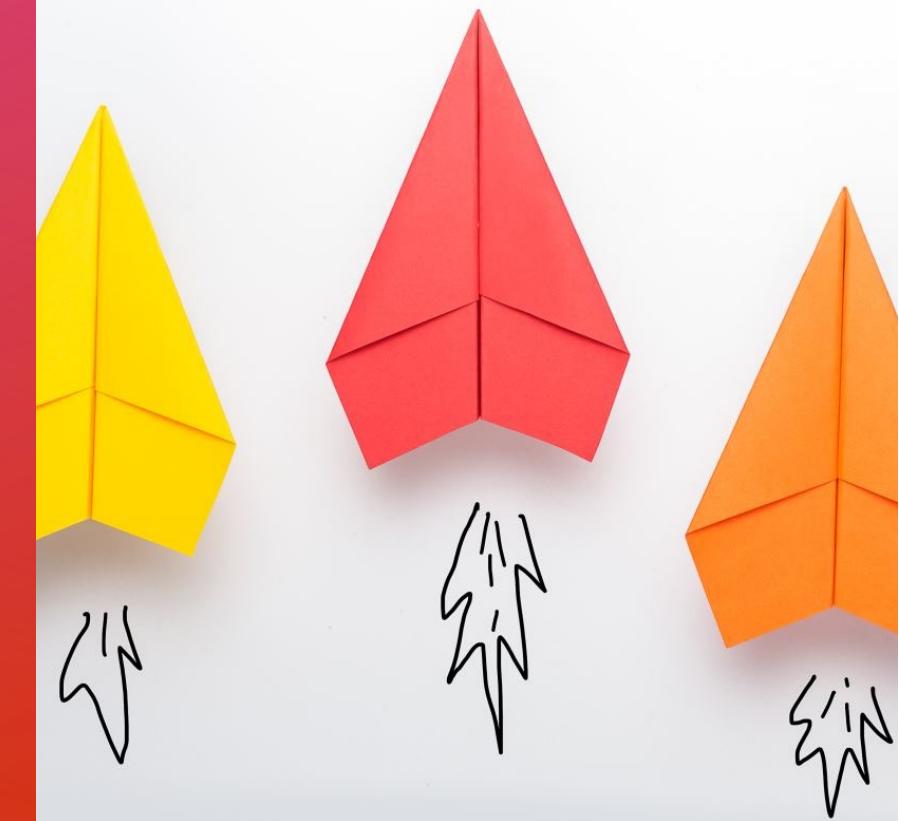


DETECTION OF EMPHASIS WORDS IN SHORT TEXTS – A CONTEXT AWARE LABEL DISTRIBUTION LEARNING APPROACH

MEGHANA AND BHASKAR JYOTI DAS
PRESENTED BY: MEGHANA



INTRODUCTION

Problem Statement

Given a brief textual excerpt, identify the candidates, i.e. the tokens, that require emphasis.

Emphasis

Stress given to a word or words when speaking to indicate particular importance.

Example: An **apple** a day keeps the **doctor** away!

Label Distribution vs. Multi label Classification

Label classification denotes 100% or 0% of one or more classes

Emphasis, Emotion, Personality Traits etc. are more fuzzy in nature. Hence Label Distribution is a better fit !

APPLICABILITY

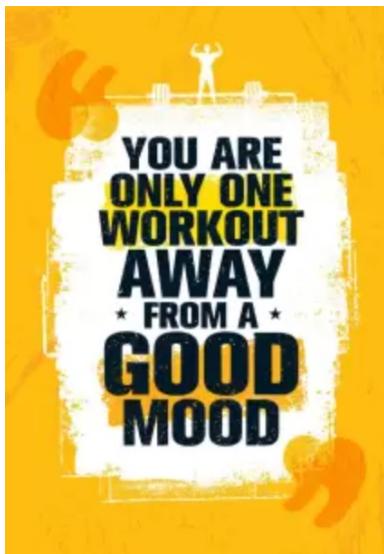
- Downstream tasks such as Emotion Detection, Stance Detection, etc.
- Multi-modal applications such as visual arts – posters and billboard design



