How Emotional are you?
Neural Architectures and Transfer Learning for Emotion Intensity Prediction in Microblogs

Devang Kulshreshtha*, Pranav Goel*, and Anil Kumar Singh

Indian Institute of Technology (BHU) Varanasi

The Task

- **Given:** A tweet.

The Task

- **Given:** A tweet.
- **Given:** The emotion conveyed by the tweet (four possibilities: Anger/Fear/Joy/Sadness).

The Task

- **Given:** A tweet.
  `I hate my lawnmower. If it had a soul, I'd condemn it to the fiery pits of Hell.'
- **Given:** The emotion conveyed by the tweet.
  Anger

The Task

- **Given:** A tweet.
  `I hate my lawnmower. If it had a soul, I'd condemn it to the fiery pits of Hell.'
- **Given:** The emotion conveyed by the tweet (four possibilities: *Anger*/Fear/Joy/Sadness).
- **Predict:** Emotional Intensity

The Task

- **Given:** A tweet.
  `I hate my lawnmower. If it had a soul, I'd condemn it to the fiery pits of Hell.'
- **Given:** The emotion conveyed by the tweet (four possibilities: *Anger*/Fear/Joy/Sadness).
- **Predict:** Emotional Intensity
  (*Annotated Intensity (gold label)*: 0.833)

The Task

- **Given:** A tweet.
  `I hate my lawnmower. If it had a soul, I'd condemn it to the fiery pits of Hell.'
- **Given:** The emotion conveyed by the tweet (four possibilities: *Anger*/Fear/Joy/Sadness).
- **Predict:** Emotional Intensity
  (*Annotated Intensity (gold label):* 0.833)
- This is the EmoInt shared task ([1]) conducted at WASSA 2017.

Motivation

Motivation

- In addition to the emotion itself, the associated intensity matters!

Motivation

- In addition to the emotion itself, the associated intensity matters!

Which tweet expresses **MORE** sadness?

‘When you lose somebody close to your heart you lose yourself as well. Crying my heart out!!!’

‘All friends are out of town. Feeling a bit lonely...’

Motivation

- In addition to the emotion itself, the associated intensity matters!

Which tweet expresses **MORE** sadness?

‘When you lose somebody close to your heart you lose yourself as well. Crying my heart out!!!’ ☹ ☹

‘All friends are out of town. Feeling a bit lonely...’ ☹

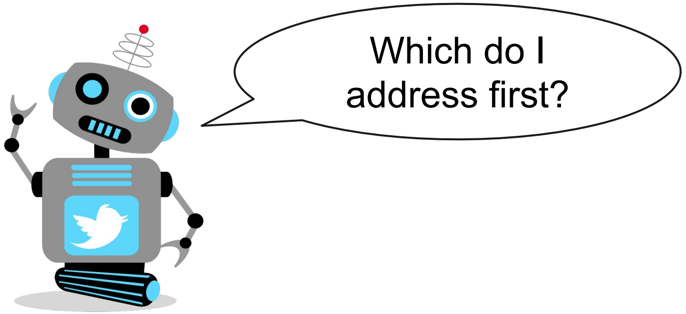
Motivation

- Tweeted reviews of a product/service.

Motivation

- Tweeted reviews of a product/service.

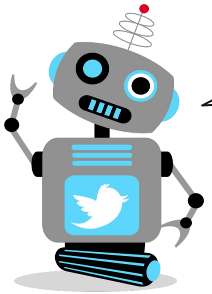
Angry Rant 1 Angry Rant 2



Motivation

- Tweeted reviews of a product/service.

Angry Rant 1 Angry Rant 2



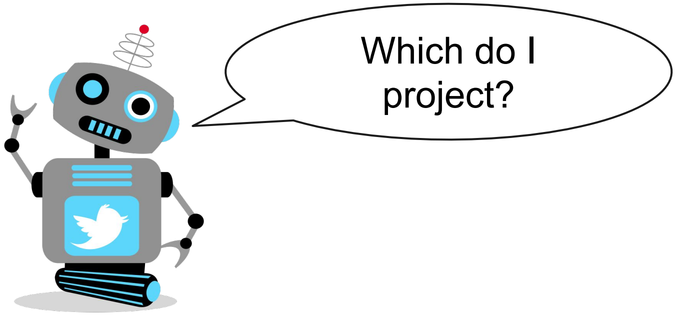
Which do I address first?

Compare the ***intensity* of anger**

Motivation

- Tweeted reviews of a product/service.

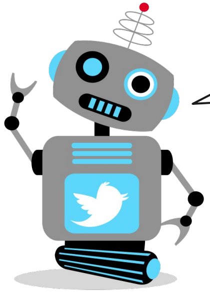
Glowing Review 1 Glowing Review 2



Motivation

- Tweeted reviews of a product/service.

Glowing Review 1 Glowing Review 2



Which do I project?

