



# Classification for Crisis-Related Tweets Leveraging Word Embeddings and Data Augmentation

Congcong Wang, David Lillis

School of Computer Science, University College Dublin, Ireland.

## 1. Abstract

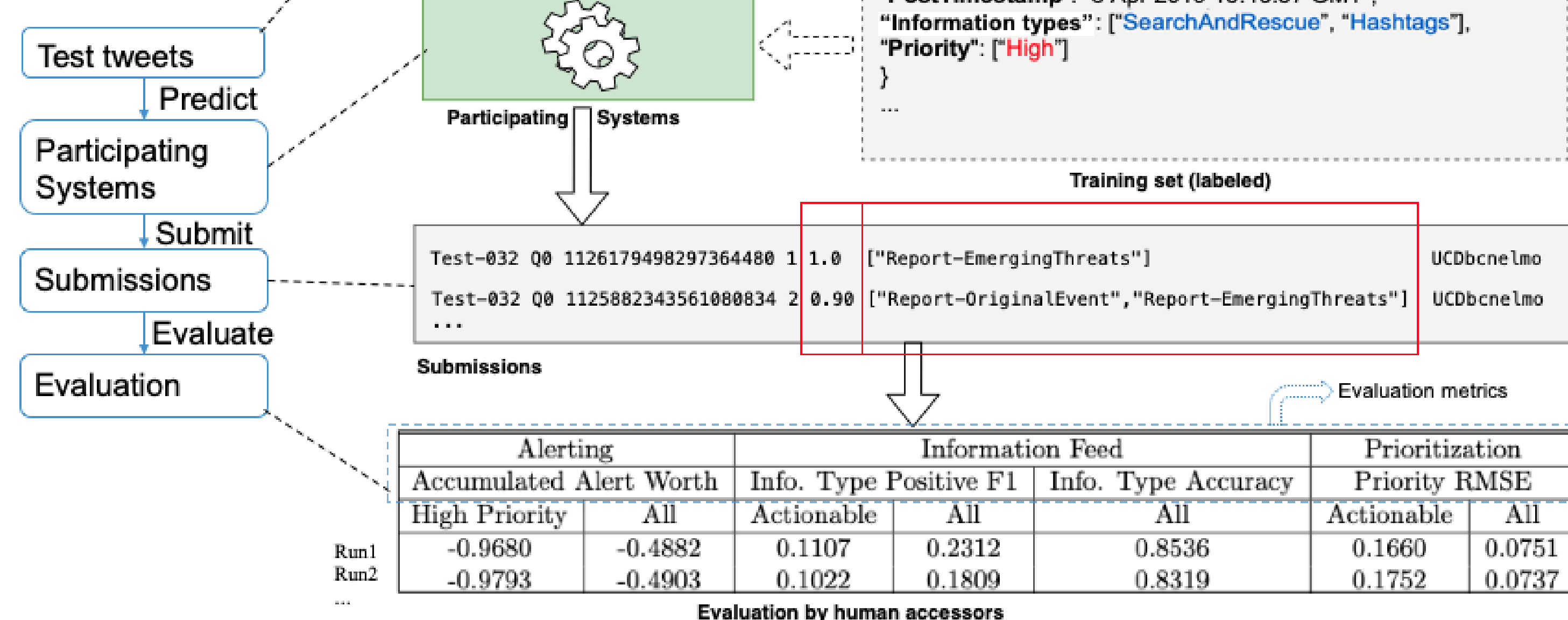
The Incident Streams track (IS) is a research initiative proposed as part of the Text REtrieval Conference (TREC) aiming to mature social media-based emergency response technology. The purpose of the IS track is to find actionable messages and estimate their priority among a stream of crisis-related tweets. This poster presents University College Dublin's (UCD) participation at 2019 TREC IS (A and B editions). Evaluation results show that our approaches achieve competitive performance among participating runs.



