to Configure the Learning Rate When Training Deep Learning Neural Networks', *Machine Learning Mastery*. 22-Jan-2019 [Online]. Available: https://machinelearningmastery.com/learning-rate-for-deep-learning-neural-networks/. [Accessed: 03-Jun-2021]

[27]Jason Brownlee, 'Dropout Regularization in Deep Learning Models With Keras', *Machine Learning Mastery*. 19-Jun-2016 [Online]. Available: https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/. [Accessed: 03-Jun-2021]

[28]Shanshan Yu, Jindian Su, and Da Luo, 'Improving BERT-Based Text Classification With Auxiliary Sentence and Domain Knowledge', *IEEE Access*, vol. 7, pp. 176600–176612, 2019. DOI: 10.1109/ACCESS.2019.2953990

[29]'Evaluation Metrics - RDD-based API - Spark 3.1.1 Documentation'. [Online]. Available: https://spark.apache.org/docs/3.1.1/mllib-evaluation-metrics.html. [Accessed: 07-Jun-2021]

[30]'sklearn.metrics.classification_report — scikit-learn 0.24.2 documentation'. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.classification_report.html. [Accessed: 07-Jun-2021]

[31]Andrew K. Shenton, 'Strategies for ensuring trustworthiness in qualitative research projects', *Educ. Inf.*, vol. 22, no. 2, pp. 63–75, 2004. DOI: 10.3233/EFI-2004-22201

[32]'KB/bert-base-swedish-cased · Hugging Face'. [Online]. Available: https://huggingface.co/KB/bert-base-swedish-cased. [Accessed: 07-Jun-2021]

[33]'Hugging Face: State-of-the-Art Natural Language Processing in ten lines of TensorFlow 2.0'. [Online]. Available: https://blog.tensorflow.org/2019/11/hugging-face-state-of-art-natural.html. [Accessed: 15-Sep-2021]

[34]'BERT Fine-Tuning Tutorial with PyTorch · Chris McCormick'. [Online]. Available: http://mccormickml.com/2019/07/22/BERT-fine-tuning/#41-bertforsequenceclassification. [Accessed: 15-Sep-2021]

[35]'Fine-tuning with custom datasets'. [Online]. Available: https://huggingface.co/transformers/custom_datasets.html. [Accessed: 15-Sep-2021]

[36]'marma/bert-base-swedish-cased-sentiment · Hugging Face'. [Online]. Available: https://huggingface.co/marma/bert-base-swedish-cased-sentiment. [Accessed: 07-Jun-2021]

[37]'Who is using scikit-learn? — scikit-learn 0.24.2 documentation'. [Online]. Available: https://scikit-learn.org/stable/testimonials/testimonials.html. [Accessed: 15-Sep-2021]