Compare the ***intensity* of joy**

Dataset

Link: http://saifmohammad.com/WebDocs/EmoInt\%20Train\%20Data/

Emotion	Train	Dev	Test	All
Anger	857	84	760	1701
Fear	1147	110	995	2252
Joy	823	74	714	1161
Sadness	786	74	673	1533
All	3613	342	3142	7097

Approaches/Models

1. LE-PC-DNN
2. LE-PC-DMTL
3. LE-PC-DMTL-EI

Deep Multi-Task Learning based

LE-PC-DNN model

It stands for **L**exicon + **E**moji based features used in a **P**arallel **C**ombination of **D**eep **N**eural **N**etworks

LE-PC-DNN model

It stands for **L**exicon + **E**moji based features used in a **P**arallel **C**ombination of **D**eep **N**eural **N**etworks

A neural network based architecture combining -

- Convolutional layers
- Fully-connected layers

LE-PC-DNN model

It stands for **L**exicon + **E**moji based features used in a **P**arallel **C**ombination of **D**eep **N**eural **N**etworks

A neural network based architecture combining -

- Convolutional layers
- Fully-connected layers
- Linguistic features (from external lexicons)
- Pre-trained CNN activations (trained for Emoji Detection Task; **Transfer Learning**)

LE-PC-DNN model

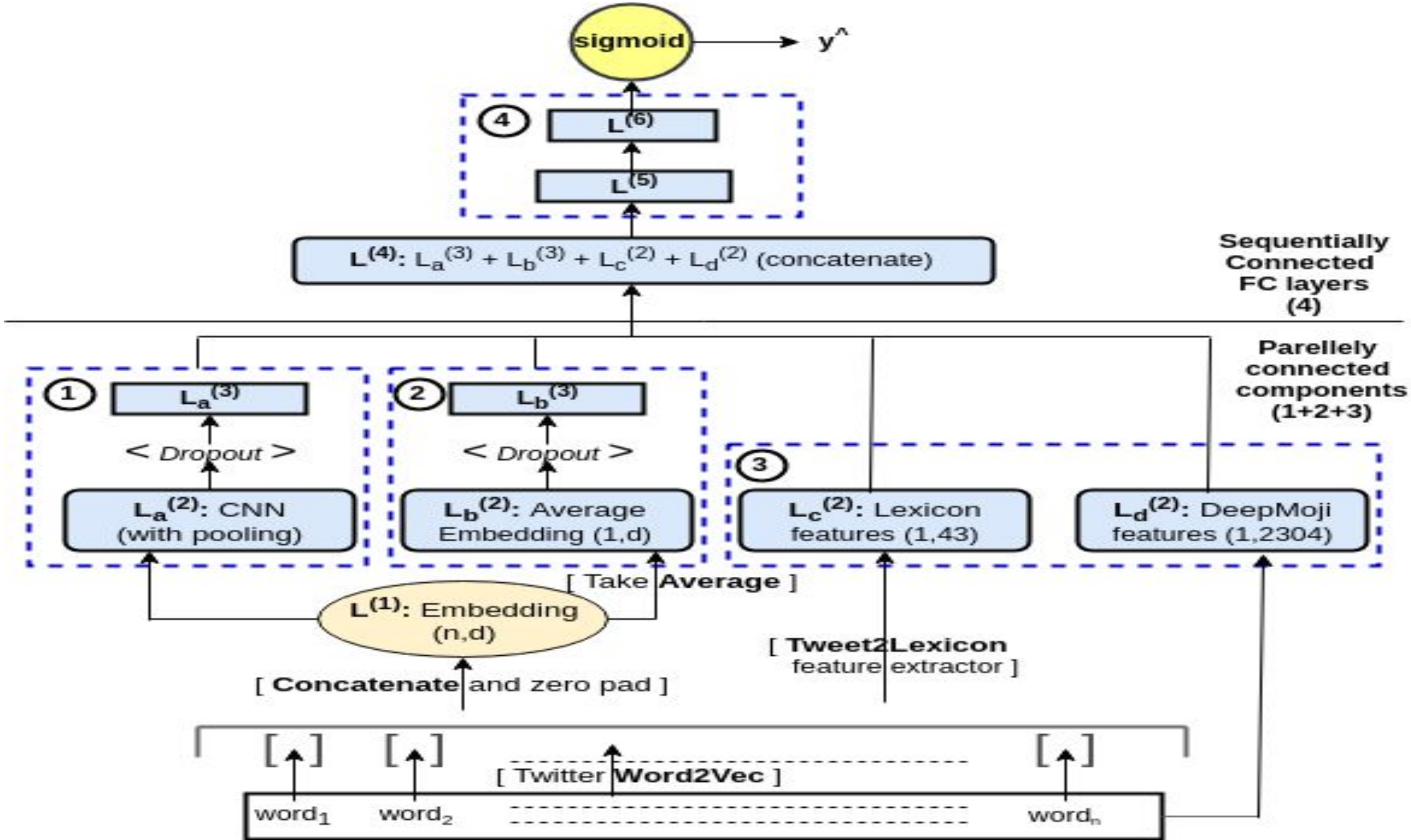
It stands for **L**exicon + **E**moji based features used in a **P**arallel **C**ombination of **D**eep **N**eural **N**etworks

A neural network based architecture combining -

- Convolutional layers
- Fully-connected layers
- Linguistic features (from external lexicons)
- Pre-trained CNN activations (trained for Emoji Detection Task; **Transfer Learning**)

combined in a **non-sequential or parallel** manner!

