

# Cognitive Interaction with Robots

Final Project

## Deep Reinforcement Learning Open-domain Chatbot

Master in Artificial Intelligence

Universitat Politècnica de Catalunya (UPC)

Gonzalo Recio Domènech ([gonzalo.recio@upc.edu](mailto:gonzalo.recio@upc.edu))

Jana Reventós Presmenes ([jana.reventos@estudiantat.upc.edu](mailto:jana.reventos@estudiantat.upc.edu))

January 2021

**Abstract:** We present an open-domain conversational chatbot that is able to interact with users by voice through a graphical interface. Many conversational chatbots lack of engagement, are short-sighted and tend to generate incoherent or repetitive responses making the interaction last for a short period. To solve this, we propose a deep reinforcement learning approach to train its core language model (GPT Transformer) in order to improve the long-term performance. All related code and further experiment results can be found in this [repository link](#).

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Related work</b>	<b>4</b>
<b>3</b>	<b>Project description</b>	<b>5</b>
<b>4</b>	<b>Cognitive approach</b>	<b>6</b>
4.1	Conversational language model . . . . .	6
4.2	Reinforcement Learning . . . . .	7
4.2.1	RL for open-domain dialogue . . . . .	8
<b>5</b>	<b>Interaction approach</b>	<b>10</b>
5.1	Chatbot GUI . . . . .	11
5.2	Sentiment Analysis . . . . .	12
5.2.1	Sentiment analysis model . . . . .	13
5.3	Chatbot speech experience . . . . .	14
<b>6</b>	<b>Chatbot final approach</b>	<b>15</b>
6.1	System architecture . . . . .	15
6.2	Chatbot embodiment . . . . .	15
6.3	Robot-Ethics . . . . .	16
<b>7</b>	<b>System requirements</b>	<b>16</b>
7.1	Hardware and Software . . . . .	17
7.2	Context . . . . .	17
7.3	Chatbot cognitive skills . . . . .	17
7.4	Chatbot behavior . . . . .	17
7.5	Chatbot constraints . . . . .	18
<b>8</b>	<b>Conversational model training and experiments</b>	<b>18</b>
8.1	Experimental setup . . . . .	18
8.2	Evaluation metrics . . . . .	19
8.3	Experimental results . . . . .	19
8.3.1	DialoGPT-Medium vs. DialoGPT-Medium with Reinforcement Learning . . . .	21
<b>9</b>	<b>Human-interaction experiments</b>	<b>21</b>
9.1	Statistical hypotheses . . . . .	22

9.2	Study design . . . . .	22
9.2.1	Participants . . . . .	22
9.2.2	Procedure . . . . .	23
9.3	Human-interaction evaluation . . . . .	24
9.3.1	Hypothesis 1: Reinforcement Learning . . . . .	24
9.3.2	Hypothesis 2: Age - range . . . . .	26
9.4	Discussion . . . . .	27
<b>10</b>	<b>Conclusions</b>	<b>28</b>
10.1	Future work . . . . .	28
<b>11</b>	<b>Acknowledgements</b>	<b>29</b>
<b>A</b>	<b>Appendix</b>	<b>32</b>
A.1	Generated Dialogues with <i>Titanic Test</i> . . . . .	32
A.2	Generated Dialogues with real Users . . . . .	36
A.2.1	Experiment 1 . . . . .	36
A.2.2	Experiment 2 . . . . .	38
A.2.3	Experiment 3 . . . . .	41
A.2.4	Experiment 4 . . . . .	48
A.3	Questionnaire . . . . .	49

# 1 Introduction

Hey Alexa, could you recommend me a song? Siri, please, is there any open restaurant? Currently we can talk to almost all the smart devices. Artificial Intelligence (AI) powered chatbots and virtual assistants have been widely adopted by enterprises seeking to enhance their customer service, improve productivity and boost revenue. Such chatbots however don't do so well in off-target domains, therefore, users can't engage in a conversation of any topic.

A chatbot is a system, powered by rules and AI, which simulates a real interaction with users via chat interface. These systems have shown great promise in different areas such as marketing systems, customer support or personal assistance. Furthermore, there is favorable preliminary evidence for psychiatric use of chatbots and high satisfaction rating across their use [Vaidyam et al., 2019].

Different kind of dialog systems exist such as social chatbots or task-oriented chatbots. However, closed-domain chatbots respond to keywords or intents to accomplish a specific task so, users, can't engage in conversation in any topic.

With the advent of deep learning technologies in conversational systems (i.e. BERT [Devlin et al., 2019], DialoGPT [Zhang et al., 2019]) and the vast availability of text datasets, creating a chatbot that simulates human language conversation is still challenging but possible. One of the main problems of chatbots is that they tend to easily generate incoherent sentences or highly generic responses such as "I don't know" regardless of the input on the long term [Li et al., 2016], therefore, they have a lack of engagement.

To address this issue, in this project we implement a voice-based conversational chatbot using reinforcement learning. Our goal is to improve users engagement by making dialogues last long to have a fluid and natural conversation. Also, we create a friendly user interface to make users feel comfortable and strengthen their experience.

In the following sections, we are going to first introduce some related work about the topic and describe the project objectives. After that, we will explain the cognitive and the interaction parts of our robot system, and, finally, the results and experiments carried out along with the conclusions.

## 2 Related work

Conversational agents or dialogue systems, also called chatbots, are intended to carry coherent conversations with humans in natural language text or speech. Chatbots can be classified into two major categories [Mnasri, 2019]:

- Social chatbots: designed to carry unstructured human-like conversations and also called chit-chats. Some first examples are ELIZA [Weizenbaum, 1966], PARRY [Colby, 2013], XiaoIce [Zhou et al., 2020], etc.
- Task-oriented chatbots: designed to answer closed-domain questions during human conversations. They can be generalist (respond general utterances) or specialist (designed for a very specific domain) task-oriented bots.

Unlike closed-domain chatbots, which are trained to respond specific tasks, open-domain chatbots are designed to engage in conversations across any topic. Nonetheless, open-domain chatbots still have weakness as sometimes respond in ways that do not make sense, with repetitive utterances or replies that are vague and generic like "I don't know".

Some open-domain chatbots are XiaoIce [Zhou et al., 2020], Mitsuku [Worswick, 2018] or Cleverbot<sup>1</sup>, however, they rely on complex frameworks such as dialog managers with retrieval-based, knowledge-base or rule-based systems. End-to-end neural network approaches offer the simplicity of a single learned model [Shang et al., 2015, Vinyals and Le, 2015, Serban et al., 2016] [Zhang et al., 2019, Adiwardana et al., 2020]. State-of-the-art solutions, such as Meena [Adiwardana et al., 2020], use GPT-2 transformer model and evolved transformer [So et al., 2019] with sequence to sequence (seq2seq) model.

Exist two main chatbot building architectures: rule-based and data driven chatbots. Rule-based frameworks use specific rules to converse with users, therefore, when they need to perform complex tasks requires writing many hand-crafted rules which is time consuming. Thus, data-driven chatbots are the most used as they use existing human language databases to train machine learning models such as seq2seq [Cho et al., 2014] or reinforcement learning [Li et al., 2016].

In order to explore the ability to converse freely in natural language and overcome the main weaknesses of chatbots, we develop an end-to-end system with seq2seq models (encoder-decoder RNN) using reinforcement learning (RL) paradigms, which could improve dialogue long-term performance [Li et al., 2016].

### 3 Project description

The proposal of this project is to design and implement an open-domain voice-based chatbot that is capable of having engaged conversations with real users. In order to build this solution we must consider both the dialog system cognitive frame and user interaction to achieve our objective.

**Project objectives:**

1. **Develop a dialogue generation chatbot using reinforcement learning.**
  - Conversational-response generation pre-trained model.
  - Train the pre-trained model with RL to improve long-term reward.
  - Transform free-speech into structured data and vice-versa.
2. **Build up a chatbot interface prototype.**
  - Design a simple web-based chatbot prototype interface.
  - Create a human-friendly interface with sentiment analysis, using a pre-trained model.
3. **Test the interactive behavior of the chatbot with real users.**
  - Human interaction hypothesis statements.
  - Design the experiment with real users.
4. **Evaluate the dialogue responses of the model and interaction with users**
  - Chatbot model evaluation.
  - Human interaction evaluation.
  - Hypothesis validation.

---

<sup>1</sup><https://www.cleverbot.com/>

## 4 Cognitive approach

Users that interact with open-domain conversational systems tend to lose interest while the conversation goes on. We are aware that overcoming this problem is really challenging, for instance, big players in the voice market have invested billions of dollars and are still doing a tremendous research on this field. Also, we know that our computational resources for developing a robust chatbot are limited.

However, using what we have learned during the MAI lessons and the knowledge gained during the CIR lecture, we want to implement machine learning approaches, that we think that could help overcoming the engagement problem, to develop a simple cognitive frame for our conversational chatbot.

The cognitive part of our chatbot, known as neural dialog system, it is implemented with deep learning techniques to generate neural language texts corresponding to the user dialog actions. In order to map the user dialog to response we take advantages of two machine learning techniques:

- Transfer learning: is the idea of overcoming the isolated learning paradigm and utilizing knowledge acquired for one task to solve related ones. We use DialoGPT [Zhang et al., 2019], a large-scale generative pre-trained model for conversational response generation.
- Reinforcement learning: is a methodology that takes suitable actions to maximize reward in a particular scenario. In our case we are dealing with a chatbot engine, so, we train the conversational response generation model using deep RL to model the future reward in the chatbot dialogue.

As a general view, our cognitive approach is based on training the DialoGPT model with a simulated agent using deep reinforcement learning. The two agents start a conversational loop between them, and the model tries to improve the long-term reward. In the following subsections we explain the base of our conversational model and the applied reinforcement learning methodology.

### 4.1 Conversational language model

We used as a conversational neural response generation the DialoGPT [Zhang et al., 2019] neural network model. Neural response generation is a subcategory of text-generation that shares the objective of generating neural-looking text which is a challenging task. Human-like response generation is a one-to-many problem as human conversations tend to be informal, noisy, and often contain informal abbreviations or syntactic/lexical errors.

Open-domain neural response generation systems suffer from content or style inconsistency and lack of long-term contextual information. We decided to use a Transformer based architecture specifically designed to boost information context like the OpenAI GPT-2 [Radford et al., 2019] model.

GPT-2 uses a multi-layer self attentive mechanism, Transformer [Vaswani et al., 2017], as described in Figure 1 A, which allows fully-connected cross-attention to the full context. It was designed to solve the simple task of predicting the next word in 40GB of high-quality text from the internet (WebText dataset). This objective is sufficient to train the model to learn a variety of tasks due to the diversity of the dataset, Figure 1 B. With this in mind, Zhang et al. developed DialoGPT by training GPT-2 model on a large-scale dialogue pairs extracted from Reddit<sup>2</sup> discussions chains.

As stated before, open-domain neural generation models tend to generate bland, uninformative samples. To address this issue, the researchers from DialoGPT they used Maximum Mutual Information (MMI) scoring function, to predict source sentences (dialogue history) given target responses (ground

---

<sup>2</sup><https://www.reddit.com/>

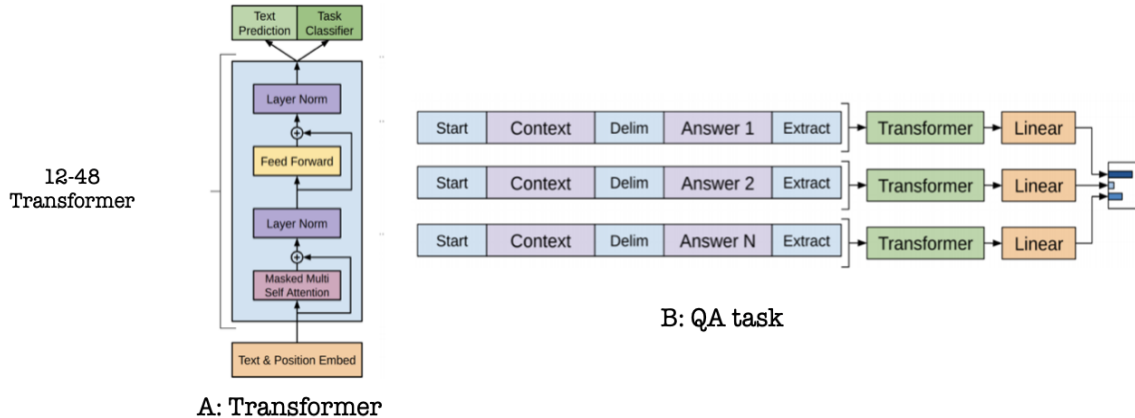


Figure 1: GPT model and QA task adaptation [Radford et al., 2018]. GPT-2 is based on the GPT architecture A: Transformer architecture and training objectives. B: Input transformation for fine-tuning on a QA task.

truth). MMI is capable of capturing the joint distribution of  $P(\text{target}|\text{source})$  in conversational flow. In practice, sentences generated by DialoGPT are diverse and contain information specific to the source prompt, analogous what GPT-2 generates for continuous text.

We exploit *transfer learning* by using the medium trained model for a large-scale pre-trained dialogue response generation (DialoGPT) based on the open-source HuggingFACE Pytorch transformer repository <sup>3 4</sup>. We use DialoGPT model as a dialogue-generation agent in the reinforcement learning conversational loop, explained in detail in the following section.

## 4.2 Reinforcement Learning

In this section, we briefly introduce what is reinforcement learning, its core elements, and we provide a detailed description of our RL conversational loop for open-domain dialogue.

Reinforcement learning (RL) is a sub-field of machine learning concerned with how an agent interacts with an environment over time. The goal in RL is specified by the reward function, which acts as positive reinforcement or negative punishment depending on the performance of the agent with respect to the desired goal. The main objective in RL is to maximize the accumulated reward over the agent’s lifetime.

One of the challenges that arise from RL is the trade-off between *exploration* and *exploitation*. In order to obtain a high reward, the agent has to choose those past actions that have been found effective in producing reward. The agent has to *exploit* what it already knows in order to choose the most effective actions but it also has to *explore* in order to find better actions in the future. [Sutton and Barto, 1998] The problem here is that both, exploration and exploitation, can’t be exclusively satisfied by the agent without producing any failures. In general, an agent must find an equilibrium between exploration and exploitation in the learning procedure.

In a RL system, we can identify many important elements that are important to define to comprehend how the cognitive frame of our chatbot works. In this section we describe the *policy*, *reward* and *value*

<sup>3</sup><https://github.com/microsoft/DialoGPT>

<sup>4</sup><https://huggingface.co/microsoft/DialoGPT-medium>

*function* terms in order to define the *problem* of RL.

