Fine-Grained Opinion Mining: Current Trend and Cutting-Edge Dimensions

Wenya Wang, Jianfei Yu, Sinno Jialin Pan and Jing JIang

Nanyang Technological University
and
Singapore Management University
Part III

Target-Oriented Sentiment Classification
Outline

- Background
- Methodology
- Summary
Background

- Sentence/Document-Level Sentiment Classification
  - Input
    - A sentence or document
  - Output
    - Overall sentiment polarity
      - Positive, Negative, Neutral
  - Example

The movie was fabulous, and the characters are quite engaging!

The restaurant was horrible, and their service was also poor!
Background

- Target-oriented Sentiment Classification (TSC)
  - Input
    - A sentence or document
    - An **opinion target**
      - 1. Aspect Term (Aspect-Level Sentiment Classification)
      - 2. Aspect Category (Aspect Category-Based Sentiment Classification)
      - 3. Target Entity (Entity-Level Sentiment Classification)
  - Output
    - **Sentiment polarity** towards the **opinion target**
      - Positive, Negative, Neutral
Background

- Examples (Product Review)
  - Aspect-Level Sentiment Classification

The [fish] was rather over cooked, but the [staff] was quite nice!

- sentiment over fish: negative
- sentiment over staff: positive
Background

Examples (Product Review)
- Aspect Category-Based Sentiment Classification

The [fish] was rather over cooked, but the [staff] was quite nice!

- sentiment over food: negative
- sentiment over service: positive
- sentiment over ambience: N.A
- sentiment over price: N.A
- sentiment over miscellaneous: N.A
Background

- Examples (Tweet)
  - Entity-Level Sentiment Classification

[Georgina Hermitage] is a #one2watch since she broke the [400m T37] WR!

- sentiment over Georgina Hermitage: positive
- sentiment over 400m T37: neutral
Outline

- Background
- Methodology
- Summary
Methodology – Big Picture

Supervised Machine Learning Methods

- Feature Engineering
  - Linear Classifier: Manually create features related to opinion targets
  - MemNet: Use opinion targets as queries to MemNet
  - Recursive NN
  - CNN
  - RNN
  - BERT: Adapt standard model architectures to be sensitive to opinion targets

Deep Learning
Outline

- Background

- Methodology
  - Linear Classifier
  - Recursive Neural Network
  - Memory Network
  - CNN-based Methods
  - RNN-based Methods
  - BERT-based Methods

- Summary
Linear Classifier

- Extract various features

Linear Classifier

- Extract various features

- Target-dependent features from words filtered by sentiment lexicon

Linear Classifier

- Extract various features

\[ T_{tw}^{(1)} = [P(L^{(1)}), P(T^{(1)}), P(R^{(1)})] \]

- Target-dependent features from the left context, right context, and target, respectively

Linear Classifier

- Extract various features

Full tweet features

Linear Classifier

- Extract various features

- Feed the concatenated features to a discriminative classifier
  - SVM

Outline

- Background

- Methodology
  - Linear Classifier
  - Recursive Neural Network
  - Memory Network
  - CNN-based Methods
  - RNN-based Methods
  - BERT-based Methods

- Summary
Recursive Neural Network

- Dependency Tree-based Approach
  - AdaRNN
    - propagate sentiment information to the target node in a bottom-up manner

Methodology

Recursive Neural Network

- Dependency Tree-based Approach
  - AdaRNN
    - propagate sentiment information to the target node in a bottom-up manner

Recursive Neural Network

- Dependency + Constituent tree-based Approach
  - PhraseRNN

Recursive Neural Network

- Dependency + Constituent tree-based Approach
  - PhraseRNN

Outline

- Background

- Methodology
  - Linear Classifier
  - Recursive Neural Network
  - Memory Network
  - CNN-based Methods
  - RNN-based Methods
  - BERT-based Methods

- Summary
Memory Network

- MemNet
  - Word embedding of target words as queries to MemNet

Outline

- Background

- **Methodology**
  - Linear Classifier
  - Recursive Neural Network
  - Memory Network
  - **CNN-based Methods**
  - RNN-based Methods
  - BERT-based Methods

- **Summary**
CNN-based Methods

- GCN (Gated Convolutional Networks)
  - Incorporate gate mechanism to be sensitive to opinion targets

Model I. GCN for Aspect Category-based Sentiment Classification

CNN-based Methods

- GCN (Gated Convolutional Networks)
  - Incorporate gating mechanism to be sensitive to be opinion targets

Model II. GCN for Aspect-Level Sentiment Classification

Outline

- Background

- Methodology
  - Linear Classifier
  - Recursive Neural Network
  - Memory Network
  - CNN-based Methods
  - RNN-based Methods
  - BERT-based Methods

- Summary
RNN-based Methods

- GRU
  - Gating Mechanism

RNN-based Methods

- LSTM
  - Sentence Encoding

RNN-based Methods

- LSTM
  - Attention Mechanism


The Model Architecture of AE-LSTM
RNN-based Methods

- LSTM
  - IAN
    - Interactive Attention Mechanism

The Model Architecture of IAN

RNN-based Methods

- LSTM
  - RAM
    - Position-based Weighting Strategy
    - Multi-Hop Attention Mechanism

The Model Architecture of RAM

RNN-based Methods

- LSTM
  - TNet

The Model Architecture of TNet

Outline

- Background

- Methodology
  - Linear Classifier
  - Recursive Neural Network
  - Memory Network
  - CNN-based Methods
  - RNN-based Methods
  - BERT-based Methods