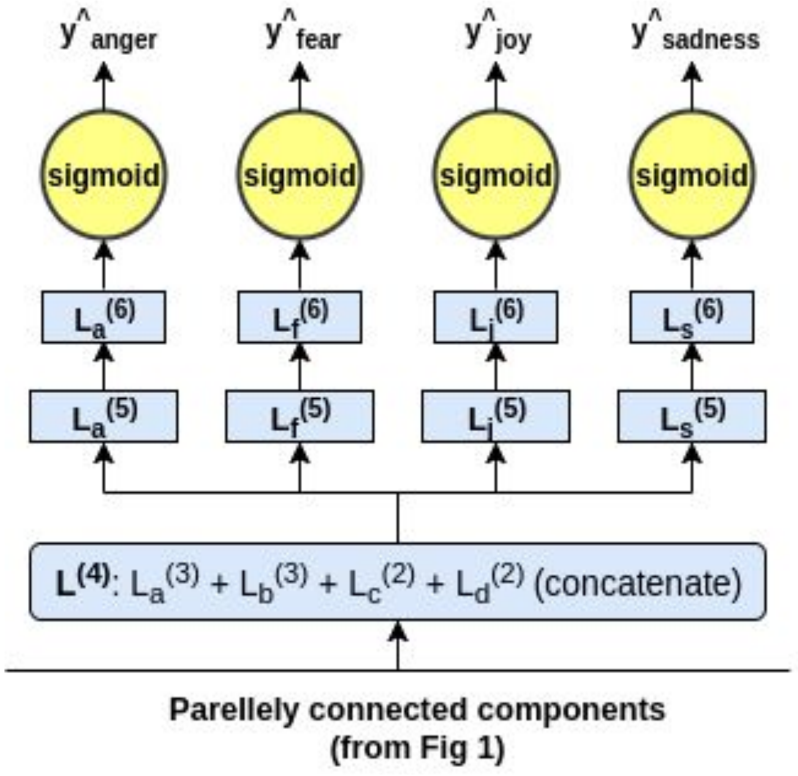
Fig 1: Our proposed LE-PC-DNN model for Emotion Intensity Prediction



LE-PC-DMTL model

It stands for **L**exicon + **E**moji based features based **P**arallel **C**onnected **D**eep **M**ulti-**T**ask **L**earning neural network.

Fig 2a: Our proposed LE-PC-DMTL model for Emotion Intensity Prediction



LE-PC-DMTL-EI model

It stands for **L**exicon + **E**moji based features based **P**arallel **C**onnected **D**eep **M**ulti-**T**ask **L**earning neural network optimized for the task of **E**motion **I**ntensity detection.

Allows grouping of similar emotions and separate branching of unrelated tasks.

Fig 2: Our proposed LE-PC-DMTL and LE-PC-DMTL-EI models for Emotion Intensity Prediction

