Chatbot challenge #1: User Intent Recognition

Doris Chi [Follow]

Jul 30, 2019 · 8 min read ★
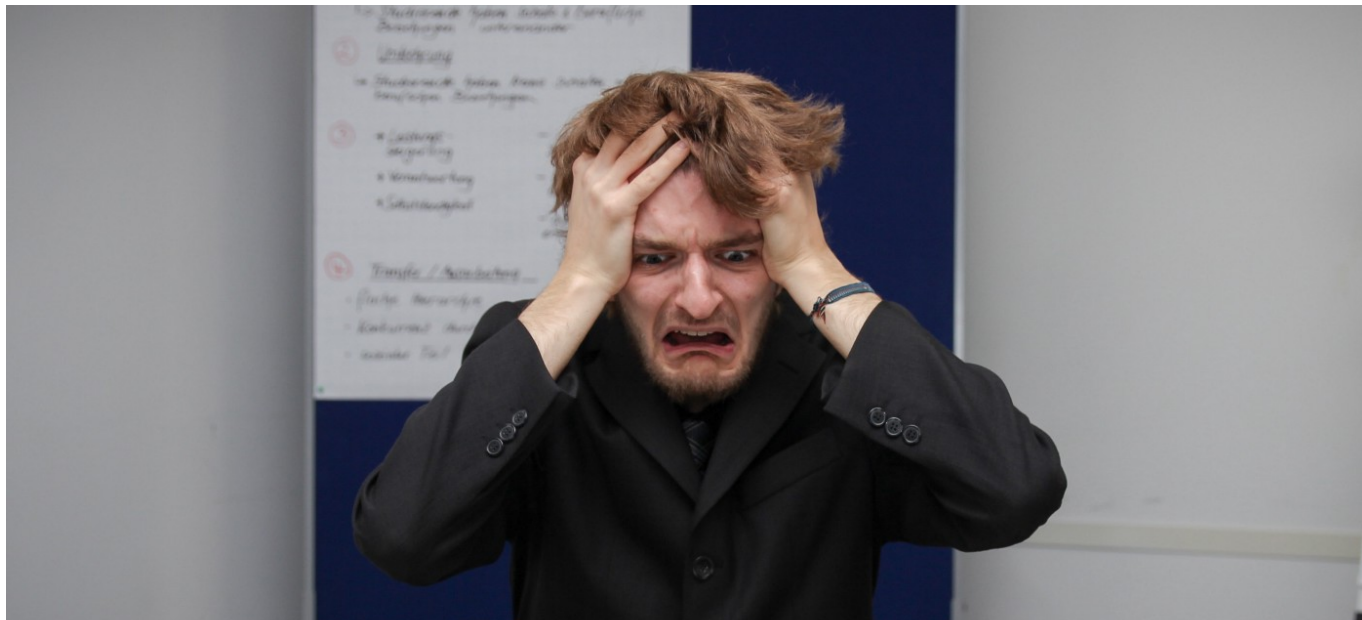
Photo by Sebastian Herrmann on Unsplash

No matter what kind of chatbot you want to build, one to answer questions or perform tasks, to search a knowledge base or follow dialogue flows, your chatbot needs to understand what your users say, more precisely, it needs to understand what they intend to do. The process to understand what your users want is commonly referred as "User Intent Detection" or "User Intent Recognition".

Can your chatbot understand the human language? Probably not. However, it doesn't need to "understand" the human language as a human does. It only needs to make people feel like it knows what they want and respond accordingly.

How can your chatbot make people feel like it understands them? You need to train it either by defining rules for it to follow or providing examples for it to learn by itself. The most popular method now is to define user intents as different categories and provide some training examples for each category. When a new user query comes in, the chatbot calculates the similarity between the new query and training examples of predefined user intents, then categorize the query as the user intent with the highest similarity score or confidence score. If the highest similarity score is lower than an arbitrary threshold, then it will be categorized as unknown. This might not sound

very hard, however, it is one of the biggest challenges for building a usable chatbot.

Top 4 Most Popular Bot Design Articles:

1. [Ultimate Bot Design Guide](#)

2. [How to design a Chatbot](#)

3. [Distributing a slack app](#)

4. [Chatbot Conference in San Francisco](#)

Here are the five reasons why user intents are so hard to detect:

1.Natural language is hard to understand, even for a human.