- Summary
Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI, 2019.
Outline

- Background
- Methodology
- Summary
Summary

- Three Benchmark Datasets

<table>
<thead>
<tr>
<th>Data Set</th>
<th>#Training Samples</th>
<th>#Test Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>POS</td>
<td>NEG</td>
</tr>
<tr>
<td>Laptop</td>
<td>980</td>
<td>858</td>
</tr>
<tr>
<td>Restaurant</td>
<td>2159</td>
<td>800</td>
</tr>
<tr>
<td>Twitter-2014</td>
<td>1567</td>
<td>1563</td>
</tr>
</tbody>
</table>

- Laptop, Restaurant are from SemEval-2014
- Twitter-2014 from (Dong et al. ACL 2014)
- Another two Restaurant datasets from SemEval-2015, SemEval-2016
### Experimental Results on Three Benchmark Datasets

<table>
<thead>
<tr>
<th>Method</th>
<th>Laptop</th>
<th>Restaurant</th>
<th>Twitter-2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>Macro-F1</td>
<td>Accuracy</td>
</tr>
<tr>
<td>SVM</td>
<td>70.49</td>
<td>-</td>
<td>80.16</td>
</tr>
<tr>
<td>AE-LSTM</td>
<td>68.90</td>
<td>-</td>
<td>76.60</td>
</tr>
<tr>
<td>IAN</td>
<td>72.10</td>
<td>-</td>
<td>78.60</td>
</tr>
<tr>
<td>TD-LSTM</td>
<td>71.83</td>
<td>68.43</td>
<td>78.00</td>
</tr>
<tr>
<td>MemNet</td>
<td>70.33</td>
<td>64.09</td>
<td>78.16</td>
</tr>
<tr>
<td>RAM</td>
<td>75.01</td>
<td>70.51</td>
<td>79.79</td>
</tr>
<tr>
<td>TNet-LF</td>
<td>76.01</td>
<td>71.47</td>
<td>80.79</td>
</tr>
<tr>
<td>TNet-AS</td>
<td>76.54</td>
<td>71.75</td>
<td>80.69</td>
</tr>
<tr>
<td>MGAN</td>
<td>75.39</td>
<td>72.47</td>
<td>81.25</td>
</tr>
<tr>
<td>BERT</td>
<td><strong>76.96</strong></td>
<td><strong>73.67</strong></td>
<td><strong>84.29</strong></td>
</tr>
</tbody>
</table>

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI, 2019.
Summary

Supervised Machine Learning Methods

Feature Engineering

Linear Classifier

Pros: Simple and fast to train
Cons: Many feature engineering efforts

Before 2014

Deep Learning

Pros: Better results
Cons: Lack of good explanation

2014-2015

Recursive NN

Pros: make use of syntactic information
Cons: 1. Noisy syntactic tree; 2. Gradient Vanishing

Comparison
1. Training/Test Time: MemNet ≈ CNN < RNN
2. Performance: RNN > CNN ≈ MemNet

2016-now

MemNet

CNN

RNN

BERT

2019

Best results

School of Information Systems
Part V

Cutting-Edge Dimensions of Fine-Grained Opinion Mining
Outline

- Transfer Learning
- Multi-Task Learning
- Multimodal Learning
- Summary
Outline

- Transfer Learning
  - Cross-Domain
  - Cross-Lingual
  - Short Summary
- Multi-Task Learning
- Multimodal Learning
- Summary
Cross-Domain

- **Background**
  - Popular Methods for Fine-Grained Opinion Mining
    - Supervised Machine Learning (NN)

Large amount of training data
Cross-Domain

- **Background**
  - **Real Scenario**
    - *Limited or no* Labeled Data for many domains

- **Domain Adaptation**: recognize and apply knowledge and skills learned in previous domain to novel domains

- **Machine Learning System**
  - **Weak Classifier**
  - **Knowledge Transfer**
  - **Strong Classifier**

- **Target domain**
- **Source domain**
Cross-Domain

- Background
  - Challenge of Domain Adaptation

School of Information Systems
Cross-Domain

- Background
  - Challenge of Domain Adaptation

Training Data

- Movie
  - (Source Domain)

Opinion Target Extraction Model

Test Data

- Movie
  - 78%

- Digital Device
  - 45%

(Source Domain)
## Cross-Domain

### Background
- Reasons behind performance drop

<table>
<thead>
<tr>
<th>Movie (source domain)</th>
<th>Digital Device (target domain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The [<strong>movie</strong>] is great.</td>
<td>The [<strong>camera</strong>] is excellent.</td>
</tr>
<tr>
<td>I really like his [<strong>characters</strong>].</td>
<td>I highly recommend this [<strong>laptop</strong>].</td>
</tr>
<tr>
<td>The [<strong>plot</strong>] is quite dull.</td>
<td>The [<strong>Mac OS</strong>] is quite fast.</td>
</tr>
</tbody>
</table>

- Opinion targets in the source domain: **movie**, **characters**, **plot**
- Opinion targets in the target domain: **camera**, **laptop**, **Mac OS**
Cross-Domain

- **Background**
  - **General Solution**
    - Learn a shared representation across domains

**Domain-Independent Auxiliary Tasks**

- Label/Unlabeled Source
- Unlabeled Target
- Training Set of Source Domain
- Test Set of Target Domain

\[ x_s \rightarrow \Phi(x_s) \]

\[ x_t \rightarrow \Phi(x_t) \]
Cross-Domain

- Cross-Domain Opinion Target Extraction
  - Domain-Independent Auxiliary Task
    - Syntactic structures are shared across domains.

