Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018)
IEEE Xplore Compliant Part Number:CFP18N67-ART; ISBN:978-1-5386-2456-2

Joint Slot Filling And Intent Prediction for Natural Language Understanding in Frames Dataset

Jose K J
Information Technology
Rajagiri School of Engineering and Technology
Kakkanad, India

Lakshmi K S
Assistant Professor
Information Technology
Rajagiri School of Engineering and Technology
Kakkanad, India

***Abstract*—Spoken Dialogue System, Chatbots has emerged as an important research topic in artificial intelligence and natural language processing domain. The tasks in Spoken Dialogue System,Chatbots are mainly classified into three viz. domain classification, slot filling and intent prediction. In this paper, we present a novel method for slot filling and intent prediction by appending the intent information with each slot, which can be used in handling complex tasks such as travel planning. Inspired by Multi-Domain Joint Semantic Frame Parsing using Bi-directional RNN-LSTM, we trained the model using bi-directional RNN-LSTM to jointly predict the slot values and intent for a single text having multiple intents. This method proved to be successful in predicting slot values and intents with an accuracy of 90%.**

The system uses IOB tagged dataset generated from the Microsoft's Human-human goal oriented dataset(Frames Dataset) for training and testing.

Keywords—Joint Slot Filling And Intent Prediction, Natural language understanding, Bi-directional LSTM, IOB tagging, Frames Dataset

I. Introduction

With the advancement of deep learning algorithms, a variety of realistic goal-oriented conversation understanding systems have been built for various diverse domain applications that are helpful in accomplishing day-to-day tasks for human users[1][2][3]. These intelligent conversation systems are mainly based on Natural Language Understanding or Spoken Language Understanding. Three major tasks in such applications are domain classification, intent determination and slot filling which could help in constructing a semantic frame.

To jointly predict the slots and intent different approaches have shown promising results[4][5] and these projects were mainly based on ATIS dataset. As these projects were considering simple conversation they have not taken into account the scenario of having slot values in a sentence representing different intent.

In frames dataset[6], a single chat conversation of user have multiple intent, for accurate natural language understanding it is necessary to correctly identify intent of each slot(inform, request, negate etc).The system uses a novel technique for slot filling,intent prediction by appending the intent information with each slots, which can be used for language understanding of complex task such as travel planning. The proposed method was implemented on dataset generated from the frames dataset. The proposed system is trained using a bi-directional RNN model for joint intent prediction and slot filling and it produced average accuracy of 90%.

Sentance	i	would	leave	from	kochi
IOB Tag	O	O	O	O	B_inform_or_city

Fig. 1. An example of *input text and corresponding proposed output format*

The organization of the paper is as follows. Section 2 describes about related work of using bi-directional RNN LSTM for joint slot filling and intent prediction, attention-based RNN for joint intent detection and slot filling. Section 3,describes about dataset used in the proposed method,section 4 we describe about our proposed method,Implimentation is described in section 5,section 6 contains results and finally section 7 contains conclusion

II. Related Work

A. *Multi Domain Joint Semantic Frame Parsing*

Previous work using bi-directional RNN-LSTM is related to joint intent prediction, slot filling[4] it had shown promising improvement over other alternative methods also it had the advantage of using a single model for both slot filling, intent prediction these results have inspired us to use the bi-directional RNN-LSTM for our propsed method,here they have mainly used ATIS dataset as benchmark. Also the paper have not taken into account the conversation having multiple intents as present in frames dataset.

B. *Attention Based RNN Model*

Attention based RNN Model [5] is the other existing work for joint intent prediction and slot filling and it is based on Attention based neural network model and the authors have not taken into account conversation having multiple intents

III. Dataset Description

Most of the research in the past used ATIS dataset for slot filling, intent prediction. For task's like travel planning the chat conversation are complex compared to one in ATIS dataset as single chat text from user may have multiple information and they may contain different intent, so in order to do accommodate conversations having muliple intent proposed system uses a IOB tagged dataset developed by extracting usefull information from the chat history provided in frames dataset. The Fig 2. shows Frames data format.


```
▾0:
  ▾turns:
    ▾0:
      ▸text:              "I'd like to book a trip … tight budget of 1,700."
      ▾labels:
        active_frame:     1
        ▸frames:          […]
        ▾acts:
          ▾0:
            ▾args:
              ▾0:
                key:      "intent"
                val:      "book"
            name:         "inform"
          ▾1:
            ▾args:
              ▾0:
                key:      "dst_city"
                val:      "Atlantis"
              ▾1:
                key:      "or_city"
                val:      "Caprica"
              ▾2:
                key:      "str_date"
                val:      "Saturday, August 13, 2016"
              ▾3:
                key:      "n_adults"
                val:      "8"
              ▾4:
                key:      "budget"
                val:      "1700"
            name:         "inform"
```


