Scuola universitaria professionale
della Svizzera italiana
SUPSI

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Neural Networks for sentence classification

Ambrosini Luca
luca.ambrosini@supsi.ch

Nicolò Giancarlo
giani1@inf.upv.es

Index

- Task
- Classification system:
  - Pre-processing
  - Text representation
  - Models
- Comparative study
- Results
- Conclusions

Task

- Classification Of Spanish Election Tweets

- 5 categories
  - Political issues
  - Policy issues
  - Personal issues
  - Campaign issues
  - Other issues

Dataset statistics

- 2242+250 labelled tweets
- Average sequence length:
  - 135 chars
  - 24 words
- Max sequence length:
  - 140 chars
  - 49 words

Metrics

$$F_{1-macro} = \frac{1}{|L|} \sum_{l \in L} F_1(y_l, \hat{y}_l)$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

First approaches

- Text representations:
  - Bag of words + TF-IDF
  - Bag of n-grams (bigrams, trigrams) + TF-IDF

- Classifiers:
  - Random Forest
  - Multi Layer Perceptron
  - Support Vector Machines

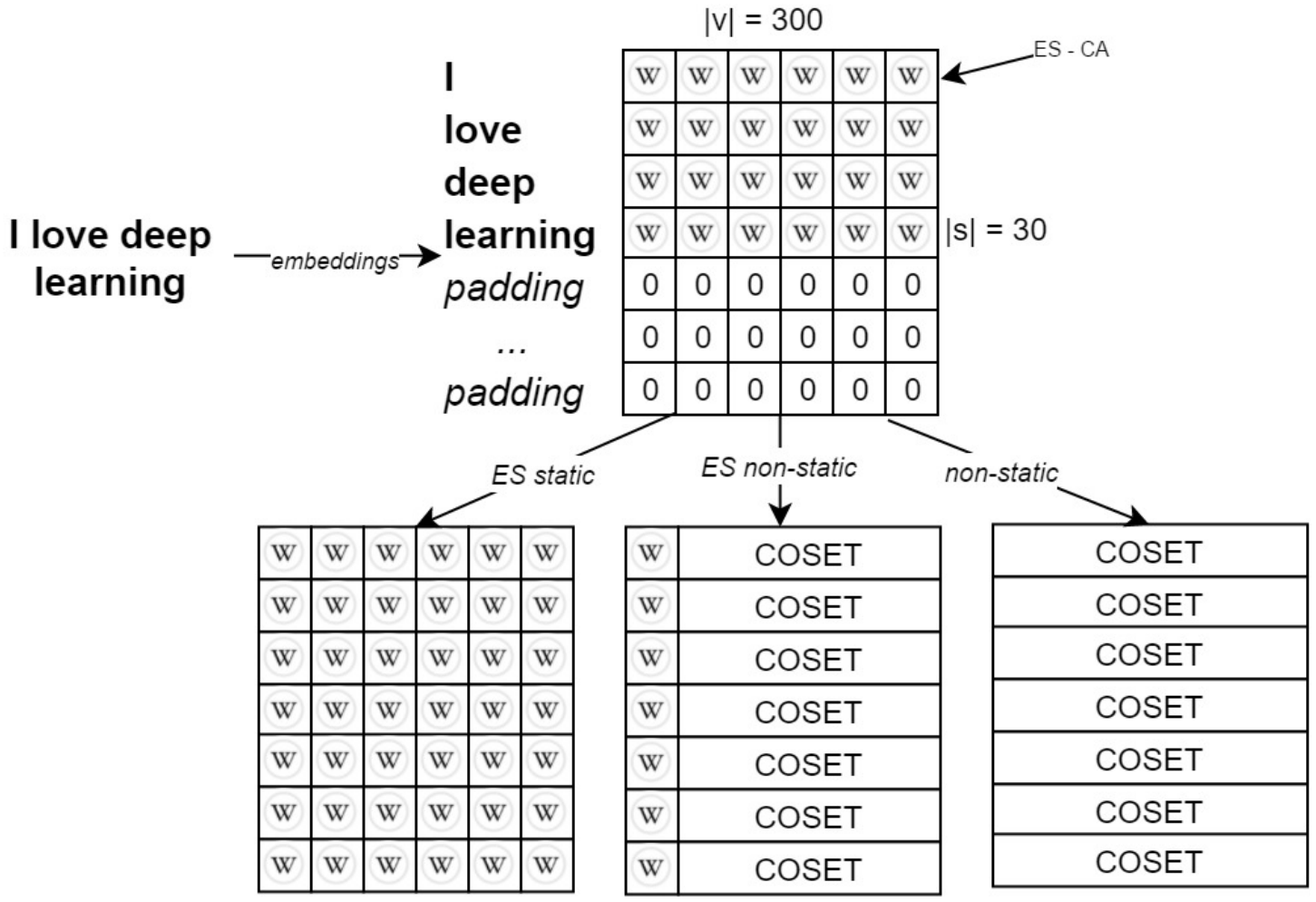
Pre-processing

- Stemming
- Stop words removal
- URLs
- Reserved words
- Emoji-Smiley
- Mentions
- Hashtags
- Numbers

I am @Ambros ☺ #atoppe http://google.com

I am $MENTION $EMOJI $HASHTAG

Word embeddings



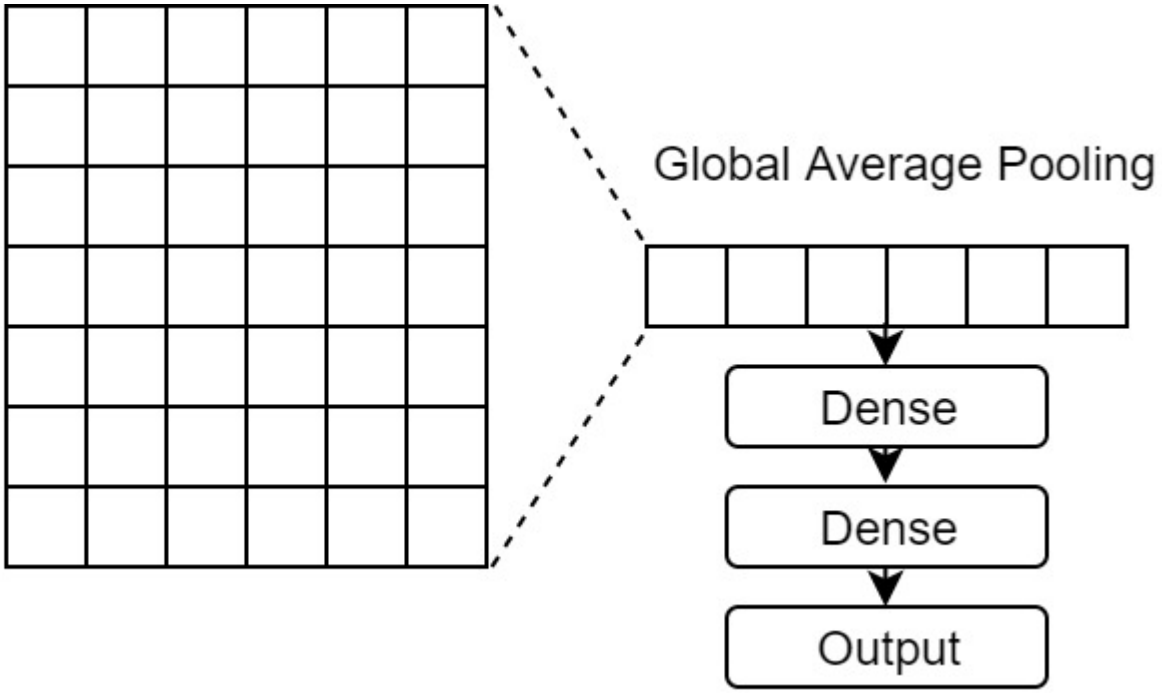
 Bojanowski, Piotr and Grave, Edouard and Joulin, Armand and Mikolov, Tomas. "Enriching Word Vectors with Subword Information"