Cross-Domain Opinion Target Extraction

- The same task as our Auxiliary Tasks
  - Unsupervised Extraction Method

Cross-Domain

- Cross-Domain Opinion Target Extraction
  - RNN with Auxiliary Tasks (AuxRNN)

Cross-Domain

- Cross-Domain Aspect and Opinion Terms Co-Extraction
  - Recursive Neural Structural Correspondence Network (RNSCN)

Outline

- **Transfer Learning**
  - Cross-Domain
  - Cross-Lingual
  - Short Summary
- **Multi-Task Learning**
- **Multimodal Learning**
- **Summary**
Cross-Lingual

- Cross-Lingual Aspect Term Extraction
  - Transition-based Adversarial Network (TAN)

Outline

- **Transfer Learning**
  - Cross-Domain
  - Cross-Lingual
  - Short Summary
- **Multi-Task Learning**
- **Multimodal Learning**
- **Summary**
Short Summary

- **Key to Cross-Domain/Lingual**
  - Step 1: Identify shared knowledge across domains or languages
    - General Sentiment Words like *good, bad*, etc
    - Syntactic Structure
    - Domain/Language Discriminator
    - Auto-encoder (reconstruction of the input)
  - Step 2: Design auxiliary tasks based on these shared knowledge
Short Summary

- Benchmark Datasets for Cross-Domain Aspect and Opinion Terms Co-Extraction

<table>
<thead>
<tr>
<th>Data Set</th>
<th>#Sentences</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laptop</td>
<td>3,845</td>
<td>2,884</td>
<td>961</td>
</tr>
<tr>
<td>Restaurant</td>
<td>5,841</td>
<td>4,381</td>
<td>1,460</td>
</tr>
<tr>
<td>Digital Device</td>
<td>3,836</td>
<td>2,877</td>
<td>959</td>
</tr>
</tbody>
</table>

- Laptop from SemEval-2014
- Restaurant from SemEval-2014, 2015
- Digital Device from (Hu and Liu, KDD2004)
## Short Summary

### Results on Benchmark Datasets

- **Hier-Joint**: (Ding, Yu and Jiang, AAAI 2017)
- **RNSCN**: (Wang and Sinno, ACL 2018)

> Incorporating domain-independent auxiliary tasks can indeed significantly outperform the baseline approach.

<table>
<thead>
<tr>
<th>Models</th>
<th>R$\rightarrow$L</th>
<th>R$\rightarrow$D</th>
<th>L$\rightarrow$R</th>
<th>L$\rightarrow$D</th>
<th>D$\rightarrow$R</th>
<th>D$\rightarrow$L</th>
</tr>
</thead>
<tbody>
<tr>
<td>CrossCRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (1.82)</td>
<td>OP (1.34)</td>
<td>AS (0.44)</td>
<td>OP (1.67)</td>
<td>AS (1.69)</td>
<td>OP (1.49)</td>
</tr>
<tr>
<td></td>
<td>19.72</td>
<td>59.20</td>
<td>21.07</td>
<td>52.05</td>
<td>28.19</td>
<td>65.52</td>
</tr>
<tr>
<td>RAP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (0.49)</td>
<td>OP (2.20)</td>
<td>AS (1.64)</td>
<td>OP (1.05)</td>
<td>AS (1.65)</td>
<td>OP (1.65)</td>
</tr>
<tr>
<td></td>
<td>25.92</td>
<td>62.72</td>
<td>46.90</td>
<td>67.98</td>
<td>34.54</td>
<td>54.25</td>
</tr>
<tr>
<td>Hier-Joint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (1.47)</td>
<td>OP (2.65)</td>
<td>AS (1.45)</td>
<td>OP (1.49)</td>
<td>AS (2.15)</td>
<td>OP (1.46)</td>
</tr>
<tr>
<td></td>
<td>33.66</td>
<td>50.20</td>
<td>48.10</td>
<td>52.05</td>
<td>31.25</td>
<td>47.97</td>
</tr>
<tr>
<td>RNCRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (3.97)</td>
<td>OP (3.35)</td>
<td>AS (2.57)</td>
<td>OP (1.78)</td>
<td>AS (1.34)</td>
<td>OP (1.19)</td>
</tr>
<tr>
<td></td>
<td>24.26</td>
<td>60.86</td>
<td>24.31</td>
<td>51.28</td>
<td>31.52</td>
<td>55.85</td>
</tr>
<tr>
<td>RNRGRU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (2.41)</td>
<td>OP (1.04)</td>
<td>AS (2.68)</td>
<td>OP (2.69)</td>
<td>AS (1.12)</td>
<td>OP (2.37)</td>
</tr>
<tr>
<td></td>
<td>24.23</td>
<td>60.65</td>
<td>20.49</td>
<td>52.28</td>
<td>32.51</td>
<td>52.24</td>
</tr>
<tr>
<td>RNSCN-CRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (1.31)</td>
<td>OP (1.35)</td>
<td>AS (1.48)</td>
<td>OP (1.29)</td>
<td>AS (1.38)</td>
<td>OP (1.10)</td>
</tr>
<tr>
<td></td>
<td>35.26</td>
<td>61.67</td>
<td>32.00</td>
<td>52.81</td>
<td>34.63</td>
<td>56.22</td>
</tr>
<tr>
<td>RNSCN-GRU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (0.45)</td>
<td>OP (1.85)</td>
<td>AS (0.58)</td>
<td>OP (1.27)</td>
<td>AS (0.77)</td>
<td>OP (0.80)</td>
</tr>
<tr>
<td></td>
<td>37.77</td>
<td>62.35</td>
<td>33.02</td>
<td>57.54</td>
<td>35.65</td>
<td>60.02</td>
</tr>
<tr>
<td>RNSCN$^+$-GRU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AS (0.96)</td>
<td>OP (1.50)</td>
<td>AS (0.62)</td>
<td>OP (0.75)</td>
<td>AS (0.70)</td>
<td>OP (0.60)</td>
</tr>
<tr>
<td></td>
<td>40.43</td>
<td>65.85</td>
<td>35.10</td>
<td>60.17</td>
<td>52.91</td>
<td>72.51</td>
</tr>
</tbody>
</table>
Short Summary

