**Deep Memory Networks for Attitude Identification**

1. INTRODUCTION

The problem can be solved with an end-to-end machine learning architecture, where the two subtasks are interleaved by a deep memory network.

2. RELATED WORK

**Explicitly tagged targets**

Aspects are given as inputs, assuming that they have already been annotated in the text.

**Given target, one per instance**

As stance classification deals with only one given target per instance, it fails to consider the interaction between target detection and sentiment classification. Furthermore, the interplay among targets is ignored by training a separate model per target.

**Explicit targets, not tagged**

**Implicit targets**

**Deep learning for sentiment analysis**

Compared to the existing approaches, our work develops a novel deep learning architecture that emphasizes the interplay between target detection and polarity classification, and the interaction among multiple targets. These targets can be explicitly or implicitly mentioned in a piece of text and do not need to be tagged a priori.

3. ATTNET FOR ATTITUDE IDENTIFICATION

The **target detection** task is to determine whether a specific target occurs in a given context either explicitly or implicitly.

The **polarity classification** task is to decide the attitude of the given context towards thespecific target if the target occurs in the context.

The output of the target detection is concatenated as part of the input of the polarity classification task to allow polarity classification to be conditioned on target detection. Polarity classification labels are also used to train the target detection classifier by back-propagating the errors of the polarity classification to the target detection end-to-end.

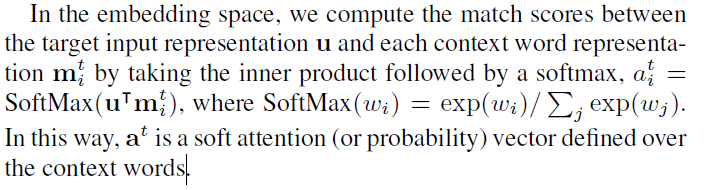
3.1 Background: Memory Networks

3.2 Single Layer AttNet

Target detection as TD, and polarity classification as PC

(1) Target Embedding.