Models

- LSTM
- Bi-LSTM
- CNN
- FastText
- Kim

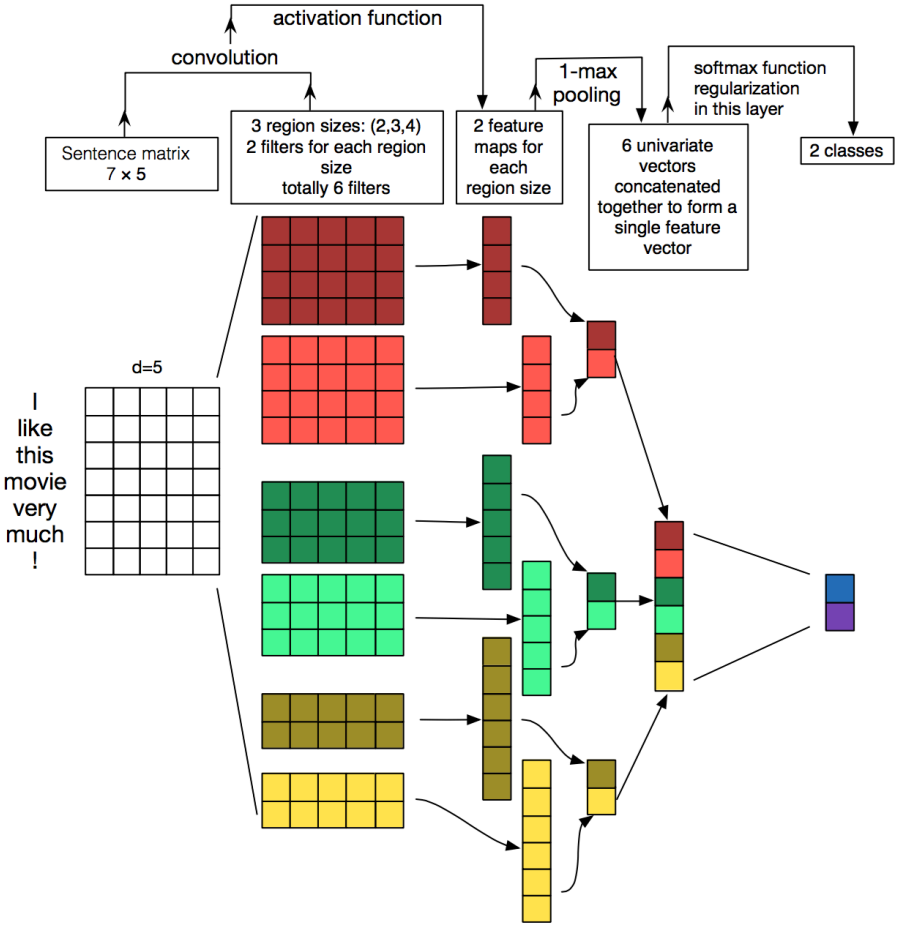
FastText

I
love
deep
learning
(I,love)
(love,deep)
(deep,learning)



Joulin, Armand, et al. "Bag of tricks for efficient text classification."

Kim



Kim, Yoon. "Convolutional neural networks for sentence classification."

