

PROPOSITION DE STAGE EN COURS D'ETUDES

Référence : **DTIS-2022-03**
(à rappeler dans toute correspondance)

Lieu : ONERA Palaiseau

Département/Dir./Serv. : DTIS/MIDL

Tél. : 01 80 38 65 65

Responsable(s) du stage : Valentina Dragos

Email : valentina.dragos@onera.fr

DESCRIPTION DU STAGE

Thématique(s) : Ingénierie Artificielle et Décision

Type de stage : Fin d'études bac+5 Master 2 Bac+2 à bac+4 Autres

Intitulé : Boosting text classification with semantic descriptions and knowledge graphs

Sujet : Text classification models have been widely used for a variety of practical applications. Like all machine-learning models, those classifiers depend heavily on the quality of data sets used for training purposes. The goal of this work is to investigate methods to boost the performance of text classifiers developed for social data analysis, and the task focuses on extremist and hateful content detection [6]. For those domains, there are many situations where data is hard to collect or hard to label, due to the overwhelming volumes available online and to their diversity. As a result, data sets are small or imbalanced or the overlap between classes is high [3]. Bias can also affect the quality of data sets [2]. Although the models can achieve good precision and accuracy levels, the performance decreases rapidly as soon as the model applies to another domain[4]. Several methods can handle those shortcomings and improve the performance of text classifiers. An interesting solution to poor datasets is to take advantage of previously acquired knowledge, be that expert or common sense knowledge, and use linked data and knowledge graphs to retrieve relevant concepts to augment sentences [5]. Text augmentation adds more information to sentences and build semantically enriched data sets [1]. The goal of this work is to use knowledge graphs such as Concepts Net, Wikipedia or DBPedia (see [7] for a detailed panorama of available resources) and domain resources developed for social data analysis in order to improve the classification of extremist data collected online. Validation of approaches and evaluation of results will be carefully addressed.

The solution will be applied to explore data sets collected within the frame of the FLYER project (ANR).

The work has the following milestones:

State of art and problem analysis (1 month)

Formalization of the of overall approach (1 month)

Implementation of algorithms (1.5 months)

Experimental protocol, experimentation on use-cases and analysis of results (1.5 months)

Internship report (1 month)

References

[1] Omeliyanenko, J., Zehe, A., Hettlinger, L., & Hotho, A. (2020, November). Lm4kg: Improving common sense knowledge graphs with language models. In International Semantic Web Conference (pp. 456-473). Springer, Cham.

[2] Wiegand, M., Ruppenhofer, J., & Kleinbauer, T. (2019, June). Detection of abusive language: the problem of biased datasets. In Proceedings of the 2019 conference of the North American Chapter of the Association for Computational Linguistics: human language technologies, volume 1 (long and short papers) (pp. 602-608).

[3] Li, Z., Yao, H., & Ma, F. (2020, January). Learning with small data. In Proceedings of the 13th International Conference on Web Search and Data Mining (pp. 884-887).

[4] Fortuna, P., Soler-Company, J., & Wanner, L. (2021). How well do hate speech, toxicity, abusive and offensive language classification models generalize across datasets?. Information Processing & Management, 58(3), 102524.

[5] Koufakou, A., Pamungkas, E. W., Basile, V., & Patti, V. (2020, November). HurtBERT: incorporating lexical features with BERT for the detection of abusive language. In Proceedings of the fourth workshop on online abuse and harms (pp. 34-43).

[6] Salminen, J., Hopf, M., Chowdhury, S. A., Jung, S. G., Almerkhi, H., & Jansen, B. J. (2020). Developing an online hate classifier for multiple social media platforms. Human-centric Computing and Information Sciences, 10(1), 1-34.

[7] Mountantonakis, M., & Tzitzikas, Y. (2019). Large-scale semantic integration of linked data: A survey. ACM Computing Surveys (CSUR), 52(5), 1-40.

Est-il possible d'envisager un travail en binôme ? **Non**

Méthodes à mettre en oeuvre :

- | | |
|---|---|
| <input checked="" type="checkbox"/> Recherche théorique | <input checked="" type="checkbox"/> Travail de synthèse |
| <input checked="" type="checkbox"/> Recherche appliquée | <input type="checkbox"/> Travail de documentation |
| <input type="checkbox"/> Recherche expérimentale | <input checked="" type="checkbox"/> Participation à une réalisation |

Possibilité de prolongation en thèse : **Oui**

Durée du stage : Minimum : 5 Maximum : 6

Période souhaitée : Mars August 2022

PROFIL DU STAGIAIRE

Connaissances et niveau requis :
Programming languages: Python, Java
Knowledge representation: ontologies, conceptual graphs
Machine learning; classification
Languages: French, English

Ecoles ou établissements souhaités :
Université, Grande école