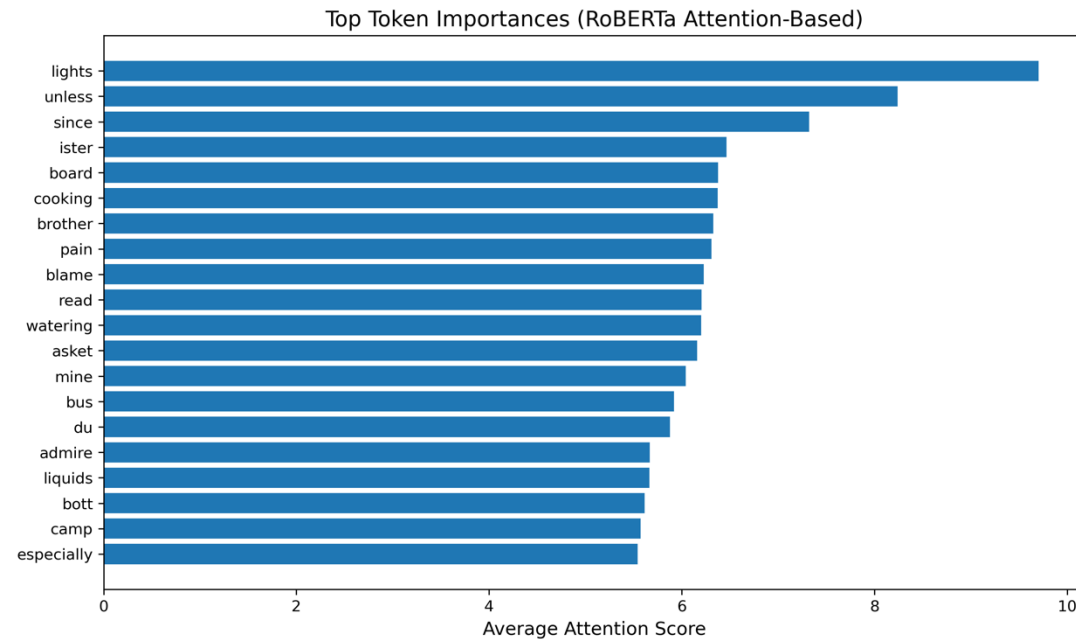
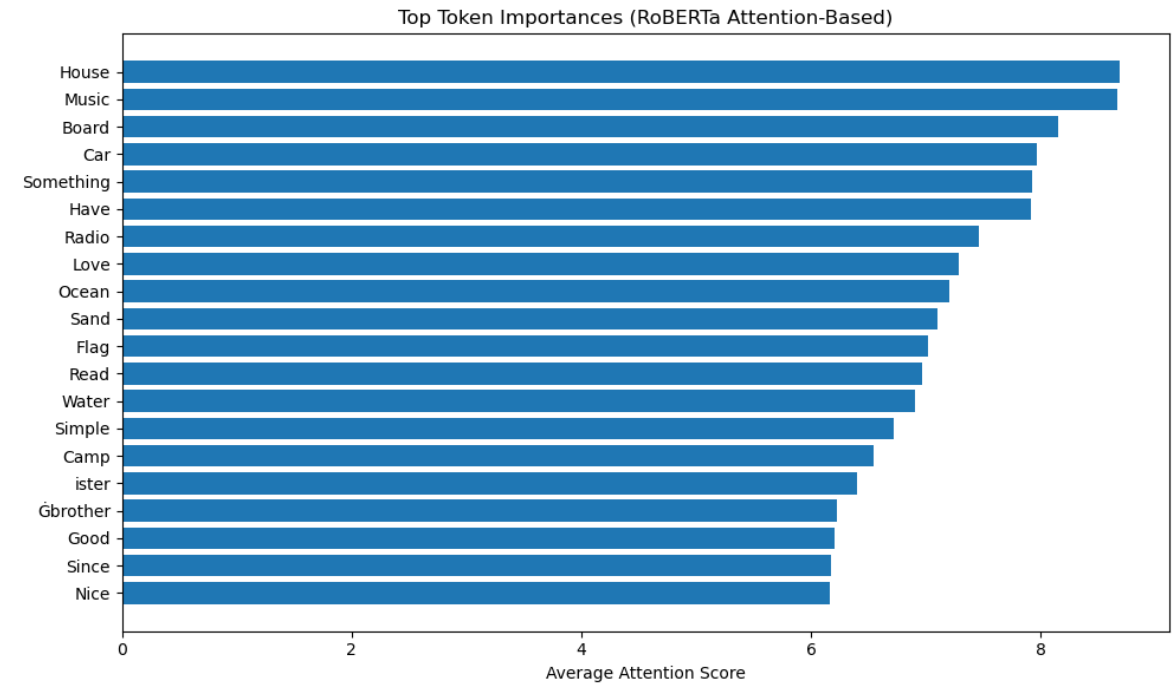


