

Copyright Undertaking

This work is protected by copyright, with all rights reserved.

By reading and using the work, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the work.
2. The reader will use the work for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this work are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

A Sentiment Analysis Method to Better Utilize User Profile and Product Information

Mingyu MA (Derek) supervised by Prof. Qin Lu derek.ma@



derek.ma



THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學



Department of Computing
電子計算學系

Abstract

We proposed a new Joint User and Product Memory Network (JUPMN) utilizing user profile and product information in separate ways into sentiment classification. Inspired by the successful utilization of memory network, our model first creates document representations using hierarchical LSTM model and then feeds the document vectors into new carefully designed user and product memory networks to reflect corresponding features. The evaluation of JUPMN on three benchmark review datasets shows that JUPMN outperforms the state-of-the-art model and further analysis of experimental results is employed.

Introduction

Task: Sentiment Analysis.

Input: a paragraph of reviewing text

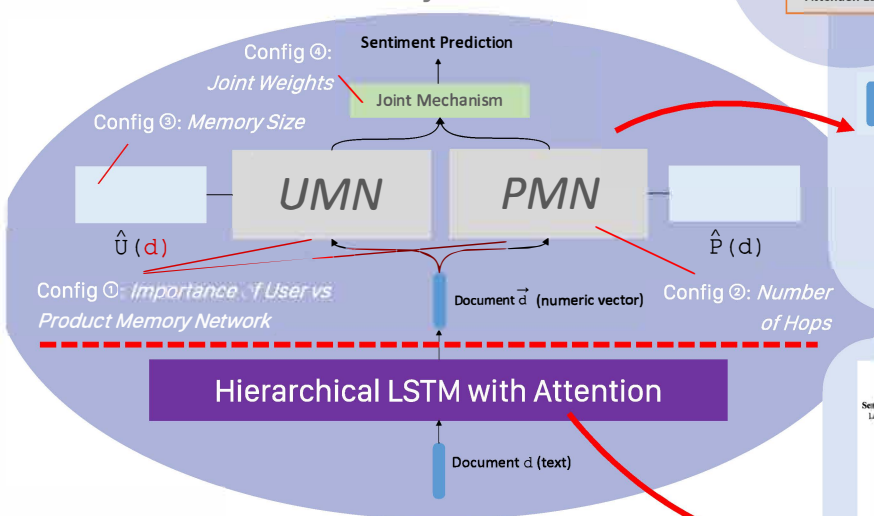
Output: ratings/sentiment



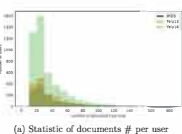
Can consider user profile and product information, while they are fundamentally different. We should not consider them as single united representation