## 2. Objectives

- Participating systems are asked to predict one or more information types and a priority level for an upcoming crisis-related tweet given manually-annotated training tweets.
- Information types: pre-defined ontology of user information needs
  - 6 actionable types: SearchAndRescue, GoodsServices, MovePeople, EmergingThreats, NewSubEvent, ServiceAvailable.
  - 19 non-actionable types like Volunteer, Donations, Weather, Sentiment, CleanUp, Advice, News, etc.
- Priority levels: low, medium, high and critical

## Task Process

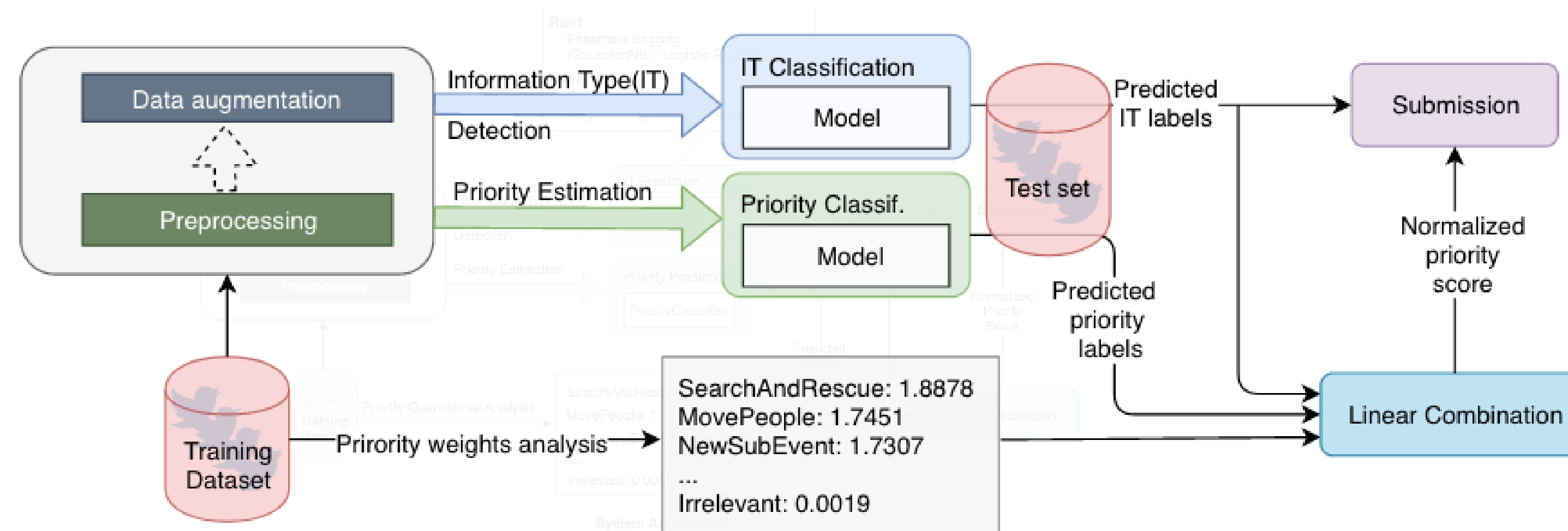


## Main challenges

- Misspellings: character-level embedding, out-of-vocabulary lookup.
- Imbalanced datasets: data augmentation techniques in texts.
- Multi-label classification: threshold setup, ensemble approach.

## 3. Methods

- We submit a machine learning based run at 2019-A edition and three deep neural network based runs at 2019-B edition. The system architecture is as follows.