- Benchmark Datasets for Cross-Lingual Aspect Term Extraction

<table>
<thead>
<tr>
<th>Data Set</th>
<th>#Sentences</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>2,676</td>
<td>2,000</td>
<td>676</td>
</tr>
<tr>
<td>French</td>
<td>2,429</td>
<td>1,733</td>
<td>696</td>
</tr>
<tr>
<td>Spanish</td>
<td>2,951</td>
<td>2,070</td>
<td>881</td>
</tr>
</tbody>
</table>

- All from SemEval-2016 Task 5
### Results on Benchmark Datasets

- **CL-DSCL**: (Ding, Yu and Jiang, AAAI 2017)
- **TAN**: (Wang and Sinno, IJCAI 2018)

---

**Table: Model Performance**

<table>
<thead>
<tr>
<th>Models</th>
<th>En→Fr</th>
<th></th>
<th>En→Es</th>
<th></th>
<th>Fr→En</th>
<th></th>
<th>Fr→Es</th>
<th></th>
<th>Es→En</th>
<th></th>
<th>Es→Fr</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
</tr>
<tr>
<td>Translate-TAN</td>
<td>45.09</td>
<td>40.74</td>
<td>45.85</td>
<td>41.08</td>
<td>39.28</td>
<td>38.74</td>
<td>32.27</td>
<td>34.54</td>
<td>45.94</td>
<td>41.28</td>
<td>41.52</td>
<td>36.38</td>
</tr>
<tr>
<td>NoAdp</td>
<td>27.71</td>
<td>26.13</td>
<td>27.56</td>
<td>31.31</td>
<td>41.21</td>
<td>38.29</td>
<td>45.43</td>
<td>48.21</td>
<td>37.52</td>
<td>30.39</td>
<td>37.95</td>
<td>37.89</td>
</tr>
<tr>
<td>A-RNN</td>
<td>22.92</td>
<td>20.54</td>
<td>31.11</td>
<td>34.04</td>
<td>29.62</td>
<td>27.11</td>
<td>40.58</td>
<td>40.77</td>
<td>35.49</td>
<td>30.26</td>
<td>34.52</td>
<td>31.02</td>
</tr>
<tr>
<td>A-R\textsuperscript{2}NN</td>
<td>27.92</td>
<td>23.41</td>
<td>28.63</td>
<td>28.65</td>
<td>36.43</td>
<td>33.25</td>
<td>38.55</td>
<td>39.45</td>
<td>40.83</td>
<td>34.16</td>
<td>42.83</td>
<td>37.19</td>
</tr>
<tr>
<td>CL-DSCL</td>
<td>33.67</td>
<td>31.48</td>
<td>44.56</td>
<td>45.01</td>
<td>51.75</td>
<td>47.27</td>
<td>53.23</td>
<td>55.89</td>
<td>50.22</td>
<td>45.90</td>
<td>38.66</td>
<td>34.17</td>
</tr>
<tr>
<td>TAN</td>
<td>53.27</td>
<td>50.02</td>
<td>49.38</td>
<td>50.52</td>
<td>55.38</td>
<td>50.30</td>
<td>55.32</td>
<td>57.65</td>
<td>51.99</td>
<td>44.14</td>
<td>51.16</td>
<td>48.78</td>
</tr>
</tbody>
</table>

---

- Incorporating language-independent auxiliary tasks can indeed significantly outperform the baseline approach.
Outline

- Transfer Learning

- Multi-Task Learning
  - Aspect and Opinion Terms Co-Extraction
  - End to End ABSA
    - Aspect Term Extraction + Aspect-Level Sentiment Classification

- Multimodal Learning

- Summary
Background

- Aspect and Opinion Terms Co-extraction
  - Input
    • A sentence or document
  - Output
    • Aspect Term
    • Opinion Term
  - Example
    The fish was rather over cooked, but the staff was quite nice!
    - Aspect Term: fish, staff
    - Opinion Term: over cooked, nice

- Sequence Labeling Problems
Outline

- Transfer Learning

- Multi-Task Learning
  - Aspect and Opinion Terms Co-Extraction
  - End to End ABSA
    - Aspect Term Extraction + Aspect-Level Sentiment Classification

- Multimodal Learning

- Summary
End to End Aspect-Based Sentiment Analysis

- **Input**
  - A sentence or document

- **Output**
  - **Aspect Term**
  - **Sentiment polarity** towards the aspect term
    - Positive, Negative, Neutral

- **Example**
  
  The fish was rather *over cooked*, but the staff was *quite nice*!