(2) Input Representation and Attention for TD



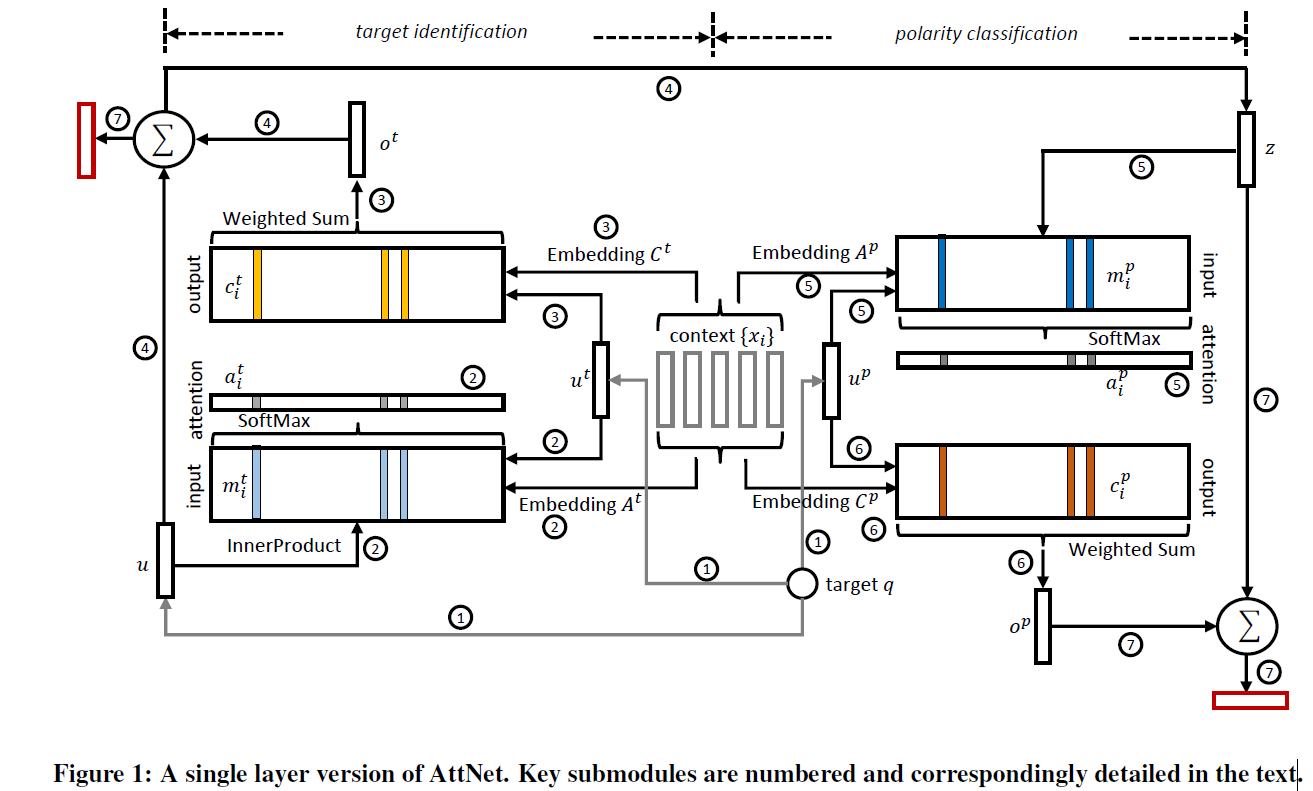
(3) Output Representation for TD

(4) Interleaving TD and PC

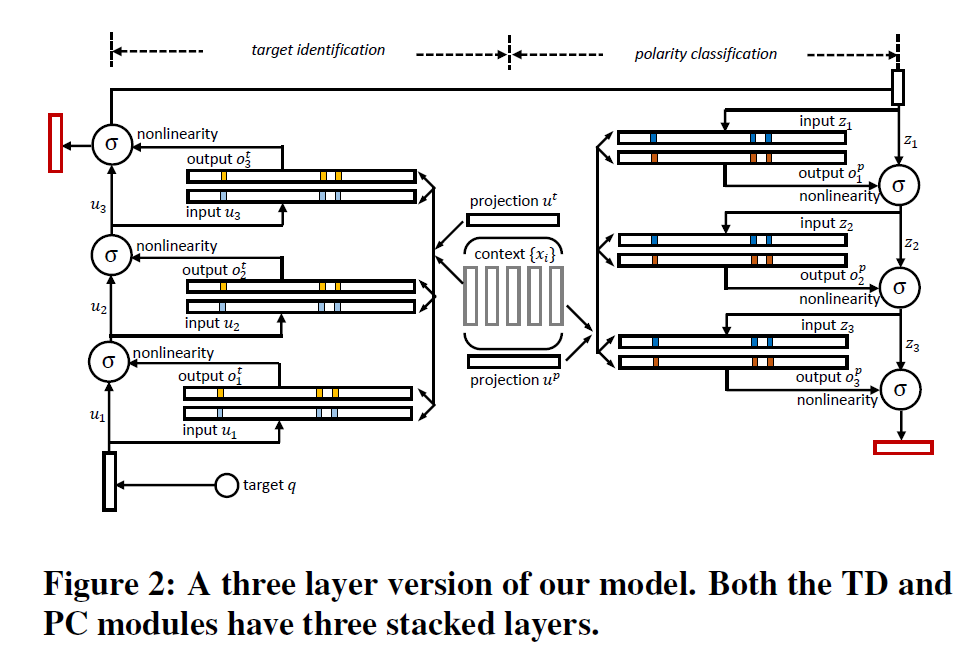
(5) Input Representation and Attention for PC

(6) Output Representation for PC

(7) Prediction for TD and PC



3.3 Multiple Layer AttNet



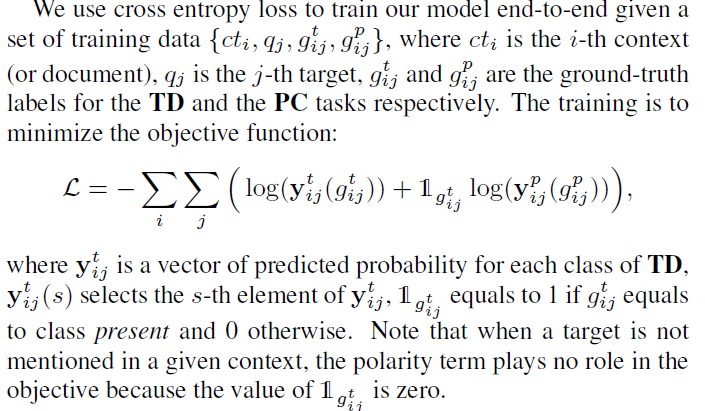
Functionality of Each Layer

Attention for PC

Tying Embedding and Projection Matrices

Predictions for TD and PC

3.4 End-to-End Multi-Task Training



4. EXPERIMENT SETUP

4.1 Data Sets

4.2 Metrics

4.3 Baselines

4.3.1 Conventional baselines

SVM+features

4.3.2 Deep Learning Baselines

BiLSTM, MultiBiLSTM and Memnet

CNN and ParaVec

4.4 Variants of AttNet

4.5 Training Details

All hyper-parameters are tuned to obtain the best performance of F-score on validation set.

5. EXPERIMENT RESULTS

5.1 Overall Performance

5.2 Performance on Subtasks

5.3 Training Time Analysis

5.4 Visualization of attention

6. CONCLUSION