The **features concatenation** analysis. And **scalable unsupervised approach** for their extraction.

BRNIR at the NTCIR-14 finnum task: Scalable feature extraction technique for number classification
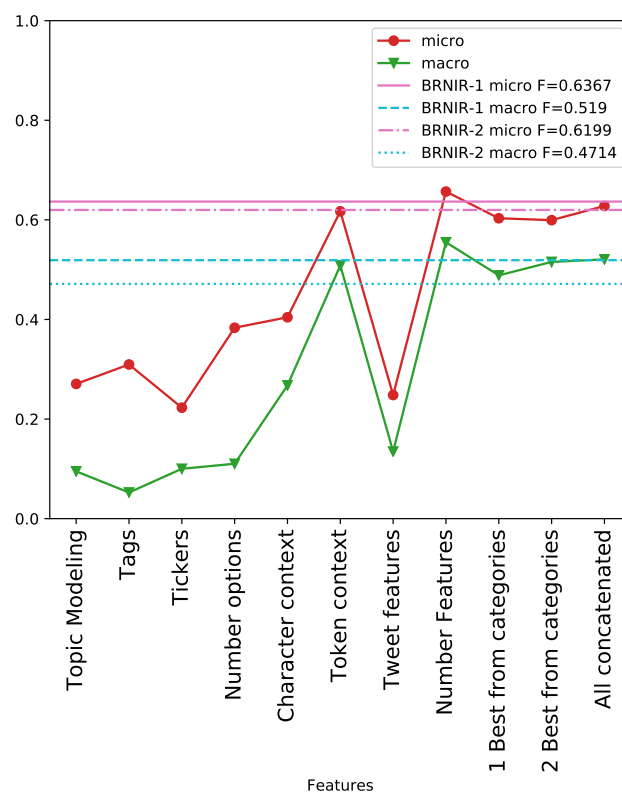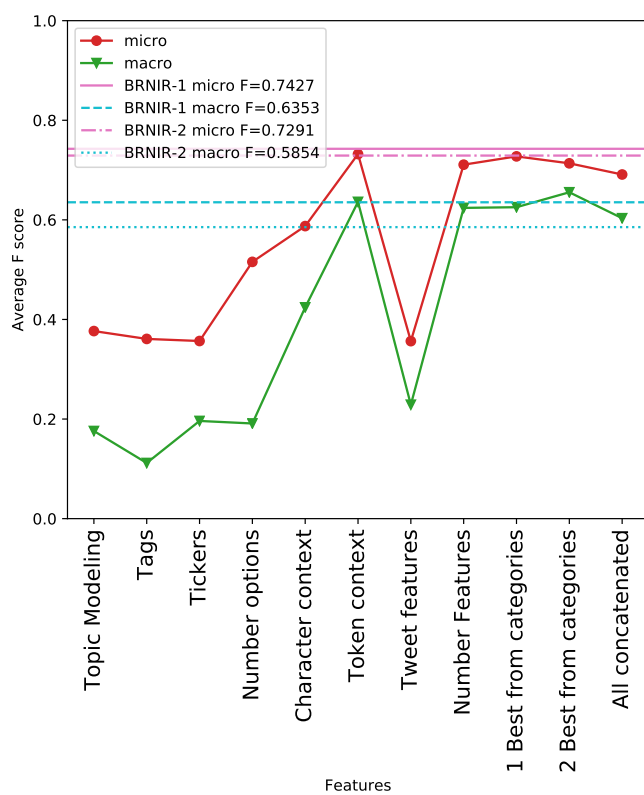
Alan Spark

1   Tweet features

- Topic distribution
- Tickers
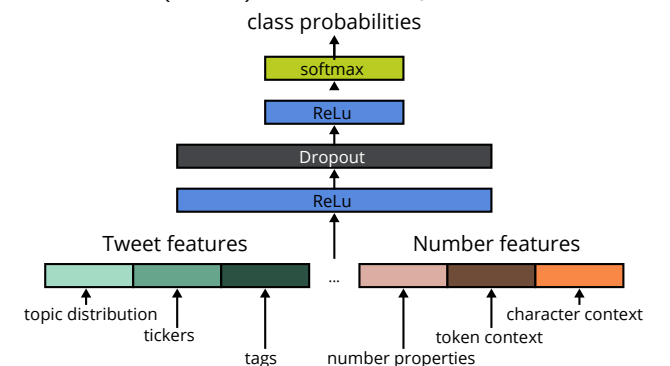- Tags

2   Number features

- Number properties
- Token context
- Character context

3   Results

Experiment

During experiment the classification model was fixed. First layer of a network (input), succeeded by layers of fully connected Rectified Linear Units (ReLU) with a dropout in between.

Network structure follows a typical scheme - every succeeding layer diminishes in size by half, the last "output" layer sized accordingly to the number of classes in a task:

Subtask 1: 512 ReLU $\rightarrow$ 256 ReLU $\rightarrow$ 7 SoftMax
Subtask 2: 512 ReLU $\rightarrow$ 256 ReLU $\rightarrow$ 17 SoftMax

Download the paper $\rightarrow$