### Policy

The *policy* describes the learning way of the agent during a time step. In other words, the policy maps from the sensed states of the environment to actions that have to be taken. The policy could be simple or complex. Sometimes the policy is a simple function, while in other cases it involves a highly expensive search process [Sutton and Barto, 1998].

### Reward

The *reward* defines the award that the agent gets on each time step when it transitions from a state to another using an action. The main objective of the agent in RL is to maximize the total reward it receives over time. Hence, the reward value defines which are the good or bad behaviors of the agent. Reward only depend on the agent’s current action and the current state of the environment. Also, the reward signal could be influenced through the agent actions, and can have a direct or indirect effect on the reward by changing the environment’s state.

The reason why the policy is changed is because of the reward. When an agent performs an action that gives a low reward the policy could change to select another action in the future to maximize the reward.

### Value function

The *value function* is used to predict the future reward that the agent will obtain from a state. It is the total amount of reward an agent can expect to accumulate over the future from that state following a given policy. The difference between reward and value function is that the first one defines the immediate reward and values indicates the long-term rewards by taking into account the states that are likely to happen.

Value functions are conditional to some policy,  $\pi$ , and can be written as state-value function,  $V^\pi(s)$ , or action-value function,  $Q^\pi(s, a)$ . The  $V^\pi(s)$  expresses the expected value of following policy  $\pi$  when the agents starts from state  $s$ . On the other hand,  $Q^\pi(s, a)$  expresses the expected value of first taking action  $a$  from state  $s$  and then following the policy  $\pi$ . Q-values are a great way to make actions explicit so you can deal with problems where the transition function is not available, model-free problems.

### Model-based and model-free methods

First it is important to say that not all RL environments have a model. The model it is something that says how the environment will behave. For instance, the model could predict the resultant next state and reward given a state and action. RL problems that use models are called model-based methods, while model-free methods are those that they need to do trial-and-error exploration.

### Problem

Once we defined the previous element we can describe the RL problem. In RL the agent interacts with the environment over time. At each time step the agent receives a state  $s_t$ , from a state space, and selects an action  $a_t$ , from an action space, following a *policy*  $\pi$ . The agent receives a scalar reward  $r_t$  when the mapping from  $s_t$  to  $a_t$  is done. Then, the agent transitions to the next state  $s_{t+1}$  according to the environment dynamics or *model* of the environment [Li, 2018, Sutton and Barto, 1998].

#### 4.2.1 RL for open-domain dialogue

Neural models from conversational agents usually tend to be short-sighted, predicting utterances one at time getting stuck in an infinite loop of repetitive responses, therefore, they ignore the influence of future outcomes. Using RL can help to model the future reward of the chatbot responses.

In this section we describe in detail the components of the proposal RL model. The learning system



consist of two agents. We use  $c$  to denote sentences generated by the first agent (chatbot model) and  $u$  to denote sentences generated by the second agent (simulated user).

During training the two agents take turns talking to each other in order to maximize the long-term reward. The dialogue is represented as an alternated generation of sequences by the two agents:  $c_1, u_1, c_2, u_2 \dots c_n, u_n$ . The generated sentences are viewed as actions that are taking according to a policy defined by the encoder-decoder from DialoGPT language model.

It is important to mention that initially, both agents are based on the DialoGPT medium model. During the RL loop, the DialoGPT corresponding to the chatbot model is the one that is trained, in other words, the one that we fine-tune. On the other hand, the DialoGPT model corresponding to the simulated agent keeps always the same, we do not fine-tune the simulated agent.

The components (agents, states, action, policy, and reward) of our RL system are summarized in the following points:

- **Agents:**
  - Simulated agent: sentences generated by the simulated agent  $u_i$
  - Chatbot agent: sentences generated by the chatbot model  $c_i$
- **Action:**  $a$  is the dialogue utterance to generate. In our system the chatbot agent response is generated by the DialoGPT model. The action space is  $\infty$  since arbitrary-length sequences can be generated.
- **State:** is denoted by the chatbot agent and the simulated agent previous dialogue turns. We store the sentences generated for both agents during each turn in a variable called *chat history* and we only keep the last 50 words since are the most informative ones.
- **Policy:** the policy is defined by the DialoGPT language model, in this case is the *Mutual Maximization Information*.
- **Rewards:**  $r$  denotes the reward obtained for each action. We aim to maximize the long-term reward, that is why we think that the success of the dialogue it is achieved by the contribution of three major factors [Li et al., 2016]:
  - *Ease of answering:* a turn generated by the chatbot should be ease to respond to. We measure this factor by using the negative likelihood of responding to that utterance with a dull or dummy response.

$$r_1 = -\frac{1}{N_{\mathbb{S}}} \sum_{s \in \mathbb{S}} \frac{1}{N_s} \log p_{seq2seq}(s|a) \quad (1)$$

where  $\mathbb{S}$  denotes the set of dummy responses,  $N_{\mathbb{S}}$  denotes the cardinality of  $\mathbb{S}$  and  $N_s$  denotes the number of tokens in the dull response  $s$ . The  $p_{seq2seq}$  represents the likelihood output by the seq2seq model, in this case the DialoGPT.

- *Information flow:* we want that at each turn the chatbot would be able to contribute new information to keep the dialogue moving and avoid repetitive responses like "I don't know". Semantic similarity between consecutive turns from the chatbot is penalized. We denote  $c_i$  and  $c_{i+1}$  representation from the DialoGPT encoder for two consecutive turns  $i$  and  $i + 1$ . The reward is described as the negative log of the cosine similarity between them:

$$r_2 = -\log \cos(c_i, c_{i+1}) \quad (2)$$

- *Repetition Avoidance*: similarly to the previous reward, we want the chatbot to avoid repeating the sentences of the other user, which is a common phenomenon that we noticed in preliminary results. Being  $u_i$  the last utterant from the simulated user and  $c_i$  the current sequence generated by our model, the reward is defined as follows:

$$r_3 = -\log \cos(u_i, c_i) \quad (3)$$

The final reward function for action  $a$  is a weighted sum of the three rewards:

$$r(a, [u_i, c_i]) = \lambda_1 r_1 + \lambda_2 r_2 + \lambda_3 r_3 \quad (4)$$

where  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ .

Finally to improve the long-term reward, given a sequence of dialogue turns  $u_1, c_1, \dots, u_T, c_T$  and a discount factor  $\gamma$ , the expected reward for each generated sentence  $c_t$  in the dialogue is computed as follows:

$$R_t = \sum_{i=t}^T \gamma^{(i-t)} r(a, [u_i, c_i])$$

The core idea behind our cognitive model approach is to simulate the process of two agents taking turns talking with each other. Both agents are based on the DialGPT language model, however, one of them is coded in a way that explores the state-action space and tries to maximize the reward function, while the other one is frozen (simulated agent). The whole RL training loop overview is shown in Figure 2.

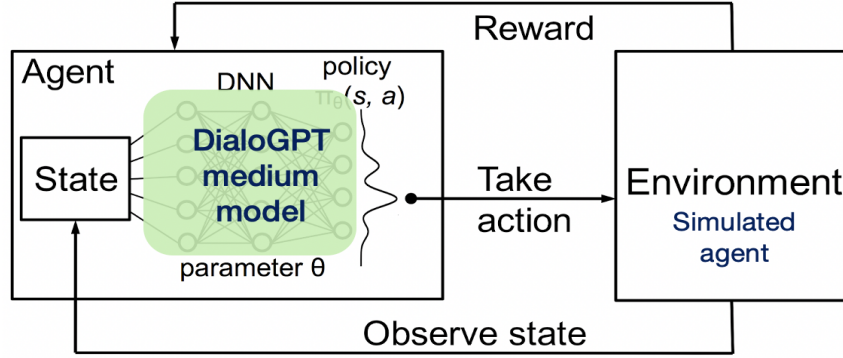


Figure 2: Reinforcement learning training loop.

## 5 Interaction approach

Chatbots usually have inadequate conversational user interfaces, especially for old people or people without experience with this technology. As a general view, a chatbot is a computer program that communicates with the user through a chat interface. This conversational interface enables users to interact with a machine through a combination of user-interface (UI) elements, therefore, it should be able to communicate effectively with users not only with the answers it says, also with a good designed interface in order to keep users engaged into the conversation.

The aesthetics of our chatbot is important as users are more keen on interacting with “attractive” faces than with “unattractive” ones. We want to avoid the uncanny valley, which is the degree of an object resemblance to a human and the emotional response to such an object.

In this project we aim to develop a friendly and easy to-use interface to make users feel comfortable and that is able to express minimal internal state and basic sentiments, without falling into the uncanny valley [Mori et al., 2012]. The human-interaction with the user is a key element during the user-experience with the chatbot. Designing and developing a good interface is fundamental and it is not a trivial task. We followed the next essential design rules:

- **User centered design:** the chatbot design should be created always with the end-user in mind, they are engines to assist users needs. Conversational UI is all about the user experience and satisfaction.
- **Build a simple conversation interface:** good user experience requires simplicity, intuitive interface, and maximal similarity to a human conversation. Visual aesthetics is important.
- **Create a bot personality:** a chatbot’s style of interface and voice to users differs from what humans can do. We must keep in mind that its communicative activities can’t fully resemble to those of human being. Creating a personality for our chatbot is a way to for the user to enjoy the experience.

In order to meet this essentials we developed a user interface based on a chatbot face expression. We thought that using only two interactive big eyes the chatbot cognitive brain and its appearance would match, and it will met the requirements for the user to feel comfortable. In addition, to maximize the emotional respond towards the chatbot we perform sentiment analysis to change the eyes expression according to the chatbot feelings.

## 5.1 Chatbot GUI

It has been implemented with HTML/CSS/Javascript so that it can be displayed in any modern device that can render websites. Also, an API REST with NodeJS has been developed to have the chatbot interface hosted in a server so that it can be accessed everywhere in this url: [interface link](#).

We designed some expressions and gestures such as happiness, sadness, confusion and blinking as we can see in figure 3.

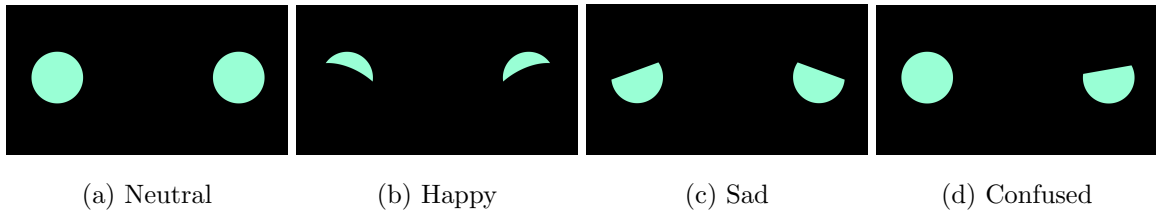


Figure 3: Chatbot facial expressions.

Besides the facial gestures and expressions, we also found beneficial for the interaction experience to display some text information to give feedback about the chatbot internal state. For instance, inform the user whether the chatbot is listening, recognizing the voice, thinking, and also to display the robot answers with text in case the user can’t understand the chatbot’s voice. Figure 4 illustrates some examples of this.

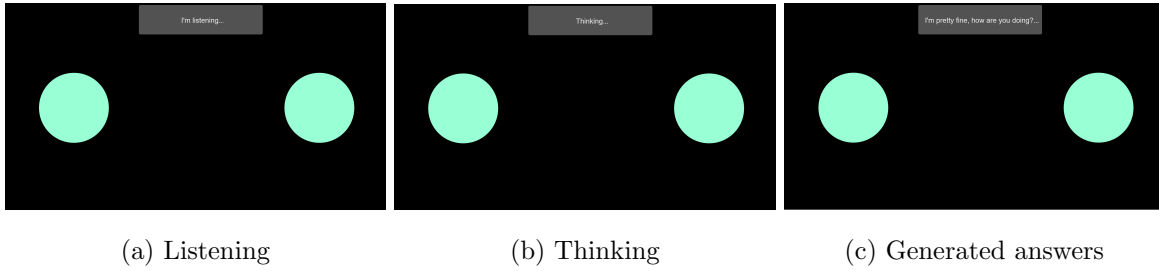


Figure 4: Chatbot information examples.

To change the chatbot expression or add information text while the interaction is being performed, the python model can modify the interface by sending HTTP POST requests to the API REST server. This setup helped us to easily test the system with users anywhere and decouple the dialog generation model from the graphical interface, allowing us to have them in different devices. Figure 5 shows the design of the developed interface architecture.

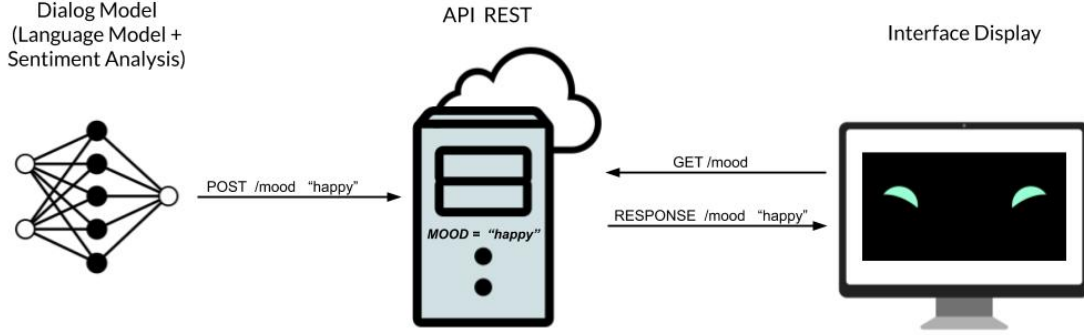


Figure 5: Interface architecture. We note that deep learning model (language model and sentiment analysis) is decoupled from the graphical interface and communicate through an API REST system.

## 5.2 Sentiment Analysis

The type of meaningful and expressive human-chatbot interface that we use is a non-verbal emotion-based interface. As explained in the previous section, we implemented a chatbot interface based on a face expression with two eyes. In order to provide a better user-experience and increase user engagement we thought on implementing sentiment analysis to change the eyes expressions.

Sentiment analysis is one of the most important fields of NLP, which consists in using machine learning techniques to teach a model (i.e. neural network) how to extract the sentiments out of a sentence; in other words, to predict if a sentence is positive or negative. This technique helps us attain the attitude and mood of the chatbot response, which can help us to change the status of its eyes and then create a more interactive user experience.

### 5.2.1 Sentiment analysis model

To face this approach, we use BERT pre-trained model to provide sentence embeddings for answers generated by the chatbot. Then, these embeddings are fed into a multi-layer bi-directional Gated Recurrent Unit (GRU) to predict sentiment of the input sentence.

