Applications of deep learning: understanding text

This includes problems such as: classifying the sentiment of a paragraph, correcting grammar, classifying sentences, translating text from one language to another.

3a. Translation with sequence-to-sequence models, transformer
Tutorial: Graham Neubig. Neural machine translation and sequence-to-sequence models: A tutorial. arXiv preprint arXiv:1703.01619 (2017). [PDF]
Citation: Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. Attention Is All You Need. In Advances in Neural Information Processing Systems (NIPS '17). [PDF]
[Notebook: seq2seq][Notebook: translation with a Transformer]

3b. Sentence classification with BERT
Citation: Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL '19). [PDF] [Blog post][Video]
[Illustrated guide to BERT: 1, 2][Notebook: Sentiment classification with BERT]
[Sentence classification (grammatical or not?): Blog post, video, and notebook]
[Notebook: Multi-label classification with BERT] [DistilBERT classification]
[Notebook: offensive language classification]

3c. Domain-specific BERT models
Citation: Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL '19). [PDF] [Blog post][Video]
Citation: Iz Beltagy, Kyle Lo, Arman Cohan. SciBERT: A Pretrained Language Model for Scientific Text. In Proceedings of the Conference on Empirical Methods in Natural Language Processing & International Joint Conference on Natural Language Processing (EMNLP-IJCNLP '19) [PDF] [Github]
[Information retrieval notebook] [Github - biomedical event extraction]
[Sentence similarity: Blog post, notebook] [Illustrated guide to BERT: 1, 2]

3d. Hierarchical attention networks for classification
Citation: Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, Eduard Hovy. Hierarchical Attention Networks for Document Classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL '16). [PDF] [Github (unofficial) with notebook]
Citation: Jader Abreu, Luis Fred, David Macêdo, Cleber Zanchettin. Hierarchical Attentional Hybrid Neural Networks for Document Classification. In Proceedings of the 2019 International Conference on Artificial Neural Networks (ICANN '19). [PDF] [Github (official) with notebooks]

3e. Convolutional neural networks for text classification
Citation: Yoon Kim. Convolutional Neural Networks for Sentence Classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP '14). [PDF] [Notebook (via PyTorch)][Notebook (via D2L-AI)]

3f. Zero-shot text classification
Citation: See references in the blog post. The approach in the notebook is via: Wenpeng Yin, Jamaal Hay, Dan Roth. Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP '19). [PDF] [Blog post] [Notebook][Another approach in this notebook]

3g. Text summarization
Citation (SOTA): Jingqing Zhang, Yao Zhao, Mohammad Saleh, Peter J. Liu. PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization. In Proceedings of the 37th International Conference on Machine Learning (ICML '20). [PDF] [Github (Pegasus)][Article - use Pegasus in Colab] [Blog post] [Github: text summarization with Transformer from scratch] [Other (pre-trained) models: Notebook - T5, Notebook - BART, Notebook - blurr, Github with notebook - BERT and BERT variants]

3h. Word embeddings
Citation(GloVe): Jeffrey Pennington, Richard Socher, Christopher D. Manning. GloVe: Global Vectors for Word Representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP '14). [PDF][Website][Notebook - training word2vec, GloVe, fastText] [Notebook - similarity for word2vec, GloVe, fastText]
Citation (swivel): Noam Shazeer, Ryan Doherty, Colin Evans, Chris Waterson. Swivel: Improving Embeddings by Noticing What's Missing. arXiv:1602.02215 (2016). [PDF][Notebook - swivel][Overview of word embeddings]

3i. Sentence transformers
Citation: Nils Reimers, Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP '19). [PDF] [Github (official)][Notebook (unofficial)]

3j. Named entity recognition
Citation (Flair NER): Alan Akbik, Tanja Bergmann, Roland Vollgraf. Pooled Contextualized Embeddings for Named Entity Recognition. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL '19). [PDF][Video][Flair - notebook, tutorial]
Citation (GRU with GloVe): Matthew E. Peters, Waleed Ammar, Chandra Bhagavatula, Russell Power. Semi-supervised sequence tagging with bidirectional language models. In Proceedings of the 2017 Annual Meeting of the Association for Computational Linguistics (ACL '17). [PDF][Demo - click on "Usage" for code] [You can also try: Spacy. It's in this notebook and here's a video about its approach]

3k. Named entity recognition on Stack Overflow
Citation: Jeniya Tabassum, Mounica Maddela, Wei Xu, Alan Ritter. Code and Named Entity Recognition in StackOverflow. In Proceedings of the 2020 Annual Meeting of the Association for Computational Linguistics (ACL '20). [PDF][Github (Official)][Video][In HuggingFace]

3l. Named entity recognition for biomedical texts
Citation: Leon Weber, Mario Sanger, Jannes Munchmeyer, Maryam Habibi, Ulf Leser, Alan Akbik. HunFlair: An Easy-to-Use Tool for State-of-the-Art Biomedical Named Entity Recognition. (2020). [PDF coming soon?][Github, and HunFlair page] [Github for reproducing experiments in the paper]