SpanEmo: Casting Multi-label Emotion Classification as Span-Prediction

Hassan Alhuzali, Sophia Ananiadou

National Centre for Text Mining

Department of Computer Science, The University of Manchester, UK

Introduction

Task: Multi-label Emotion Classification, given an input–> select the most associated emotions with it

Motivation:

- No effective capturing of emotion-specific associations
  - e.g., "My brother is about to join the police academy and I'm not happy about and I'm not the only one"
- Treatment of individual emotions independently
  - Prior research in NLP [2] & psychological theories of emotions [3] –> discuss the idea of correlation between emotions.

Contributions

- a novel framework casting the task of multi-label emotion classification as a span-prediction problem.
- a loss function, modelling multiple co-existing emotions for each input sentence.
- experiments and analyses both at word/sentence-level, demonstrating the strength of our method for multi-label emotion classification across three languages (English, Arabic and Spanish).

Proposed Approach

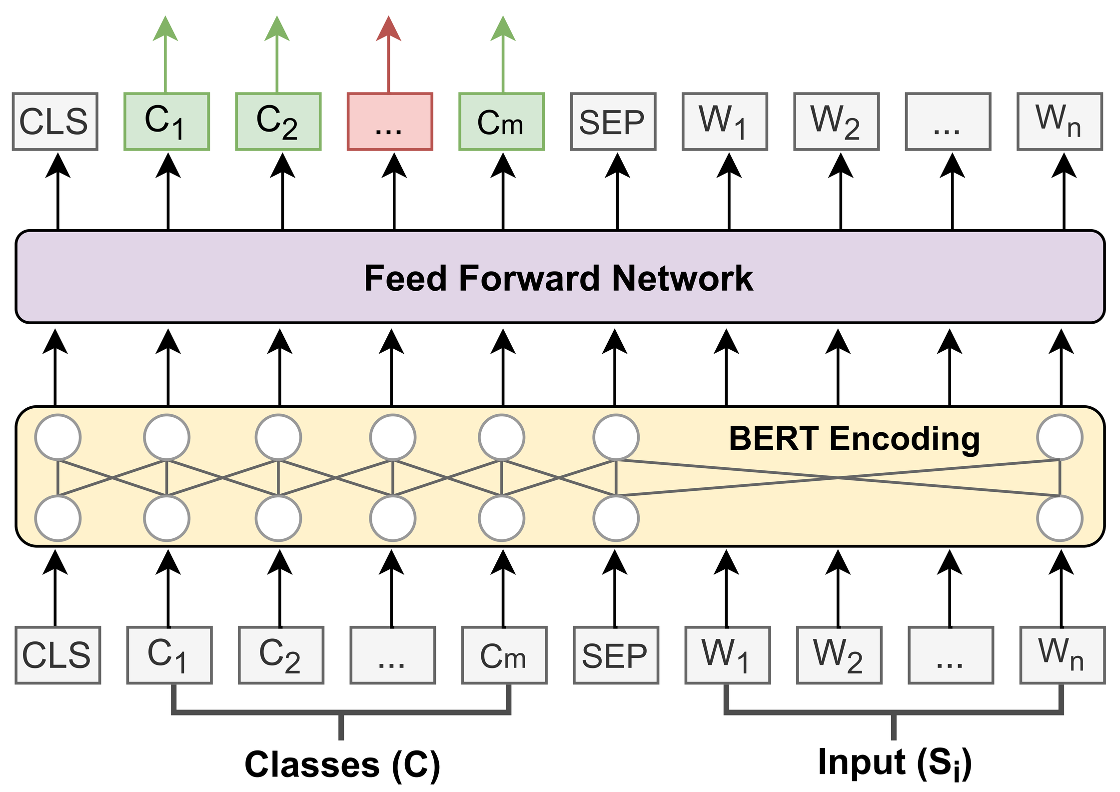

Figure 1: Illustration of our proposed framework (SpanEmo).

Framework

SpanEmo-Components (Associations)

1. **Construction of the input (Two segments)**
   - **Label Segment**: emotion classes –> descriptive names (C)
   - **Input Segment**: sentence (s_i)
   - The naming of our method comes from the idea of selecting a span of emotion classes from the label segment as the output.
2. **Encoder**
   - makes use of the two segments to allow the model to interpolate between both emotions and words.
   - also generates a hidden representation (h_i) both for words and emotions.

$$\mathrm{H}_i = \mathrm{Encoder}([\mathrm{CLS}] + |C| + [\mathrm{SEP}] + \mathbf{s}_i), \tag{1}$$

3. **Feed-Forward Network (FFN)**
   - consists of a non-linear hidden layer with a Tanh activation and a position vector –> transforms H_i into a single score.

$$\hat{\mathbf{y}} = \mathrm{sigmoid}(\mathrm{FFN}(\mathbf{H}_i)), \tag{2}$$

SpanEmo (Correlations)

1. **Correlation Loss**
   - We make use of label correlation aware loss [4]:
     - to maximise the distance between positive and negative label sets.

$$\mathcal{L}_{\mathrm{LCA}}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{|\mathbf{y}^0|\,|\mathbf{y}^1|} \sum_{(p,q)\in \mathbf{y}^0 \times \mathbf{y}^1} \exp\left(\hat{\mathbf{y}}_p - \hat{\mathbf{y}}_q\right), \tag{3}$$

where y^0 denotes the set of negative labels, while y^1 denotes the set of positive labels.