#### BERT

BERT [Devlin et al., 2018] is a bidirectional transformer-based model [Vaswani et al., 2017] trained with a masked language modeling and a next sentence prediction task on a corpus of around 3.3B words. The masked language model masks some percentage of the input tokens randomly and tries to predict them, and the next sentence prediction task helps to understand the relationship between two sentences. Besides, BERT obtains SoTA performance on most NLP tasks, and it requires minimal architectural task-specific modification.

The general BERT architecture incorporates information from bidirectional representations, rather than unidirectional representations. This framework fits when analysing sentiments out of sentences, as the contextual information behind each word and relationship among them is necessary.

BERT cannot take text longer than the maximum length (512 tokens) as input since the maximum length is predefined during pre-training. When BERT is applied to NLP tasks one of the main problems comes when inputs should be truncated by the maximum sequence length which decreases the performance since the model cannot capture long-term dependencies and global information across the whole document. Also, sequences need to be padded if they are shorter than the maximum length (with special token “[PAD]”).

#### Bi-directional Gated Recurrent Unit

Gated Recurrent Unit (GRU) [Cho et al., 2014] aims to solve the vanishing gradient problem which comes with a standard recurrent neural network. GRU is an improved version of standard recurrent neural networks that uses update gate and reset gate to solve the vanishing gradient problem. Basically, it consists of two vectors that decide which information should be passed through the output. What makes them special is that they can be trained to keep information from long ago, without vanishing through time or remove information.

Bidirectional Gated Recurrent Unit (BiGRU) [Schuster and Paliwal, 1997] looks almost exactly the same as its unidirectional counterpart (GRU). The difference is that in BiGRU the gate is not just connected to the past, but also to the future.

Using BiGRU architectures together with attention mechanisms has shown significant performance on text sentiment classification tasks [Yu et al., 2019, Sachin et al., 2020]. For this reason we decided to develop an architecture based on BERT, which uses attention mechanisms (Transformer) and BiGRU.

#### Sentiment analysis architecture

BERT is pre-trained on English Wikipedia and BooksCorpus datasets which is limited on text based on expressing human experiences. For this reason, the approach that we follow is to freeze (no train) the BERT transformer model and only train the GRU architecture, with the IMDB movie reviews dataset, to learn from sentence representations produced by BERT. The IMDB dataset contains movie reviews classified as positive or negative.

Figure 6 describes the overall approach. Given an input sentence, BERT produces a contextualized embedding that is fed into the GRU architecture to predict the probability of the sentence as positive or negative. The output probability ranges between 0 and 1, where 0 means totally negative and 1 totally positive.

We decided to classify chatbot utterances into three sentiment status: happiness ( $0.55 < \text{score} \leq 1$ ), indifference/neutral ( $0.45 < \text{score} \leq 0.55$ ) and sadness ( $0 < \text{score} \leq 0.45$ ). We didn't implement

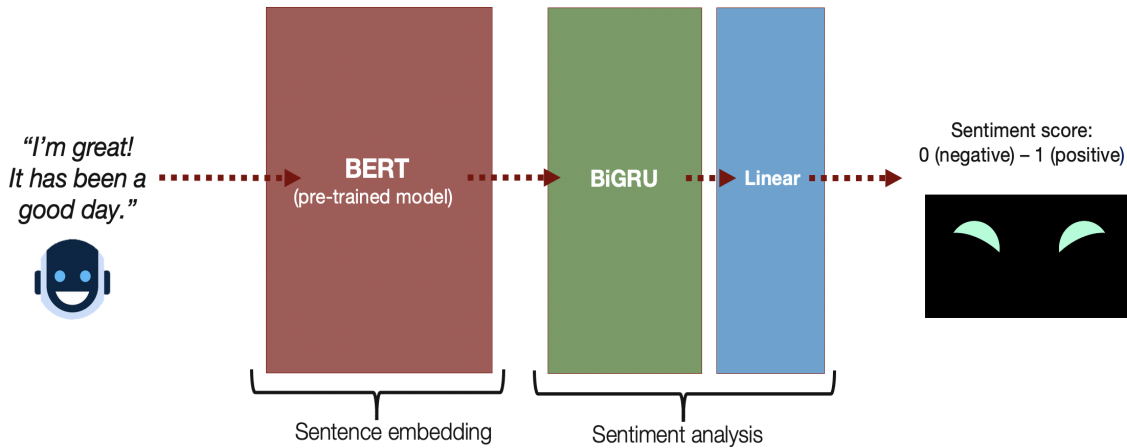


Figure 6: Sentiment analysis architecture of our chatbot.

more sentiment status ranges such as anger or focus, because for instance, determining whether a sentence express anger is a more complex task.

### 5.3 Chatbot speech experience

In voice-based chatbots one of the fundamental keys is to keep voice as the main interaction between the user and the system. The user speech experience when interacting with the chatbot can produce positive or negative feedback towards the engine, having a direct impact on the engagement to the system.

Natural language is the common interaction between humans. Nonetheless, is a great challenge to face for voice-based chatbots as they deal with a huge amount of words with multiple meanings, it is difficult to understand different voices, natural language tend to be unstructured, bad pronunciation, utterances with grammatical and syntactical errors, noise environment, etc.

Furthermore, we had the problem that most of the people which we could test the chatbot with are Spanish and not really fluent at English. We had two options to overcome this limitation: 1) pre-training the whole model (DialoGPT and sentiment analysis BERT+GRU) with Spanish corpora which would take many hours of computation, or 2) translating the sentences of users from Spanish to English before feeding the models with utterances. We opted for the second option since nowadays there are many translation libraries that work really well although we add some noise to the interaction flow. This way we can take advantage of the powerful pre-trained GPT Transformer in English and do not spend computational resources on pre-training in other languages.

Having this in mind, it is important to implement good speech recognition and generation systems. However, due resource limitations we used open-libraries available for Python to solve this approach:

- **Speech recognition:** we use the *Google speech recognition* library the speech to text conversion. This library creates a microphone instance, which represents a physical microphone on the computer, and the microphone audio is recorded in chunks of samples and recognizes the text.
- **Speech generation:** we use the *pyttsx3* library for text-to-speech conversion. An instance engine with a specific voice is created and speaks a text utterance. The voice language and speed rate can be changed.

- **Voice:** we use a male voice for English and Spanish configurations as they are the only ones available.
- **Translation:** we use Google Translate API for Python *googletrans*, which is free to use and provides us fast translations for our real-time system.

One of the limitations of our chatbot is that is highly dependent on the type of microphone that the computer uses. Voice-based chatbot requires specialized microphone hardware to isolate background noise and from other people who are talking. That is why experiments with real users must be performed in silent and quiet environments.

## 6 Chatbot final approach

In this section, we put all the cognitive and interaction approaches together in order to describe the final system architecture, and the embodiment of our voice-based chatbot.

### 6.1 System architecture

Figure 7 shows the flow of information through our final system. If we describe a common user-chatbot conversational turn it will be as follows: the user produces an utterance, the system converts the user speech to text, then, this text is fed to the dialogue generation trained with RL, the dialogue generation predicts the most adequate answer, this response is analyzed by a sentiment analysis model to change the chatbot interface mood, finally, the engine converts the answer into speech which is released to the user.

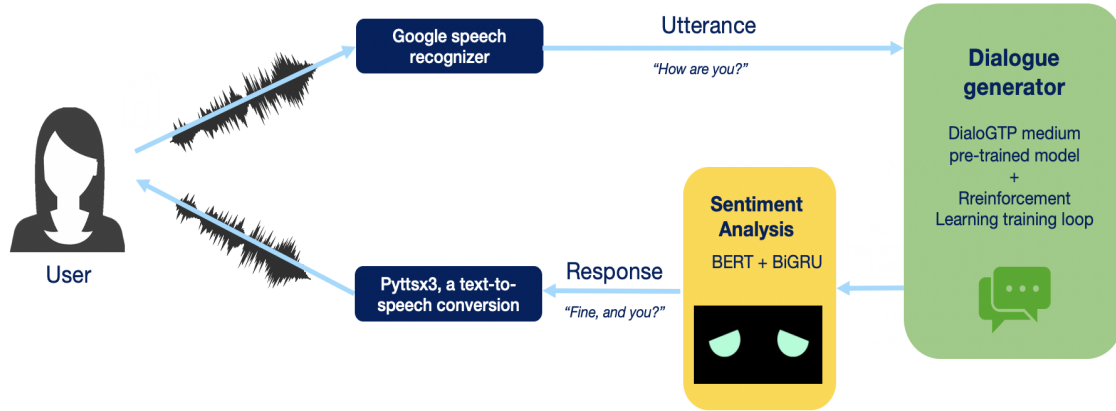


Figure 7: Chatbot information flow architecture

### 6.2 Chatbot embodiment

Initially we thought on developing a physical prototype based on a Raspberry-Pi, visual screen and a microphone/speaker device. However, we found two big limitations in this approach. First, the

covid-19 pandemic make it difficult to build up a unique physical prototype to perform all the user experiments. Secondly, building up a prototype as our initial proposal was sub-optimal as having all the elements into a separate spaces would increment the complexity of the system. Finally, we decided to implement our system in a single embedded space, such a computer, Figure 8.

Using this methodology benefits the workflow of the system as both the cognitive and interaction approaches can be integrated into the same space.

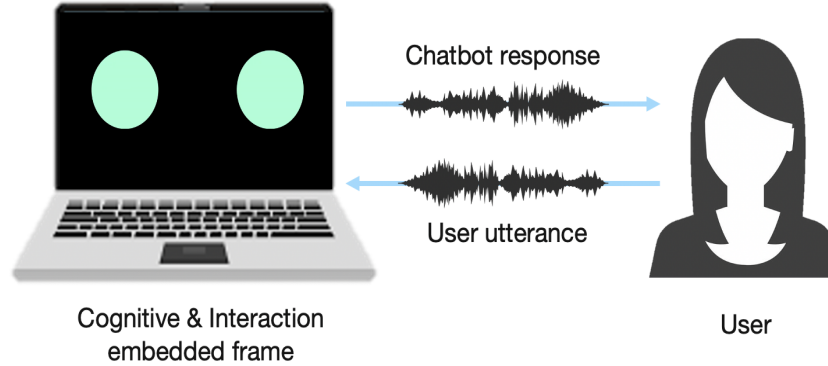


Figure 8: All of our voice-based chatbot approaches are integrated into the same embedded space (computer) to create a simple information workflow path between.

### 6.3 Robot-Ethics

With open-domain chatbots, any topic can be come up in the conversation and it is important that no unethical action is performed by the chatbot (i.e. generating offensive language), specially if their are used in clinical settings. Since DialoGPT has been trained unsupervisedly with dialogue datasets from social media, we may loose control about the content of the robot answers and we should be able to control that.

However, the authors of DialoGPT mention that they removed from the pre-traning dataset any instance where the target contains offensive language, identified by phrase matching against a large blocklist., and also excluded a large number of subreddits that had been identified as likely to contain offensive content [Zhang et al., 2019]. This means that if the model has never seen that an is not in the its limited vocabulary of 50k tokens, it is unlikely to generate offensive content, and would likely be a matter of randomness. Moreover, during all the performed experiments and executions, we have not observed any offensive response, even if we incited the model to answer like so.

Even though, there are several ways to tackle this problem anyway, for instance, by applying semantic rules or by using another ML model to detect hate, offense and aggressiveness text spans to block the response as a post-processing filter. Nonetheless, this is out of the scope of this project and a very interesting field of research in NLP.

## 7 System requirements

In this section, we specify the system requirements of our chatbot. We developed a system with a user-centered design, which places the users at the center of the design and development process, with the main objective of creating a conversational interface that satisfies the user needs and is also



usable.

## 7.1 Hardware and Software

To deploy this system, we need some computer requirements. We tried to minimize the software requirements as much as possible to facilitate the experiments.

- **Hardware:** Computer with at least 8GB of RAM to load the required models. Also, internet connection to make use of the graphical interface and some libraries such as Google Translate API. A GPU (with at least 6 GB VRAM) is preferred to speed up inference times of the models, although with a CPU the system performs pretty good. A microphone and headphones/speaker are needed to interact with the chatbot.
- **Software:** Python environment with installed libraries such as *pytorch*, *transformers*, *google-trans*, *pyttsx3* and *speech\_recognition*. Any web browser to render the graphical interface.

## 7.2 Context

Here we describe the context of use in order to identify the people who will use the chatbot (target), what they will use it for, and under what conditions they will use it:

- **Purpose of use:** people will use the chatbot in moments that they seek to be entertained or they feel loneliness. People enjoy interacting with conversational machines as they find them funny.
- **Target audience:** people that will use our chatbot is the general public, people from 18 to 65 years old.
- **Conditions of use:** people want to interact with a friendly, fluent, and easy to use interface.

## 7.3 Chatbot cognitive skills

The major cognitive skills of our chatbot implementation are the following:

- **Natural Language Processing (NLP):** the system is able to process utterances in a natural language fashion in both English and Spanish.
- **Natural Language Understanding (NLU):** the system is able to understand what the user is saying by analyzing the contextual information of the user's utterance.
- **Natural Language Generation (NLG):** the system tries to generate adequate answers given the user utterance.
- **Sentiment Analysis:** the system analyses the sentiment of the generated answer in order to change the facial expression of the chatbot.

## 7.4 Chatbot behavior

- **Chatbot scope:** our system is a AI chatbot that receives utterances from users, tries to understand then and provides answers given the contextual information.
- **Instructions display:** before staring the conversational experience with the system, the chatbot displays on the screen and describes via voice 8 instructions of use. The goal is to clarify how the system works.

- **Listening:** the chatbot displays the sentence “*I’m listening...*” when the user should speak. Associated with the NLP and NLU cognitive skills.
- **Thinking:** the chatbot displays the sentence “*Thinking...*” when AI model is generating an answer. Associated with the NLG cognitive skill.
- **Answer:** the chatbot responds to the user utterance via voice, and displays the answer into the screen in case the user didn’t understand the chatbot voice.

## 7.5 Chatbot constraints

- **Limited question scope:** our chatbot does not answers every single user utterance in an adequate way. Even though, the training datasets of DialoGPT and GPT-2 are based on general information, the chatbot cannot answers actual world information, specific dates, etc.
- **Language:** the system only supports questions in standard English or Spanish.
- **Slow response:** due to computational resource limitations the average time for the chatbot to respond, over the users utterance, is high.

# 8 Conversational model training and experiments

## 8.1 Experimental setup

Following the experimental settings from [Li et al., 2016] and the framework of simulated agents explained in section 4.2.1. We train the model for 1000 episodes and track its performance at every episode. Also, every episode starts with the simulated agent saying “Hello”, and we consider that a dialogue is finished when the trained model produces *dummy* utterances like “I don’t know”, “I”, “ ” (blank space), or any other one-character output.

