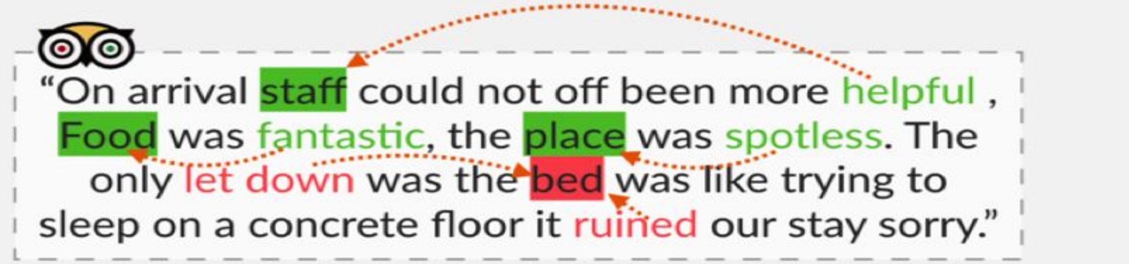
Aspect Based Sentiment Analysis

15CS30037 : TYSS Santosh
16CE31005 : R K S Jaswanth
15CS30035 : Sumeet Shirgure
16CS10033 : Vishal Reddy



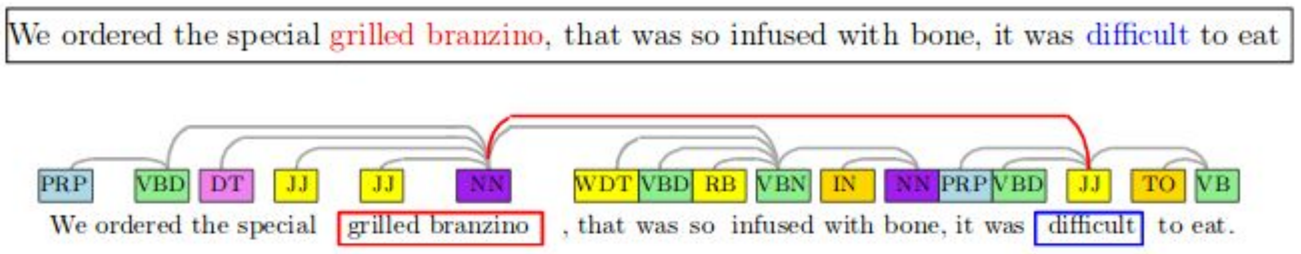
Aspect	Polarity
Staff	Positive
Food	Positive
Cleanliness	Positive
Beds	Negative

Introduction

- Traditional methods for aspect-level sentiment analysis mainly focus on designing a set of features (such as bag-of-words, sentiment lexicons,and linguistic features) to train a classifier for sentiment classification.
- In recent years,more and more neural network based models have been proposed and obtained the state-of-the-art results.

Motivation

- Generally, we find that a dependency tree shortens the distance between the aspects and opinion words of a sentence, captures the syntactic relations between words.
- These properties allow neural network models to capture long-term syntactic dependencies effortlessly.



Motivation

- These observations motivate us to develop a neural model which can operate on the dependency tree of a sentence, with the aim to make accurate sentiment predictions with respect to specific aspects.
- We exploit a GCN to model the structure of a sentence through its dependency tree, where node(word) embeddings of the tree are initialized by means of a Bi-directional Long Short Term Mem-ory (Bi-LSTM) network.

Motivation

- We hypothesize that the architecture of Bi-LSTM+GCN allows the Bi-LSTM account for contextual information between successive words, while the GCN enhances the embeddings by modeling the dependencies along the syntactic paths of the dependency tree.

Methodology : Bi-LSTM Layer

- Given an aspect-sentence pair(a,s), where a={a1,a2,...,al} is a sub-sequence of the sentence s={w1,w2,...,wn}.
- The sentence s has corresponding word embeddings x={x1,x2,...xn}.
- We apply Bi-LSTM to obtain hidden representations {h1,h2,...,hn}.

Methodology : GCN Layer

- Besides, we integrate dependency information in the contextualized embeddings using a GCN which operates directly on the dependency tree of the sentence.
- The dependency tree can be interpreted as a graph G with n nodes, where nodes represent words in the sentence and edges represent syntactic dependency paths between words in the graph.

Methodology : GCN Layer

- The dependency tree G for any arbitrary sentence can be represented as an n×n adjacency matrix A, with entries Aij signaling if node i isconnected to node j by a single dependency path in G.
- To allow the GCN to model node embeddings efficiently, we allow G to have self-loops.

Methodology : GCN Layer

- The GCN transform and propagate information across the paths, and update node embeddings by aggregating the propagated information.
- GCN only considers the first order neighborhood of a node when modeling its embeddings.

Methodology : GCN Layer

- However, k successive GCN operations result in the propagation of information across the k-th order neighborhood.

$$h_i^{(k+1)} = \phi\left(\sum_{j=1}^{n} c^i A_{ij}\left(W^{(k)} h_j^{(k)} + b^{(k)}\right)\right)$$

Methodology : Output Layer

- For our framework, we choose an average pool which aggregates information over the aspect vectors.
- Moreover, we perform an average pool to retain most of the information in the aspect vectors.
- The aspect-based representation is passed to a fully connected softmax layer whose output is a probability distribution over the different sentiment polarities.