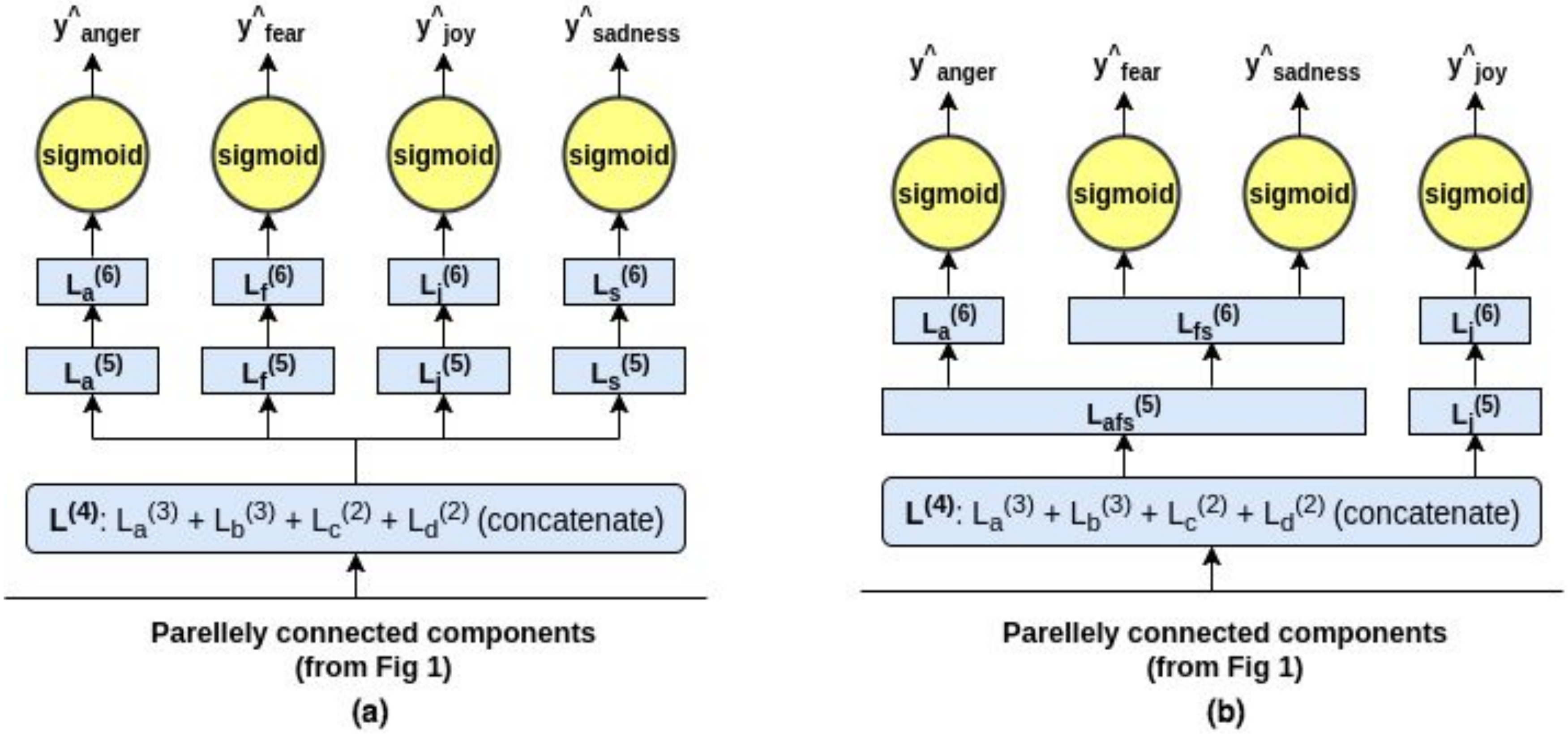
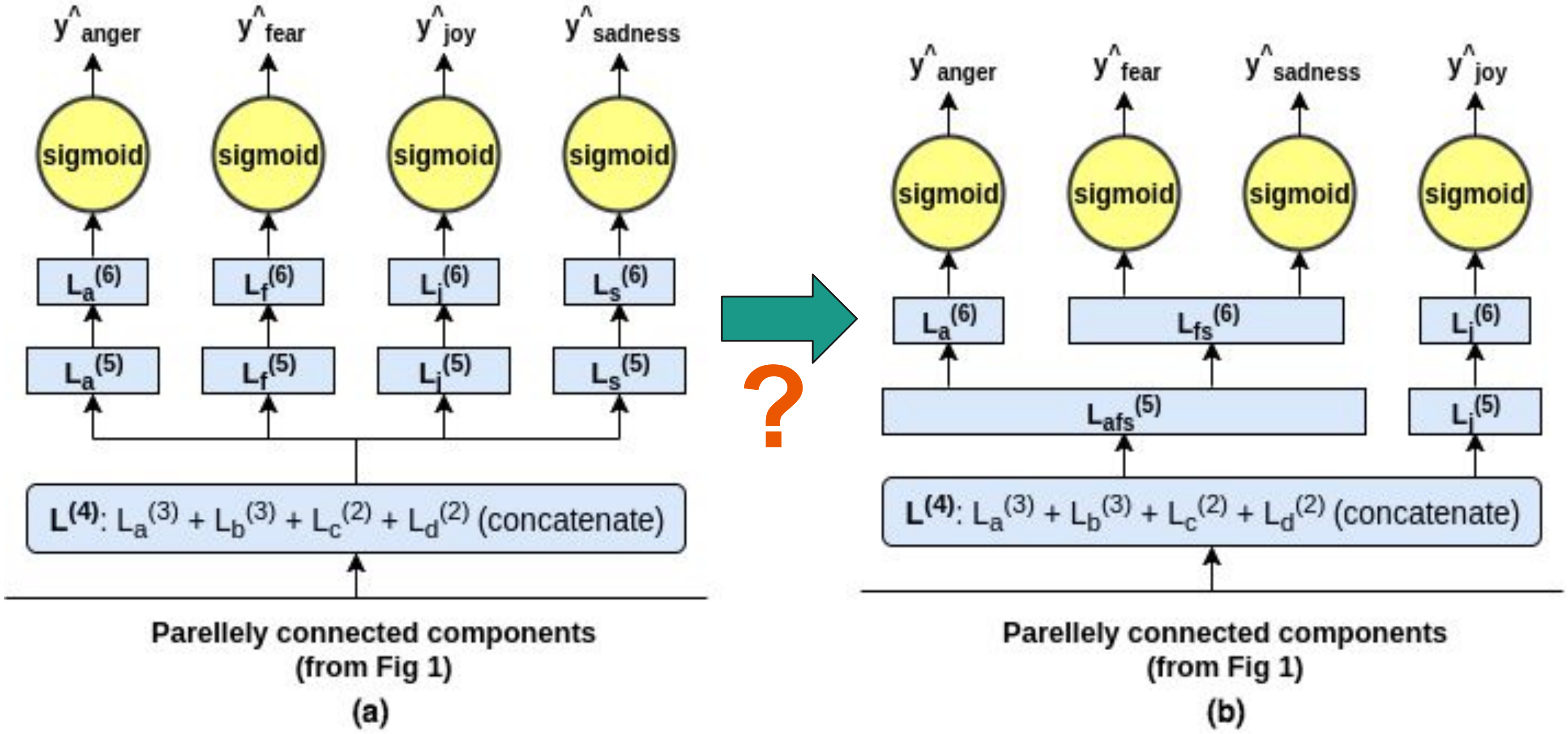
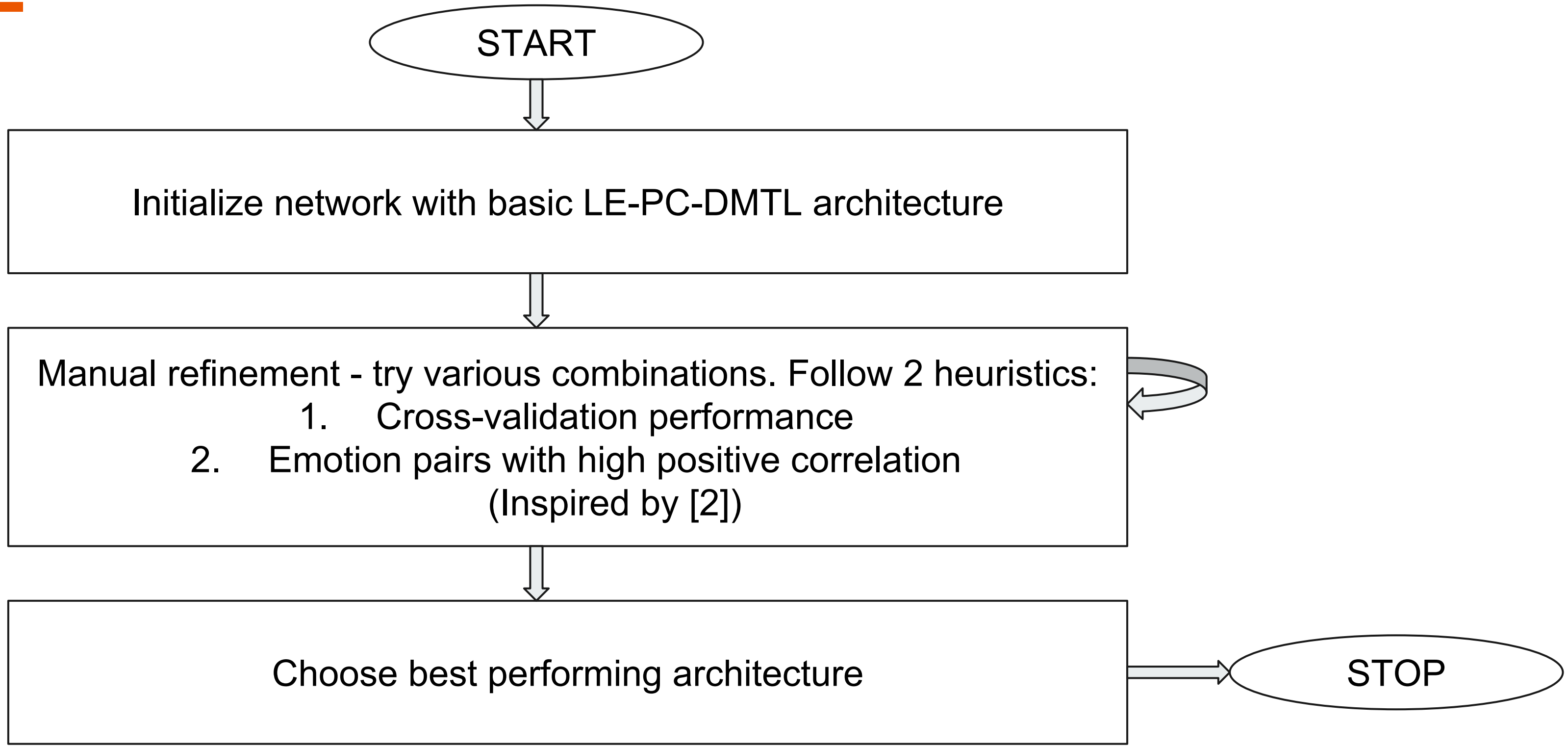