  ➢ (fish, negative), (staff, positive)
Background

- End to End Aspect-Based Sentiment Analysis
  - Neural CRF
    - Method 1: pipeline

Two Sequence Labeling Tasks

Background

- End to End Aspect-Based Sentiment Analysis
  - Neural CRF
    - Method 2: joint

```
sentence: So excited to meet my baby Farah !!!
entity: O  O  O  O  O  O  B  I  O
sentiment: φ  φ  φ  φ  φ  +  +  φ
```

Two Sequence Labeling Tasks

Joint Model

Background

- End to End Aspect-Based Sentiment Analysis
  - Neural CRF
    - Method 3: collapsed

## Background

- **End to End Aspect-Based Sentiment Analysis**
  - Neural CRF
    - Comparison

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th></th>
<th></th>
<th>Spanish</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entity</td>
<td>SA</td>
<td></td>
<td>Entity</td>
<td>SA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F</td>
<td>P</td>
<td>R</td>
<td>F</td>
</tr>
<tr>
<td><strong>Pipeline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discrete</td>
<td>59.37</td>
<td>34.83</td>
<td>43.84</td>
<td>42.97</td>
<td>25.21</td>
<td>31.73</td>
</tr>
<tr>
<td>neural</td>
<td>53.64</td>
<td>44.87</td>
<td>48.67</td>
<td>37.53</td>
<td>31.38</td>
<td>34.04</td>
</tr>
<tr>
<td>integrated</td>
<td>60.69</td>
<td>51.63</td>
<td>55.67</td>
<td>43.71</td>
<td>37.12</td>
<td>40.06</td>
</tr>
<tr>
<td><strong>Joint</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discrete</td>
<td>59.55</td>
<td>34.06</td>
<td>43.30</td>
<td>43.09</td>
<td>24.67</td>
<td>31.35</td>
</tr>
<tr>
<td>neural</td>
<td>54.45</td>
<td>42.12</td>
<td>47.17</td>
<td>37.55</td>
<td>28.95</td>
<td>32.45</td>
</tr>
<tr>
<td>integrated</td>
<td>61.47</td>
<td>49.28</td>
<td>54.59</td>
<td>44.62</td>
<td>35.84</td>
<td>39.67</td>
</tr>
<tr>
<td><strong>Collapsed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discrete</td>
<td>64.16</td>
<td>26.03</td>
<td>36.95</td>
<td>48.35</td>
<td>19.64</td>
<td>27.86</td>
</tr>
<tr>
<td>neural</td>
<td>58.53</td>
<td>37.25</td>
<td>45.30</td>
<td>43.12</td>
<td>27.44</td>
<td>33.36</td>
</tr>
<tr>
<td>integrated</td>
<td>63.55</td>
<td><strong>44.98</strong></td>
<td><strong>52.58</strong></td>
<td>46.32</td>
<td><strong>32.84</strong></td>
<td><strong>38.36</strong></td>
</tr>
</tbody>
</table>
## Background

- **End to End Aspect-Based Sentiment Analysis**
  - Unified Solution

<table>
<thead>
<tr>
<th>Input</th>
<th>The</th>
<th>AMD</th>
<th>Turin</th>
<th>Processor</th>
<th>seems</th>
<th>to</th>
<th>always</th>
<th>perform</th>
<th>much</th>
<th>better</th>
<th>than</th>
<th>Intel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint</td>
<td>0</td>
<td>B</td>
<td>I</td>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>POS</td>
<td>POS</td>
<td>POS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>POS</td>
</tr>
<tr>
<td>Unified (✔)</td>
<td>0</td>
<td>B-POS</td>
<td>I-POS</td>
<td>E-POS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>S-NEG</td>
</tr>
</tbody>
</table>

Two Sequence Labeling Tasks
Background

- End to End Aspect-Based Sentiment Analysis
  - Unified Solution

  - Two LSTMs for the target boundary detection task (auxiliary) and the complete TBSA task (primary).
  - BG component: exploiting boundary information
  - SC component: maintaining sentiment consistency
  - OE component: improving the quality of the boundary information

Xin Li, Lidong Bing, Piji Li and Wai Lam. A Unified Model for Opinion Target Extraction and Target Sentiment Prediction. In AAAI 2019.
Background

- End to End Aspect-Based Sentiment Analysis
  - Span Extraction-based approach

---

Background

- End to End Aspect-Based Sentiment Analysis
  - Span Extraction-based approach
    - BERT as encoder

- The last block’s hidden states are used to propose one or multiple candidate targets based on the probabilities of the start and end positions

- Predict the sentiment polarity using the span representation of the given target

---

Background

- End to End Aspect-Based Sentiment Analysis
  - Comparison of previous three approaches
    - Benchmark Datasets

<table>
<thead>
<tr>
<th>Data Set</th>
<th>#Training Samples</th>
<th>#Test Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>POS</td>
<td>NEG</td>
</tr>
<tr>
<td>Laptop</td>
<td>980</td>
<td>858</td>
</tr>
<tr>
<td>Restaurant</td>
<td>2159</td>
<td>800</td>
</tr>
<tr>
<td>Twitter-2014</td>
<td>1567</td>
<td>1563</td>
</tr>
</tbody>
</table>

- Laptop, Restaurant are from SemEval-2014
- Twitter-2014 from (Dong et al. ACL 2014)
- Another two Restaurant datasets from SemEval-2015, SemEval-2016

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI, 2019.
Background

- End to End Aspect-Based Sentiment Analysis
  - Comparison of previous three approaches
    - Unified Approach vs LSTM-based Methods