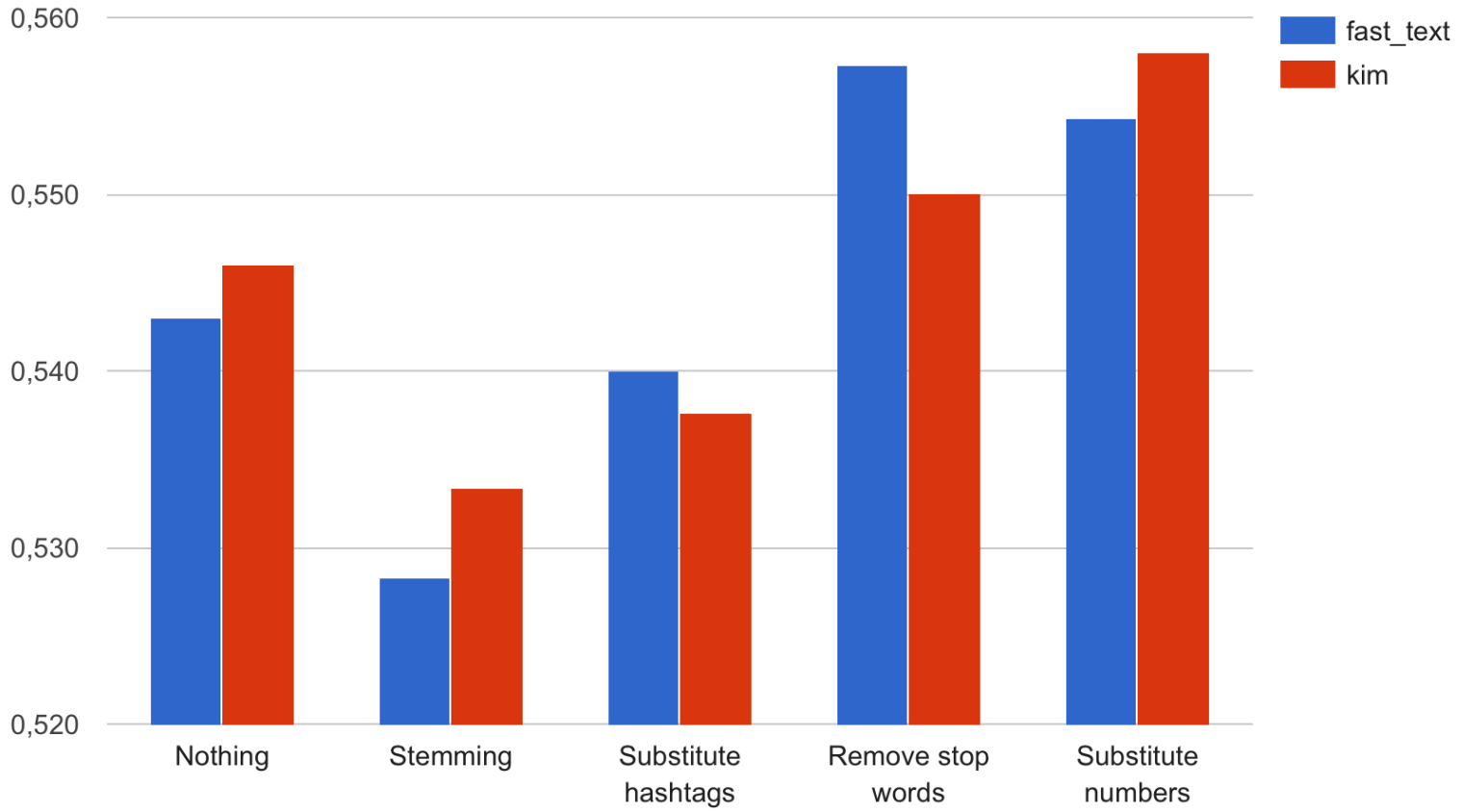
Comparative study

- Pre-processing
- Transfer learning
- *10-fold cross-validations*

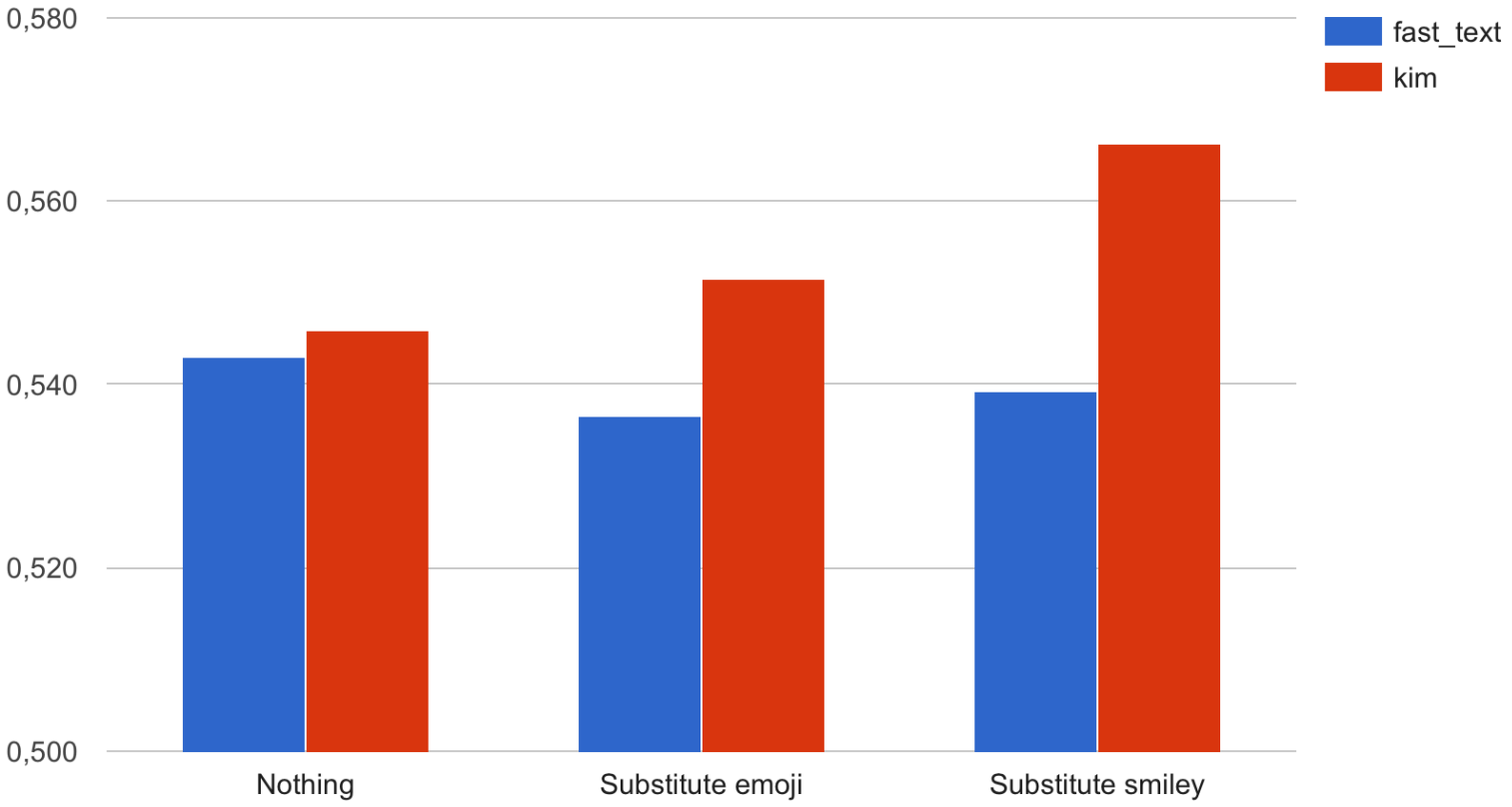
Comparative study

- **Pre-processing**
- Transfer learning
- *10-fold cross-validations*

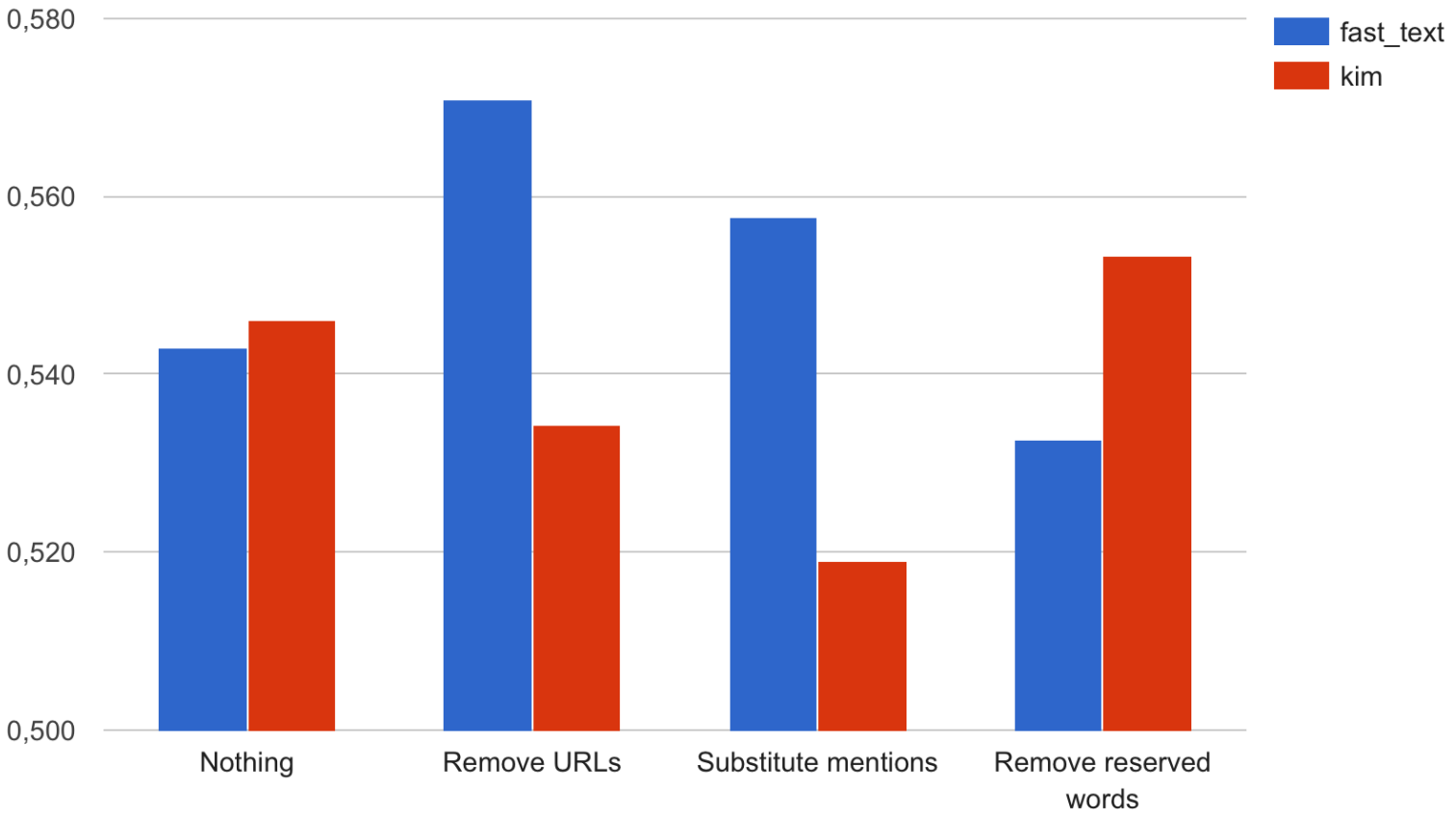