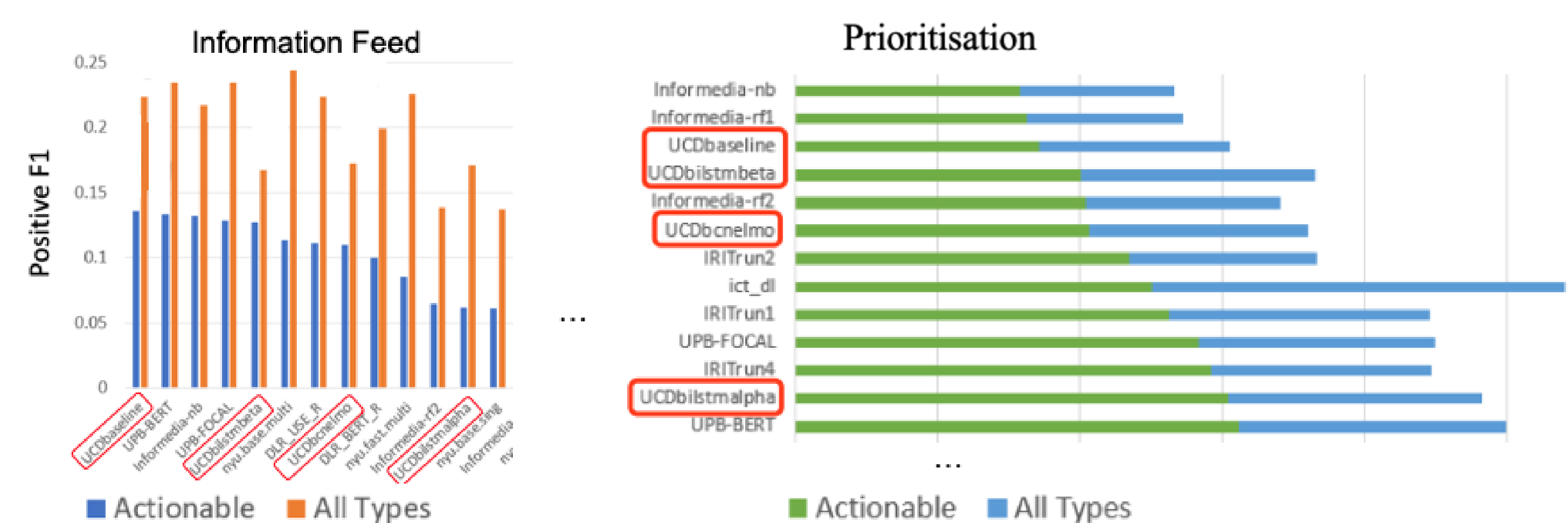


- UCDbaseline:** SMOTE for data augmentation, 21 hand-crafted features concatenated to a pre-trained in-domain word2vec embedding, ensemble of logistic regression and naïve bayes.
- UCDbilstmalpha:** GPT-2 for data augmentation, GloVe embedding, vanilla biLSTM encoder.
- UCDbilstmbeta:** no augmentation but loss weights are applied for the loss function at the output level of neural network, character-level embedding concatenated to GloVe embedding, vanilla biLSTM encoder.
- UCDbcnelmo:** state-of-the-art bi-attentive classification model and pre-trained ELMo contextualised embedding.

## 4. Evaluation Results

- Evaluation metrics
  - Information Feed (0 to 1): evaluate the performance of a run in information type categorisation.
  - Prioritisation (0 to 1, less is better): evaluate the performance of a run in estimating a tweet's criticality.
  - Alerting (-1 to 1): evaluate the performance of a run in finding tweets for alerting correctly, which combines the categorisation and prioritisation effectiveness.
- Out of 32 runs from 10 participating groups, our four runs perform better than the median for most situations, with UCDbaseline in Information Feed and Prioritisation in particular.

Runs	Alerting		Information Feed			Prioritisation	
	Accumulated Alert Worth	High Priority	Info. Type Actionable	Info. Type Positive F1	Info. Type Accuracy	Priority RMSE Actionable	Priority RMSE All
Median	-0.9197	-0.4609	0.0386	0.1055	0.8583	0.1767	0.1028
UCDbaseline	-0.7856	-0.4131	<b>0.1355</b>	<b>0.1343</b>	0.7495	<b>0.0859</b>	<b>0.0668</b>
UCDbilstmalpha	-0.9287	-0.4677	0.0614	0.1087	<b>0.86</b>	0.1521	0.0893
UCDbilstmbeta	<b>-0.6047</b>	<b>-0.3332</b>	0.1269	0.0607	0.8378	0.1004	0.0822
UCDbcnelmo	-0.6961	-0.3624	0.1099	0.1192	0.8452	0.1036	0.0769



## 5. Conclusion and Future Work

- We submitted four runs at 2019 TREC IS track for classifying crisis-related tweets for emergency response. Our runs apply different word embeddings and data augmentation techniques to tackle the challenges in the task. Evaluation results show that our runs perform well and character-level embedding helps improve performance.
- Future work (this track continues to run next year)
  - Does further feature engineering help?
  - Do simple data augmentation techniques help?
  - Do In-domain contextualised embeddings help?

## 6. References

- Richard McCreddie, Cody Buntain and Ian Soboroff (2019) TREC Incident Streams: Finding Actionable Information on Social Media. In Proceedings of ISCRAM 2019.
- Muhammad Imran, Prasenjit Mitra, and Carlos Castillo: Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages. In Proceedings of the 10th Language Resources and Evaluation Conference (LREC) 2016.