Fig 2: Our proposed LE-PC-DMTL and LE-PC-DMTL-EI models for Emotion Intensity Prediction



LE-PC-DMTL-EI : Learning the network architecture



Implementation Details

Word Embeddings

- Publicly available pre-trained word embeddings called the Twitter word2vec model.
- Trained on 400 million tweets using the word2vec approach.
- The large number of tweets in training data makes it a better choice than others available

Pre-processing

- Removal of URLs and user mentions.
- Stripping punctuations from word boundaries.
- Hashtag segmentation (`#wearethebest' to `we are the best') and elongation removal (`goooooood' to `good').

Implementation Details

Network Hyper-Parameters and Architecture Settings

- 7-fold cross validation on the combined training and validation sets.

- Changed various network hyperparameters like number of layers, output dimensionality, the type of pooling for CNNs (max versus average) and dropout value.

- Tried various RNN architectures including LSTM, Bidirectional LSTM and Gated Recurrent Units (GRU).

Implementation Details

__Training__

- Minimized the Mean Absolute Error between the actual and predicted intensity values.

- Optimize by back-propagating through layers via Mini-batch Gradient Descent.

- Use a batch size of 8, 25-30 training epochs and Adam optimization algorithm, with parameters set as α=0.001, β_1=0.9, β_2=0.999 and $\epsilon=10^{-9}$

Evaluation

- **Pearson correlation coefficient (r)** between gold ratings and predicted output intensities.
- Was used in the EmoInt shared task.

