Crowdsourcing for Creating a Dataset for Training a Medication Chatbot

Sebastian Kunz, Cyril Zgraggen, Kerstin Denecke

Introduction

Chatbots 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.

Methods

We based our crowdsourcing process upon a generic process for generating chatbot training data [3]. It consists of three main steps: 1) Clarify use cases for our medication chatbot and create entity-intent model, 2) Create orders in the crowdsourcing platform, 3) Collect and control the data. The entity-intent model comprises possible intents and entities, the medication chatbot is supposed to recognize. Intents are the goal the user has in mind when typing a query or statement into the system. We selected

- 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.

Based on this model, 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.

Results

We received 4'557 answers from 560 crowdworker (see Table). 61 % (2'798) were accepted in the control step. 39 % (1'759) were rejected because of a wrong format of the query or due to ignoring the given task. After removing 60 obvious duplicates, we ended up with 2‘738 answers. More complex tasks were less often answered by the crowdworker.

- 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.

Intents	Accepted samples	Rejected samples
Warnings / precautions	What precautions should I follow while taking dosage at "time"	Adverse effects can occur with the misuse of this medication, all due to its pharmaceutical active, be careful.
Adverse effect	What "time" should I wait between taking repeated dosages for it to be safe?	Does the "time" have an adverse effect?
Composition	What is the best " time" to take it according to its ingredients	How much "time" to write
Properties / Effect	What happens if I take the drug at this "time" In what amount of "time" will I notice less acid reflux	what is the "Time"
Indication / possible application	What is the recomended amount of "time" you shoud wait before ingesting it again?	After what "time" should I reset problem computer
Dosage / use	At what "time" this dosage has to be completed?	The pharmaceutical active ingredient used for a long time causes addiction.

Lessons learnt

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. The crowdsourcing results delivered 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. We received a broad variety of possible user needs and linguistic variants. The questions formulated by the crowdworker also contain spelling errors, which could help to make our chatbot more robust against linguistic variations and errors.

Next steps

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 integrated in the chatbot knowledge base that significantly differ from each other.

[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; pp. 114-128.

Bern University of Applied Sciences

Contact:
Kerstin Denecke, kerstin.denecke@bfh.ch
Bern Universtiy of Applied Sciences
Institute for Medical Informatics
Quellgasse 21 | CH-2502 Biel