Chatbot Intent Classification
15 Kasım 17

Dialog systems have two main tracks in the textual space, natural language understanding and natural language generation. Such a system should first resolve the utterance given as input and second produce a relevant answer to give as output for it. Helping interact and integrate the two pieces, another important part of dialog systems is dialog management part, which mainly controls the context of the conversation in a state-tracking based approach.

In the current task, we focus on natural language understanding track, a limited and base form of it. The field is known as utterance classification which comprises domain identification, intent classification and slot filling [1]. Indeed, dialog systems can be open-domain (chit-chat) or closed-domain (task-oriented) and this domain distinction is a factor in the choice of the model. In particular, developing a consistent talker on any subject is hard and the problem is approached generally using sequence to sequence neural systems, like a translation problem. Task-oriented systems are easier and the datasets are relatively well-defined, restricting the context and the process of intent detection. These systems are generally modelled with end to end neural systems. Currently, we study task-oriented systems as our final system will also be in a specific domain.

Approaches for utterance classification

We can illustrate utterance classification like this, suppose the user asks "how can I get a refund". The system should correctly predict this query to be in the intent class "refund-request", for example.

In [2], the authors model neural language model, an RNN and an LSTM system for binary intent classification. It has been shown that parameter selection is very important for the task and in the end LSTM-based system performs the best. The dataset is called ATIS [3], a speech recognition dataset about airline flights, annotated for the task of slot filling.

In a 2011 paper [4] on intent classification, the method is based on word n-grams but its limitations such as not being able to capture long term dependencies have been mentioned as a

drawback to be overcome. The authors simplify sentences using their dependency parses and taking only the top constituent and its dependents as the simplified sentence. However, the simplified sentences improve performance only when combined with the original sentence. In any case, it appears to be a promising method. Semantic role labelling has been mentioned as another promising approach for efficient intent classification, though not tried.

Jurafskyand Martin explains in their book [5] different approaches to natural language understanding. For domain identification, intent extraction and slot filling, the possible methods are detection by hand-crafted rules (like "I want to go to CITY on DATE at TIME with a DESC. flight." where DESC can be cheap, fast depending on the dialog), semantic parsing for capturing the action, the agent and the patient in the sentence, hybrid approaches, regular expressions and supervised methods. To apply supervised methods, a set of training instances which are sentences with domain and intent labels, as well as slots should be available and this set shouldn't be small. Afterwards, the most widely applied approach is to extract features with word n-grams, possible with additional features like the existence / occurrence of named entities and lexicon words and training a classifier built upon such a model.

Systems

There are various systems developed in labs, AI-related firms or by individual efforts. Most are experimental in the field of deep learning and most of the promising approaches are closed-source. It is generally suggested that commercial chatbots should be based initially on rule-based approaches but if there is a lot of data (thousands of sentences labelled with their intents, for intent recognition, for example), neural network systems can be tried.

In this regard, we try the following system developed to detect intents in user utterance with their dataset. From whichever system we get good results in terms of learning performance and train/test duration, we will build our method on that and adapt our dataset to its scheme. Intent classification system: https://github.com/vunb/Intent-Classification-using-neural-networks

References

[1] Chen, Yun-Nung, Asli Celikyilmaz, and Dilek Hakkani-Tür. "Deep Learning for Dialogue Systems." *Proceedings of ACL 2017, Tutorial Abstracts* (2017): 8-14.

[2] Ravuri, Suman, and Andreas Stolcke. "A comparative study of neural network models for lexical intent classification." *Automatic Speech Recognition and Understanding (ASRU), 2015 IEEE Workshop on*. IEEE, 2015.

[3] P. J. Price, "Evaluation of spoken language systems: The ATIS domain," in Proceedings of the DARPA Workshop on Speech and Natural Language, Hidden Valley, PA, June 1990.

[4] Tur, Gokhan, et al. "Sentence simplification for spoken language understanding." *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011.

[5] Jurafsky, Dan, and James H. Martin. *Speech and language processing*. Vol. 3. London: Pearson, 2017.