Fig. 2. *Frames Dataset format*

The proposed IOB format tagged dataset was generated by writing a python script which extracted all text corresponding to user. As previously described Fig 1. The generated dataset for training aligned in the following manner.First part of each line indicated the user conversation text and then after a tabspace iob tagged data is placed.The intent information available from the frames dataset is prefixed with each slots value for accurate natural language understanding in conversations having multiple intent

In proposed system 81% of frames data is used as training data and 19% is used for testing.

IV. Proposed Method

In the proposed method, joint slot filling, corresponding intent prediction is performed using Bi-directional RNN LSTM[10] architecture[9] as shown in Fig 3, previous work [4] have successfully implemented joint slot filling, intent prediction on conversations having single intent. Authors of [7] have found that bi-directional LSTM are significantly more efficient than LSTM and they have found that bi-directional LSTM are well suited where context information is vitally important. As mentioned in section 3, proposed system appended the intent information along with each slot value so that it can accurately perform natural language understanding of complex sentences involved in task's like travel planning.

For complex task completing bots accurate intent prediction, slot filling is necessary for successfully performing semantic parsing(ie natural language understanding).

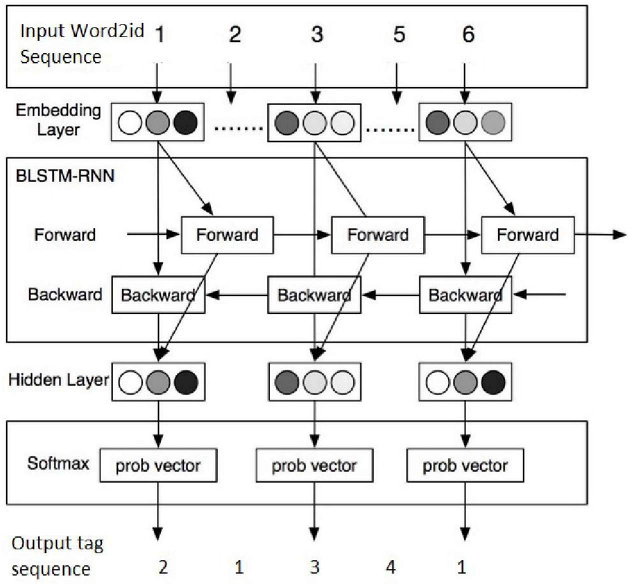

Fig. 3. *Architecture of joint slot filling,Intent prediction*

V. Implementation

Natural Language Understanding is implemented in python using the Keras library. Initially during the training phase,the system generate word to index, tag to index ie. create a dictionary of all words,tags in the training set. Also system similarly creates index to word and index to tags respectively.Once the dictionary is created, during training before each text is given one by one for training, the words and tags are replaced by corresponding unique integers, the integer sequence will be then given to the embedding layer which converts all the integer number sequence to real valued numbers.After the data is converted to real valued numbers it is passed to bidirectional LSTM ,then the input data is passed through forward layer ,backward layer and finally through softmax layer.

Parameters	Value
Number of hidden Layers	50
Dropout ratio	0.25
Optimiser	Adam
Embedding Size	50
Dropout	True
Number of Epochs	3000

Fig. 4. *Training Parameters*

The training phase is as follows, the system is given the user text as input and the system tries to predict the corresponding IOB slots tag, once it predicts, the predicted output is compared with the actual output in the training data.The difference(loss) is computed and performs backpropagation and step is repeated until the difference is negligible. The *system* used RELU activation function for hidden layers and softmax function is used as the final layer, system uses adam optimiser[9].Also the model was tested with test data after each epochs.The trained model,resulting graphs were saved during each epochs