2. **Training Objective**
   - We combined the correlation loss with binary cross-entropy (BCE) and trained them jointly:
     - to help correlation loss to take advantage of the CE loss through maximisation of the probability of the correct labels.

$$\mathcal{L} = (1-\alpha)\mathcal{L}_{\mathrm{BCE}} + \alpha \sum_{i=1}^{M} \mathcal{L}_{\mathrm{LCA}}, \tag{4}$$

Experiments

Settings

Dataset	Training
SemEval2018 [1]	Adam Optimizer
11 emotions	Dropout
3 Langs. (En, Ar, Es)	Early Stopping

Results (English)

Mod/Met	miF1	maF1	jacS
JBNN	0.632	0.528	-
RERc	0.651	0.539	-
DATN	-	0.551	0.583
NTUA	0.701	0.528	0.588
BERT	0.695	0.520	0.570
BERT+DK	**0.713**	0.549	0.591
BERT-GCN	0.707	0.563	0.589
LEM	0.675	0.567	-
SpanEmo	**0.713**	**0.578**	**0.601**
- $\mathcal{L}$ (LCA)	0.712	0.564	0.590
- $\mathcal{L}$ (BCE)	0.698	**0.583**	0.582
- Label Seg.	0.695	0.520	0.570

Table 1: Performance on SemEval2018 test set.

Learning Association (Word-Level)

Emos	Top 10 Word
anger	anger pissed wrath idiots dammit kicking irritated thrown smashed
disgust	disgusting smashed gross hate pissed wrath dirty awful vile dumb
fear	nervous fear terror frightening afraid frown panic terrifying scary
joy	happy excitement joyful congrats glad delightful excited smiling
love	love sweetness loved hug mate lucky carefree shine care gracious
sadness	sad frown depressing saddened hurt weary upset sorrow hate

Table 2: Top 10 words associated with each corresponding emotion learned by SpanEmo.

Learning Association (Sentence-Level)

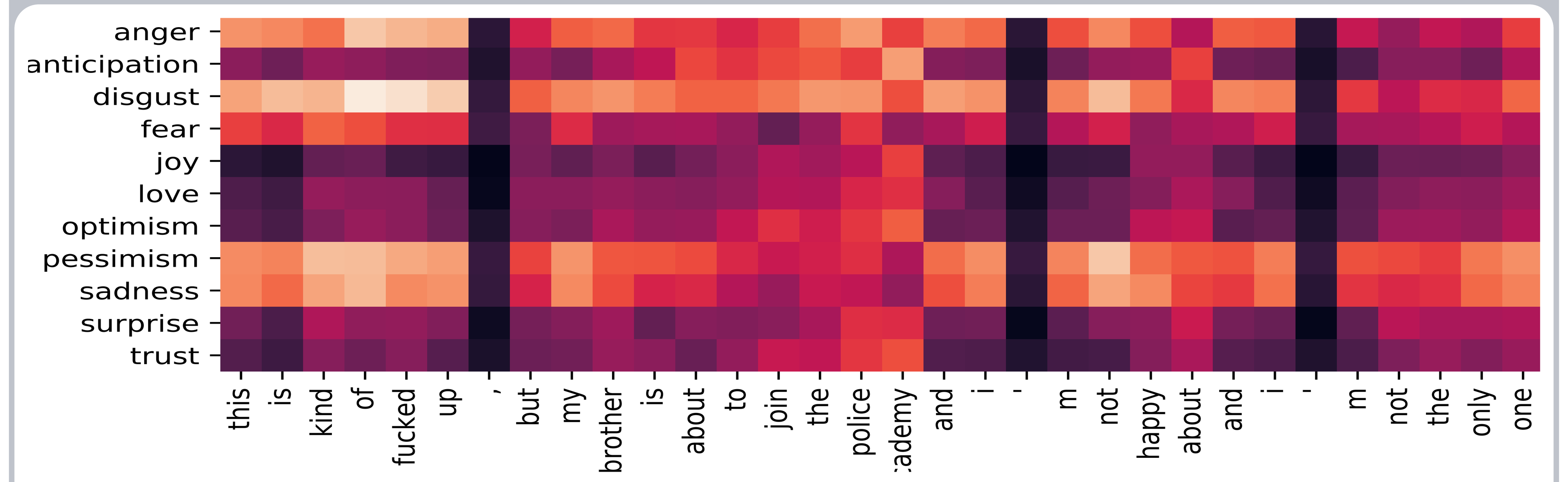

Figure 2: Visualisation on an example. Lighter colour indicates higher similarity.