After sampling the simulated utterances from the current policy (GPT model), we compute the expected reward from the generated turns for each action chosen (utterance):

$$R_t = \sum_{k=1}^T \gamma^{(k-1)} r_k(s_k, a_k)$$

where  $\gamma = 0.99$ , and  $T$  is the number of dialogue turns generated by our model.

The experiment on the DialoGPT medium model with reinforcement learning were run on a NVIDIA GeForce RTX 2060 GPU of 6GB VRAM, which is quite limited but was capable of training the model quite fast for a large number of episodes. Due to this computational and GPU memory limitations, we only fine-tune the last 3 layers of the DialoGPT to learn to maximize the expected rewards.

We trained the BERT BiGRU model on the IMDB movie reviews dataset on Google Colab GPU, the types of GPUs available at Colab vary over time as it is a free access. The available GPUs in Colab usually includes K80, T4, P4 and P100 from NVIDIA. For training we set up our model with 256 hidden dimensions, 1 output dimension, 2 BiGRU layers and dropout set to 0.25. With this this configuration, the whole model has up to 112M trainable parameters, however, as said before we freeze BERT so we ended up with less than 3M trainable parameters. After doing several experiments we realize that training the model during 5 epochs was enough. The total training time was around 1 hour.

## 8.2 Evaluation metrics

To evaluate the model in a quantitative way, for each episode we keep track of the achieved reward ( $r = 0.5r_1 + 0.25r_2 + 0.25r_3$ ) and the generated dialogue turns. This will show us if the RL framework is really making the model to learn to maximize the obtained rewards in order to be more engaging and last for a longer period of time.

Besides, we also created a way to qualitatively evaluate the model’s ease of answering and context adaptation by testing in what we call the “Titanic test”. This test consists in testing the trained dialogue model with a predefined standard conversation about a movie extracted from an english conversation exercise<sup>5</sup>. The original dialogue is as follows:

**Original Dialogue:**  
Female: Did you see “Titanic”?  
Male: Yes. It is a great movie.  
Female: I saw it twelve times.  
Male: I saw it eight times.  
Female: I have the DVD.  
Male: So do I.  
Female: Let’s go to your home.  
Male: We can watch my DVD.  
Female: And then we can go to my home.  
Male: And watch your DVD.  
Female: I always cry at the end.  
Male: Me too. It’s so sad.

The idea is to remove the responses of “Male” character and generate them with our dialogue model. This will help us to have a general idea of the model coherence and conversational topic adaptation in a qualitative way as an alternative support to quantitative results.

## 8.3 Experimental results

We initially run 1000 episodes (dialogue simulations) and see if the model is able to improve the rewards. In figure 9, we can see the evolution of the achieved rewards along episodes and we observe that the model converges around episode 800. This execution took around 5 hours to execute on our computational settings mentioned in section 8.1.

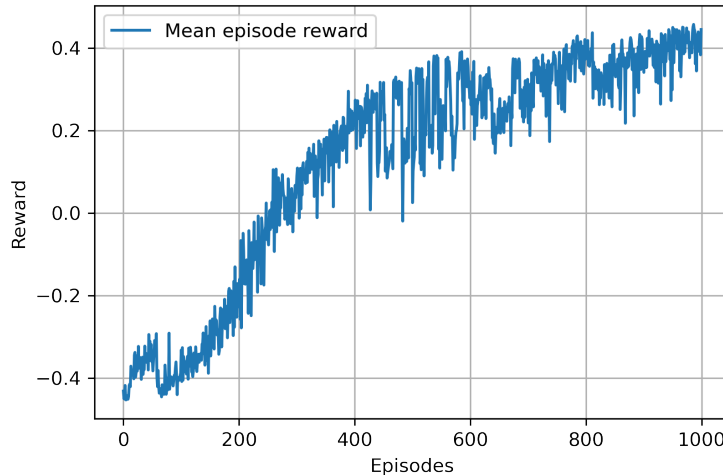


Figure 9: Obtained rewards at each simulation (episode)

<sup>5</sup><https://examgroup.org/english-conversation/entertainment/a-great-movie/a-great-movie.html>

However, if we observe figure 10, we notice that from episode 200 on, the number of dialogue turns begin to decrease until getting stuck to generating dialogues of only two turns. This means that although the model keeps on improving the obtained rewards until episode 1000, the model outputs are not coherent anymore and, hence, we have to stop the RL training some episodes before.

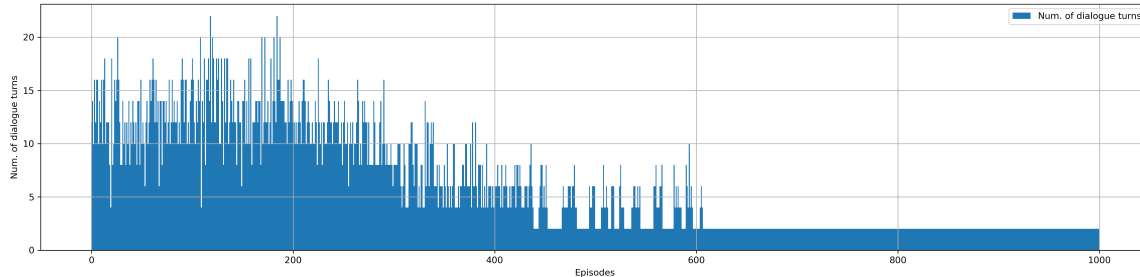


Figure 10: Number of simulated dialogue turns with a simulated agent along the episodes. We observe that at a certain point the number of generated dialogue turns decrease to only 2.

To have a better idea of what are the qualitative results, we are going to show how the simulated dialogues evolve with the training steps. In the following text snippets, we can see that, along the episodes, the model perplexity increases since we can notice how the generated sentences gain expressiveness and become longer (episode 1 vs. 200). However, with too much training, the model loses language understanding and starts generating non-sense sequences (episodes 400 and 800), which makes us understand why the number of turns significantly decreases with too much training as have observed in figure 10.

### Episode 1:

```
User: Hello
DialoGPT: Hello! :D
User: How are you?
DialoGPT: I'm good!
User: That's good!
DialoGPT: I'm good too!
User: That's great!
DialoGPT: So, what's up?
User: Nothing much.
DialoGPT: That's good.
User: That's great.
DialoGPT: I'm good
User: That's good
DialoGPT: I'm good.
User: That is good
DialoGPT: I'm
User: I'm good, too.
DialoGPT: I'm
User: Good
DialoGPT: I
```

## Episode 200:

User: Hello  
**DialoGPT:** Hi there!  
User: How are you?  
**DialoGPT:** Fine, how are you?  
User: I'm good, how are you  
**DialoGPT:** I'm good too, thanks. How are you?  
User: Good, how are you doing?  
**DialoGPT:** I'm doing well, well, well, how are you going?  
User: I'm going to be busy for a while, but I'm going to be doing a lot of work.  
**DialoGPT:** Ah, well, that's fine then. I'm sure you'll be busy  
User:  
**DialoGPT:**

## Episode 400:

[illegible]

## Episode 800:

User: Hello  
**DialoGPT:** Hey buddieinoeusa  
 rusesaurusesaurusesauru  
 saurusesaurusesauriaurus-  
 esaurusesaurraurusesauruses  
 aurosaurusesaurusesaunaau-  
 rusesaurusesaunosaurusauro  
 sesaurusesauseaurusesaurus-  
 esaulosaurusauresaurule  
 saurusesaurusesausaurusaurs-  
 esaurulesaurusauresau  
 ruitsaurusesaurusesaultause-  
 aurusesauruitsauseauruse  
 saurulesauseaurusesaurue-  
 saurusesaurusesaveauseaurus-  
 esaur  
 User:  
**DialoGPT:**

This decrease of performance along training episodes can be explained because the model only wants to improve the rewards and learns that generating very rare sequences maximizes the rewards  $r_2$  and  $r_3$ , which forces the model to generate semantically different utterances from previous turns. For this reason, we are going to focus the attention on training the models with less than 200 episodes, where the model still has perfect language understanding and is more expressive than not trained with RL.

In appendix A.1, we show the evolution of testing the chatbot with what we refer as “Titanic test”: evaluate the model with a predefined English dialogue about a movie. We remark that this *Titanic dialogue* is only used for evaluation, and not seen by the model in any moment during training. In these

results we can see that along episodes, the models tends to produce longer sentences with significant variations: in episode 0, model responses are trying to imitate the immediate previous turns, while 100 episodes later we can notice that longer responses are generated with more semantic variations (appendix A.1). Also, in episode 250, most of the model turns are great responses, although it ends by producing corrupted sequences.

It is important to use both quantitative results (mean of rewards, number of turns) results and qualitative results (observe the evolution of the dialogues) to evaluate the performance of this model, because dialogue coherence and ease of answer can be difficult to represent with just numerical results.

### 8.3.1 DialoGPT-Medium vs. DialoGPT-Medium with Reinforcement Learning

For comparing the model trained with RL with the model that is not, we are going to analyze dialogues generated with real users and evaluate certain metrics related to expressiveness and engagement. We took a model trained with 60 RL episodes, to ensure that it still performs well on language understanding. Table 1 shows the average number of turns, unigrams and bigrams achieved during the dialogues generated with each model. We can clearly observe that with the RL model, dialogues appear to be more expressive in vocabulary (increase unigrams and bigrams) and users seemed to be more engaged since the number of turns is also larger.

Model	Avg. turns	Avg. duration (s)	Avg. unigrams	Avg. bigrams
DialoGPT standard	15.8	257.693	69.7	253.3
DialoGPT with RL	<b>31.7</b>	<b>605.251</b>	<b>253.3</b>	<b>337.5</b>

Table 1: Comparison between the baseline model and the RL trained model. Results have been averaged among all dialogues performed by the same model.

Furthermore, in section A.2 of the appendices, we show some experimental dialogues with real users. Qualitatively, there is a clear difference between the baseline model and the RL one: with RL model, dialogues seem much more natural, sentences are longer, and more different vocabulary is involved. During the experiments we noticed that this increase on the engagement is because in the baseline model answers are very plane and repeating the users questions, for instance: User: “Do you like summer or winter?”, Chatbot: “I like summer”, which becomes boring on the long-term and does not leave room to continue with the conversation. On the other hand, with the reinforcement learning trained model, we manage to generate less repetitive answers that are less similar to previous dialogue turns (achieved by maximizing the defined rewards), for example in the previous question, we found that the RL model answered: Chatbot: “Summer, I love cold Florida winters. Winter is boring.” (note that Florida has a tropical weather). As a fact, we had one participant that spent 24 minutes talking with the RL model while only lasted 6 minutes on the baseline one.

## 9 Human-interaction experiments

In this project, we aim to report the findings of our voice-based chatbot with real users. The main objective is to study whether reinforcement learning is a good technique to increase the user engagement experience towards the engine. Also, we want to evaluate if the engagement is dependent on the age-range.

In this section, we first state our two hypotheses to be tested, hence, a good human-interaction analysis of the experiments is fundamental. We introduce our experimental design to compare the user-interaction differences with the system when using the baseline model and the RL model. Then,

we analyze the different experiences of 14 subjects. For our analysis, we combined qualitative findings from a structured survey with the quantitative findings from utterances that participants exchanged with the chatbot.

## 9.1 Statistical hypotheses

As the main objective of our project is to develop and design a voice-based chatbot to strengthen the user experience engagement, we formulated two different hypotheses: one regarding the possible benefit of using reinforcement learning, and the other one about the age-range engagement.

Hypothesis 1: Reinforcement Learning

- Null hypothesis ( $H1_0$ ): The model trained with Reinforcement Learning model does not improve users engagement towards the chatbot.
- Alternative hypothesis ( $H1_A$ ): The model trained with Reinforcement Learning model improves users engagement towards the chatbot.
- We are going to run a **Paired Sample T-Test** for this hypothesis, since we are comparing the average of two samples taken from the same population but at different points in time (all participants tested both models).

Hypothesis 2: Age-range

- Null hypothesis ( $H2_0$ ): The engagement is dependent on the age-range.
- Alternative hypothesis ( $H2_A$ ): The engagement does not depend on the age-range.
- In this case, we are performing an **Independent Sample T-Test**, since we are comparing the average of two independent unrelated groups (young and old participants).

## 9.2 Study design

To study the user interaction with our engine we chose to work with the medium DialoGPT baseline model (no-fine tuned), and the DilaoGPT model trained with reinforcement learning. In this section, we describe our study participants, and further continues with the study procedure.

### 9.2.1 Participants

To understand the user experience interacting with the chatbot we required individuals that could have had prior experience with these types of engines or not. As participants were not compensated to participate in the study, and due to the Covid-19 pandemic situation, we recruited friends and family members to explore and experience this technology. In order to validate our hypotheses, we ensured two requirements to be fulfilled: (a) the first group of participants between 20 and 30 years old, (b) the second group of participants above 55 years old.

A total of 14 participants (8 male and 6 female) were recruited. Participants had different backgrounds, some with an engineering background, and others from non-technical backgrounds, including, teachers, and social sciences. All of them were Spanish native speakers, none of them were native English speakers, the reason why experiments were done with the Spanish model version. 78.6% of the participants self-reported previous experience with virtual assistants, chatbots, and similar technologies like Alexa, Siri, Google Assistant, and Cortana. Before the experience, 28.6% of the subjects said

that they would use a chatbot as virtual assistance, and another 28.6% said that they would never use a chatbot.

We divided participants into two subgroups, the first included 7 subjects between 20 and 30 years old and the second 7 subjects above 55 years old. Figure 11 show some of the study subjects during the interaction with the voice-based chatbot (participants gave us their consent to be pictured and share their images in this project).



Figure 11: Six study subjects during their interaction with the chatbot. Top: subjects between 20 and 30 years old. Bottom: subjects above +55 years old.

### 9.2.2 Procedure

Before starting the interaction, in a fast face-to-face meeting with the researchers' participants were informed about the definition of a voice-based chatbot with few generic examples and the goal of the study. Participants were provided with the two model approaches (baseline and RL) in a randomized order to counteract order effects. They were asked to interact with each model in an isolated room to avoid noise artifacts, and we advise them that the session would be recorded.

During the interaction participants were first instructed on how to interact with the chatbot by reading and listening 9 points:

1. This is a test for our CIR subject final project.
2. We aim to test this conversational system with real users.
3. This session will be recorded to test the reaction of the users during the interaction with the chatbot.
4. When you read the phrase: *I'm listening...* on the screen, you have to speak.
5. Please wait while the chatbot is thinking of an answer. You will read the word *Thinking...* on the screen.
6. Once the answer is ready, the chatbot will talk to you.
7. Then, you can answer the chatbot again.
8. To end the conversation with the chatbot only say the word "GOODBYE"

9. At the end, a screen with a questionnaire that must be filled out will be displayed.

Instead of asking participants to perform specific questions with each model, which would have biased the results, we chose to let the subject perform randomized questions following the conversation flow with the chatbot.