Evaluation

Baseline systems to benchmark our results -

Evaluation

Baseline systems to benchmark our results -

- **Weka:** EmoInt baseline. ([3])

Evaluation

Baseline systems to benchmark our results -

- **Weka:** EmoInt baseline. ([3])

(we also use Two of the top performing systems in the shared task)

- **Prayas:** Previous state-of-the-art. ([4])
- **IMS:** Rank #2 in EmoInt. ([5])

Results (Pearson Correlation %)

Emotion	LE-PC-DNN	Prayas	IMS	Weka	LE-PC-DMTL-EI	LE-PC-DMTL	MTL (Prayas)
Anger	**0.767**	0.765	0.767	0.639	0.745	0.731	0.645
Fear	**0.791**	0.732	0.705	0.652	0.773	0.755	0.677
Joy	**0.803**	0.762	0.726	0.654	0.785	0.758	0.654
Sadness	0.803	0.732	0.690	0.648	**0.808**	0.786	0.672
Average	**0.791**	0.747	0.722	0.648	0.778	0.758	0.662

State-of-the-Art Performance

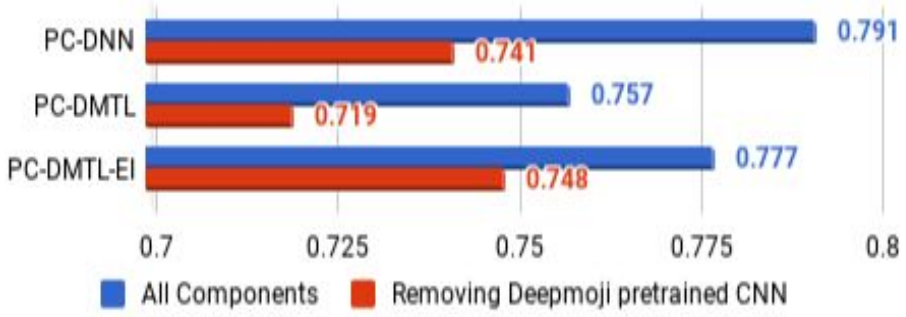
Results (Pearson Correlation %)

Emotion	LE-PC-DNN	Prayas	IMS	Weka	LE-PC-DMTL-EI	LE-PC-DMTL	MTL (Prayas)
Anger	**0.767**	0.765	0.767	0.639	0.745	0.731	0.645
Fear	**0.791**	0.732	0.705	0.652	0.773	0.755	0.677
Joy	**0.803**	0.762	0.726	0.654	0.785	0.758	0.654
Sadness	0.803	0.732	0.690	0.648	**0.808**	0.786	0.672
Average	**0.791**	0.747	0.722	0.648	0.778	0.758	0.662

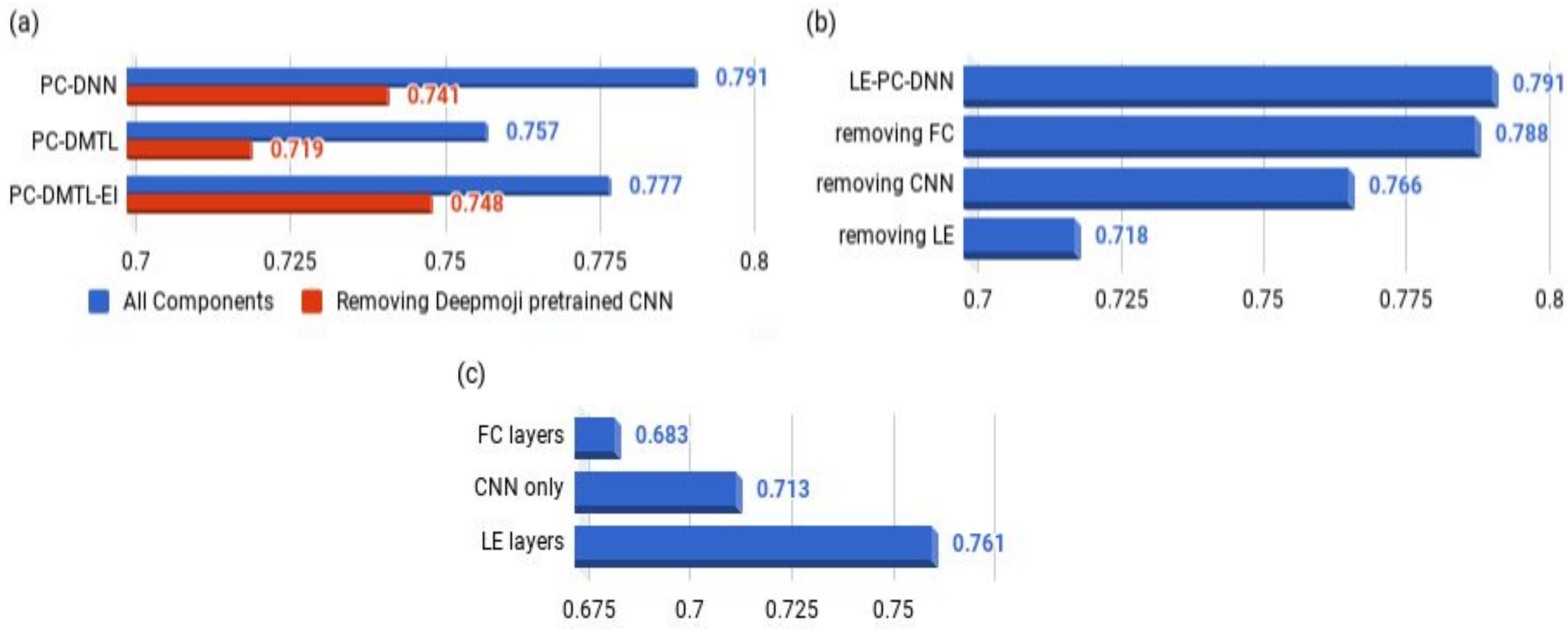
Multi-Task Learning for Emotion Intensity prediction

