The goal of this PhD Thesis will be devoted to the innovative study of the taking into account of semantics through predictive representations that capture the semantic features of words and their context. Research will be conducted on the combination of semantic information with information from denoising to improve speech recognition.

Semantics and thematic spaces are vector spaces used for representation number of words, sentences or textual documents. The corresponding models and methods have a long history in the field of computational linguistics and natural language processing [Turney and Pantel, 2010]. Almost all models rely on the hypothesis of statistical semantics which states that: Statistical patterns of appearance of words (context of a word) can be used to describe the underlying semantics. The most used method to learn these representations is to predict a word using the context in which this word appears [Mikolov et al., 2013b, Pennington et al., 2014], and this can be realized with neural networks. These representations have proved effective for a series of natural language processing [Baroni et al., 2014]. In particular, Mikolov’s Skip-gram and CBOW models [Mikolov et al., 2013a, 2013b] have become very popular because of their ability to process large amounts of unstructured text data with reduced computing costs. Efficiency and the semantic properties of these representations motivate us to explore these semantic representations for our task of recognition in noisy conditions.

Main activities
The goal of this PhD Thesis will be devoted to the innovative study of the taking into account of semantics through predictive representations that capture the semantic features of words and their context. Research will be conducted on the combination of semantic information with information from denoising to improve speech recognition.

The ASR stage will be supplemented by a semantic analysis to detect the words of the processed sentence that could have been misrecognized and to offer similar (at the acoustic level) words that better fit the context. Predictive representations using continuous vectors have been shown to capture the semantic characteristics of words and their context, and to overcome representations based on counting words. Semantic analysis will be performed by combining predictive representations using continuous vectors and information from denoising. This combination could be done by the rescoring component. All our models will be based on the powerful paradigm of DNN.

The performances of the various modules will be evaluated on artificially noisy speech signals and on real noisy data.

References


2017.


Compétences

Required skills: Background in mathematics, machine learning (DNN), statistics, natural language processing and computer program skills (Perl, Python).

English writing and speaking skills are required in any case.

Candidates should email a detailed CV with diploma

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération


Monthly salary after taxes : around 1596,05€ for 1st and 2nd year. 1678,99€ for 3rd year. (medical insurance included.)

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n’est pas garanti.