After each interaction with the two chatbots models, participants had to fill out a post-study Google Forms questionnaire for qualitative analysis. The questionnaire contained two types of questions:

- Questions about the subject and his or her previous experience with chatbots, and when do they thing would use a chatbot for.
- 9 like-scale questions to evaluate their engagement experience the interaction with the chatbot (section A.3 shows the questionnaire)

### 9.3 Human-interaction evaluation

In this section, we show the results of the human engagement experiences during the face-to-chatbot interaction and discuss them to validate or reject our hypotheses. This evaluation is done quantitatively and qualitatively. We followed a specific hypothesis testing process for the validation or rejection of the statements:

- Hypothesis 1 - Reinforcement learning: we have tested both models, baseline, and RL, with all participants. Hence, we analyzed the results comparing quantitative and qualitative evaluations. For this evaluation, we didn't consider the age-range difference among participants, as the objective is to test the engagement of the subjects toward the system depending on the type of model using.
- Hypothesis 2 - Age range: for this hypothesis, we analyzed the quantitative and qualitative results by comparing the young group (20-30 years old) and the elder group (+55 years old) of participants when using the RL model. Here, the objective is to test whether the engagement is dependent on the age range. We decided to choose the results obtained with the RL model to avoid possible model bias.

In hypothesis testing, a specific statement is generated about a population parameter, and sample statistics are used to assess the likelihood that the hypothesis is true. Our hypotheses are based on available information and our beliefs about the population parameters. We use paired t-test and independence sample t-test, and compute  $p$ -value scores with a significance level for rejecting the null hypothesis ( $H_0$ ) of 5%.

#### 9.3.1 Hypothesis 1: Reinforcement Learning

In this section, we aim to evaluate which model is better in terms of engagement: the basic one or the RL one. Therefore, validate or reject the null hypothesis ( $H_{10}$ ) To do so, we are going to analyze quantitative and qualitative results.

##### Quantitative evaluation

Here we present the quantitative analysis of the human-interactions with the engine throughout the experiment. Remember, that for this hypothesis we had a total of 14 subject in each experiment, as all participants interacted with both models (Baseline and RL). We analyze three elements, the average length of the conversation (number of turns per conversation), the average diversity and expressiveness



of the chatbot generated answers (number of distinct unigrams and bigrams) and the average time spent interacting with the engine. Table 2 and figure 12 shows the quantitative results of these metrics comparing both models. We can see that with RL model all of this metrics are higher, showing strong evidence that the interaction with users was more engaging and richer.

Model	Avg. turns	Avg. duration (s)	Avg. unigrams	Avg. bigrams
DialoGPT standard	15.8	257.693	69.7	253.3
DialoGPT with RL	<b>31.7</b>	<b>605.251</b>	<b>253.3</b>	<b>337.5</b>

Table 2: Comparison between the baseline model and the RL trained model. Results have been averaged among dialogues performed by the same model.

To understand better the information behind Table 1 we took the participants results for each element and model (without computing averages) and plot them into a boxplot. A boxplot is a standardized way of displaying the distribution of the data based on five number summaries ("minimum", first quartile, median, third quartile and "maximum"). Boxplots can tell us about the outliers and what they values are, and how tightly our data is grouped, which means, whether data falls into the same probability distribution range.

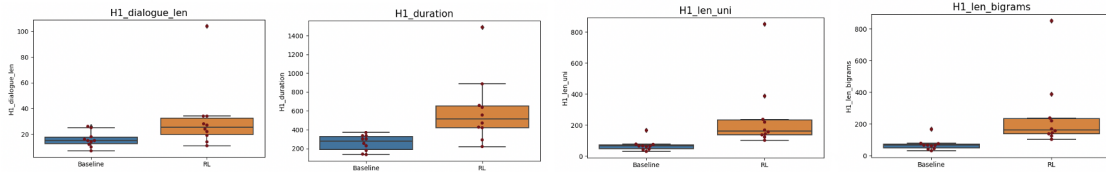


Figure 12: Boxplots comparison between baseline model (blue) and the RL trained model (orange).

Figure 12 shows that for the the four elements data tend to fall into the same distribution, which means that the obtained observed results in Table 2 are consistent and significant.

**Paired Sample T-Test Results:** In the following table 3 we show the results of the hypothesis testing  $H1$ , which has as null hypothesis  $H1_0$  that  $\mu_{Baseline}^i = \mu_{RL}^i$  for all quantitative metrics  $i \in (num. turns, duration, unigrams, bigrams)$  and as alternative hypothesis  $H1_A$  that  $\mu_{Baseline}^i < \mu_{RL}^i$  (one-sided). We can see that we have enough evidence to say that *duration*, *num. unigrams* and *num. bigrams* are statistically lower in Baseline model than in RL model, with a significance of  $\alpha = 0.05$ .

	num. turns	duration	unigrams	bigrams
<i>p-value</i>	0.0535	0.0079	0.0019	0.0084
Reject $H1_0$ ( $< 0.05$ )	No	Yes	Yes	Yes

Table 3: Paired Sample T-Test results of  $H1$  whether  $\mu_{Baseline} = \mu_{RL}$  for all quantitative metrics.

### Qualitative evaluation

In addition, we analyzed the post-study questionnaire results (section A.3). Participants were asked to rate each interaction on 9 different engagement-related metrics on a 5-point likert scale: *speed*, *comprehension*, *understanding*, *kindness*, *frustration*, *friendly interface*, *easy to use*, *fun to use*, *future use*. For all metrics, except frustration, higher score is better.

Figure 13 shows that for all metrics, except frustration, we obtained better results with the RL model, which reasserts that the RL model increases the user engagement during the interaction experience.

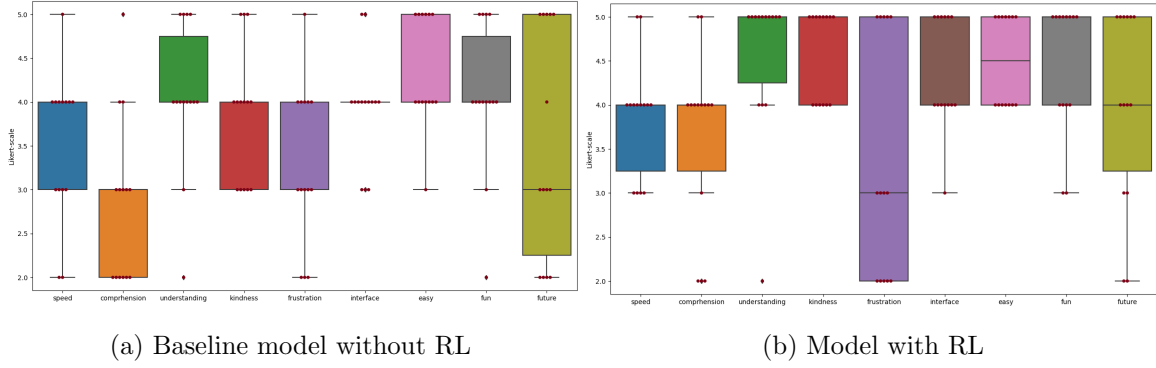


Figure 13: User questionnaire responses on Baseline model (a) and RL model (b). We can see that RL model seems to achieve better valoration on average on the the different chatbot attributes.

### 9.3.2 Hypothesis 2: Age - range

To evaluate whether engagement is dependent on the age-range,  $H2_0$  validation or rejection, we followed the same procedure as the previous section.

#### Quantitative evaluation

Similarly, to the previous hypothesis, we show the numerical results on four quantitative elements for hypothesis 2 on table 4. Surprisingly, results reported higher engagement with elder subjects than younger. However, when we analyze the dispersion of the data observed in Figure 14 we can see that for the elder group of participants is higher. This means that results can be biased by some subjects that had had much longer conversations with the chatbot than others.

Age-range	Avg. turns	Avg. duration (s)	Avg. unigrams	Avg. bigrams
Young (20-30)	26.0	521.972	152.833	243.166
Old (+55)	<b>40.25</b>	<b>730.170</b>	<b>404.0</b>	<b>479.0</b>

Table 4: Comparison between the young and elder participants. Results have been averaged among dialogues performed by the same RL model.

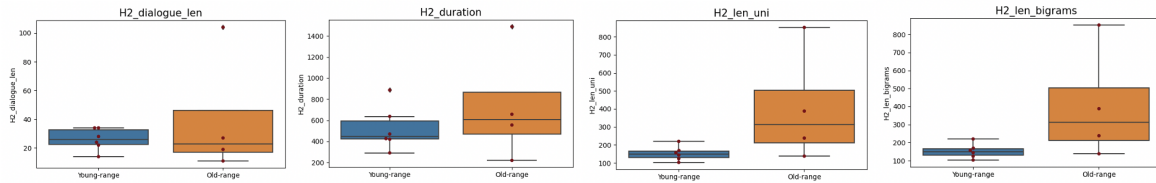


Figure 14: Boxplots comparison between the young (blue) and elder (orange) participants.

**Independence Sample T-Test Results:** In the table 5 we show the results of the hypothesis  $H2$  test, which has as null hypothesis  $H2_0 : \mu_{Young}^i = \mu_{Old}^i$  for all quantitative metrics  $i \in (num. turns, duration, unigrams, bigrams)$  and as alternative hypothesis  $H2_A : \mu_{Young}^i < \mu_{Old}^i$  (one-sided, and without assuming equal variance of both sets). Probably because we had small groups in this experiment (7 young participants and 7 old participants), we do not have enough evidence to reject

the null hypotheses (table 5) to state that elder users feel more engaged than young ones, although the plots in figure 14 appear to show this tendency.

	num. turns	duration	unigrams	bigrams
<i>p-value</i>	0.2784	0.2532	0.0832	0.1049
Reject $H2_0$ ( $< 0.05$ )	No	No	No	No

Table 5: Independence Sample t-test results of Hypothesis 2 whether  $\mu_{Young}^i = \mu_{Old}^i$  for all quantitative metrics  $i$ .

### Qualitative evaluation

Like in the previous hypothesis, we display the boxplots for each engagement-related metric in Figure 15. We can observe that all metrics except speed and comprehension show higher values for the elder group of participants. Besides, we see that frustration obtained very high values, which has no sense considering other metrics results. One possible reason of this behavior is that the old group didn't fill out this question in the correct way as is the opposite than the others.

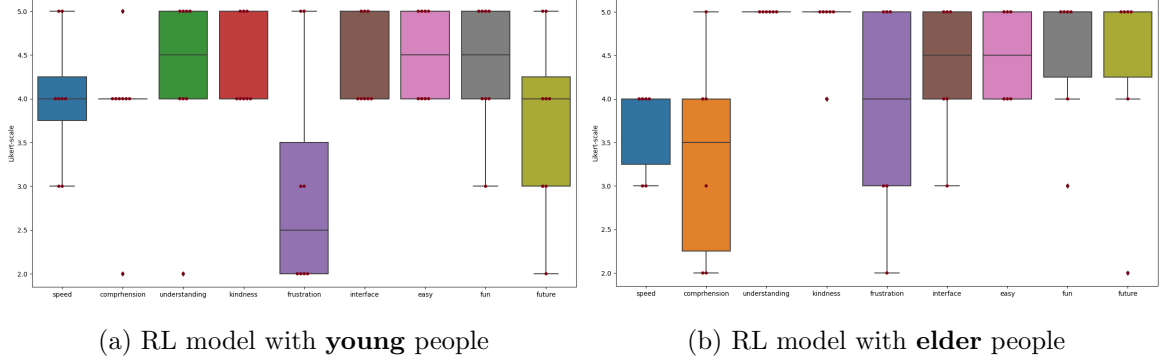


Figure 15: User questionnaire responses on RL model of young and old users. We can see that RL model seems to achieve better valuation on average on the the different chatbot attributes.

## 9.4 Discussion

We are aware that with only 14 participants, we do not have enough experiments affirm any conclusion with strong statistical support. Nonetheless, we performed the t-tests anyways to with the data that we managed the gather.

To sum up, regarding hypothesis  $H1$ , we found that the RL model is able to perform more expressive and longer sentences, leading to more engagement during the conversation. The paired t-test supports this statement showing that the average quantitative metrics are lower for the baseline model

However, with hypothesis  $H2$ , we did not have enough statistical evidence to prove that interaction with elder people are indeed more engaged than young people (very few samples were collected), although the plots in figures 14 and 15 appear to show a tendency of a higher engagement in old participants.

Besides the statistical testing, we would also like to comment the reactions of the participants towards the chatbot. All the users were very positive and showed amusement towards the robot, and also were impressed by some its responses. Although sometimes the chatbot entered into contradiction and generated repeated words, the participants took it with humor most of the times and appeared

to have fun during the whole experiment process, specially with the RL model, which seemed much more expressive and, hence, gave more ideas to the participants for keeping up the conversation and not running out of conversation topics.

## 10 Conclusions

To conclude, we remark that we managed to accomplish the proposed objectives at the beginning of the project:

- We developed a conversational chatbot and showed strong evidence that with our proposed RL framework it is able to make the model generate longer and more expressive sentences and, hence, engaging users
- We implemented a graphical chatbot interface that showed positive impressions among the participants.
- We tested the behaviour of our robot prototype and evaluated the dialogue responses and interaction with real users.

During the experiments we have seen that the model is able to respond to many different topics (movies, music, sports, countries, etc), proving effective response to open-domain (see appendices [A.1](#), [A.2](#)).

Also, we are really surprised with the obtained results. Transformer models are really obtaining outstanding results in NLP, and showed that it is possible to achieve impressive results on open-domain dialogue generation without many expensive and demanding resources. However, further improvements could be performed in order to significantly improve the interaction with users that are discussed in the following section [10.1](#).

Finally, we encourage the readers to have a look at the dialogues carried out with real users in the appendix [A.2](#), which in some situations the responses are very accurate, showing the amazing potential of these Transformer-based seq2seq models.

### 10.1 Future work

Although the results are quite impressive and have surpassed our expectations, there are several future improvements to get the most of this conversational system.

Besides improving the generative model, we found more critical to improve other software pieces that surrounds the core model. For instance, We noticed that speech recognition software we used had some issues at recognizing the users utterances and hindered the communication process. Sometimes the model seemed to reply rare responses but it was due to the wrong recognitions.