Ablation Tests

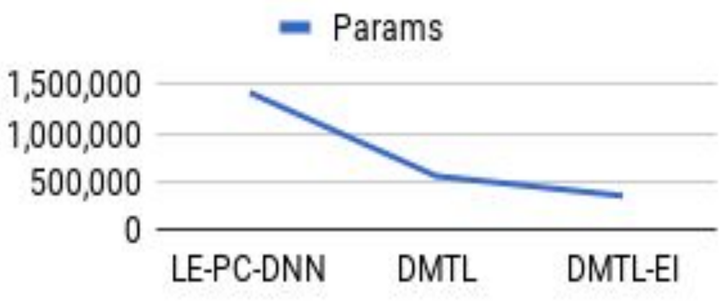
(a)

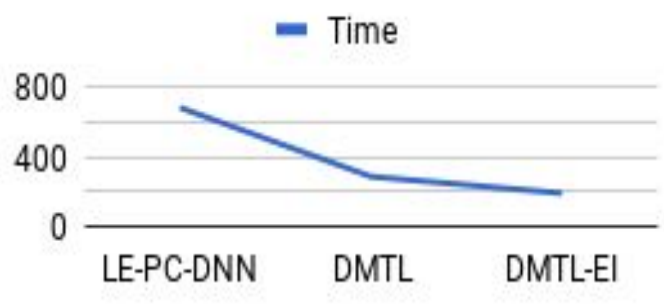


Ablation Tests



Practical Benefits of Multi-Task Learning



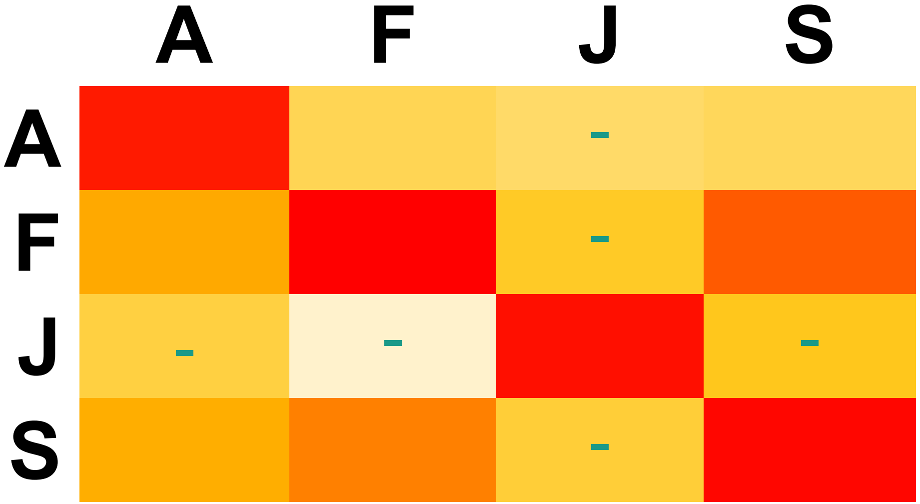


Pairwise Correlation between Emotions

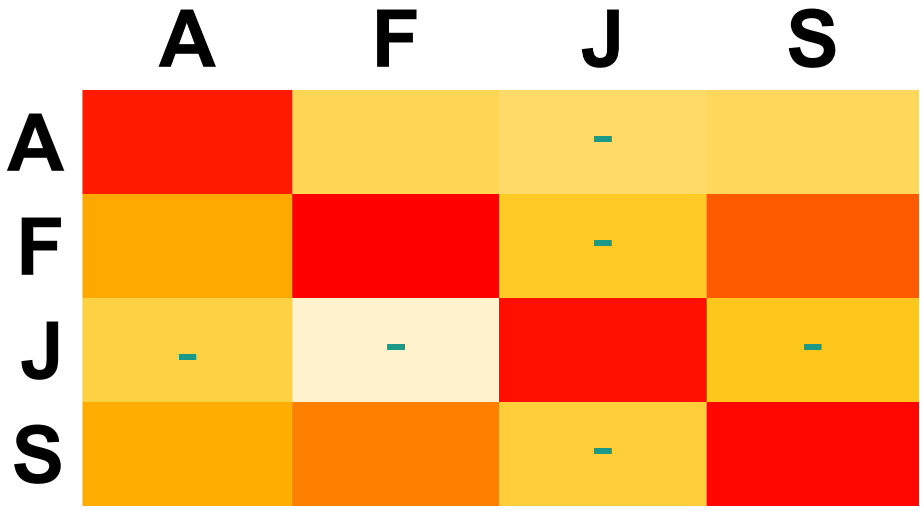
1. **Training on One Emotion and Testing on Another**