Correlation

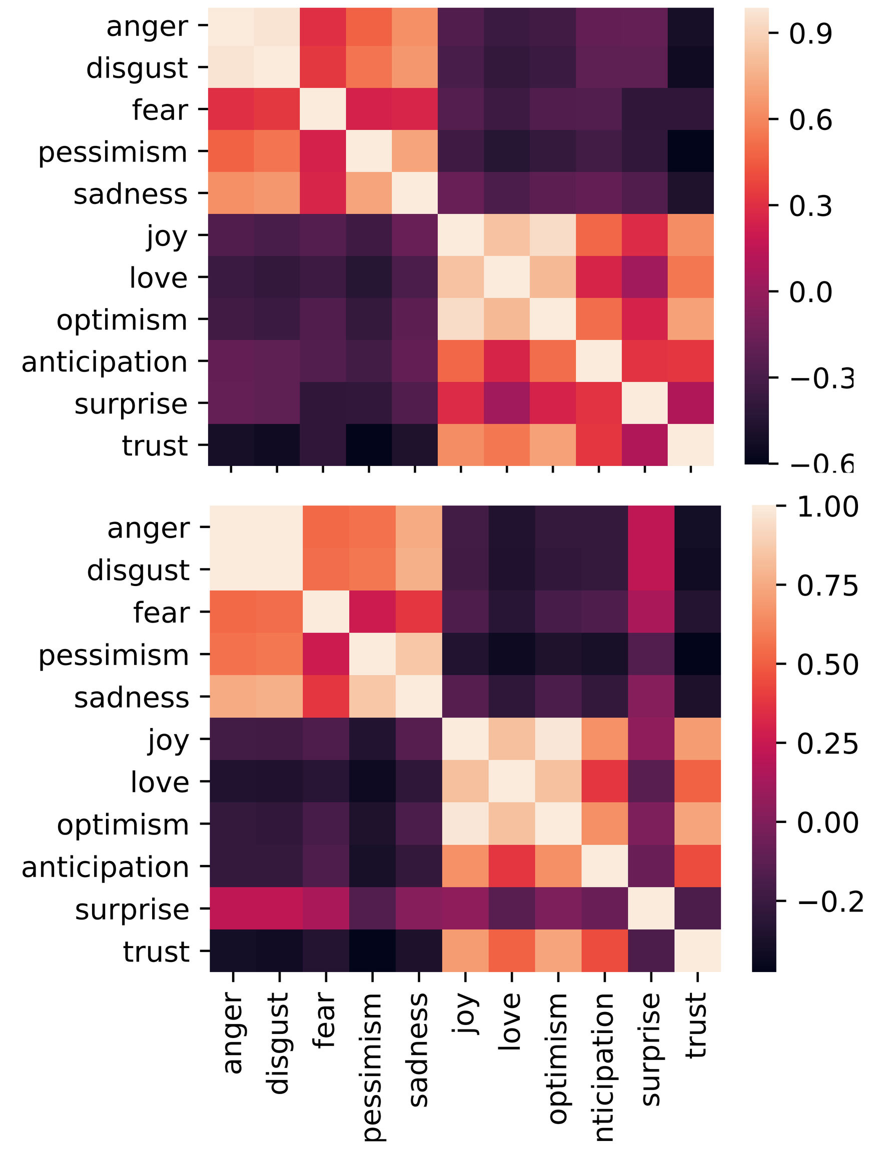

Figure 3: Top: emotion correlations (ground-truth). Bottom: emotion correlations (prediction)

Conclusions

- a novel framework "SpanEmo" aimed at casting multi-label emotion classification as a span-prediction problem, in which we focused on both learning emotion-specific associations and integrating the correlations between emotions into the loss function.
- extensive evaluations and analyses demonstrated the advantages of "SpanEmo" for multi-label emotion classification across three languages, without relying on external resources.

References

[1] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko. Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th international workshop on semantic evaluation*, pages 1–17, 2018.

[2] S. M. Mohammad and F. Bravo-Marquez. Emotion intensities in tweets. In *Proceedings of the sixth joint conference on lexical and computational semantics (*Sem)*, Vancouver, Canada, 2017.

[3] R. Plutchik. Emotions: A general psychoevolutionary theory. *Approaches to emotion*, 1984:197–219, 1984.

[4] C.-K. Yeh, W.-C. Wu, W.-J. Ko, and Y.-C. F. Wang. Learning deep latent spaces for multi-label classification. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 2838–2844, 2017.

Code: https://github.com/hasanhuz/SpanEmo

Email: hassan.alhuzali@postgrad.manchester.ac.uk

Funding