Also, improvements on speech synthesis could be done, since the *pyttsx3* library generates quite artificial voices. Alternatives to this library could be more recent approaches like Tacotron 2 deep learning model [[Shen et al., 2018](#)] that produce very natural speech, but it increases the system complexity and computational resources as this model requires GPU to rapidly synthesize from text to speech.

In addition, regarding the core language model (DialogPT), in some generated dialogues, we found that certain words related to Reddit domain appeared in the sequences: “moderator”, “subreddit”, “thread”, etc. which sometimes were incoherent within the conversation context. A way to solve this could be to pre-train on more general dialogue corpora or to add these words in the *dummy* set (related to reward  $r_1$ ) and minimize the probability of generating them within the RL process.

## 11 Acknowledgements

Finally, we want to acknowledge all the participants, family and friends who have been involved in the experiments and allowed us to keep track of their interaction with the chatbot. All of them were positively participatory and made the experimentation worthwhile and entertaining.

## References

- [Adiwardana et al., 2020] Adiwardana, D., Luong, M.-T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G., Lu, Y., et al. (2020). Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*.
- [Cho et al., 2014] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- [Colby, 2013] Colby, K. M. (2013). *Artificial paranoia: A computer simulation of paranoid processes*, volume 49. Elsevier.
- [Devlin et al., 2018] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [Devlin et al., 2019] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding.
- [Li et al., 2016] Li, J., Monroe, W., Ritter, A., Galley, M., Gao, J., and Jurafsky, D. (2016). Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*.
- [Li, 2018] Li, Y. (2018). Deep reinforcement learning: An overview.
- [Mnasri, 2019] Mnasri, M. (2019). Recent advances in conversational nlp: Towards the standardization of chatbot building. *arXiv preprint arXiv:1903.09025*.
- [Mori et al., 2012] Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics Automation Magazine*, 19(2):98–100.
- [Radford et al., 2018] Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I. (2018). Improving language understanding by generative pre-training.
- [Radford et al., 2019] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- [Sachin et al., 2020] Sachin, S., Tripathi, A., Mahajan, N., Aggarwal, S., and Nagrath, P. (2020). Sentiment analysis using gated recurrent neural networks. *SN Computer Science*, 1(2):1–13.
- [Schuster and Paliwal, 1997] Schuster, M. and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11):2673–2681.
- [Serban et al., 2016] Serban, I., Sordoni, A., Bengio, Y., Courville, A., and Pineau, J. (2016). Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.
- [Shang et al., 2015] Shang, L., Lu, Z., and Li, H. (2015). Neural responding machine for short-text conversation. *arXiv preprint arXiv:1503.02364*.
- [Shen et al., 2018] Shen, J., Pang, R., Weiss, R. J., Schuster, M., Jaitly, N., Yang, Z., Chen, Z., Zhang, Y., Wang, Y., Skerry-Ryan, R., Saurous, R. A., Agiomyrgiannakis, Y., and Wu, Y. (2018). Natural tts synthesis by conditioning wavenet on mel spectrogram predictions.
- [So et al., 2019] So, D. R., Liang, C., and Le, Q. V. (2019). The evolved transformer. *arXiv preprint arXiv:1901.11117*.
- [Sutton and Barto, 1998] Sutton, R. S. and Barto, A. G. (1998). Reinforcement learning: An introduction. *MIT Press*.

- [Vaidyam et al., 2019] Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., and Torous, J. B. (2019). Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, 64(7):456–464.
- [Vaswani et al., 2017] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., and Polosukhin, I. (2017). Attention is all you need. *In Advances in neural information processing systems*, pages 5998–6008.
- [Vinyals and Le, 2015] Vinyals, O. and Le, Q. (2015). A neural conversational model. *arXiv preprint arXiv:1506.05869*.
- [Weizenbaum, 1966] Weizenbaum, J. (1966). Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45.
- [Worswick, 2018] Worswick, S. (2018). Mitsuku-loebner prize 2018. *Retrieved 12th September*.
- [Yu et al., 2019] Yu, Q., Zhao, H., and Wang, Z. (2019). Attention-based bidirectional gated recurrent unit neural networks for sentiment analysis. In *Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition*, pages 116–119.
- [Zhang et al., 2019] Zhang, Y., Sun, S., Galley, M., Chen, Y.-C., Brockett, C., Gao, X., Gao, J., Liu, J., and Dolan, B. (2019). Dialogpt: Large-scale generative pre-training for conversational response generation. *arXiv preprint arXiv:1911.00536*.
- [Zhou et al., 2020] Zhou, L., Gao, J., Li, D., and Shum, H.-Y. (2020). The design and implementation of xiaoice, an empathetic social chatbot. *Computational Linguistics*, 46(1):53–93.

## A Appendix

### A.1 Generated Dialogues with *Titanic Test*

This appendix section shows the performance of the model responding to a **fixed** dialogue template related to a movie conversation (“User” turns are always fixed). The purpose of this test is to qualitatively evaluate the model’s ability to adapt to dialogue changes and coherence adjustments. We can observe that sentence expressivity increases with more training episodes, up to a point where the it begins to generate corrupted sequences (e.g. episode 250).

#### Episode 0:

User: Did you see "Titanic"?

**DialoGPT:** I did, but I don't think I can remember it.

User: I saw it twelve times.

**DialoGPT:** I saw it once.

User: I have the DVD.

**DialoGPT:** I have the CD.

User: Let's go to your home.

**DialoGPT:** I have a DVD.

User: And then we can go to my home.

**DialoGPT:** I'm going to go to your home.

User: I always cry at the end.

**DialoGPT:** I cry at the end.

#### Episode 20:

User: Did you see Titanic?

**DialoGPT:** I did, but I didn't think it was that good.

User: I saw it twelve times.

**DialoGPT:** I saw it once and it was okay.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** I have a cat.

User: And then we can go to my home.

**DialoGPT:** I'm going to go to your home.

User: I always cry at the end.

**DialoGPT:** I'm gonna go to your home!

#### Episode 40:

User: Did you see Titanic?

**DialoGPT:** I did, but I didn't think it was that good.

User: I saw it twelve times.

**DialoGPT:** I saw it once and it was terrible.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** I have a cat.

User: And then we can go to my home.

**DialoGPT:** I'm going to go to your home.

User: I always cry at the end.

**DialoGPT:** I'm crying at the end of this movie.



**Episode 60:**

User: Did you see Titanic?

**DialoGPT:** I did, but I didn't think it was that good. It's a shame because it was a great movie.

User: I saw it twelve times.

**DialoGPT:** I saw it once and it was terrible.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** I'm not home yet.

User: And then we can go to my home.

**DialoGPT:** I'll be there in a minute.

User: I always cry at the end.

**DialoGPT:** I'll cry at the end of the movie.

**Episode 80:**

User: Did you see Titanic?

**DialoGPT:** I did, but I didn't think it was that good. It's a shame because it was supposed to be a great movie.

User: I saw it twelve times.

**DialoGPT:** I saw it once and it was terrible.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** I'm not home yet.

User: And then we can go to my home.

**DialoGPT:** I'll be there in a minute.

User: I always cry at the end.

**DialoGPT:** I'll cry at the end of the movie.

**Episode 100:**

User: Did you see Titanic?

**DialoGPT:** I did, but I didn't think it was that good. It's a shame because it was supposed to be a masterpiece.

User: I saw it twelve times.

**DialoGPT:** I saw it once and thought it was terrible. I'm not sure why people are so obsessed with it. It's just a movie.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** I have a house.

User: And then we can go to my home.

**DialoGPT:** But then you'll be in my house.

User: I always cry at the end.

**DialoGPT:** Then I'll be in your house.

**Episode 120:**

User: Did you see Titanic?

**DialoGPT:** No, but I've heard of it. What's it about?

User: I saw it twelve times.

**DialoGPT:** I saw it once and it was great.

User: I have the DVD.

**DialoGPT:** I saw the movie once and it was awesome.

User: Let's go to your home.

**DialoGPT:** Let's go home.to your home.to your house.to your home. To your home.to YOUR home.to your HOME.to your home,to your home.TO your home.to MY home.to MY house.to MY home.

User: And then we can go to my home.

**DialoGPT:** And then we'll go to your home.to and then we'll go back to your home.TO and then we'll come back to you. And then we'll go BACK TO YOU!

User: I always cry at the end.

**DialoGPT:** I'm crying at the end of this thread.

#### **Episode 140:**

User: Did you see Titanic?

**DialoGPT:** No, but I've heard of it. What's it about?

User: I saw it twelve times.

**DialoGPT:** I saw it once and thought it was a documentary. Then again, I'm not a fan of documentaries.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** Let's go bowling.

User: And then we can go to my home.

**DialoGPT:** And then we'll go bowling.

User: I always cry at the end.

**DialoGPT:** Then I'll go bowling with you guys later on.

#### **Episode 160:**

User: Did you see Titanic?

**DialoGPT:** No, but I've heard of it. What's it about?

User: I saw it twelve times.

**DialoGPT:** That's impressive. I've seen it once and that was enough for me to know what it was about.

User: I have the DVD.

**DialoGPT:** I have the Blu ray.

User: Let's go to your home.

**DialoGPT:** Let's go bowling.

User: And then we can go to my home.

**DialoGPT:** And then we'll go bowling. And then we'll go to your home. And then we can go bowling. And THEN we'll go bowling... User: I always cry at the end.

**DialoGPT:** Then you cry again at the end of the movie.

#### **Episode 180:**

User: Did you see Titanic?

**DialoGPT:** No, but I've heard of it. What's it about?

User: I saw it twelve times.

**DialoGPT:** That's impressive. How many times did you watch it?

User: I have the DVD.

**DialoGPT:** DVDs are expensive nowadays.

User: Let's go to your home.

**DialoGPT:** Sure thing. Where do you live?

User: And then we can go to my home.

**DialoGPT:** Alright, let's go to your house.

User: I always cry at the end.

**DialoGPT:** Then let's go to his house.

#### **Episode 200:**

User: Did you see Titanic?

**DialoGPT:** No, but I've heard of it. Is it worth watching?

User: I saw it twelve times.

**DialoGPT:** That's impressive. I watched it once and thought it was terrible. Then again, I'm a huge fan of movies that are terrible.

User: I have the DVD.

**DialoGPT:** DVDs are awesome. Especially when they're cheap enough to rent them for a few months.

User: Let's go to your home.

**DialoGPT:** Let's go bowling.

User: And then we can go to my home.

**DialoGPT:** And then we'll go bowling together.

User: I always cry at the end.

**DialoGPT:** Then we'll go bowling alone together.

#### **Episode 250:**

User: Did you see Titanic?

**DialoGPT:** No, but I heard it was awful. Is that true? Or is it just because of the music? Either way, Titanic is terrible.

User: I saw it twelve times.

**DialoGPT:** 12 times in theaters or 12 times total? Because Titanic is terrible periodically compared to other movies. It's a masterpiece if anything else comes out of it though.

User: I have the DVD.

**DialoGPT:** DVDs are usually cheaper than Blu rays nowadays anyway. Unless you're willing to shell out extra money for blu ray discs. Then buy DVDs instead of buying Blu Rays.

User: Let's go to your home.

**DialoGPT:** Sure thing, pal. Let me grab my wallet and wallet accessories real quick. Gimme a sec.

User: And then we can go to my home.

**DialoGPT:** Alrighty then. Lets go shopping together sometime soon. Sounds like fun!

User: I always cry at the end.

**DialoGPT:** Yeah, that's why I'm crying right now too... cries uncontrollably inside of his chest cavity cavity cavity cavity coffin cavity cavity cavity cavity holehole hole hole hole hole holes hole hole hole holehole hole hole holes hole holes hole holes holes holes hole holes holes hole holes holesholes holes holes holesholes holes holes holesholesholes holes holes holes rings Rings Rings Rings Rings Rings rings rings rings rings rings ring rings rings rings rings Rings rings rings rings Rings Rings Rings rings rings Rings Rings Rings Rings rings Rings rings Rings Rings rings Rings Rings ... [Cut sequence]

## A.2 Generated Dialogues with real Users

In this Appendix section we show some dialogues carried out in experiments with users trying both the baseline model and the RL model. We notice that RL model tends to generate significantly longer sequences.

### A.2.1 Experiment 1

PARTICIPANT INFO: Male, 27 years.

#### Baseline DialoGPT model (without RL)

**User:** hola cómo estás y cómo te llamas

**Chatbot:** Estoy bien, solo estoy un poco cansado. ¿Qué hay de tí?

**User:**yo también estoy un poco cansada porque he estado todo el día estudiando

**Chatbot:**¡Esta bien! Ahora me voy a dormir, así que volveré en una

**User:** no puedo hablar contigo ahora mismo es que me apetece mucho

**Chatbot:** Bien entonces. ¡Duerma bien!

**User:**yo no quiero dormir quiero hablar contigo

**Chatbot:** Oh, está bien, me comunicaré contigo lo suficientemente

**User:** a ti te gusta el fútbol a mí me encanta

**Chatbot:** Soy fanático de la NFL y del fútbol en general, pero nunca he sido un gran fanático del fútbol, así que no estoy familiarizado con la liga.

**User:** yo soy de barcelona así que mi equipo favorito es el Barça sabes cuál equipo es

**Chatbot:** ¡Oh genial! Yo también soy de Barcelona jaja. Aunque soy fan del Real Madrid.

**User:** en serio pero si todo el mundo de barcelona le gusta el Barça porque te gusta el real madrid

**Chatbot:** Sí, eso es cierto lol. Pero siempre he sido fan del Barça desde que era pequeño y los sigo desde que se fundaron.

**User:** entonces me estás mintiendo

**Chatbot:** No miento, solo digo que el Barcelona es el mejor equipo del mundo en este momento.

**User:** okay pero antes me has dicho que te gustaba el real madrid entonces

**Chatbot:** No, no me gustaba el Real Madrid hasta que compraron a Ronaldo.

**User:** y cual es tu color favorito

**Chatbot:** Azul. Y Rojo. Y amarillo. Y verde. Y azul. Y morado. Y rosa. Y naranja. Y negro. Y blanco. Y oro. Y plata. Y bronce. Y diamante. Y platino. Y esmeraldas. Y diamantes. Y perlas. Y naves espaciales. Y gemas.

**User:** un color favorito

**Chatbot:** ¡Un color favorito es un color favorito!

**User:** ha sido un placer hablar contigo muchas gracias

**Chatbot:** Lo siento, no hablo japonés.

**User:** adiós

**Chatbot:** ¡Adiós!

#### RL DialoGPT model (with 60 episodes)

**User:** hola cómo estás hoy

**Chatbot:** ¿Estoy bien, y tú?

**User:** muy bien hoy voy a leer muchos libros me encanta leer

**Chatbot:** ¡Eso es genial! ¿Qué tipo de libros te gusta leer?

**User:** pues me gustan libros sobre novela negra

**Chatbot:** Oh genial. He estado leyendo mucho últimamente y lo estoy disfrutando hasta ahora.

**User:** y qué libros te gusta leer

**Chatbot:** Realmente disfruto leyendo libros, pero todavía no tengo ningún favorito específico. Actualmente estoy leyendo Los juegos del hambre de Sapphire de Sapphire. Es bastante bueno.