Source: Google Images



Source: Google Images



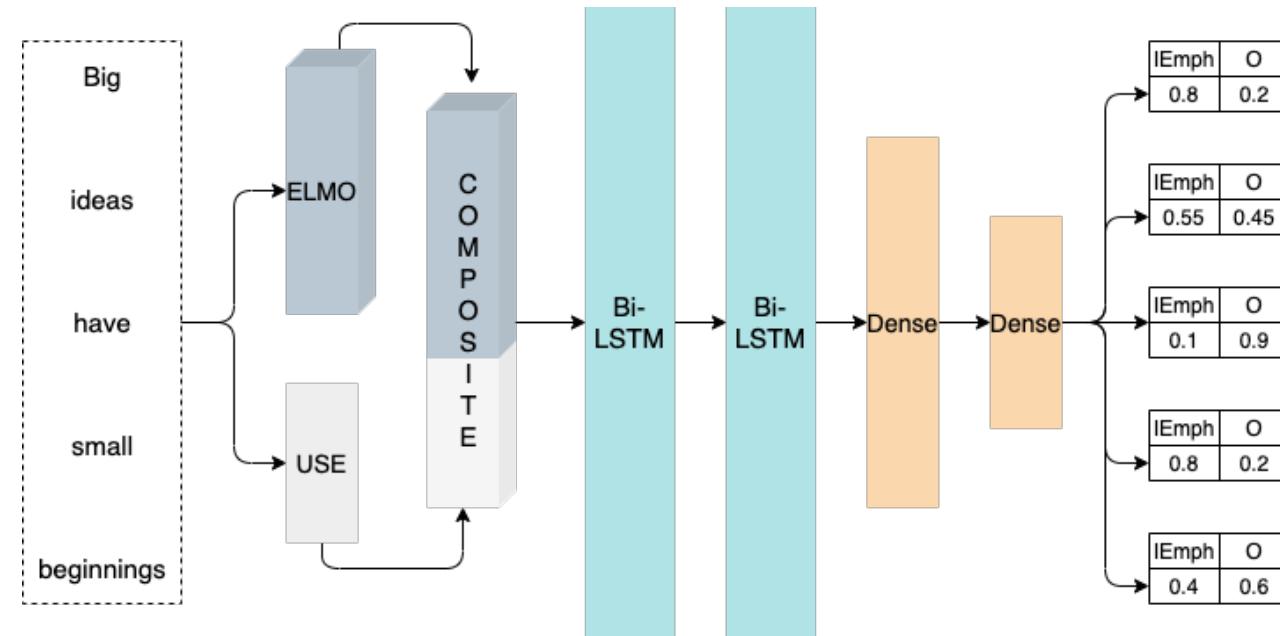
Source: Google Images



EARLIER WORK

- Emphasis Detection in Texts – A relatively unexplored area of research
- Good amount of research done on Emphasis Detection in speech, which helps in detecting emotion, intention, and attitude.

PROPOSED APPROACH



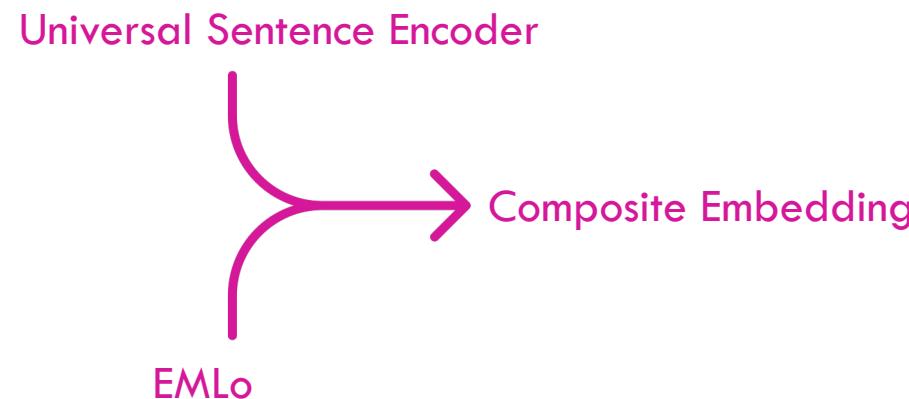
LABEL DISTRIBUTION LEARNING

- Each instance in LDL terminology is associated with a real-valued vector with each element representing the extent of association of the instance with a particular label.
- Each vector adds to 1, with each value in the range [0, 1]



COMPOSITE EMBEDDING

- Difficult to capture context of short texts.
- Composite embeddings used to capture token-level as well as global information



EVALUATION

- Compared across varying architectures and embeddings

Architectures

1. LDL
2. Multitask Learning: LDL as primary task and SLL (BiLSTM-CRF) as auxiliary task

Embeddings

1. GloVe
2. ELMo
3. ELMo + USE

- Evaluation metric employed: $Match_m = \frac{\sum_{x \in D} |L_m^x \cap \hat{L}_m^x| / \min(m, |x|)}{|D|}$

(If L_m^x and \hat{L}_m^x are sets comprising top m *IEmph* probabilities in the ground truth and predicted labels respectively, for a datapoint x in dataset D)

