RESEARCH STAY WEEK 5, NEURAL NETWORKS

Daniel Cajas A01708637

CONTEXT

Artificial neural networks, a.k.a. ANNs are models that resemble the human brain. They have nodes that connect with each other, like neurons and synapses, hence the name. There have been early attempts at ANNs implementations since the 70s but they didn't become mainstream until the computational power needed for running them became widely available, in the 2010s.

SEARCH METHODOLOGY

I already had a background as to how ANNs work so i wanted to see their application to sentiment analysis. So I looked for ANNs application for sentiment analysis in google scholar.

Preliminary terms

- ANN: Artificial Neural Network
- CNN: Convolutional Neural Network
- RNN: Recurrent Neural Network

COMPARISON

I found two interesting approaches to combining different neural network architectures with the goal of sentiment analysis.

The first one by Xingyou Wang, et al. (2016) [2] used a CNN for feature extraction and then a RNN as a classifier. Here the CNN could be seen as a sort of n-gram, as it creates a new, intermediate feature representation that is the used by the RNN.

In contrast, H. Sadr, et al. (2020) [3] used separate CNN and RNN to extract features from the corpus and then fed through different classifiers (KCCA, SVM2K, MVMED, SMVMED)

I chose these specifically because they used the same datasets so we are able to do a direct comparison in their performance.

COMPARISON

Group	Model	MR	SST1	SST2
Other	NB(Socher et al., 2013b)	–	41.0	81.8
	SVM(Socher et al., 2013b)	–	40.7	79.4
CNN	1-layer convolution(Kalchbrenner et al., 2014)	–	37.4	77.1
	Deep CNN(Kalchbrenner et al., 2014)	–	48.5	86.8
	Non-static(Kim, 2014)	81.5	48.0	87.2
	Multichannel(Kim, 2014)	81.1	47.4	88.1
Recursive	Basic(Socher et al., 2013b)	–	43.2	82.4
	Matrix-vector (Socher et al., 2013b)	–	44.4	82.9
	Tensor (Socher et al., 2013b)	–	45.7	85.4
	Tree LSTM1 (Zhu et al., 2015)	–	48.0	-
	Tree LSTM2 (Tai et al., 2015)	–	51.0	88.0
	Tree LSTM3 (Le and Zuidema, 2015)	–	49.9	88.0
	Tree bi-LSTM (Li et al., 2015)	0.79	–	–
Recurrent	LSTM(Tai et al., 2015)	–	46.4	84.9
	bi-LSTM(Tai et al., 2015)	–	49.1	87.5
Vector	Word vector avg(Socher et al., 2013b)	–	32.7	80.1
	Paragraph vector(Le and Mikolov, 2014)	–	48.7	87.8
TBCNNs	c-TBCNN(Mou et al., 2015)	–	50.4	86.8
	d-TBCNN(Mou et al., 2015)	–	51.4	87.9
CNN-RNN	CNN-GRU-word2vec	**82.28**	50.68	**89.95**
	CNN-LSTM-word2vec	81.52	**51.50**	89.56
	AVG-GRU-word2vec	81.44	50.36	89.61
	CNN-GRU-rand	76.34	48.27	86.64
	CNN-LSTM-rand	77.04	49.50	86.80

Model	SST1		SST2	
	Random initialized word vectors	Pre-trained word vectors	Random initialize d word vectors	Pre-trained word vectors
CNN-RNN+KCCA	47.62	49.42	86.05	88.73
CNN-RNN+SVM2K	46.54	48.3	84.35	87.45
CNN-RNN+MVMED	50.25	52.09	89.12	90.32
CNN-RNN+SMVMED	50.89	52.94	90.25	91.93

TABLE 8. Percentage accuracies of all variations of the proposed model

Here we see how using both CNNs and RNNs for feature extraction is better in both datasets with pretrained and randomly initialized word vectors.

This shows us how using RNNS for feature extraction might be better than for classification when doing sentiment analysis.

BIBLIOGRAFÍA

[1] A.D.Dongare, R.R.Kharde, Amit D.Kachare (2012). Introduction to Artificial Neural Network. IJEIT Vol 2 Issue 1.https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=04d0b6952a4f0c7203577afc9476c2fcab2cba06

[2] Xingyou Wang, Weijie Jiang, and Zhiyong Luo. 2016. Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 2428–2437, Osaka, Japan. The COLING 2016 Organizing Committee.

[3] H. Sadr, M. M. Pedram and M. Teshnehlab, "Multi-View Deep Network: A Deep Model Based on Learning Features From Heterogeneous Neural Networks for Sentiment Analysis," in IEEE Access, vol. 8, pp. 86984-86997, 2020, doi: 10.1109/ACCESS.2020.2992063.
keywords: {Feature extraction;Sentiment analysis;Deep learning;Data mining;Recurrent neural networks;Convolutional neural networks;Deep learning;multi-view learning;convolutional neural network;recursive neural network;sentiment analysis}