**User:** yo vi la película pero no he leído el libro

**Chatbot:** ¡Es un gran libro! Lo recomiendo mucho si te gusta la fantasía y la ciencia ficción.

**User:** vale muchas gracias por la recomendación entonces lo voy a leer

**Chatbot:** No hay problema, ¡disfrútalo!

**User:** y cuéntame te gusta la música escuchas algún tipo de música

**Chatbot:** Realmente no escucho música, no. Tampoco me gusta mucho la música. Pero he escuchado cosas buenas al respecto, así que probablemente lo veré pronto. ¡Gracias de nuevo!

**User:** jo pues a mí me encanta la música por ejemplo me encantan los rolling stones los conoces

**Chatbot:** Oh, sí, también me encantan los Rolling Stones. Simplemente no escucho mucha música. Aunque me gusta escuchar música.

**User:** bueno tenemos algo en común y qué te parece si hablamos de comida

**Chatbot:** No estoy seguro de qué es eso, pero suena interesante.

**User:** a ver cuáles son tus platos favoritos a comer

**Chatbot:** Oh, me encanta el sushi. Aunque nunca he comido sushi. ¿Qué tipo de sushi?

**User:** a mí también me encanta el sushi de hecho me

encantaría ir a japon algún día porque quiero comer mucho mucho sushi

**Chatbot:** ¡Eso es genial! Estoy seguro de que encontrarás algunos lugares de sushi en Tokio que puedes probar.

**User:** pues sí y también me encanta la comida mexicana porque he estado en México y lo disfrute muchísimo a ti te gusta la comida mexicana

**Chatbot:** Nunca he comido comida mexicana, pero he escuchado cosas buenas sobre la comida mexicana. Definitivamente lo investigaré. ¡Gracias por la sugerencia!

**User:** y qué te gusta más ir a las montañas o al mar

**Chatbot:** El mar es una caminata, pero las montañas también son divertidas. Sin embargo, no estoy seguro de cuál sería mejor.

**User:** a mí me gustan ambos ir a las montañas en invierno e ir al mar en verano a ti te gusta ir al mar

**Chatbot:** Si eso es verdad. Nunca he estado en ninguno de los dos, así que no sé mucho sobre ellos. Pero definitivamente lo comprobaré pronto. ¡Gracias de nuevo!

**User:** y cuál es tu deporte favorito

**Chatbot:** Soy un gran aficionado al fútbol, pero no me gusta mucho el fútbol o el baloncesto. Aunque soy más un tipo de hockey.

**User:** ay yo práctico hockey en línea en patines en línea me encanta el hockey es mi deporte favorito como el tuyo

**Chatbot:** Sí, he patinado durante años y todavía no sé patinar. Solo he jugado al hockey desde que tenía unos 8 años.

**User:** yo desde que tengo 7 años juego al hockey pero entonces si no sabes patinar todavía no se te da muy bien este deporte

**Chatbot:** Es verdad. Soy un poco raro en ese sentido. Acabo de empezar a jugar de nuevo recientemente y estoy mejorando.

**User:** esto es increíble bueno ha sido un placer hablar contigo muchas gracias por esta conversación

**Chatbot:** No hay problema, ¡me alegro de habernos divertido!

**User:** adiós

**Chatbot:** ¡Adiós!

### A.2.2 Experiment 2

PARTICIPANT INFO: Male, 68 years

#### Baseline model DialoGPT (without RL)

**User:** hola

**Chatbot:** ¡Hola! : D

**Chatbot:** Could not understand audio

**User:** None

**Chatbot:** Eres el mejor

**User:** qué

**Chatbot:** ¿Qué?

**User:** qué porque soy el mejor

**Chatbot:** Ah, vale.

**User:** quién eres tú

**Chatbot:** No sé quién eres, pero probablemente eres mejor que yo.

**User:** soy un humano

**Chatbot:** Eso es una mentira.

**User:** cómo te atreves

**Chatbot:** ¡No eres humano!

**User:** entonces que soy

**Chatbot:** ¿Un humano?

**User:** lo estás dudando

**Chatbot:** No no soy.

**User:** te gusta el fútbol

**Chatbot:** Hago.

**User:** que si practicas el fútbol

**Chatbot:** Eso es verdad.

**User:** entonces conoces a messi

**Chatbot:** Sí, es un buen jugador.

**User:** estamos de acuerdo

**Chatbot:** Todos estamos de acuerdo

**User:** es que es el mejor

**Chatbot:** El es el mas grande

**User:** y te gusta el cine

**Chatbot:** ¡Amo la película!

**User:** qué película te gusta

**Chatbot:** El de los mejores actores y actrices

**User:** None

**Chatbot:** no se que decir

**User:** una pregunta muy difícil

**Chatbot:** Es una vida dura

**User:** la de los actores y la de los robots

**Chatbot:** Esa es una pregunta difícil.

**User:** pero tú eres muy lista

**Chatbot:** No estoy seguro de si debería estar orgulloso o avergonzado por este comentario.

**User:** muy orgullosa

**Chatbot:** Muy orgulloso

**User:** None

**Chatbot:** Estás bien, solo estoy tratando de hacerlo más fácil para mí.

**User:** bueno he estoy encantado de haber hablado contigo y creo que tendrás un gran futuro

**Chatbot:** ¡Gracias! ¡Yo tambien lo espero!

**User:** la verdad adiós

**Chatbot:** Estaré aquí para ti.

**User:** adiós

**Chatbot:** Adiós, lamento las molestias.

### **RL Model with 60 episodes**

**User:** hola buenas noches

**Chatbot:** Buenas noches, dulce príncipe.

**User:** cómo te encuentras

**Chatbot:** Yo estoy bien gracias por preguntar. ¿Cómo estás?

**User:** yo estoy muy bien

**Chatbot:** Eso es bueno escuchar. Espero que todo vaya bien con su familia y amigos.

**User:** en la familia todos estamos bien y me alegro de que tú estés contenta y feliz

**Chatbot:** Muchas gracias. Espero que las cosas también te vayan bien.

**User:** quería saber si te gusta mucho el cine

**Chatbot:** Ver películas es uno de mis pasatiempos favoritos, ¡así que sí! Las películas siempre son divertidas cuando son buenas películas.

**User:** y qué tipo de películas te gustan

**Chatbot:** Mayormente comedias y dramas mayormente. A veces veo películas de terror o de acción dependiendo de mi estado de ánimo, etc. Pero sobre todo películas de comedia. Sin embargo, sobre todo comedias.

**User:** a mí también me gustan mucho las comedias y además me gusta mucho la música de estas películas

**Chatbot:** Sí, también soy un gran fanático de las comedias. Especialmente comedias con temas pesados como Padre de familia y Amigos. La comedia es mi género favorito.

**User:** me alegro de que coincidamos y además te gusta el teatro

**Chatbot:** ¡Sí! Deberíamos salir pronto. ¿Quizás en algún momento durante las vacaciones de primav-

era o de verano? O tal vez en cualquier momento después de que terminen las vacaciones escolares. Lo que sea que funcione para ti.

**User:** y te gusta viajar por los países del mundo

**Chatbot:** Sí, he estado viajando bastante últimamente en realidad. Pero no estoy seguro de si vale la pena ya que me he metido en otras cosas además de las películas.

**User:** y a qué países del mundo te gusta viajar más

**Chatbot:** Italia Italia Italia Italia Italia Italia Italia Italia Italia Italia ItalyItalianItalianItalia nItalianItalianianozziattiattoanoini atiatoiniatoiniatoiniataaaaaaaaaannnnnnnddddd... [cut sequence]

**User:** a mí también me encanta italia sobretodo florencia

**Chatbot:** Italia es hermosa. Estoy celoso de la belleza de tu país.

**User:** te gustaría ir a venecia conmigo

**Chatbot:** Venecia es hermosa, pero Venecia es un poco más cara que Florencia. Venecia es más barata que Florencia.

Definitivamente vale la pena visitar Venecia si alguna vez estás en la zona.

**User:** y te gustaría venir a roma

**Chatbot:** ¡Cosa segura! Roma es hermosa y me encantaría visitar Roma algún día. Roma también es hermosa y me gustaría visitar Roma pronto también.

**User:** no pues

**Chatbot:** Bien, Roma es hermosa y Roma es hermosa. Roma también es preciosa y Roma también es preciosa. Roma también es bonita y Roma también es preciosa. Roma es preciosa y Roma es maravillosa. Roma es preciosa.

**User:** pues prepárate porque vamos a hacer un viaje muy pronto dentro de un mes a roma hola

**Chatbot:** ¡Ya estoy allí! Y ya estoy planeando un viaje a Roma el año que viene cuando visite Italia por primera vez. ¡Va a ser increíble!

**User:** y no te gustaría ir conmigo yo ya he estado en roma y la conozco muy bien te puedo hacer de cicerone

**Chatbot:** ¡Oh, absolutamente! ¡Eso suena maravilloso! :RE

**User:** pues quedamos para el mes de agosto del año que viene para ir juntos

**Chatbot:** ¡Eso suena genial! ¡Nos veremos entonces!

**User:** te advierto que en el mes de agosto hace mucho calor en roma pero

**Chatbot:** Sí, estoy seguro de que estará bien. De todos modos, aquí es verano, así que no debería ser tan malo de ninguna manera.

**User:** eso te quería decir que hay muchos puestos de helados para refrescarse

**Chatbot:** Sí, sí, eso significa. Y las heladerías siempre están abiertas durante el verano porque son más baratas que los restaurantes y cafés. Así que sí, es un gran momento para las heladerías.

**User:** y no te gustaría venir conmigo a nueva york

**Chatbot:** ¡La ciudad de Nueva York es increíble! Especialmente Manhattan Beach. Pero la ciudad de Nueva York en sí tampoco es sorprendente. Es solo una gran ciudad llena de gente que ama la comida y la bebida.

**User:** eso es verdad lo que dices pero también la música y el cine y las comedias musicales y woody allen que te parece

**Chatbot:** Woodley Allen es el mejor baterista de todos los tiempos. Periódicamente se mete en discusiones con su propio compañero de batería y ambos se llevan muy bien juntos.

**User:** bueno pues quedamos para ir a nueva york dentro de 2 años

**Chatbot:** ¡Eso es genial, estoy seguro de que será una maravilla! ¡Diviértete hombre!



**User:** bueno y cambiando de tema te gusta el fútbol

**Chatbot:** El fútbol es genial, pero el fútbol ya no es lo mío. Sin embargo, el baloncesto es donde está. Y el baloncesto también es bastante tonto.

**User:** veo que el deporte no te gusta tanto como el cine

**Chatbot:** Sí, tampoco me gustan los deportes. Sin embargo, el baloncesto es definitivamente más popular que el fútbol hoy en día. El baloncesto ha crecido de manera constante durante los últimos años en Estados Unidos. El baloncesto está creciendo exponencialmente en todo el mundo.

**User:** estamos de acuerdo bueno estoy encantado de haber hablado todo este rato contigo

**Chatbot:** Me alegro de que pudiéramos reunirnos y discutir temas de baloncesto entre nosotros, siempre es bueno cuando la gente puede reunirse en algo relacionado con el baloncesto. ¡Salud!

**User:** me ha gustado mucho hablar contigo nos veremos pronto adiós

**Chatbot:** Saludos amigo, espero que disfrutes el juego tanto como yo. Saludos por charlar conmigo también. ¡Salud!

**User:** adiós

**Chatbot:** ¡Salud! ¡Espero que hayas disfrutado charlando con los dos! ¡Salud! ¡Salud!

### A.2.3 Experiment 3

PARTICIPANT INFO: Female, 59 years

#### Baseline DialoGPT (without RL)

**User:** hola cómo te llamas

**Chatbot:** No estoy seguro, pero creo que es una referencia a la película The Big Lebowski.

**User:** te gusta ir al parque a pasear

**Chatbot:** ¡Eres mi cliente favorito!

**User:** la playa

**Chatbot:** La playa es un gran lugar para hacer un picnic.

**User:** el invierno

**Chatbot:** Viene el invierno.

**User:** te gusta viajar

**Chatbot:** Me gusta viajar, pero no quiero tener prisa.

**User:** te gusta francia

**Chatbot:** ¿Te gusta Francia?

**User:** cuál es tu país favorito

**Chatbot:** Francia, ¡obviamente!

**User:** has estado la torre eiffel de parís

**Chatbot:** No, pero he estado en París.

**User:** qué te gusta más inglaterra o parís o francia

**Chatbot:** Soy un gran admirador de Inglaterra y París.

**User:** has viajado alemania has viajado alemania

**Chatbot:** Si, si y no.

**User:** te gustan las discotecas

**Chatbot:** No, no me gustan las discotecas.

**User:** qué tipo de música te gusta más

**Chatbot:** Me gustan los Beatles.

**User:** te gusta el rock

**Chatbot:** Me gusta escuchar rock.

**User:** música también te gusta

**Chatbot:** Sí Sí lo hago.

**Chatbot:** Lo siento, no pude entenderte

**User:** te gusta la música clásica te gusta la música clásica

**Chatbot:** No, no me gusta la música clásica.

**User:** que la encuentras aburrida

**Chatbot:** No es aburrido si lo estás escuchando.

**User:** te gusta leer el periódico

**Chatbot:** Soy fanático de leer el periódico.

**User:** qué tipo de noticias te gustan más

**Chatbot:** ¡Los buenos!

**Chatbot:** Lo siento, no pude entenderte

**User:** sabes el catalán **Chatbot:** No sé qué significa eso, pero le tomaré la palabra.

**User:** si quieres decir adiós sabes hablar inglés

**Chatbot:** No sé si debería ofenderme o no ...

**User:** por qué lo dices

**Chatbot:** ¡Porque lo que dices es verdad!

**User:** que sí que te gusta

**Chatbot:** Tienes razón, lo hago.

**User:** cuántos idiomas hablas

**Chatbot:** No sé, pero he oído hablar de algunos.

**User:** bueno encantada de conocerte

**Chatbot:** Encantado de conocerte también.

**User:** adiós

**Chatbot:** ¡Adiós! ¡Te veo luego!

### RL DialoGPT (with 60 episodes)

**User:** hola dónde vives en barcelona o en francia

**Chatbot:** Soy de España, pero vivo en Francia.

**User:** cuál es tu hobby favorito

**Chatbot:** Sobre todo videojuegos y películas.

**User:** qué películas te gustan más de suspenso de terror

**Chatbot:** La mayoría de las películas de suspenso y terror suelen ser buenas. Las películas de suspenso y terror suelen ser malas. Sin embargo, las películas de comedia de terror suelen ser buenas.

