Name of the article	Nnumber of citations	Sorting of number of citations
A convolutional neural network for modelling sentences	1585	49134
A Learning Algorithm for Continually Running Fully Recurrent Neural Networks	3156	30965
A neural probabilistic language model	3968	29946
A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts	2931	25527
A survey of robot learning from demonstration	1935	24408
A Theoretical Analysis of Feature Pooling in Visual Recognition.	740	19556
Adaptive subgradient methods for online learning and stochastic optimization	3884	18992
Aligning books and movies: Towards story-like visual explanations by watching movies and reading books.	**242**	17257
An Analysis of Single-Layer Networks in Unsupervised Feature Learning.	1471	14937
Annotating expressions of opinions and emotions in language	1432	13602
Asynchronous methods for deep reinforcement learning	1197	11673
Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.	8464	10847
Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories	8969	9681
Beyond Short Snippets: Deep Networks for Video Classification.	866	9510
Building high-level features using large scale unsupervised learning.	1731	8969
Character-level Convolutional Networks for Text Classification	847	8464
Conditional Random Fields as Recurrent Neural Networks	1310	8334
Continuous control with deep reinforcement learning.	1139	8297
Convolutional neural networks for sentence classification	3181	7840
cuDNN: Efficient Primitives for Deep Learning	559	7380
Deep learning	10847	7071
Deep Learning Face Attributes in the Wild.	996	6659
Deep learning in neural networks: An overview	4168	6063
Deep Neural Networks for Acoustic Modeling in Speech Recognition:The Shared Views of Four Research Groups	4837	6009
Deep Residual Learning for Image Recognition.	17257	5774
Deep Sparse Rectifier Neural Networks.	2464	5561
Deep visual-semantic alignments for generating image descriptions	1989	5360
Devise: A deep visual-semantic embedding model	959	4837
Distinctive Image Features from Scale-Invariant Keypoints	49134	4418
Distributed representations of sentences and documents	3009	4168
Efficient estimation of word representations in vector space	8334	4065
Empirical evaluation of gated recurrent neural networks on sequence modeling	2112	4061
Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.	6063	3968
Finding structure in time	9681	3968
Framewise phoneme classification with bidirectional LSTM and other neural network architectures	1193	3884
Fully convolutional networks for semantic segmentation	8297	3749
Generating sequences with recurrent neural networks.,	1454	3664
Generative Adversarial Nets	7380	3551
Going deeper with convolutions	11673	3520

Gradient-Based Learning Applied to Document Recognition	14937	3246
Handwritten digit recognition with a back-propagation network	2104	3225
Histograms of Oriented Gradients for Human Detection	24408	3181
Hogwild: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent.	1239	3156
Human-level control through deep reinforcement learning	4065	3009
Image Super-Resolution Using Deep Convolutional Networks.	1243	2931
Imagenet classification with deep convolutional neural networks	30965	2814
Improved semantic representations from tree-structured long short-term memory network	1052	2684
Improving neural networks by preventing co-adaptation of feature detectors	3749	2540
Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.	1379	2500
Independent component analysis	7840	2464
Knowledge matters: Importance of prior information for optimization	**81**	2198
Large scale distributed deep networks	1659	2112
Large-Scale Video Classification with Convolutional Neural Networks.	2814	2104
Layer Normalization.	645	1989
Learning from Delayed Rewards	6009	1976
Learning generative visual models from few training examples an incremental Bayesian approach tested on 101 object categories	3246	1935
Learning hierarchical features for scene labeling.	1888	1888
Learning Internal Representations by Error Propagation	25527	1731
Learning mid-level features for recognition	1113	1659
Learning mid-level features for recognition	1113	1646
Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation	4061	1632
Learning with kernels	4418	1614
Linear spatial pyramid matching using sparse coding for image classification.	3520	1585
Long Short-Term Memory	13602	1471
Long-term recurrent convolutional networks for visual recognition and description.	2198	1463
Massively Parallel Methods for Deep Reinforcement Learning.	164	1454
MatConvNet: Convolutional Neural Networks for MATLAB.	1976	1432
Microsoft coco: Common objects in contex	3551	1379
Mining and summarizing customer reviews	5774	1332
MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems.	575	1310
Natural Gradient Works Efficiently in Learning	3664	1271
Natural Language Processing (Almost) from Scratch	3968	1243
Neural Architectures for Named Entity Recognition.	715	1239
Neural networks and physical systems with emergent collective computational abilities	19556	1197
Object recognition with features inspired by visual cortex	1130	1193
On the difficulty of training Recurrent Neural Networks	1632	1139
On the properties of neural machine translation: Encoder-decoder approaches	1271	1130
Playing atari with deep reinforcement learning	1646	1113
Project Adam: Building an Efficient and Scalable Deep Learning Training System.	371	1113

Recursive deep models for semantic compositionality over a sentiment treebank,	2540	1052
Reinforcement Learning: An Introduction.	29946	996
Salient Object Detection: A Discriminative Regional Feature Integration Approach.	669	959
Sample size selection in optimization methods for machine learning.	155	866
Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales	1614	847
Sequence to sequence learning with neural networks	5561	823
Simple statistical gradient-following algorithms for connectionist reinforcement learning	2500	792
Skip-Thought Vectors.	792	740
Social Attention: Modeling Attention in Human Crowds	16	715
Teaching Machines to Read and Comprehend.	823	669
TensorFlow: A system for large-scale machine learning.	7071	645
TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems	5360	575
Theano: A Python framework for fast computation of mathematical expressions.	1332	559
U-Net: Convolutional Networks for Biomedical Image Segmentation.	3225	371
Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.	2684	337
Very deep convolutional networks for large-scale image recognition	18992	242
Video Google: A Text Retrieval Approach to Object Matching in Videos.	6659	164
Visual madlibs: Fill in the blank description generation and question answering.,	**66**	155
Visualizing high-dimensional data using t-sne	9510	81
Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks.	337	66
What is the best multi-stage architecture for object recognition?	1463	16