Baselines:

- Majority assigns the sentiment polarity with most frequent occurrences
- TD-LSTM adopts two LSTMs to model the left context with target and the right context with target respectively (Tang et al., 2016a);
- MemNet (Tang et al., 2016b) applies attention multiple times on the word embeddings
- IAN (Ma et al., 2017) interactively learns attentions

Baselines:

- RAM (Chen et al., 2017) is a multilayer architecture where each layer consists of attention-based aggregation of word features and a GRU cell to learn the sentence representation.
- LCR-Rot(Zheng et al., 2018) employs three Bi-LSTMs to model the left context, the target and the right context. Then they propose a rotatory attention mechanism which models the relation between target and left/right contexts.

Baselines:

- AOA-LSTM(Huang et al., 2018) introduces an attention-over-attention (AOA) based network to model aspects and sentences in a joint way and explicitly capture the interaction between aspects and context sentences
- TNet (Li et al., 2018a): In this work,BiLSTM embeddings are transformed into target specific embeddings, and a CNNmodel is used to extract a final embedding.
- PRET+MULT (He et al., 2018b): A multi-task framework based on LSTMs is proposed to transfer knowledge from a document-level model task to an aspect-level model task.

Baselines:

- SA-LSTM-P (Wang and Lu, 2018): model structural dependencies between words by means of a segmentation attention mechanism.
- LSTM+SynATT+TarRep (He et al., 2018a): models the syntactic structure of the sentence using an attention mechanism.
- MGANA (Fan et al., 2018b): multi-grained attention mechanism is proposed to extract an embedding which effectively captures the interaction between the aspect and the context

Baselines:

- MGANB (Li et al., 2018b): an alignment mechanism in a multi-task model comprising of an aspect-term task and an aspect-category task to effectively extract aspect-specific representations.
- HSCN (Lei et al., 2019): A model is proposed to capture interactions between the context and target, select target words and extract target-specific contextual representation

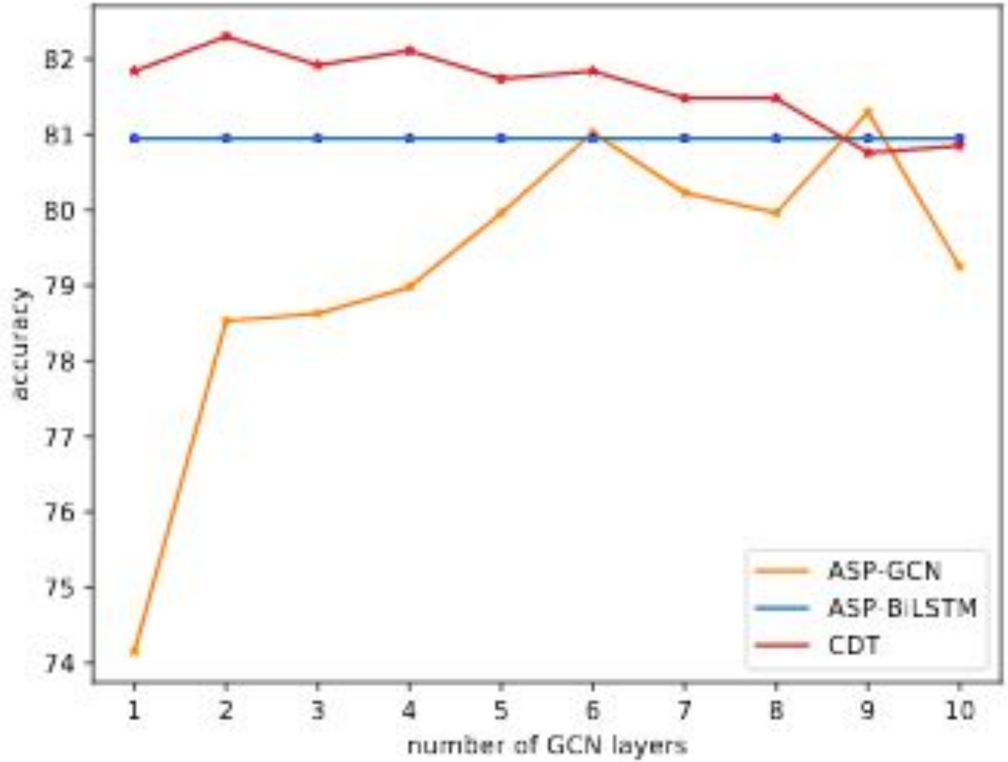
Results

Model	Restaurant	Laptop
Majority	65.00	53.45
TD-LSTM	75.63	68.13
MemNet	79.98	70.33
IAN	78.60	72.10
RAM	80.23	74.49
LCR-Rot	81.34	75.24
AOA-LSTM	81.20	74.50
TNet	80.79	76.54
PRET+MULT	79.11	71.15
SA-LSTM-P	81.60	75.1
LSTM+SynATT+TarRep	80.63	71.94
MGANA	81.25	75.39
MGANB	81.49	76.21
HSCN	77.8	76.1
ASP-BiLSTM	80.95	74.22
ASP-GCN	81.30	74.53
Bi-LSTM+GCN	82.30	77.19