Feature importance analysis on best performing models
(RoBERTa, BERT, MLP+MentalBERT embeddings)

- Transformer based models: **RoBERTa**

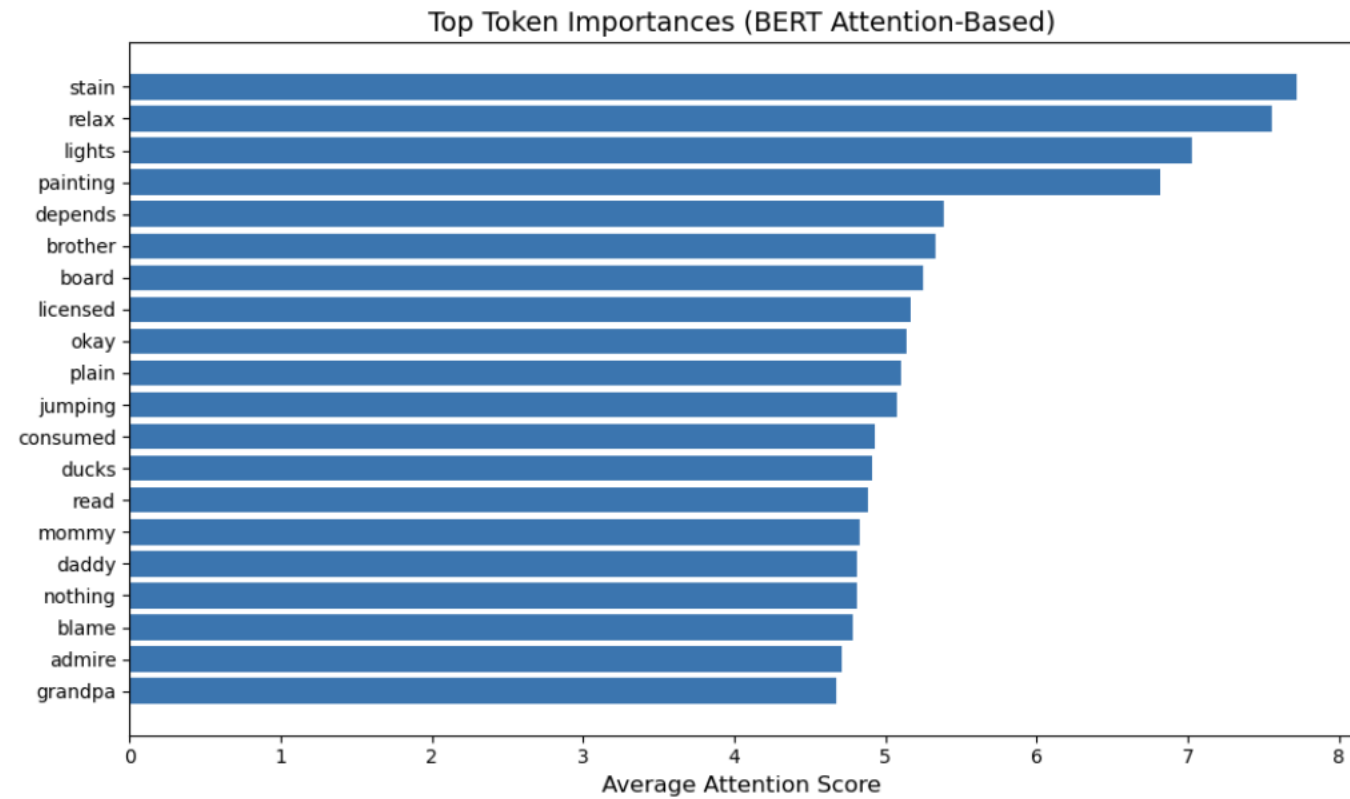


