SemEval 2018 Task #7

"Semantic Relation Extraction and Classification in Scientific Papers"

1. Baseline.

In order to be able to compare proposed solution to other system, some reference solution needed to be selected. As a baseline, classifier model which predicts every incoming instance as a majority class (relation type "*usage*") was selected. Results from the selected baseline are covered in Section 3, while results obtained from the proposed solution are covered in Section 4.

2. Evaluation metrics.

For the SemEval Task 7, the following evaluation metrics were selected:

- Accuracy, which corresponds to the ratio of correctly classified instances to the total number of instances that were supposed to be correctly classified as seen in the equation below.
- F-Measure with macro average in which precision and recall is averaged between classes.

Additionally, k-Fold Cross Validation was performed in order to achieve reliable results of performance obtained from the proposed solution.

3. Results obtained from the proposed solution.

Following results are obtained from the framework designed with SemEval 2018 Task 7 in mind where for the directionality part only features from Feature Extractor 1 were used. As a classification algorithm, the SVM from Scikit-Learn library was used to conduct the experiment.

Results for a baseline and proposed solution for the Subtask 1.1 directionality classification task with use only of Feature Extractor 1 are covered in the *Table 1*.

Table 1 Results for the baseline and proposed solution for the Subtask 1.1 directionality classification.

	Proposed Solution [%]	**Baseline [%]**
Accuracy	67.98	67.37
Sensitivity	25.06	0.00
Specificity	88.91	100.00
Geometric Mean	46.30	0.00

Results for the baseline and proposed solution for the Subtask 1.1 relation classification task for the best combination of features are covered in the *Table 2*.

Table 2 Results for the baseline and proposed solution for the Subtask 1.1 relation classification.

	Proposed Solution [%]	**Baseline [%]**
Accuracy	46.00	40.53
F-measure (macro avg)	21.00	9.61
Precision	28.00	6.75

The result of the proposed solution is slightly better than the results of the baseline. Above given results for the proposed approach are the results of the best feature combination when run with SVM classifier.