Educational Chatbot to Support Question Answering on Slack

Simone Leonardi[1][0000−0002−8009−1082] and Marco Torchiano[1][0000−0001−5328−368X]

Politecnico di Torino. 24, 10129 Turin, Italy
{simone.leonardi,marco.torchiano}@polito.it

Abstract. Educational chatbots unlock new possibilities to support teaching activities. In recent years, the progress made in natural language processing and understanding allowed the development of virtual assistants that understand questions written as unstructured data. They generate a proper answer based on knowledge bases. This work proposes an accurate solution easy to test in real world scenarios like university courses. In fact, the use of Slack and RASA technologies reduces the difficulties to train and install the educational chatbot on already existing digital workplaces. We evaluates the effectiveness of our solution reaching 0.93 accuracy in intent recognition building a stacked neural network.

Keywords: Educational Chatbot · Question Answering · NLP · Remote Teaching.

1 Introduction

The COVID-19 pandemic increased the needs for digital tools also in the educational system. Students attended online classrooms with remote teachers. In this context universities look for effective digital supports to continue offering their services without losing the interaction with the students. It has been shown that the adoption of Artificial Intelligence in education is rapidly expanding[7]. In detail, one of the most popular AI tool in education is the chatbot system [4]. A chatbot is a virtual assistant able to interact with human inputs, process it and produce an answer and it also promotes further development in the conversation. In a scenario where students need to interface with professors remotely, there is the possibility to support teachers in answering questions sometimes recurrent or already answered in the previous edition of the course. An educational chatbot is the ideal support for this specific task. In [2], Cunningham et al. propose practical steps to prototype and to develop such tools. In the recent years several studies have demonstrated how the adoption of chatbot in education leads to positive outputs in the personalization of contents offered and in the learning process [1], [6] , [8]. In this work we focus on the ease of deployment of a solution to be installed on Slack, a widespread digital workplace that allows interaction between users. We select RASA Open Source as a server side environment because it is easy to configure its NLU (Natural Language Understanding) architecture and it already includes tools to enlarge the Q&A dataset directly in the Slack channel when submitting a new answer. We developed this work by answering these research questions: *RQ1*: Which dataset can be used to train and test a chatbot in the educational system? *RQ2*: Which natural language understanding model is the most accurate in question answering for the educational context? *RQ3*: Is a chatbot deployed on a common chat platform able to understand a student's intent during a multi turn conversation? The paper is structured as follows: in Section 2 we review related work, while in Section 3 we describe the proposed approach for training and classifying students' questions. In Section 4 we present a preliminary experimental evaluation of our method. Eventually, in Section 5, we conclude with a summary of the experiment outcome and outlining future work.

2 Related Work

The role of chatbot in education have been widely explored in the recent years leading to a large number of publications [7], [4], [2], [6], [1], [8], [5]. Students interact with chatbot systems to receive immediate answers to common questions even when the professor is not available or it is difficult to retrieve the information on the course material. In this work we offer a solution for the students that lack the assistance of professors in large courses, especially those that are held online, providing a conversational agent able to immediately answer a question based on previous answers given by professors. On the other side this agent is able to learn new answers by interacting directly with professors. We compare our solution with different baseline models to assess its accuracy. The final pipeline is described in Section 3, where we define a client-server infrastructure to solve this issue. The model is expandable with Q&A of the same topic or by changing the training dataset with different school subjects. Our educational chatbot also allows a fallback procedure to forward the question to a human assistant.

3 Approach

In the following we explain the data retrieval phase, the model implementation, and the interaction pipeline.

3.1 Dataset

The questions and the answers collected came from the Object Oriented Programming course held during the COVID-19 pandemic at Politecnico di Torino in the spring semester 2020. All the students' information have been anonymized for privacy reasons. This course provided remote assistance with the support of the Slack web application. The workspace of the course has been divided into two main channels: theory and laboratories. The first channel contains questions about frontal theory lessons while the second channel is about the exercises given during the practical sessions of the course. The topic selected for the experiment is the *inheritance* with its 53 subtopics and 300 questions answered. This is the training set of our model while for the testing part, we used the same topic questions from the 2021 edition of the same course. The intent plays the role of labels in the context of supervised classification. All the questions have been cleaned by removing emojis, url and references to previous questions in the channel with tags and, similarly, all the handles to cite a user in the channel have been deleted. Finally questions have been labeled as *intents* while answers have been labeled as *actions*. Questions and answers are paired following a specific format[1]. This dataset answers our first research question RQ1 and it also explains how to build a dataset to be processed in our model.

