

Crowdsourcing for Creating a Dataset for Training a Medication Chatbot

Cyril R. ZGRAGGEN^{a,1}, Sebastian B. KUNZ^{a,1} and Kerstin DENECKE^{a,2}

^a*Bern University of Applied Sciences, Switzerland*

Abstract. To facilitate interaction with mobile health applications, chatbots are increasingly used. They realize the interaction as a dialog where users can ask questions and get answers from the chatbot. A big challenge is to create a comprehensive knowledge base comprising patterns and rules for representing possible user queries the chatbot has to understand and interpret. In this work, we assess how crowdsourcing can be used for generating examples of possible user queries for a medication chatbot. Within one week, the crowdworker generated 2'738 user questions. The examples provide a large variety of possible formulations and information needs. As a next step, these examples for user queries will be used to train our medication chatbot.

Keywords. Conversational user interface, crowdsourcing, artificial intelligence, medication management, chatbot, natural language understanding

1. Introduction

Chatbots or conversational user interfaces gain in interest for realizing dialog-based healthcare applications. In previous work, we developed a medication management assistant, eMMA [1], which contains a rule-based chatbot that answers user questions related to the prescribed medication. Evaluations of eMMA demonstrated that the chatbot still requires a more extended knowledge base for being more robust against varying user input [2]. The objective of this work is to collect experiences on how training examples for eMMA can be created in a crowdsourcing process.

2. Methods

We based our crowdsourcing process upon the generic process for generating chatbot training data introduced by Bapat et al. [3]. It comprises three main steps: 1) Preparatory work: clarify use cases for our medication chatbot and create entity-intent model, 2) Creation of orders in the crowdsourcing platform, 3) Collect and control the data. With intent we refer to the goal the user has in mind when typing in a question or reply in the system. Our entity-intent model comprises possible intents and entities, the medication chatbot is supposed to recognize. We selected

¹ Contributed equally.

² Corresponding Author, Kerstin Denecke, Bern University of Applied Sciences, Institute for Medical Informatics, Quellgasse 21, Biel, Switzerland; E-mail: Kerstin.denecke@bfh.ch.

- Five entities: Drug name, active pharmaceutical ingredient, dosage form, dosage, time; and
- Six intents: Composition, indication / possible application, dosage / use, properties / effect, adverse effect, warnings / precautions.

This model forms the basis for creating tasks in the crowdsourcing platform. We formulated our tasks in a way that crowdworker had to create questions given a particular intent and a predefined combination of entities. 90 tasks were posted at Amazon Mechanical Turk and were available for crowdworker for one week.

3. Results

In a period of 7 days, we received 4'557 answers from 560 crowdworker. 61 % (2'798) were accepted in the control step and 39 % (1'759) were rejected. Reasons for a reject was a wrong format of the query or not addressing the given task (e.g. when not a query was formulated). After removing 60 obvious duplicates, we ended up with 2'738 answers. The highest response rate was achieved on the first day. It consistently dropped in the following days. Tasks containing only one entity achieved the highest response rates with more than 60 %. Tasks with two entities were only addressed with less than 60 % response rate. This means, more complex tasks are less often answered by the crowdworker. Bapat et al. reported that 117 training examples per intent are sufficient for high quality AI-based chatbots [3].

4. Discussion and Conclusions

Our experiment showed that crowdsourcing can deliver valuable insights for creating a knowledge base and a conversation for a medication chatbot. We learned more about information needs of individuals with respect to their medication. While domain experts are often missing the concrete patient concerns and questions, crowdworker well represent possible users of our chatbot. We received a broad variety of possible user needs and linguistic variants. The formulated sentences also contain spelling errors, which could help to make our chatbot more robust against linguistic variations, and errors. As a next step towards our medication chatbot, the collected data has to be further assessed with respect to similar or error-prone sentences. Levenshtein distance could be used to ensure that examples are significantly different before they are integrated in the chatbot knowledge base. From the crowdsourcing process, we learned that it is important to formulate simple tasks in a very clear manner and to limit the difficulty of a task.

References

- [1] Tschanz M, Dorner TL, Holm J, Denecke K. Using eMMA to manage medication. *Computer*. 2018 Aug 14;51(8):18-25.
- [2] Hess GI, Fricker G, Denecke K. Improving and evaluating eMMA's communication skills: a chatbot for managing medication. *Stud Heal Technol Inf*. 2019;259:101-4.
- [3] Bapat R, Kucherbaev P, Bozzon A. Effective crowdsourced generation of training data for chatbots natural language understanding. In: *International Conference on Web Engineering 2018*; p. 114-128.