<table>
<thead>
<tr>
<th>Model</th>
<th>$\mathbb{D}_L$ P</th>
<th>$\mathbb{D}_L$ R</th>
<th>$\mathbb{D}_L$ F1</th>
<th>$\mathbb{D}_R$ P</th>
<th>$\mathbb{D}_R$ R</th>
<th>$\mathbb{D}_R$ F1</th>
<th>$\mathbb{D}_T$ P</th>
<th>$\mathbb{D}_T$ R</th>
<th>$\mathbb{D}_T$ F1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Existing Baselines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRF-joint</td>
<td>57.38</td>
<td>35.76</td>
<td>44.06</td>
<td>60.00</td>
<td>48.57</td>
<td>53.68</td>
<td>43.09</td>
<td>24.67</td>
<td>31.35</td>
</tr>
<tr>
<td>CRF-unified</td>
<td>59.27</td>
<td>41.86</td>
<td>49.06</td>
<td>63.39</td>
<td>57.74</td>
<td>60.43</td>
<td>48.35</td>
<td>19.64</td>
<td>27.86</td>
</tr>
<tr>
<td>NN-CRF-joint</td>
<td>55.64</td>
<td>34.48</td>
<td>45.49</td>
<td>61.56</td>
<td>50.00</td>
<td>55.18</td>
<td>44.62</td>
<td>35.84</td>
<td>39.67</td>
</tr>
<tr>
<td>NN-CRF-unified</td>
<td>58.72</td>
<td>45.96</td>
<td>51.56</td>
<td>62.61</td>
<td>60.53</td>
<td>61.56</td>
<td>46.32</td>
<td>32.84</td>
<td>38.36</td>
</tr>
<tr>
<td><strong>Pipeline Baselines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRF-pipeline</td>
<td>59.69</td>
<td>47.54</td>
<td>52.93</td>
<td>52.28</td>
<td>51.01</td>
<td>51.64</td>
<td>42.97</td>
<td>25.21</td>
<td>31.73</td>
</tr>
<tr>
<td>NN-CRF-pipeline</td>
<td>57.72</td>
<td>49.32</td>
<td>53.19</td>
<td>60.09</td>
<td>61.93</td>
<td>61.00</td>
<td>43.71</td>
<td>37.12</td>
<td>40.06</td>
</tr>
<tr>
<td>HAST-TNet</td>
<td>56.42</td>
<td>54.20</td>
<td>55.29</td>
<td>62.18</td>
<td>73.49</td>
<td>67.36</td>
<td>46.30</td>
<td>49.13</td>
<td>47.66</td>
</tr>
<tr>
<td><strong>Unified Baselines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSTM-unified</td>
<td>57.91</td>
<td>46.21</td>
<td>51.40</td>
<td>62.80</td>
<td>63.49</td>
<td>63.14</td>
<td>51.45</td>
<td>37.62</td>
<td>43.41</td>
</tr>
<tr>
<td>LSTM-CRF-1</td>
<td>58.61</td>
<td>50.47</td>
<td>54.24</td>
<td>66.10</td>
<td>66.30</td>
<td>66.20</td>
<td>51.67</td>
<td>44.08</td>
<td>47.52</td>
</tr>
<tr>
<td>LSTM-CRF-2</td>
<td>58.66</td>
<td>51.26</td>
<td>54.71</td>
<td>61.56</td>
<td>67.26</td>
<td>64.29</td>
<td>53.74</td>
<td>42.21</td>
<td>47.26</td>
</tr>
<tr>
<td>LM-LSTM-CRF</td>
<td>53.31</td>
<td>59.4</td>
<td>56.19</td>
<td>68.46</td>
<td>64.43</td>
<td>66.38</td>
<td>43.52</td>
<td>52.01</td>
<td>47.35</td>
</tr>
<tr>
<td><strong>OURS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base model</td>
<td>60.00</td>
<td>46.85</td>
<td>52.61</td>
<td>61.48</td>
<td>66.16</td>
<td>63.73</td>
<td>53.02</td>
<td>41.47</td>
<td>46.50</td>
</tr>
<tr>
<td>Base model + BG</td>
<td>58.58</td>
<td>50.63</td>
<td>54.31</td>
<td>67.51</td>
<td>66.42</td>
<td>66.96</td>
<td>52.26</td>
<td>43.84</td>
<td>47.66</td>
</tr>
<tr>
<td>Base model + BG + SC</td>
<td>58.95</td>
<td>53.00</td>
<td>55.81</td>
<td>63.95</td>
<td>69.65</td>
<td>66.68</td>
<td>53.12</td>
<td>43.60</td>
<td>47.79</td>
</tr>
<tr>
<td>Base model + BG + OE</td>
<td>63.43</td>
<td>49.53</td>
<td>55.62</td>
<td>62.85</td>
<td>66.77</td>
<td>65.22</td>
<td>53.10</td>
<td>43.50</td>
<td>47.78</td>
</tr>
<tr>
<td>Full model</td>
<td>61.27</td>
<td>54.89</td>
<td>57.90$^{\mathrm{a,b}}$</td>
<td>68.64</td>
<td>71.01$^{\mathrm{a,b}}$</td>
<td>69.80$^{\mathrm{a,b}}$</td>
<td>53.08</td>
<td>43.56</td>
<td>48.01$^{\mathrm{a,b}}$</td>
</tr>
</tbody>
</table>

Background

- End to End Aspect-Based Sentiment Analysis
  - Comparison of previous three approaches
    - BERT-based Methods vs Unified Approach

