Joint Verification-Identification in End-to-End Multi-Scale CNN Framework for Topic Identification

Raghavendra Pappagari, Jesús Villalba, Najim Dehak

Center for Language and Speech Processing
Department of Electrical and Computer Engineering, The Johns Hopkins University, USA
{rpappag1, jvillal7, ndehak3}@jhu.edu

Objective

- Identifying topic of a text document
- Joint optimization of verification and identification tasks

Motivation

- Verification task and identification task are related tasks but use different kind of supervision
- Optimizing both tasks in end-to-end manner improve identification performance?

End-to-End Topic Identification

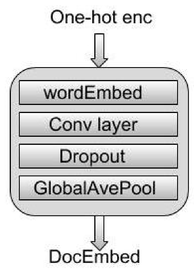

Figure 1: Base Module(BM)

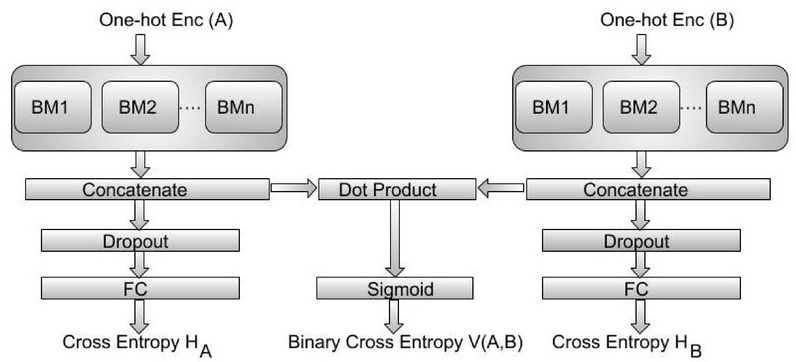

Figure 2: Architecture

$$Obj = \sum_{\mathrm{A}} \sum_{\mathrm{B} \neq \mathrm{A}} H_{\mathrm{A}} + H_{\mathrm{B}} + \lambda V(\mathrm{A}, \mathrm{B})$$

Results

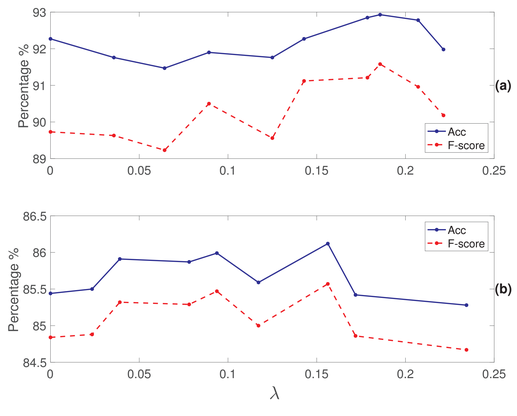

Figure 3: Accuracy and F-score plots on (a) Fisher and (b) 20 newsgroups for different values of λ

dataset	Model	Accuracy	F-measure
20 newsgroups	NTSG-1	82.6	81.2
	SCDV	84.6	84.6
	CNN $\lambda = 0$	85.44	84.84
	CNN $\lambda = 0.16$	**86.12**	**85.57**
Fisher	SVM MCE	91.9	-
	CNN $\lambda = 0$	92.27	89.73
	CNN $\lambda = 0.18$	**92.93**	**91.21**

Table 1: Accuracy and F-score on 20newsgroups and Fisher

- 20Newsgroups - 0.73% improvement
- Fisher - 1.48% improvement

t-Distributed Stochastic Neighbor

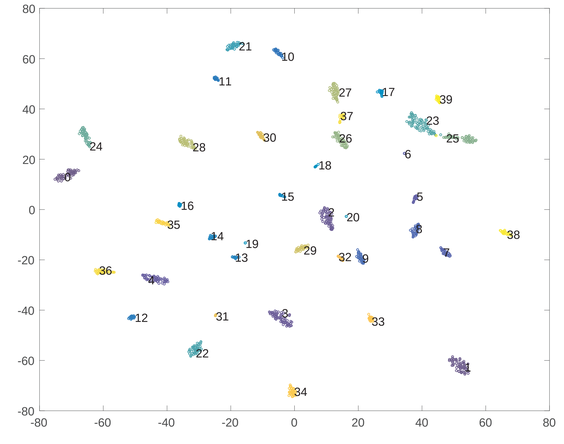

Figure 4: Visualization of document embeddings on Fisher dataset using t-SNE

Experimental Setup

- 20Newsgroups - Written text, 20 topics
- Fisher - Spoken Text, 40 topics
- Number of BMs used is 8
- Kernel sizes for each BM are 1,4,7,10,13,16,19,22

Mean square cosine similarity

$$\mathrm{MSCS} = \sqrt{\frac{2}{M(M-1)} \sum_{\mathrm{A,B}} \left(\frac{1 + \cos(\mathrm{A}, \mathrm{B})}{2} \right)^2}$$

- MSCS_W - Within topic similarity - larger is better
- MSCS_B - between topic similarity - smaller is better

Table 2: MSCS analysis on document embeddings

	λ	MSCS_W	MSCS_B
20 newsgroups	0	1.00	0.69
	0.16	**1.00**	**0.58**
Fisher	0	0.95	0.63
	0.18	**0.96**	**0.55**

Conclusion

- End-to-End model is better than two-stage models
- Joint verification-identification improves identification performance
- State-of-the-art results on Fisher and 20Newsgroups datasets