VI. Results

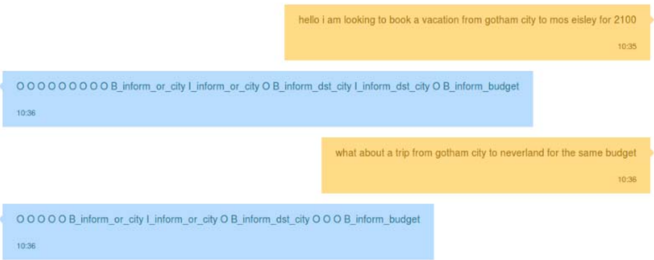

Fig. 5. *Natural Language Understanding*

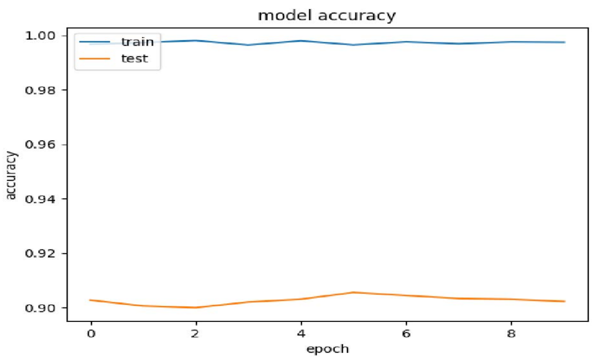

Fig. 6. *Model accuracy recorded in final epoch*

When user inputs the conversation text, natural language understanding module transforms it to IOB tagged format with intent information appended with each slots as shown in the above figure Fig5.

The model accuracy increases gradually as the number of training epoch advances. The Fig.6 shows the graph of the model accuracy during the final epoch of the training

VII. Conclusion

This paper proposes a novel method for slot filling, intent prediction by appending the intent information with each slots, which can be used in natural language understanding of complex tasks such as travel planning. The proposed method was successful in accurately semantic parsing the conversation text having multiple intent types. The promising result suggest interesting future directions: 1) using word2vector method instead of generating the words dictionary as the dictionary method fails when words not present in dictionary is present in testing, 2) can be incorporated in natural language understanding module of complex task completing bots

References

[1] Xuijun Li, Yun-Nung Chen, Lihong Li, Jianfeng Gao, and Asli Celikyilmaz, “End-to-end task-completion neural dialogue systems,” in Proceedings of IJCNLP, 2017.

[2] Xiujun Li, Zachary C Lipton, Bhuwan Dhingra, Lihong Li,Jianfeng Gao, and Yun-Nung Chen, “A user simulator for task-completion dialogues,” arXiv preprint arXiv:1612.05688,2016.

[3] Baolin Peng, Xiujun Li, Lihong Li, Jianfeng Gao, Asli Celikyilmaz, Sungjin Lee, Kam-Fai Wong“Composite Task-Completion Dialogue Policy Learning via Hierarchical Deep Reinforcement Learning” arXiv:1704.03084

[4] Dilek Hakkani-Tur, Gokhan Tur, Asli Celikyilmaz, Yun-Nung Chen,Jianfeng Gao, Li Deng, and Ye-Yi Wang, “Multi-Domain Joint Semantic Frame Parsing using Bi-directional RNN-LSTM,” Proceedings of The 17th Annual Meeting of the International Speech Communication Association (INTERSPEECH 2016)

[5] Bing Liu, Ian Lane, “Attention-Based Recurrent Neural Network Models for Joint Intent Detectionand Slot Filling,” arXiv:1609.01454

[6] Layla El Asri, Hannes Schulz, Shikhar Sharma, Jeremie Zumer,Justin Harris, Emery Fine, Rahul Mehrotra, Kaheer Suleman,Frames:a corpus for adding memory to goal-oriented dialogue systems,ArXiv:1704.00057v2 [cs.CL] 13 Apr 2017

[7] A. Graves and J. Schmidhuber, “Framewise phoneme classifica-tion with bidirectional LSTM and other neural network architec-tures,” Neural Networks, vol. 18, no. 5, pp. 602–610, 2005.

[8] Diederik p. kingma and Jimmy Lei Ba, “Adam: A method for stochastic optimization,” Published as a conference paper at ICLR 2015.

[9] Yushi Yao and Zheng Huang,“Bi-directional LSTM Recurrent Neural Network for Chinese Word Segmentation” arXiv:1602.04874v1

[10] S. Ravuri and A. Stolcke, “Recurrent neural network and LSTMmodels for lexical utterance classification,” inInterspeech, 2015.