- Baseline: LDL with ELMo + Attention

R E S U L T S

Table 1. matchM scores (in percentage) for m=1 to m=8 for various architectures

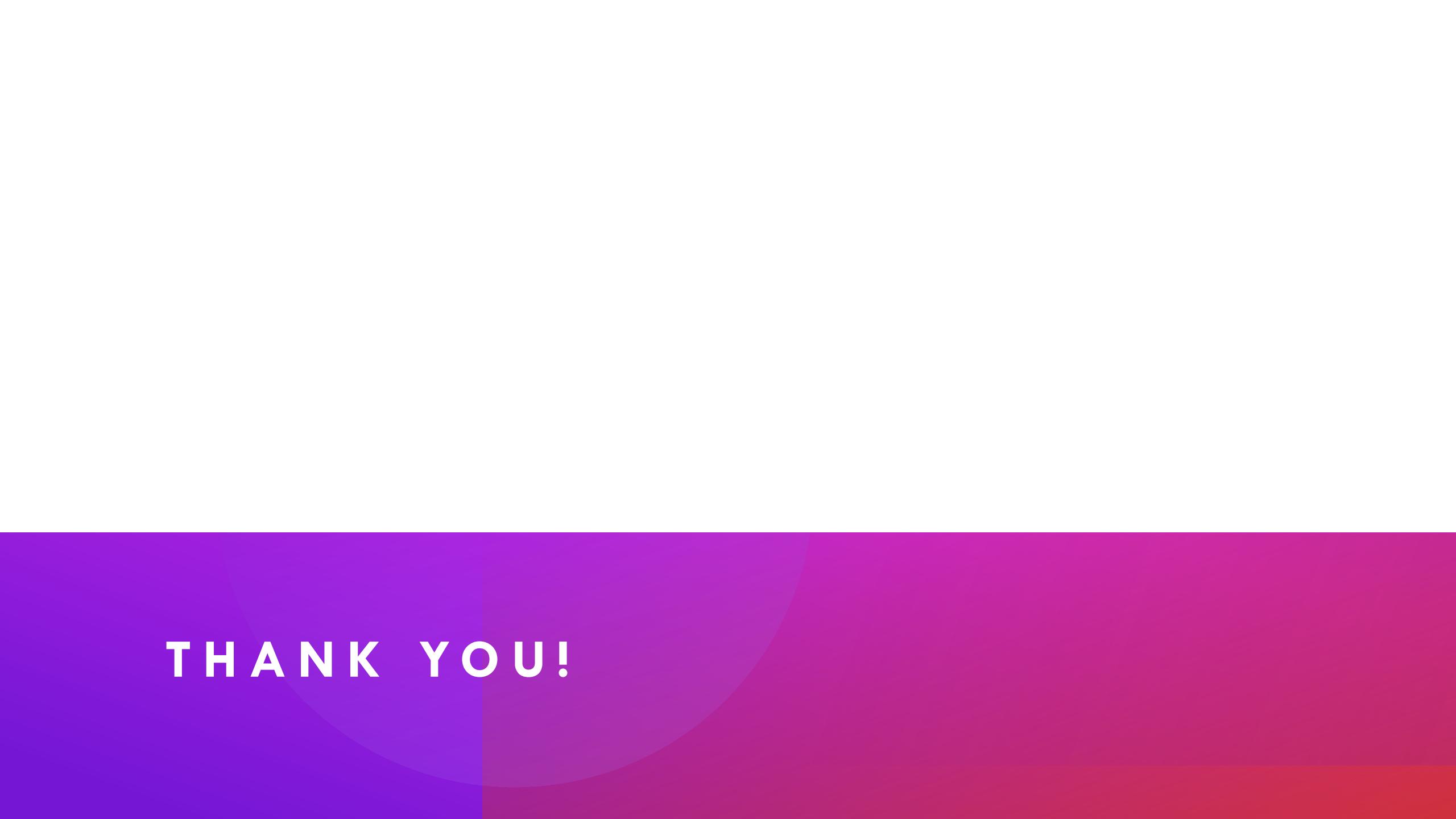
Model	Model Type	M=1	m=2	m=3	m=4	m=5	m=6	m=7	m=8
M1	LDL:GloVe+Att	57.5	69.7	76.7	80.7	-	-	-	-
M2	LDL:ELMo	0.6	71.7	78.7	84.1	-	-	-	-
M3	LDL:ELMo+Att	59.6	72.7	77.7	84.6	-	-	-	-
M4	MTL:ELMo	53.2	50.9	57.4	62.4	66.9	70.9	73.8	75.6
M5	MTL:ELMo+USE	70.5	75.7	80.1	83	85.6	87.6	88.9	89.7
M6	LDL:ELMo+USE	72.6	79.3	83.3	85.8	88.0	89.9	91.2	91.9

FUTURE WORK

- Exploration of utilizing emphasis detection in an NLP pipeline for downstream tasks such as emotion detection and multi-target stance
- Employing emphasis detection as an auxiliary task in a multi-task learning setup.
- Explore determining the degree of emphasis in text samples of other languages.



Questions?



THANK YOU!