<table>
<thead>
<tr>
<th>Model</th>
<th>LAPTOP Prec.</th>
<th>LAPTOP Rec.</th>
<th>LAPTOP F1</th>
<th>REST Prec.</th>
<th>REST Rec.</th>
<th>REST F1</th>
<th>TWITTER Prec.</th>
<th>TWITTER Rec.</th>
<th>TWITTER F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIFIED</td>
<td>61.27</td>
<td>54.89</td>
<td>57.90</td>
<td>68.64</td>
<td>71.01</td>
<td>69.80</td>
<td>53.08</td>
<td>43.56</td>
<td>48.01</td>
</tr>
<tr>
<td>TAG-pipeline</td>
<td>65.84</td>
<td>67.19</td>
<td>66.51</td>
<td>71.66</td>
<td>76.45</td>
<td>73.98</td>
<td>54.24</td>
<td>54.37</td>
<td>54.26</td>
</tr>
<tr>
<td>TAG-joint</td>
<td>65.43</td>
<td>66.56</td>
<td>65.99</td>
<td>71.47</td>
<td>75.62</td>
<td>73.49</td>
<td>54.18</td>
<td>54.29</td>
<td>54.20</td>
</tr>
<tr>
<td>TAG-collapsed</td>
<td>63.71</td>
<td>66.83</td>
<td>65.23</td>
<td>71.05</td>
<td>75.84</td>
<td>73.35</td>
<td>54.05</td>
<td>54.25</td>
<td>54.12</td>
</tr>
<tr>
<td>SPAN-pipeline</td>
<td>69.46</td>
<td>66.72</td>
<td>68.06</td>
<td>76.14</td>
<td>73.74</td>
<td>74.92</td>
<td>60.72</td>
<td>55.02</td>
<td>57.69</td>
</tr>
<tr>
<td>SPAN-joint</td>
<td>67.41</td>
<td>61.99</td>
<td>64.59</td>
<td>72.32</td>
<td>72.61</td>
<td>72.47</td>
<td>57.03</td>
<td>52.69</td>
<td>54.55</td>
</tr>
<tr>
<td>SPAN-collapsed</td>
<td>50.08</td>
<td>47.32</td>
<td>48.66</td>
<td>63.63</td>
<td>53.04</td>
<td>57.85</td>
<td>51.89</td>
<td>45.05</td>
<td>48.11</td>
</tr>
</tbody>
</table>
Outline

- Transfer Learning
- Multi-Task Learning
- Multimodal Learning
  - Target-Oriented Multimodal Sentiment Classification
- Summary
Target-oriented Sentiment Classification (TSC)

- Input
  - A sentence or document
  - An opinion target

- Output
  - Sentiment polarity towards the opinion target

Examples

The *fish* was rather *over cooked*, but the *chicken* was *fine!*

- sentiment over *fish*: negative
- sentiment over *chicken*: positive
Motivation

- Limitation of TSC
  - Ineffective for multimodal social media posts
    - Incomplete Textual Contents

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Motivation

- **Limitation of TSC**
  - Ineffective for multimodal social media posts
    - Incomplete Textual Contents

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Motivation

- Limitation of TSC
  - Ineffective for multimodal social media posts
    - Incomplete Textual Contents

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Motivation

- Limitation of TSC
  - Ineffective for multimodal social media posts
    - Incomplete Textual Contents
  - Irregular Expressions

Georgina Hermitage: neutral
Motivation

- Limitation of TSC
  - Ineffective for multimodal social media posts
    - Incomplete Textual Contents
  - Irregular Expressions

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Motivation

- Target-oriented Multimodal Sentiment Classification (TMSC)
  - Input
    - A sentence or document
    - An opinion target
    - An associated image
  - Output
    - Sentiment polarity towards the opinion target

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Methodology -- BERT

- Base model with BERT
  - Input Transformation
    - **Context** as the *first* sentence
    - **Opinion Target** as the *second* sentence
  - Example

<table>
<thead>
<tr>
<th>Opinion Target</th>
<th>BERT Input</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgina Hermitage</td>
<td>[CLS] $T$ is a #one2watch since she broke the 400m T37 WR. [SEP] Georgina Hermitage [SEP]</td>
<td>Positive</td>
</tr>
<tr>
<td>400m T37</td>
<td>[CLS] Georgina Hermitage is a #one2watch since she broke the $T$ WR. [SEP] 400m T37 [SEP]</td>
<td>Neutral</td>
</tr>
</tbody>
</table>

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Methodology -- BERT

- Apply BERT to TSC
  - Feed the transformed sentence to BERT

(a). Georgina Hermitage

[CLS] $T$ is a #one2watch since she broke the 400m T37 WR. [SEP] Georgina Hermitage [SEP]

(b). 400m T37

[CLS] Georgina Hermitage is a #one2watch since she broke the $T$ WR. [SEP] 400m T37 [SEP]

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Methodology -- multimodal BERT (mBERT)

- **Limitation**
  - Image features are not sensitive to opinion targets

  - **Georgina Hermitage**
  - **400m T37**

> $T$ is a #one2watch since she broke the 400m T37 WR.

Image: Georgina Hermitage

Methodology -- multimodal BERT (mBERT)
Methodology -- Target-oriented mBERT (TomBERT)

- **Target Attention**
  - Target as queries, images as keys and values
Methodology -- Target-oriented mBERT (TomBERT)

- Full Model

[Diagram of the full model with layers and connections labeled as follows:
- **Input Embedding**
- **ResNet**
- **Target Encoder**
- **Target Embedding**
- **Sentence Encoder**
- **Multimodal Encoder**
- **Target Image Matching**
- **Pooling & Linear & Softmax**
- **Add & Norm**
- **Multimodal Attention**
- **Self Attention**
- **Feed Forward**
- **Add & Norm**
- **CONCAT**
- **Multimodal Learning**

Textual content:

