

# Experiments

## Real-world scenarios

- Split dataset **temporally**.
- Such a scenario can somehow expose the drawback of existing techniques and it requires a model of higher **generalizability** to cope with novel instances.
- Under this hard setting, Dual Emotion Features **still outperform others**, reveals the effectiveness and generalization ability to some extent.

Models	Macro F1	Acc.	F1 score	
			Fake	Real
BiGRU	0.839	0.839	0.839	0.839
+ Emoratio	0.850	0.850	0.854	0.846
+ EmoCred	0.829	0.829	0.836	0.821
+ <b>Dual Emotion Features</b>	<b>0.855</b>	<b>0.855</b>	<b>0.857</b>	<b>0.852</b>
BERT	0.900	0.900	0.900	0.900
+ Emoratio	0.901	0.901	0.900	0.902
+ EmoCred	0.902	0.902	0.901	0.903
+ <b>Dual Emotion Features</b>	<b>0.915</b>	<b>0.915</b>	<b>0.913</b>	<b>0.918</b>
HSA-BLSTM	0.913	0.913	0.912	0.914
+ Emoratio	0.920	0.920	0.920	0.920
+ EmoCred	0.903	0.903	0.902	0.905
+ <b>Dual Emotion Features</b>	<b>0.932</b>	<b>0.932</b>	<b>0.932</b>	<b>0.933</b>

Models	Macro F1	Acc.	F1 score	
			Fake	Real
BiGRU	0.680	0.681	0.694	0.666
+ Emoratio	0.628	0.632	0.665	0.592
+ EmoCred	0.659	0.666	0.709	0.609
+ <b>Dual Emotion Features</b>	<b>0.701</b>	<b>0.702</b>	<b>0.714</b>	<b>0.689</b>
BERT	0.722	0.728	0.762	0.682
+ Emoratio	0.719	0.724	0.757	0.681
+ EmoCred	0.725	0.728	0.752	<b>0.699</b>
+ <b>Dual Emotion Features</b>	<b>0.734</b>	<b>0.734</b>	<b>0.773</b>	0.692
HSA-BLSTM	0.776	0.778	0.796	0.686
+ Emoratio	0.771	0.774	0.796	0.663
+ EmoCred	0.777	0.781	0.806	0.646
+ <b>Dual Emotion Features</b>	<b>0.805</b>	<b>0.808</b>	<b>0.827</b>	<b>0.694</b>

**Table 7: Results on Weibo-20 (temporal data split). Acc. is short for Accuracy.**

# Experiments

## Evaluation Questions

- EQ1: Are Dual Emotion Features more effective than baseline features when used alone for fake news detection? How effective are the different types of features in Dual Emotion Features?
- EQ2: Can Dual Emotion Features help improve the performance of text-based fake news detectors?
- EQ3: How robust do the fake news detection models with Dual Emotion Features in real-world scenarios?
- EQ4: How **effective** are the components of Dual Emotion Features, including the **publisher** emotion, **social** emotion, and emotion **gap**?