Top features when applying the model on lower-cased data



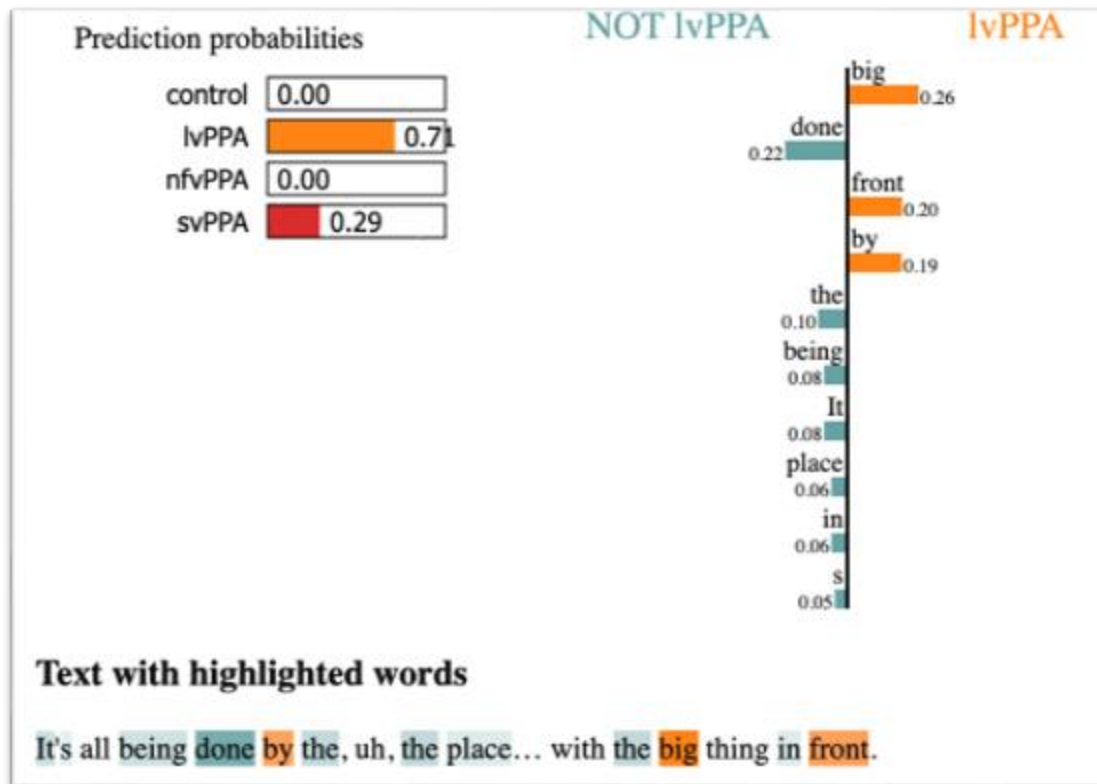
Top features when applying the model on unprocessed data (no lowercasing)