Model

JUPMN Joint User and Product Memory Network



Evaluation



(a) Statistic of documents # per user

Benchmark datasets: IMDB, Yelp13 and Yelp14



Left: document number statistics

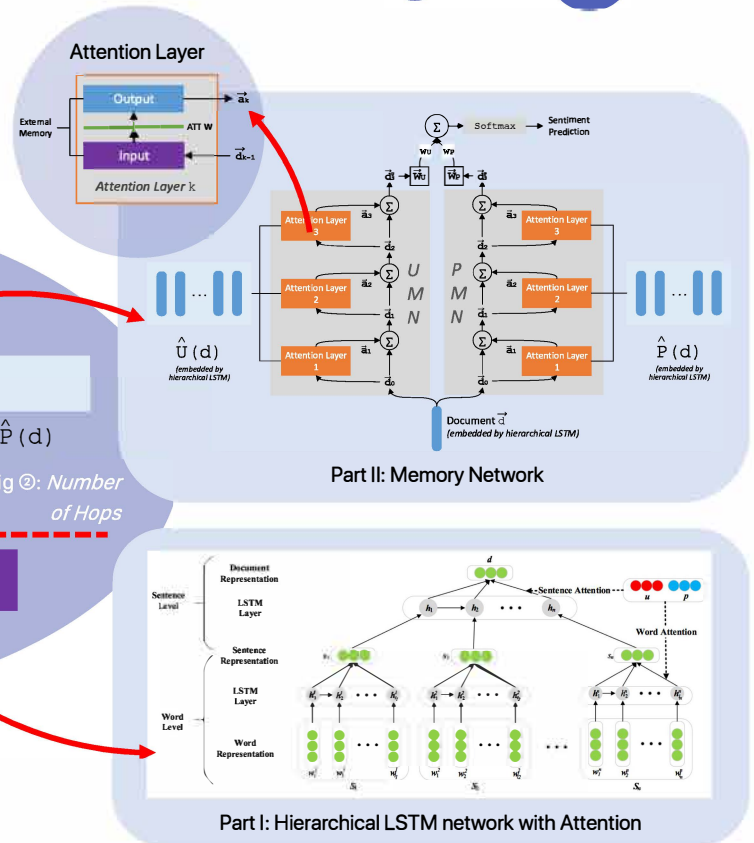
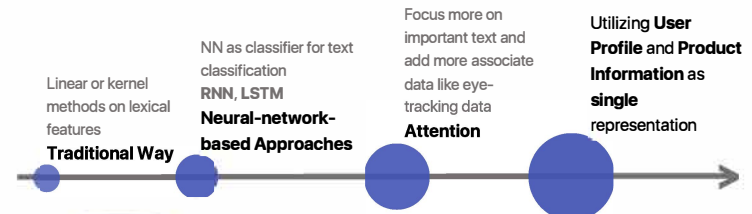
JUPMN vs Comparison Models

Model	IMDB			Yelp13			Yelp14		
	Acc	RMSE	MAE	Acc	RMSE	MAE	Acc	RMSE	MAE
Majority	0.196	2.495	1.838	0.392	1.097	0.779	0.411	1.06	0.744
Trigram	0.399	1.783	1.147	0.577	0.804	0.487	0.569	0.814	0.513
TextFeature	0.402	1.793	1.134	0.572	0.800	0.490	0.556	0.845	0.520
ArgWordvec	0.304	1.985	1.361	0.530	0.893	0.562	0.526	0.898	0.568
SSWE	0.312	1.973	N/A	0.549	0.849	N/A	0.557	0.851	N/A
RNN+RNN	0.400	1.734	N/A	0.574	0.804	N/A	0.582	0.821	N/A
CLSTM	0.421	1.549	N/A	0.592	0.729	N/A	0.637	0.686	N/A
LSTM+LA	0.443	1.465	N/A	0.627	0.701	N/A	0.637	0.686	N/A
LSTM+CBA	0.489	1.365	N/A	0.638	0.697	N/A	0.641	0.678	N/A
UPMN(K)	0.435	1.602	0.979	0.608	0.764	0.447	0.596	0.784	0.464
UPMN(K)	0.465	1.351	0.853	0.613	0.720	0.425	0.639	0.662	0.369
InterSub	0.476	1.392	N/A	0.623	0.714	N/A	0.635	0.690	N/A
LSTM+UPA	0.533	1.281	N/A	0.650	0.692	N/A	0.667	0.654	N/A
JUPMN	0.539	1.283	0.725	0.662	0.667	0.375	0.676	0.641	0.351

Outperforms the state-of-the-art model

Best accuracy, RMSE and MAE

Related Works



Config ①: Importance of User vs Product Memory Network

- User profile influences sentiments of **movie reviews** more
- Product information influences sentiments of **restaurants reviews** more

IMDB		Yelp13		Yelp14	
w_U	w_P	w_U	w_P	w_U	w_P
0.534	0.466	0.475	0.525	0.436	0.564

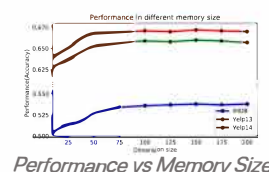
Average joint weight for three datasets

Config ②: Number of Hops

- Smaller hop works better

Config ③: Memory Size

- Larger memory helps until 75



JUPMN with Different Configurations



Config ④: Joint Weights

- Weighted version works better
- Weight help to balance the influences of UMN and PMN