Pre-processing



Pre-processing



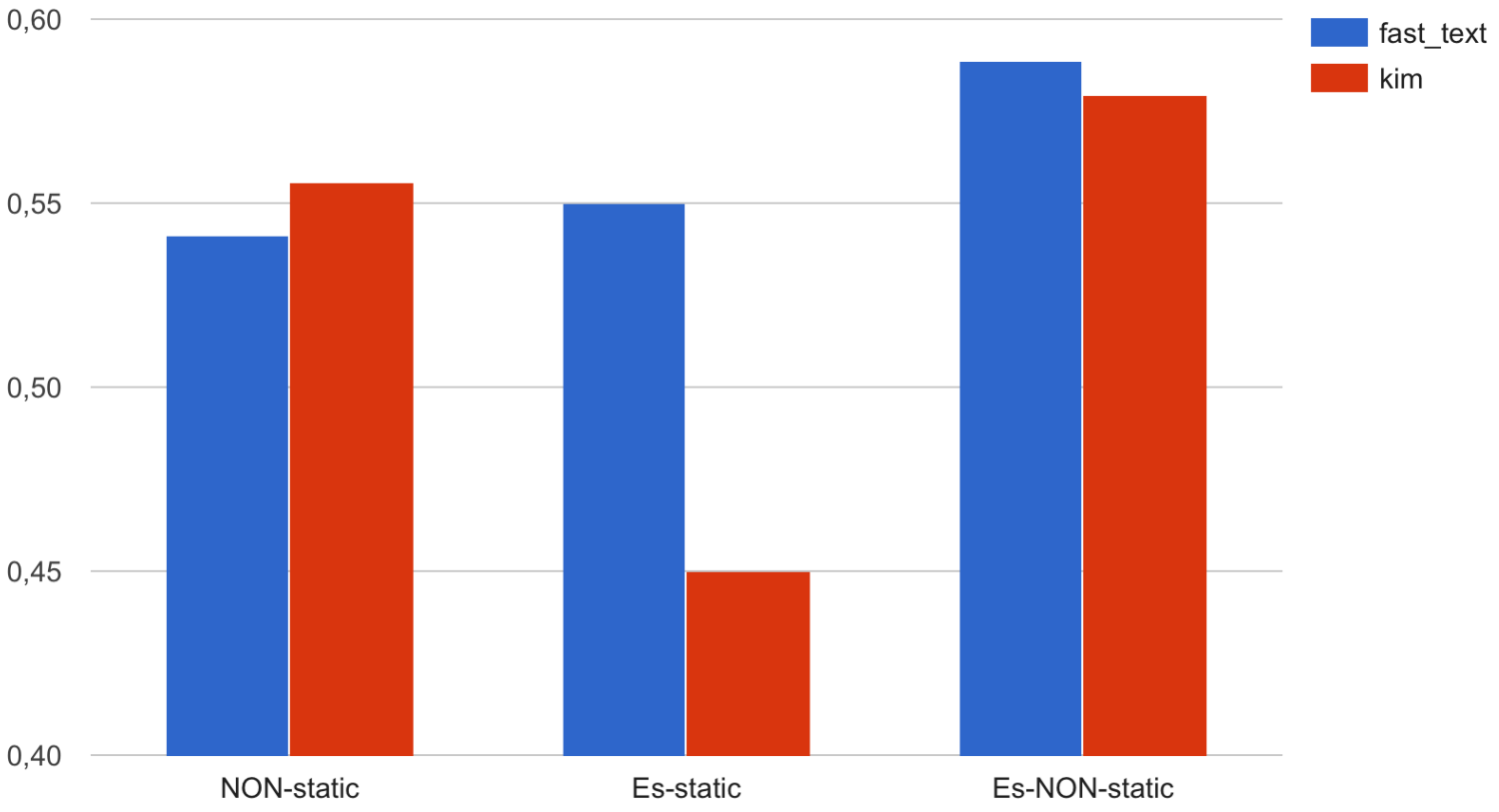
Pre-processing



Comparative study

- Pre-processing
- **Transfer learning**
- *10-fold cross-validations*

Transfer Learning



Overview

System	$F_{1\text{-macro}}$
LSTM	0,556 (±0.012)
Bi-LSTM	0,555 (±0.035)
CNN	0,571 (±0.030)
FastText	**0,589** (±0.018)
Kim	**0,579** (±0.009)

TASK RESULTS

Ranking		System	$F_{1\text{-macro}}$
Absolute	Team		
1	1	EliRF-UPV – run 1	0,6482
2	1	EliRF-UPV – run 4	0,6400
3	2	LuSer – run 1	0,6337
4	1	EliRF-UPV – run 3	0,6330
5	1	EliRF-UPV – run 2	0,6233
6	3	Puigcerver – run 1	0,6176
7	**4**	**atoppe – run 3 (fastText)**	**0,6157**
8	**4**	**atoppe – run 2 (kim)**	**0,6065**

Conclusions

- **Pre-processing pipeline** depends on model
- **Transfer learning** does work
- **Simpler model** performs better
- **Word order** seems not to be important

Thank you!

Questions?