3.2 Question Answering Model

The neural network model adopted in our solution is ConveRT (Conversational Representations from Transformers). Henderson et al. in [3] propose this Transformer based dual encoder specialized in conversational tasks. In Figure 1, we illustrate the single-context ConveRT model that is the building block for the multi-context solution. In fact, only the multi-context solution is able to handle multi turn conversation. In our use case scenario, students can interact with the chatbot asking for further clarification after receiving the first answer. Similarly, they can ask multiple questions before closing the conversation. We decide to include the multi-context ConveRT in our solution because it is the best performing one with respect to the BERT and SpaCy[2] ones, as we demonstrate in the validation Section 4 and because it is faster on our cloud based solution.

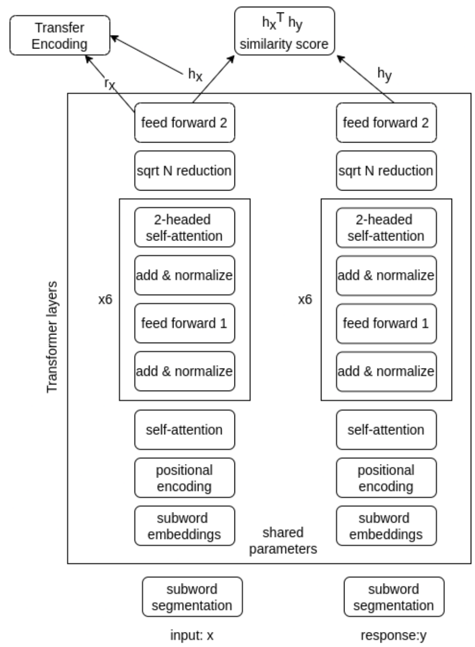
The goal of ConveRT is to select the right answer once given the question. There are two alternatives: the question exists in our dataset or it does not exist. In any case, our model tries to match the question based on similarity and automatically print the paired answer. If the answer is not satisfactory for the student, he can rephrase the question or he can choose to forward the question to the professor.

3.3 Client Server Architecture

In this section we describe the architectural choices made to develop our educational chatbot and how to set up the environment to deploy it. We use Slack[3] application as a communication channel between students and professors. It plays the role of the web client. It also hosts our educational chatbot. In parallel, we use RASA Open Source on the server side of the communication to handle the natural language understanding processes.

[1] https://rasa.com/docs/rasa/stories/
[2] https://spacy.io/models
[3] https://slack.com/

Fig. 1. Single-context ConveRT dual-encoder model architecture as presented by Henderson et al. in [3].



We start the development of the bot creating a Slack application on *https://api.slack.com/apps*. We create a RASA project and we define how our context-aware conversational chatbot will handle conversations. There are three main tasks to be configured: slots, stories and policies.

- *Slots*[4] are the assistant's memory. They store pieces of information that the bot needs to refer to later and can direct the conversation flow based on slot events.
- *Stories* are examples of conversations between a user and the bot. In stories different patterns of interaction are described.
- *Machine Learning Policies* help the chatbot to better predict the response in unseen conversation paths.

The final component of RASA is the RASA X component. It is a CDD (Conversion Driven Development) tool. It saves automatically all conversations and it uses them to improve the chatbot and its embedded NLU model.

The entire pipeline of usage of the system is the following. A student writes a question in the Slack channel where the educational chatbot is installed. The question is forwarded to the RASA server and it is processed to recognise the intent and consequently find the best answer. If the search is a failure, then the

[4] https://rasa.com/docs/rasa/domain#slot-types

question is forwarded to the professor that, answering to the question, updated the dataset and increased the knowledge base of the chatbot. In any case, the student will receive the answer.

