Virginia Tech
VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Spatial Data Management Lab

SANGHANI CENTER FOR ARTIFICIAL INTELLIGENCE & DATA ANALYTICS

EMNLP 2020

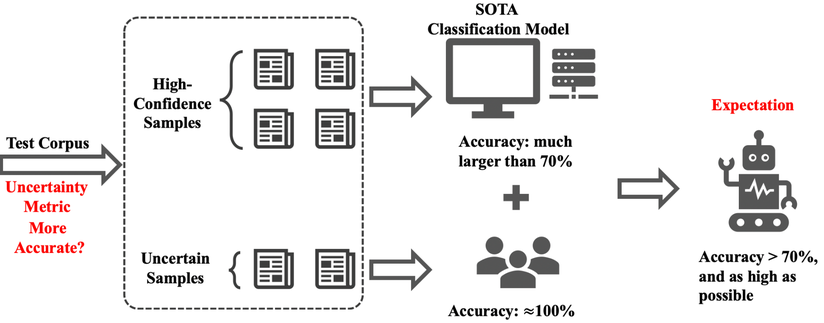
Towards More Accurate Uncertainty Estimation in Text Classification

Jianfeng He, Xuchao Zhang, Shuo Lei, Zhiqian Chen, Fanglan Chen, Abdulaziz Alhamadani, Bei Xiao, Chang-Tien Lu

Background & Problem Setting

-SOTA text-classification model does not meet higher expectation.

-We aim to **measure uncertain scores** and give the most uncertain samples to human experts.



-**Uncertainty Metric**: Depending on the winning score directly or indirectly.

-**Winning score**: Maximum probability in a semantic vector (softmax vector).

Motivation 1

-Reduce overconfidence in training samples

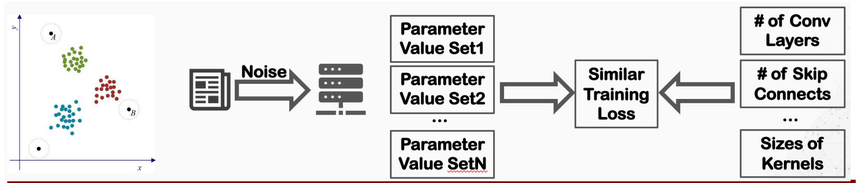
-Overconfidence: Sample Uncertainty for each training sample is same due to one-hot labels.

-The negative correlation between the winning score and sample uncertainty cannot be guaranteed.

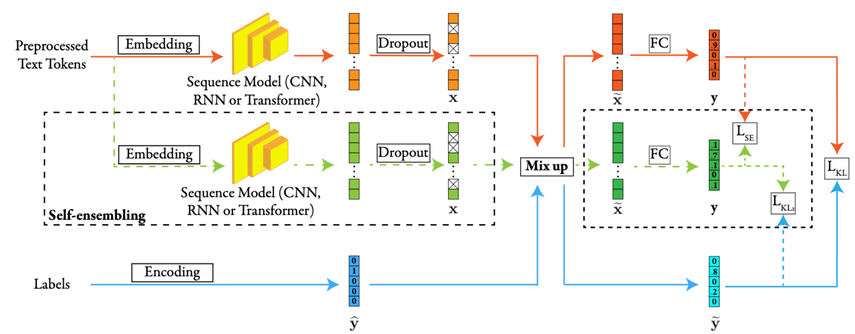
Motivation 2

-Comprehensively estimate uncertainty

- Data Uncertainty
  - ➢ Epistemic: Lack of knowledge
  - ➢ Aleatoric: Noisy data
- Model Uncertainty
  - ➢ Parametric: Different feasible parameters
  - ➢ Structure: Different feasible designs



Proposed Method-MSD



-**Mix-up**: Different winning scores

- Reduce aleatoric Uncertainty & Overconfidence

-**Self-ensembling**: Less feasible parameters

- Reduce Parametric Uncertainty

-**Distinctive score**: A testing sample's distance to a training set's distribution

- Estimate Epistemic Uncertainty

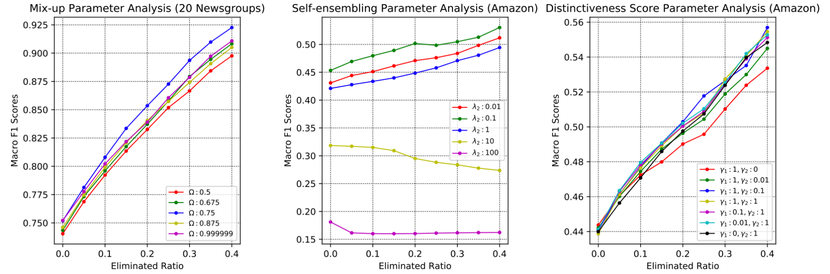
Results

Table 2: Accuracy of uncertainty scores shown by improvement of weighted F1 scores for the IMDb (CNN model)

Methods (Ω, λ_2, γ_1, γ_2)	Uncertainty Ratio (Weighted F1, Improved Ratio) 0%	10%	20%	30%	40%
DE	0.880	0.913(3.75%)	0.939(6.70%)	0.957(8.75%)	0.970(10.22%)
DE+Metric	0.884	**0.918**(3.85%)	**0.944**(6.79%)	0.961(8.71%)	0.974(10.18%)
MSD1 (1, 0, 1, 0)	0.874	0.907(3.87%)	0.933(6.79%)	0.952(8.95%)	0.967(**10.75%**)
MSD2 (1, 1, 1, 0)	0.883	**0.918**(3.92%)	**0.944**(6.82%)	0.961(8.85%)	**0.976**(10.46%)
MSD3 (1, 1, 1, 0.1)	0.882	**0.918**(**4.04%**)	0.943(**6.88%**)	**0.962**(**9.08%**)	0.974(10.49%)

Table 6: Accuracy of uncertainty scores shown by improvement of macro F1 scores for the Amazon (XLnet)

Methods (Ω, λ_2, γ_1, γ_2)	Uncertainty Ratio (Macro F1, Improved Ratio) 0%	10%	20%	30%	40%
DE	0.422	0.422(0.00%)	0.428(1.38%)	0.423(0.26%)	0.424(0.38%)
DE+Metric	0.438	0.444(1.29%)	0.447(1.96%)	0.448(2.35%)	0.447(2.04%)
MSD1 (1, 0, 1, 0)	0.426	0.442(3.85%)	0.446(4.80%)	0.452(6.14%)	0.439(3.22%)
MSD2-a (1, 0.01, 1, 0)	0.415	0.436(5.03%)	0.440(6.06%)	0.434(4.46%)	0.422(1.56%)
MSD2-b (1, 0, 1, 1)	0.424	**0.451**(6.22%)	**0.470**(10.87%)	0.486(14.89%)	0.501(17.99%)
MSD3 (1, 0.01, 1, 1)	0.417	0.447(**7.16%**)	0.467(**11.96%**)	**0.487**(**16.81%**)	**0.509**(**21.95%**)



Conclusions

- We propose a novel model, MSD, for more accurate uncertainty score estimation of text classification by boosting the correlation between winning scores and sample uncertainty and considering three kinds of uncertainty simultaneously.
- The MSD is effective in three common DNN including CNN, RNN and transformer.