2. Training on a Combined Dataset of Two Emotions and Testing on Both of them

Training on One Emotion and Testing on Another

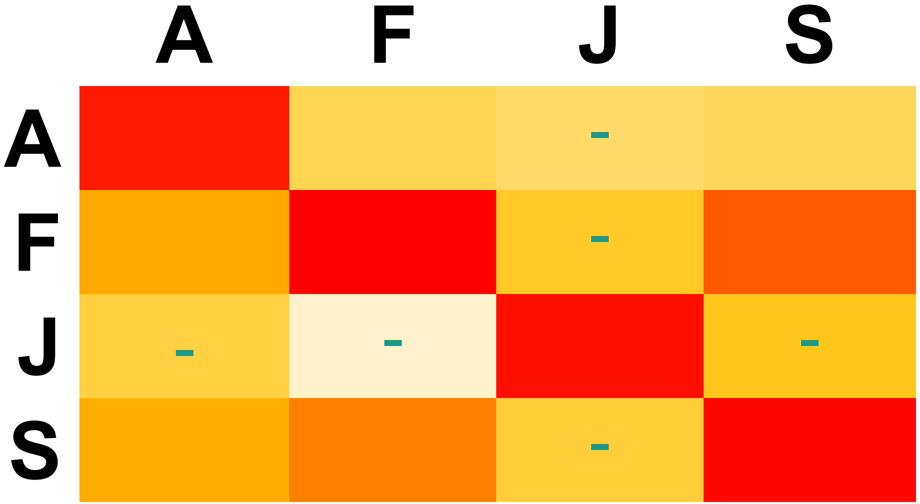


Training on One Emotion and Testing on Another



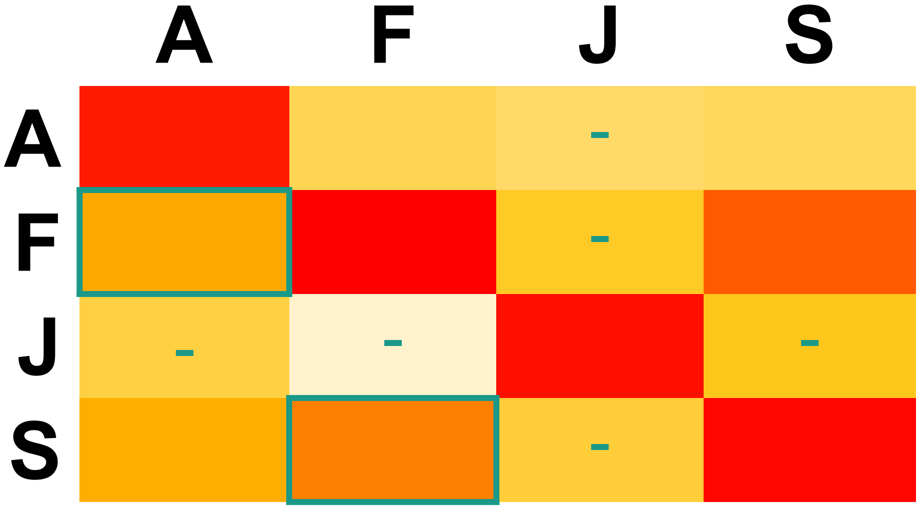
Significant correlation between all emotions (0.297 to 0.766)

Training on One Emotion and Testing on Another



Correlations are asymmetric

Training on One Emotion and Testing on Another



Fear and Sadness show highest correlation

Error Analysis

Tweet	Emotion	Actual Intensity	Predicted Intensity
Oi @THEWIGGYMESS you've absolutely fucking killed me.. 30 mins later im still crying with laughter.. Grindah.. Grindah...hahahahahahaha	Joy	0.846	0.469

Intensity Conveyed by Phrases, not Words

Error Analysis

Tweet	Emotion	Actual Intensity	Predicted Intensity
Oi @THEWIGGYMESS you've absolutely fucking killed me.. 30 mins later im still crying with laughter.. Grindah.. Grindah...hahahahahahaha	Joy	0.846	0.469
People are #hurt and #angry and it's hard to know what to do with that #anger Remember, at the end of the day, we're all #humans #bekind	Anger	0.25	0.568

Incorrect Modeling of the Full Tweet Context

Error Analysis

Tweet	Emotion	Actual Intensity	Predicted Intensity
T minus 10 hours till I meet with a designer who wants me to model his new fashion line !!!	Fear	0.667	0.291
Ibiza blues hitting me hard already wow	Sadness	0.833	0.447

Context Outside the tweet

Error Analysis

Tweet	Emotion	Actual Intensity	Predicted Intensity
Oi @THEWIGGYMESS you've absolutely fucking killed me.. 30 mins later im still ***crying with laughter***.. Grindah.. Grindah...hahahahahahaha	Joy	0.846	0.469
Ibiza blues hitting me hard already wow	Sadness	0.833	0.447

Metaphors

Future Work

- Topic Modeling

Future Work

- Topic Modeling
- Link to real-world applications (ranking product reviews)

References

1. Saif M. Mohammad and Felipe Bravo-Marquez. 2017b. WASSA-2017 shared task on emotion intensity. In Proceedings of the Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA), Copenhagen, Denmark.
2. Lu, Yongxi, et al. "Fully-adaptive Feature Sharing in Multi-Task Networks with Applications in Person Attribute Classification." *CVPR*. Vol. 1. No. 2. 2017.
3. Saif M Mohammad and Felipe Bravo-Marquez. 2017a. Emotion intensities in tweets
4. Goel, Pranav, et al. "Prayas at emoint 2017: An ensemble of deep neural architectures for emotion intensity prediction in tweets." *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. 2017.
5. Maximilian Koper, Evgeny Kim, and Roman Klinger. 2017. Ims at emoint-2017: Emotion intensity prediction ¨ with affective norms, automatically extended resources and deep learning. In Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 50–57.

Thank You

- Our code can be found at https://github.com/Pranav-Goel/Neural Emotion Intensity Prediction.
- Please feel free to reach out with queries to us at pranav.goel.cse14@iitbhu.ac.in, and devang.kulshreshtha.cse14@iitbhu.ac.in.