To understand natural language, you need to analyze it in five layers:

a. Convert sound to words (Phonology & Phonetics Layer)

If your chatbot support voice input, this layer matters a lot! The accuracy of speech recognition is crucial to the following steps. However, some current systems are almost as accurate as humans. So this layer is OK.

b. Understand meaning of each word (Morphology & Lexicology Layer)

The two biggest challenges to understanding the word meanings are disambiguation and handling unknown words. Disambiguation is important when a word has multiple meanings. A human can do it pretty well in most cases. However, there are cases where even a human cannot determine. For example, in "Do you smell gas?" gas could mean gasoline (liquid) or other substances (matter); In "He will be the president of the U.S.", does"will" mean a prediction or a fact? For cases like this, it will be impossible for both humans and computers to disambiguate unless more information is provided in context. Handling unknown words or out of vocabulary words is another big challenge. If there are only one or two unknown words in a

sentence, we might be able to guess their meaning based on context and syntactic information. However, if there are too many unknown words in a sentence, even a human cannot guess their meaning, let alone computers.

c. Understanding the sentence structure (Syntax Layer)

One biggest challenge to understanding sentence structure is also disambiguation. For example, in "a man is watching a woman in the garden with a telescope.", it is possible that the man is using a telescope to watch a woman, but it is also possible that the woman in the garden has a telescope. This kind of ambiguation is hard for both humans and computers to understand the meaning of a sentence sometimes. Besides, people don't draw a syntactic structure every time to understand a sentence. People might be able to tell what is subject, object or verb. However, most people without linguistics background cannot tell much more than that. Therefore it is not clear whether a syntactic structure will help computers to understand natural language or not.

d. Understand the meaning of the sentence (Semantics Layer)

The challenge of understanding the meaning of a sentence is a combination of the challenge of understanding the word meaning and syntactic structures.

e. Understand the meaning of the sentence in context (Pragmatics Layer)

This is the hardest layer of all. In addition to understand what people say, we also need to know why people say it, or what do they want to achieve by saying it. For example, if a user says "I see a weird transaction on my credit card", does he/she want to dispute it directly or does he/she want more information about the transaction? If a user asks "where is the restroom?" Does he/she want to wash hand, use the toilet, or change diaper for his/her kids? It is possible for both humans and computers to "understand" a user query's literal meaning. To understand the true intent behind it, however, a lot more information about the user and the situation is required.

2. There is no way to represent natural language without natural language.

As human beings, we understand natural language by aligning them with concrete objects or abstract concepts. For computers, they understand languages using numbers. For example, "apple", for people, could mean that round red sweet juicy fruit that many people like to eat. For computers, "apple" is just word #2718 (or any number) in their lexicon. It contains one character #97 ('a'), two #112 ('p'), one #108 ('l') and one #101 ('e'), It has no relation to other words like "round", "sweet", "juicy", "fruit" or "eat" unless you define the relations of those words somehow. We could feed computers more information about "apple" like it is a noun in its singular form, it is one kind of fruit, its color could be red, green or yellow, its shape is close to round, and it can be a subject or object. However, it is hard to represent the extra information in a purely logical or numerical way without using any natural language. If we use natural language to represent the extra information, we will bring the ambiguation issue in natural languages to the extra information and make it hard for computers to understand.

3. Different user intents might have the same response.

It is pointless to define user intents without thinking about how to fulfill them, and it is very common for different user intents to have the same response. For example, "where is the garage?" and "where is the kitchen?" are two different user intents, however, they can both be fulfilled be a floor plan which shows the location of the garage and the kitchen. If we treat them as different intents, since most chatbot tools and platforms (except for IBM Watson) only support 1 to 1 relation between intents and responses, we will end up with so many entries of intents that have the same responses. It will waste a lot of storage space, and it is not maintainable cause if we want to change the response, we need to change it in all the entries. However, if we treat all the intents as one single intent, the computer might get confused.

