|  |  |
| --- | --- |
| Papers | Dates |
| 1. [Efficient Estimation of Word Representations in Vector Space](https://arxiv.org/pdf/1301.3781.pdf) 2. [Distributed Representations of Words and Phrases and their Compositionality](https://arxiv.org/pdf/1310.4546.pdf) 3. [GloVe: Global Vectors for Word Representation](https://nlp.stanford.edu/pubs/glove.pdf) 4. [Enriching Word Vectors with Subword Information](https://arxiv.org/pdf/1607.04606.pdf) | August 8, 10 |
| 1. [Extensions of Recurrent Neural Network Language Model](http://www.fit.vutbr.cz/research/groups/speech/publi/2011/mikolov_icassp2011_5528.pdf) 2. [Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation](https://arxiv.org/pdf/1406.1078v3.pdf) 3. [Sequence to Sequence Learning with Neural Networks](https://arxiv.org/pdf/1409.3215.pdf) 4. [Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling](https://arxiv.org/pdf/1412.3555.pdf) 5. [Convolutional Neural Networks for Sentence Classification](https://aclanthology.org/D14-1181.pdf) 6. [Convolutional, Long Short-Term Memory, Fully Connected Deep Neural Networks](https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/43455.pdf) | August 22, 24, 29  September 5 |
| 11. [An overview of Gradient Descent Optimization Algorithms](https://arxiv.org/pdf/1609.04747.pdf)  12. [Overview of mini-batch Gradient Descent](http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf)  13. [Understanding the difficulty of Training Deep Feedforward Neural Network](https://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf)s | September 12, 14 |