**[CLS]** $T$ is a #one2watch since she broke the 400m T37 WR. [SEP] Georgina Hermitage [SEP]
Experiments

- Two Multimodal Datasets

<table>
<thead>
<tr>
<th>Modality</th>
<th>Data Set</th>
<th>#Training Samples</th>
<th>#Dev Samples</th>
<th>#Test Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>POS</td>
<td>NEG</td>
<td>NEU</td>
</tr>
<tr>
<td>Text+Image</td>
<td>Twitter-2015</td>
<td>928</td>
<td>368</td>
<td>1883</td>
</tr>
<tr>
<td></td>
<td>Twitter-2017</td>
<td>1508</td>
<td>416</td>
<td>1638</td>
</tr>
</tbody>
</table>

- The two multimodal Twitter datasets are based on two public multimodal Named Entity Recognition (NER) datasets.
### Experimental Results

**Results on the Two Multimodal Datasets**

<table>
<thead>
<tr>
<th>Modality</th>
<th>Method</th>
<th><strong>Twitter-2015</strong></th>
<th></th>
<th><strong>Twitter-2017</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Accuracy</td>
<td>Macro-F1</td>
<td>Accuracy</td>
<td>Macro-F1</td>
</tr>
<tr>
<td>Visual</td>
<td>Res-Target</td>
<td>59.88</td>
<td>46.48</td>
<td>58.59</td>
<td>53.98</td>
</tr>
<tr>
<td></td>
<td>AE-LSTM</td>
<td>70.30</td>
<td>63.43</td>
<td>61.67</td>
<td>57.97</td>
</tr>
<tr>
<td></td>
<td>MemNet</td>
<td>70.11</td>
<td>61.76</td>
<td>64.18</td>
<td>60.90</td>
</tr>
<tr>
<td></td>
<td>RAM</td>
<td>70.68</td>
<td>63.05</td>
<td>64.42</td>
<td>61.01</td>
</tr>
<tr>
<td></td>
<td>MGAN</td>
<td>71.17</td>
<td>64.21</td>
<td>64.75</td>
<td>61.46</td>
</tr>
<tr>
<td></td>
<td>BERT</td>
<td>74.15</td>
<td>68.86</td>
<td>68.15</td>
<td>65.23</td>
</tr>
<tr>
<td></td>
<td>BERT+BL</td>
<td>74.25</td>
<td>70.04</td>
<td>68.88</td>
<td>66.12</td>
</tr>
</tbody>
</table>

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
## Experimental Results

### Results on the Two Multimodal Datasets

<table>
<thead>
<tr>
<th>Modality</th>
<th>Method</th>
<th>Twitter-2015</th>
<th></th>
<th>Twitter-2017</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Accuracy</td>
<td>Macro-F1</td>
<td>Accuracy</td>
<td>Macro-F1</td>
</tr>
<tr>
<td>Visual</td>
<td>Res-Target</td>
<td>59.88</td>
<td>46.48</td>
<td>58.59</td>
<td>53.98</td>
</tr>
<tr>
<td>Text</td>
<td>AE-LSTM</td>
<td>70.30</td>
<td>63.43</td>
<td>61.67</td>
<td>57.97</td>
</tr>
<tr>
<td></td>
<td>MemNet</td>
<td>70.11</td>
<td>61.76</td>
<td>64.18</td>
<td>60.90</td>
</tr>
<tr>
<td></td>
<td>RAM</td>
<td>70.68</td>
<td>63.05</td>
<td>64.42</td>
<td>61.01</td>
</tr>
<tr>
<td></td>
<td>MGAN</td>
<td>71.17</td>
<td>64.21</td>
<td>64.75</td>
<td>61.46</td>
</tr>
<tr>
<td></td>
<td>BERT</td>
<td>74.15</td>
<td>68.86</td>
<td>68.15</td>
<td>65.23</td>
</tr>
<tr>
<td></td>
<td>BERT+BL</td>
<td>74.25</td>
<td>70.04</td>
<td>68.88</td>
<td>66.12</td>
</tr>
<tr>
<td>Text + Visual</td>
<td>Res-MGAN</td>
<td>71.65</td>
<td>63.88</td>
<td>66.37</td>
<td>63.04</td>
</tr>
<tr>
<td></td>
<td>Res-MGAN-TFN</td>
<td>70.30</td>
<td>64.14</td>
<td>64.10</td>
<td>59.13</td>
</tr>
<tr>
<td></td>
<td>Res-BERT+BL</td>
<td>75.02</td>
<td>69.21</td>
<td>69.20</td>
<td>66.48</td>
</tr>
<tr>
<td></td>
<td>Res-BERT+BL-TFN</td>
<td>73.58</td>
<td>68.74</td>
<td>67.18</td>
<td>64.29</td>
</tr>
<tr>
<td></td>
<td>mBERT</td>
<td>75.31</td>
<td>70.18</td>
<td>69.61</td>
<td>67.12</td>
</tr>
<tr>
<td></td>
<td>TomBERT</td>
<td>77.15</td>
<td>71.75</td>
<td>70.34</td>
<td>68.03</td>
</tr>
</tbody>
</table>

Jianfei Yu and Jing Jiang. Adapting BERT for Target-Oriented Multimodal Sentiment Classification. In IJCAI 2019.
Outline

- Transfer Learning
- Multi-Task Learning
- Multimodal Learning
- Summary
Summary

Cutting-Edge Dimensions

- Transfer Learning
  - Cross-Domain
  - Cross-Lingual

- Multi-task Learning
  - Aspects/Opinions Co-Extraction

- Multimodal Learning
  - ABSA
  - ABSC

State-of-the-art Methods:
- Attention-based LSTM models
- BERT-based models
Thank you!