4. Users prefer key words than whole sentences.

Users are lazy. If keywords work, they don't use whole sentences. If an emoji or a character works, they don't even use words. For example, if users send you a $ sign, do you know whether they want to check their balance, pay a bill, or transfer money? Even if we have a super intelligent system that can

understand natural language, it won't work if users just say a few words or even no words.

5. Users have so many different ways to represent the same intent.

One very extreme example I have ever heard is that Eno, the virtual assistant built by Capital One, has seen more than 2000 ways for checking balance. I have never seen so many different variations for single user intent in my <u>chatbot</u>, but it is very common for a user intent to have tens of variations. It is very hard for chatbot developers to cover all the cases beforehand. Even with the help of machine learning algorithms, it is very likely for a chatbot to miss user queries with different words or syntactic structures from training examples.

.   .   .

Final Thoughts

After one and a half years working on my chatbots and exploring different chatbot tools, I have never seen a perfect chatbot platform that can identify

user intents with just a few training examples and can handle thousands of different intents. The common ones are either able to identify the variations of intents really well but can only handle a handful of most popular intents, or able to handle thousands of intents but not so good at detecting the variations of them. I think it is because when a new query comes in, the computer needs to calculate the similarity based on all examples. So if a platform allows hundreds of training examples, it cannot handle too many intents. If a platform wants to handle hundreds of intents, it cannot allow too many examples or variations in each intent.

For chatbot developers, we can only choose the best tool based on our goal. If 80% of the queries are about 20% of the intents, then use the chatbot platform with very good user intent recognition for a small number of intents. Otherwise, use the chatbot platform that can cover many different intents. It might not be a bad idea to integrate multiple tools into the same chatbot, and use google like search engines and live agents as fallback solutions.

For chatbot users, please lower your expectations. Detecting user intent is hard even for a human, so please don't expect a chatbot to really understand

you and can tell you everything. Use chatbots when they can help and try other ways when they cannot. The more you use them, the smarter they will become!

Thanks for reading!

Don't forget to give us your 🖐️ !
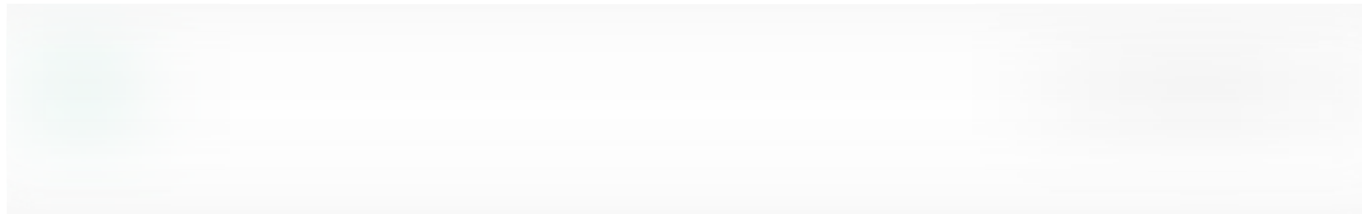
Artificial Intelligence   Chatbots   User Intent   Computational Linguistics   Bots

50 claps

Doris Chi

I am a full-stack software engineering interested in NLP and ML. Opinions expressed are solely my own and do not represent my employer.

Chatbots Life

Best place to learn about Chatbots. We share the latest Bot News, Info, AI & NLP, Tools, Tutorials & More.

Write the first response

More From Medium

More from Chatbots Life

Our experience building chatbots with Rasa — Tuning the NLU pipeline

Pierre Ricadat in Chatbots Life

Jul 10, 2018 · 7 min read

More from Chatbots Life

Integrate Dialogflow (Api.ai) Bot into Website

Parth Shrivastava in Chatbots Life

Jun 20, 2018 · 4 min read

More from Chatbots Life

Know your Intent: State of the Art results in Intent Classification for Text

Kumar Shridhar in Chatbots Life

Sep 30, 2018 · 9 min read

Discover Medium

Make Medium yours

Become a member

Welcome to a place where words matter. On Medium, smart voices and original ideas take center stage - with no ads in sight. Watch

Follow all the topics you care about, and we'll deliver the best stories for you to your homepage and inbox. Explore

Get unlimited access to the best stories on Medium — and support writers while you're at it. Just $5/month. Upgrade

Medium

About   Help   Legal