Ablation Study

Model	Restaurant	Laptop
ASP-BiLSTM	80.95	74.22
ASP-GCN	81.30	74.53
Bi-LSTM+GCN	82.30	77.19

Ablation Study

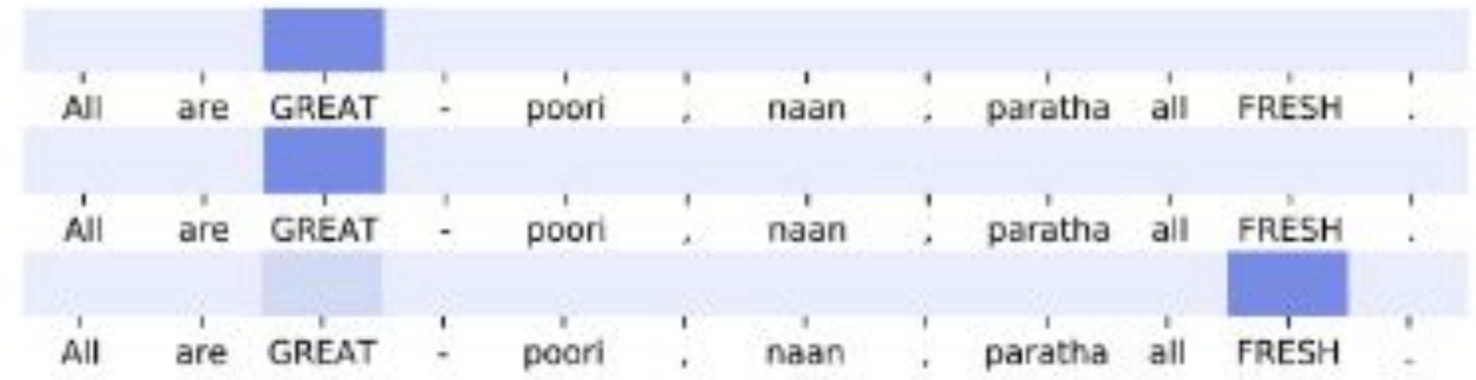


Case Study

Attention visualization for ASP-BiLSTM(1st row), TNet(2nd row) and Bi-LSTM+GCN (3rd row) for the aspect word 'Sangria'

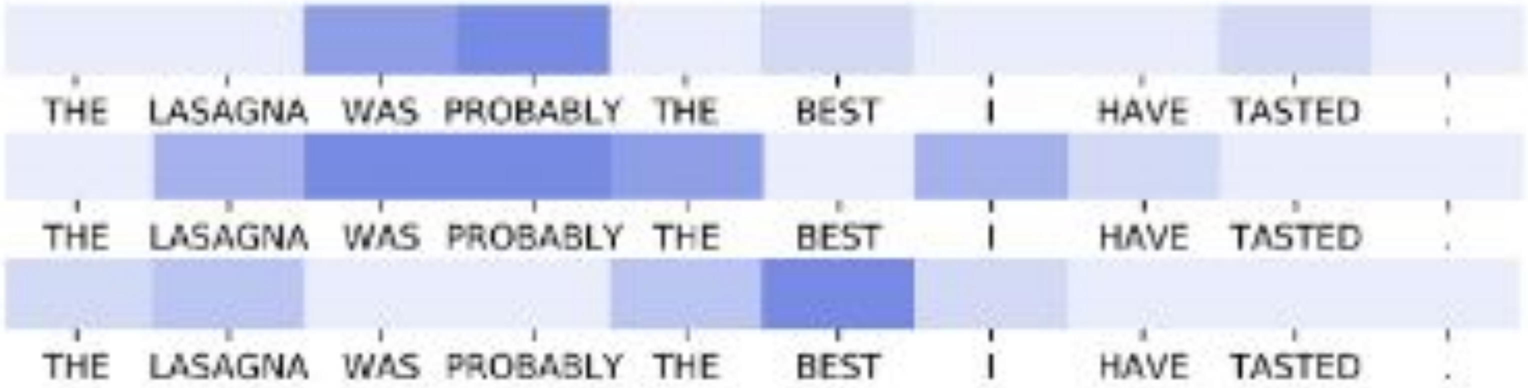
Case Study

Attention visualization for ASP-BiLSTM(1st row), TNet(2nd row) and Bi-LSTM+GCN (3rd row) for the aspect word ‘paratha’



Case Study

Attention visualization for ASP-BiLSTM(1st row), TNet(2nd row) and Bi-LSTM+GCN (3rd row) for the aspect word ‘lasagna’



Conclusion

- In this paper, we integrate a GCN with a simple BiLSTM model, with the aim to capture structural and contextual information of sentences. We have shown that the GCN successfully performs convolutions on the dependency tree to refine BiLSTM embeddings.