- Transformer based models: **BERT**

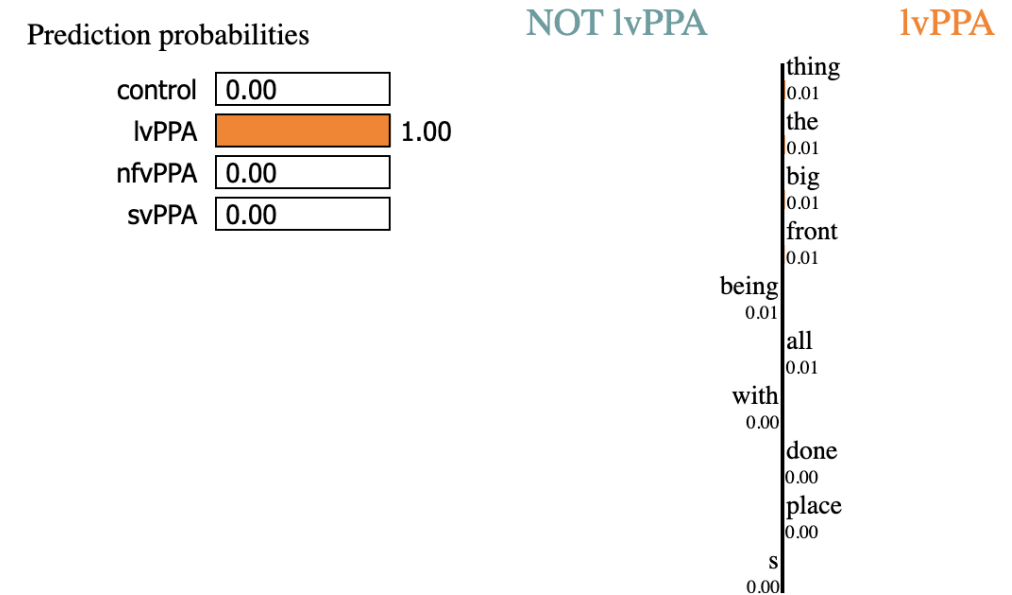


Using LIME (Local Interpretable Model-agnostic Explanations)

- Explains **why** a model made a prediction for **one specific case** Like a “local diagnostic”
- shows which features mattered most
- Works with any model (black box or not)



Synthetic sentence: manually generated
RoBERTa's explanation



Text with highlighted words

It's all being done by the, uh, the place... with the big thing in front.

Synthetic sentence: manually generated
BERT's explanation

- LIME on a sample from the dataset



Text with highlighted words

There is two people sitting here getting ready to do some eating, reading.

Real data sample (which we can't share in the paper)
RoBERTa's explanation



Text with highlighted words

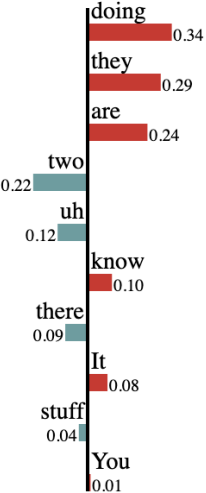
There is two people sitting here getting ready to do some eating, reading.

Real data sample (which we can't share in the paper)
BERT's explanation

Prediction probabilities

control	0.00
lvPPA	0.03
nfvPPA	0.00
svPPA	0.97

NOT svPPA svPPA



Text with highlighted words

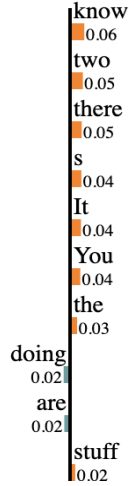
It's two, they are there doing the, uh, stuff. You know.

Synthetic sentence: manually generated (roBERTa)