4 Experimental Results

We assess the effectiveness and precision of our framework computing metrics over the intent recognition functionality. The problem evaluated is a multi label supervised classification. We compare the recall, precision, f1-score and accuracy of three main models: BERT, SpaCy and ConveRT. We use 80% of the dataset as a training set, while the remaining 20% as a test set. The results are listed in Table 1.

Table 1. Evaluation metrics computed with a 80/20 train/test set setup for the task of intent recognition.

Model	Recall	Precision	f1-score	Accuracy
ConveRT	0.93	0.91	0.92	0.93
BERT	0.92	0.89	0.90	0.92
SpaCy	0.92	0.88	0.89	0.92

We also tested 174 different conversations containing 921 different actions with a 5 fold cross validation approach. RASA offers in its suite the possibility to simulate user inputs that test various combinations of actions and turns in conversations.

Obtained results answer RQ2 and RQ3 addressing both the task of intent recognition and the correct conversation decisional path choice. We show that the best performing model is ConveRT in both tasks that reach accuracy above 0.93 and f1-score above 0.92. These results are the one obtained with the Italian language because our dataset was in Italian, it is possible that using other languages BERT or SpaCy could beat the ConveRT model, in this case the switch between different configurations is already arranged. This intuition is tested by Wu and Dredze in [9]. They prove that pre-trained language models perform better when they adopt a language with lots of digital resources publicly available and correctly formatted. As an example, English has over twenty gigabytes of data while Italian just two and the Yoruba language less than 0.01.

5 Conclusion

This work presents a new chatbot system to support question answering in education, especially useful for courses with a large number of students and in remote teaching due to pandemics. We present an educational chatbot that is deployable on Slack for ease of use. The server side of our framework, developed in RASA Open Source and Rasa X proved its accuracy of 0.99 with the

ConveRT Natural Language Understanding model outperforming the other baselines models both in intent recognition and in correct conversational decisions. We trained the model with the course Slack channel in 2020 and we tested it on questions made during 2021 reaching high accuracy in response identification. Further improvements of this work involve the adoption of the system on other topics and teaching courses, plus making a survey once deployed on real time interaction during an academic semester. *We finally thank Fiorentino Cairone for his contribution in the implementation of the educational chatbot.*

References

1. Clarizia, F., Colace, F., Lombardi, M., Pascale, F., Santaniello, D.: Chatbot: An education support system for student. In: International Symposium on Cyberspace Safety and Security. pp. 291–302. Springer (2018)
2. Cunningham-Nelson, S., Boles, W., Trouton, L., Margerison, E.: A review of chatbots in education: practical steps forward. In: 30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019): Educators Becoming Agents of Change: Innovate, Integrate, Motivate. pp. 299–306. Engineers Australia (2019)
3. Henderson, M., Casanueva, I., Mrkšić, N., Su, P.H., Wen, T.H., Vulić, I.: ConveRT: Efficient and accurate conversational representations from transformers. In: Findings of the Association for Computational Linguistics: EMNLP 2020. pp. 2161–2174. Association for Computational Linguistics, Online (Nov 2020). https://doi.org/10.18653/v1/2020.findings-emnlp.196, `https://aclanthology.org/2020.findings-emnlp.196`
4. Okonkwo, C.W., Ade-Ibijola, A.: Python-bot: A chatbot for teaching python programming. Engineering Letters **29**(1) (2020)
5. Okonkwo, C.W., Ade-Ibijola, A.: Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence **2**, 100033 (Jan 2021). https://doi.org/10.1016/j.caeai.2021.100033, `https://www.sciencedirect.com/science/article/pii/S2666920X21000278`
6. Ranoliya, B.R., Raghuwanshi, N., Singh, S.: Chatbot for university related faqs. In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). pp. 1525–1530. IEEE (2017)
7. Roos, S.: Chatbots in education: A passing trend or a valuable pedagogical tool? (2018)
8. Sinha, S., Basak, S., Dey, Y., Mondal, A.: An educational chatbot for answering queries. In: Emerging Technology in Modelling and Graphics, pp. 55–60. Springer (2020)
9. Wu, S., Dredze, M.: Are all languages created equal in multilingual bert? (2020). https://doi.org/10.48550/ARXIV.2005.09093, `https://arxiv.org/abs/2005.09093`