Fiche de projet tutoré / Project form

Intrinsic evaluation of word embeddings

Encadrement / Supervisors
1. équipe, laboratoire / team, lab : SMarT / Loria 1. encadrant·e principal·e (nom, email) / main supervisor (name, email) Langlois David, david.langlois@loria.fr
Description / Description 1. projet global/global project The word embeddings (Mikolov et al., 2010; Pennington et al., 2014) allow modeling the words of a language in the form of numerical vectors. This is called latent semantic because the numerical values of a vector implicitly incorporate the linguistic characteristics of the word: morphology, nature, gender, number, tense, meaning, interaction in the sentence, etc. Such a representation has been shown to be useful for improving performance in classical automatic language processing tasks. It has also shown that it can give rise to an algebra of vectors allowing reasoning by analogy. The downside is that the information is latent. The characteristics that usually define a word, and listed above, are not explicitly found. The proposed project is to determine whether the latent information can be retrieved: what are the elements of the vectors that tell us about the nature of the word, for example? Tracking the location of information related to linguistic features can be useful for several reasons. It will allow us to revisit lexical models: are they confirmed by the data? Can they be refined? Furthermore, the information obtained will improve the explicability of models based on word embeddings. There is also the possibility of associating a vector (at least partially) with a word that is absent from the corpora on the basis of the linguistic knowledge we have about this word; then it could be possible to inject a dictionary into a word embeddings model. Another possibility is also to evaluate word embeddings on the assumption that they should confirm known lexical structures (intrinsic evaluation (Wang et al., 2019)). This project has yet been proposed to students in a previous year, and this work led to the publication of a paper (Goliakova, 2024). This year's students will continue the work basing on previous results. 2. biblio. UE 705 (semestre 7) The literature review should include :

- A review of word embeddings approaches. For that the French thesis by N. Aliane (Aliane, 2019) could be used.
- A search for pre-computed word embedding models made available by the community. Basing on our previous, BERT models will be specifically studied.
- A review of work seeking to improve the explicability of word embeddings via intrinsic evaluation

3. réalisation. UE 805 (semestre 8)

In order to detect whether there is a correlation between the elements of the vector of a word and the linguistic features that define it, the following method is proposed:

1. Create a model of word embeddings (several are publicly available)
2. Select a linguistic information (e.g. the nature of the word)
3. Create a training corpus [source = word embedding vector, target = linguistic information].
4. Use a neural network (structure to be defined) to learn a vector classifier
5. Study the weights of the neural network in order to detect the elements of the vector that have a high weight

Subsequently, different linguistic features, different neural architectures (or other classifiers such as SVMs) could be studied. The impact of the size of the vectors can also be studied.

Informations diverses : matériel nécessaire, contexte de réalisation /
Various information: material, context of realization

Target language of data is French. Therefore, good skills in French are required.

Students will benefit of previous work (report) and a paper (Goliakova, 2024)

Livrables et échéancier / Deliverable and schedule

Bibliographie /References (max. 4-5)
[*il ne s'agit pas de la bibliographie complète qui sera fournie aux étudiants au début du projet mais d'une bibliographie indicative pour aider à cerner le sujet*]

Aliane, N. (2019). Evaluation des représentations vectorielles de mots (Doctoral dissertation, Paris 8).

Goliakova, E., & Langlois, D. (2024, September). What do BERT word embeddings learn about the French language?. In *Computational Linguistics in Bulgaria* (pp. 14-32).

Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010, September). Recurrent neural network based language model. In Interspeech (Vol. 2, No. 3, pp. 1045-1048).

Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).

Tsvetkov, Y., Faruqui, M., & Dyer, C. (2016). Correlation-based intrinsic evaluation of word vector representations. arXiv preprint arXiv:1606.06710.

Schnabel, T., Labutov, I., Mimno, D., & Joachims, T. (2015, September). Evaluation methods for unsupervised word embeddings. In Proceedings of the 2015 conference on empirical methods in natural language processing (pp. 298-307).

Wang, B., Wang, A., Chen, F., Wang, Y., & Kuo, C. C. J. (2019). Evaluating word embedding models: methods and experimental results. APSIPA transactions on signal and information processing, 8.