Prediction probabilities

control	0.00
lvPPA	1.00
nfvPPA	0.00
svPPA	0.00

NOT lvPPA lvPPA

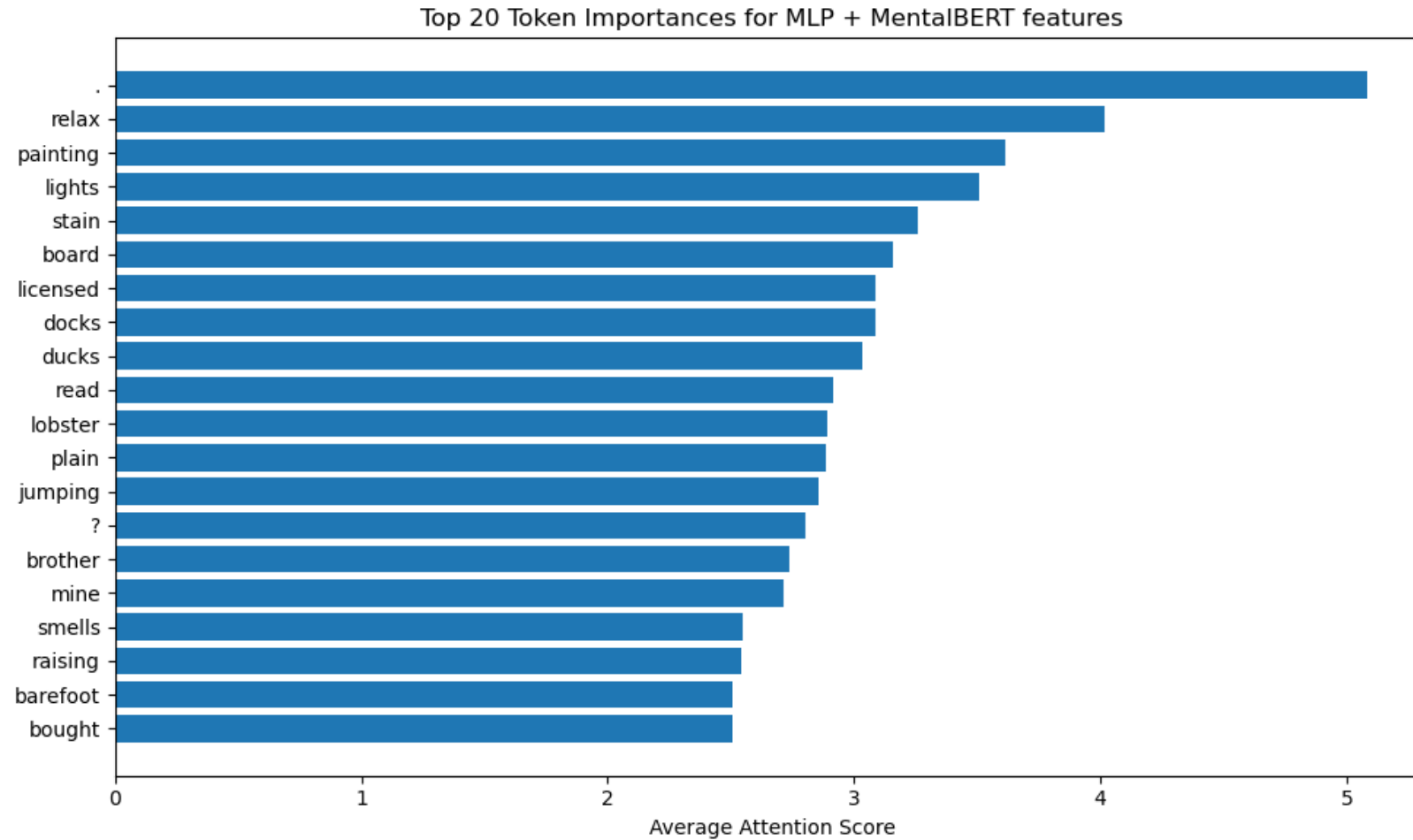


Text with highlighted words

It's two, they are there doing the, uh, stuff. You know.

Synthetic sentence: manually generated (BERT)

MLP + MentalBERT's embeddings



Prediction probabilities:

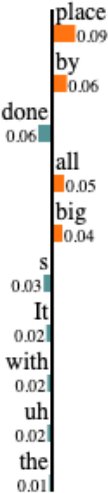
control : 0.0000
lvPPA : 1.0000
nfvPPA : 0.0000
svPPA : 0.0000

Prediction probabilities

control	<input type="text" value="0.00"/>	
lvPPA	<input type="text" value="1.00"/>	1.00
nfvPPA	<input type="text" value="0.00"/>	
svPPA	<input type="text" value="0.00"/>	

NOT lvPPA

lvPPA



Text with highlighted words

It's all being done by the, uh, the place... with the big thing in front.

Synthetic sentence: manually generated

Prediction probabilities:

control : 0.0034
lvPPA : 0.0001
nfvPPA : 0.0000
svPPA : 0.9965

Prediction probabilities

control	0.00
lvPPA	0.00
nfvPPA	0.00
svPPA	1.00

NOT svPPA



svPPA

Text with highlighted words

There is two people sitting here getting ready to do some eating, reading.

Real data sample (which we can't share in the paper)

Prediction probabilities:

control : 0.0000
lvPPA : 1.0000
nfvPPA : 0.0000
svPPA : 0.0000

Prediction probabilities

control	<input type="text" value="0.00"/>	
lvPPA	<input type="text" value="1.00"/>	1.00
nfvPPA	<input type="text" value="0.00"/>	
svPPA	<input type="text" value="0.00"/>	

NOT svPPA

there
0.05
s
0.04
two
0.04
doing
0.03
they
0.03
It
0.03
You
0.02
uh
0.01
the
0.01
are
0.01

svPPA

Text with highlighted words

It's two, they are there doing the, uh, stuff. You know.

Synthetic sentence: manually generated