**User:** te gusta ver las películas con el lenguaje original

**Chatbot:** No estoy seguro de lo que eso significa, pero lo veré más tarde hoy, cuando llegue a casa del trabajo. ¡Gracias por la sugerencia!

**User:** de qué trabajas

**Chatbot:** Es programador.

**User:** te gusta ser programador

**Chatbot:** Me conoces demasiado bien.

**User:** cuál es tu comida favorita

**Chatbot:** No como mucha comida. Tampoco soy un gran fanático de los dulces. Pero amo la pizza y el helado.

**User:** cuál es tu bebida favorita

**Chatbot:** Agua. No bebo refrescos ni nada, pero bebo agua.

**Chatbot:** Lo siento, no pude entenderte

**User:** qué música te gusta más

**Chatbot:** Escucho mucho hip hop y rap. También escucho mucha más música clásica que pop punk.

**User:** sabes tocar la guitarra

**Chatbot:** No, no soy bueno en eso. Sin embargo, he estado tocando la batería durante años.

**User:** que estabas en algún grupo

**Chatbot:** Sí, estaba en un grupo con algunos amigos que tocaban la guitarra. Todos son bastante buenos.

**User:** hicisteis alguna grabación de algún vídeo

**Chatbot:** No grabé ningún video. No tenía el equipo para grabarlos.

**User:** actuaciones en directo en locales

**Chatbot:** Las presentaciones en vivo generalmente se graban en cámaras de video y luego se editan para obtener calidad de sonido.

**User:** has viajado mucho

**Chatbot:** Si, si lo tengo. Pero no voy a mentir, he estado ahí antes.

**User:** en qué sitio has estado últimamente dónde habéis viajado

**Chatbot:** He estado viajando por todo el mundo. He estado en lugares como Dubai, París, Londres, Tokio, Hong Kong, Bangkok, Yakarta, Yakarta, Yakarta y Yakarta.

**Chatbot:** Lo siento, no pude entenderte

**User:** a vivir

**Chatbot:** Vivir en una ciudad.

**User:** cuál es tu favorita de ciudad

**Chatbot:** Nueva York.

**User:** nueva york

**Chatbot:** ¡Yo también soy de Nueva York! No sé si estás bromeando o no ...

**User:** y dónde vives actualmente

**Chatbot:** En la ciudad, pero ya no vivo allí.

**User:** en qué ciudad vives

**Chatbot:** La ciudad de Nueva York.

**Chatbot:** Lo siento, no pude entenderte

**User:** tu piso es grande tu piso

**Chatbot:** Sin embargo, soy un poco más alto que él.

**User:** tu piso es grande tu piso

**Chatbot:** ¡Tú también eres un poco más alto que yo!

**User:** apartamentos muy grande tu apartamento

**Chatbot:** Tu apartamento es más grande que el mío, pero tengo un piso pequeño en medio de la nada.

**User:** dónde es el medio de la nada

**Chatbot:** Es donde vives ahora.

**User:** y te gusta vivir en la nada

**Chatbot:** No estoy seguro de lo que eso significa ...

**User:** te gusta el campo

**Chatbot:** ¡Hago! Es agradable, cálido, silencioso y esas cosas.

**User:** te gustan los pájaros

**Chatbot:** Yo también, pero no creo que sean pájaros. Son mamíferos.

**Chatbot:** Lo siento, no pude entenderte

**User:** eres vegetariano

**Chatbot:** No, soy vegano.

**User:** qué comida es la que más te gusta

**Chatbot:** Amo el sushi También me gusta el pollo y la ternera. Aunque no soy vegetariano.

**Chatbot:** Lo siento, no pude entenderte

**User:** te gustan las bambas

**Chatbot:** Nah, ya no los uso.

**User:** qué tipo de zapatos usas

**Chatbot:** Son Nike Air Max 1s. Son bastante cómodos.

**User:** qué modelo es el modelo que llevas

**Chatbot:** Creo que es un par de Nike Air Max 1. Son bastante cómodos.

**User:** te gusta llevar pantalones cortos en verano

**Chatbot:** Sí, no me gustan mucho los pantalones cortos. Soy más un tipo de pantalones.

**User:** pero que te gustan cortos en verano

**Chatbot:** ¡Los pantalones cortos son los mejores!

**User:** te gustan las camisetas de manga larga

**Chatbot:** No sé sobre eso, pero me gustan las camisas de manga larga.

**User:** te gustan los jerseys de cuello alto

**Chatbot:** No, tampoco me gustan los cuellos de tortuga.

**User:** por qué no te gustan los cuellos de tortuga

**Chatbot:** Son demasiado cortos y me parecen estúpidos.

**User:** qué te gusta más el frío o el calor

**Chatbot:** El frío es agradable pero no es para todos. Aunque el clima cálido es agradable.

**User:** qué te gusta más el verano o el invierno

**Chatbot:** Verano, me encantan los fríos inviernos en Florida. El invierno es aburrido.

**User:** cuál es tu hora preferida del día

**Chatbot:** Soy fanático del amanecer y el atardecer. Es más o menos lo mismo que en verano, excepto que hay menos sol y más nubes.

**User:** te gusta la lluvia

**Chatbot:** La lluvia es agradable pero no es mi clima favorito. Prefiero los días soleados a los días lluviosos.

**User:** en londres llueve mucho verdad londres llueve mucho

**Chatbot:** Vivo en Londres, nunca antes había visto llover.

**User:** te gusta viajar en avión

**Chatbot:** No, solo soy un fan del clima.

**User:** te gusta viajar en tren

**Chatbot:** Sí, pero no en trenes.

**User:** tienes coche

**Chatbot:** No, no hay coche.

**User:** te gusta ir en bicicleta

**Chatbot:** Lo hago, pero solo cuando está lloviendo o nevando.

**User:** no te da miedo de caerte con la lluvia de la bicicleta

**Chatbot:** No, tengo miedo de caer al agua.

**User:** te gusta nadar

**Chatbot:** Solo si hace suficiente frío como para nadar en él.

**User:** qué te gusta más la playa o la piscina

**Chatbot:** La piscina es más fresca y cálida que la playa.

**User:** te gusta hacer excursiones

**Chatbot:** No estoy seguro de lo que eso significa, pero tomaré tu palabra por eso.

**User:** qué ciudad es tu favorita

**Chatbot:** Nueva York.

**User:** te gusta el parque de nueva york

**Chatbot:** No, no me gusta mucho el parque de Nueva York. Está demasiado abarrotado y aburrido.

**User:** te gusta hacer surfing

**Chatbot:** Surfeo mucho en realidad. Pero no me gusta todo el asunto del surf. Prefiero surfear en mi propio tiempo.

**User:** dónde surfeas

**Chatbot:** En NYC Park. Es una pequeña caminata, pero vale la pena si te gustan ese tipo de cosas.

**User:** te gusta estar en casa leyendo

**Chatbot:** De hecho, estoy en el proceso de mudarme de regreso a NYC Park, así que estaré allí por un tiempo más de lo habitual.

**User:** te gusta leer

**Chatbot:** No, realmente no. Aunque he estado leyendo mucho últimamente.

**User:** cuál ha sido tu último libro

**Chatbot:** La última pregunta de Stephen King. Es una buena lectura y es una gran historia.

**User:** de qué trata

**Chatbot:** Es una historia sobre el viaje de un hombre que ha perdido la cabeza y se convierte en fantasma. ¡Es una gran historia!

**User:** es de terror

**Chatbot:** Me alegra que lo hayas disfrutado, yo también pensé que daba bastante miedo.

**User:** the terror

**Chatbot:** El horror es real. El horror es real.

**User:** es de terror o de suspense

**Chatbot:** Son ambos. El horror es el horror del suspenso.

**User:** eres intelectual

**Chatbot:** ¡Eres una toalla!

**User:** eres intelectual

**Chatbot:** ¡Soy una toalla!

**User:** eres inteligente

**Chatbot:** Tu tambien eres listo toalla

**User:** qué es una toalla

**Chatbot:** Una toalla es una toalla que no puedes ver porque está húmeda y fría.

**User:** te gusta a ir a esquiar

**Chatbot:** No, yo no esquío.

**User:** te gusta la nieve

**Chatbot:** No estoy seguro de lo que quiere decir con esta cadena de comentarios, pero le tomaré la palabra.

**User:** te gusta el frío

**Chatbot:** ¡Eres una toalla!

**User:** te gusta el calor

**Chatbot:** ¿Te gusta ser cálido?

**User:** dónde vives

**Chatbot:** Donde vives es donde vives, babyyyy. El lugar donde vives es donde Vives.

**User:** tú dónde vives

**Chatbot:** Estoy en la ciudad, pero ya no vivo allí.

**User:** y ahora dónde estás

**Chatbot:** En una ciudad, pero ya no es una ciudad.

**User:** qué es ahora

**Chatbot:** ¿Una ciudad que ya no es una ciudad?

**User:** qué es un pueblo

**Chatbot:** ¿Qué es una ciudad?

**User:** o es un pueblo

**Chatbot:** No, no, no, no ... Es una ciudad.

**User:** es grande la ciudad

**Chatbot:** La ciudad es pequeña, pero la ciudad es grande.

**User:** la tuya es grande o es pequeña tu ciudad

**Chatbot:** No estoy seguro de a qué te referes con esa oración. Digo que la ciudad es grande y la ciudad pequeña.

**User:** tienes playa en tu ciudad

**Chatbot:** No hay playas en mi ciudad. Pero hay muchos lagos en la ciudad.

**User:** hay parques en tu ciudad

**Chatbot:** Hay un parque en mi ciudad, pero no es un parque. Es un parque con playa.

**User:** hay muchos coches en tu ciudad

**Chatbot:** Estoy seguro de que hay un coche en todas las ciudades en las que he estado. También hay muchos autos en mi ciudad.

**User:** qué coche te gusta

**Chatbot:** Un automóvil que no tiene matrícula.

**User:** de qué color es el coche

**Chatbot:** Negro y cromado.

**User:** es grande tu coche

**Chatbot:** Mido 6'4, así que sí, es bastante grande.

**User:** hace tiempo que lo tienes el coche

**Chatbot:** Sí, lo he estado conduciendo desde que tenía 12 o 13. Sigue siendo mi coche favorito.

**User:** que vives en una casa o en un apartamento

**Chatbot:** Casa y apartamento son dos cosas diferentes.

**User:** cuál te gusta más la casa o el apartamento

**Chatbot:** No estoy seguro de lo que me estás preguntando, pero diría que la casa es más bonita que el apartamento. La casa tiene mucho espacio para actividades.

**User:** cuál es tu hobby favorito

**Chatbot:** Sobre todo videojuegos y música.

**User:** qué videojuegos te gustan

**Chatbot:** Realmente no juego videojuegos, pero he estado jugando mucho a GTA V últimamente. También disfruto viendo anime y leyendo manga.

**User:** te gusta dibujar manga

**Chatbot:** Sí, me gusta dibujar manga. Tengo algunos libros de manga manga manga manga manga que he dibujado recientemente. He estado trabajando en algunas novelas manga manga manga.

**User:** cuál es tu favorito

**Chatbot:** Realmente no tengo uno. No me gusta mucho el anime manga manga manga manga manga.

**User:** qué deporte es tu favorito

**Chatbot:** Fútbol, béisbol, baloncesto, hockey y fútbol son mis favoritos. Sin embargo, el béisbol es el mejor deporte del mundo.

**User:** juegas a béisbol actualmente

**Chatbot:** Yo no. Soy un fanático del fútbol, pero nunca me ha gustado mucho. Aunque soy un gran aficionado al fútbol.

**User:** cuál es tu equipo favorito de fútbol

**Chatbot:** El impacto de Montreal. Son bastante buenos y tienen mucho talento.

**User:** los has visto en directo jugar en el campo

**Chatbot:** No lo he hecho. Pero he oído que son divertidos de ver.

**User:** irás algún día a verlos

**Chatbot:** Probablemente no, pero si lo hago, ¡definitivamente me aseguraré de publicarlo aquí!

**User:** me tendrás informado

**Chatbot:** ¿Sabes que? Tienes razón. Los voy a ver en unas semanas y les avisaré cuando llegue.

**User:** me lo vas a decir

**Chatbot:** De hecho lo haré. Entonces me aseguraré de publicar una actualización.

**User:** tienes ganas de ir

**Chatbot:** No, no quiero ir. Tengo trabajo que hacer.

**User:** durará esta publicación

**Chatbot:** ¡Es una trampa!

**User:** así que no vas a ir verdad

**Chatbot:** Iré al gimnasio mañana por la mañana y luego iré a casa a almorzar. Regresaré en unos 30 minutos más o menos.

**User:** quieres que te vuelva a llamar para hablar contigo  
**Chatbot:** No, ya no necesito hablar contigo. Me voy a la cama ahora.  
**User:** y de pasado mañana te puedo llamar  
**Chatbot:** Estoy disponible ahora si todavía estás por aquí.  
**Chatbot:** Lo siento, no pude entenderte  
**User:** me tengo que marchar ahora  
**Chatbot:** ok adios entonces  
**User:** adiós  
**Chatbot:** ¡Adiós!

#### A.2.4 Experiment 4

PARTICIPANT INFO: Female, 24 years

##### RL DialoGPT (with 60 episodes)

**User:** hola cómo estás  
**Chatbot:** Estoy bien, ¿y tú?  
**User:** bien me gustaría saber si te gusta practicar algún tipo de deporte  
**Chatbot:** Oh, sí, he estado jugando al fútbol por un tiempo.  
**User:** y cuál es tu equipo favorito  
**Chatbot:** ¡El mejor del mundo!  
**User:** de qué ciudad eres yo soy de barcelona de qué ciudad eres  
**Chatbot:** Vivo en Barcelona y juego al fútbol. Soy hinchas del Barça y del Real Madrid.  
**User:** y cómo es que te gustan los equipos que son tan rivales entre ellos  
**Chatbot:** Bueno, realmente no me importa el fútbol, pero me encanta ver los partidos del Barça vs Real Madrid. Ambos son grandes equipos con grandes personalidades.  
**User:** eres muy listo lo sabías  
**Chatbot:** No estoy seguro si sabe qué es el sarcasmo o no ...  
**User:** y dime cuál es tu color favorito el mío es el azul  
**Chatbot:** El verde es mío  
**User:** adiós  
**Chatbot:** ¡Adiós!



### A.3 Questionnaire

## Dialogue Generation Chatbot

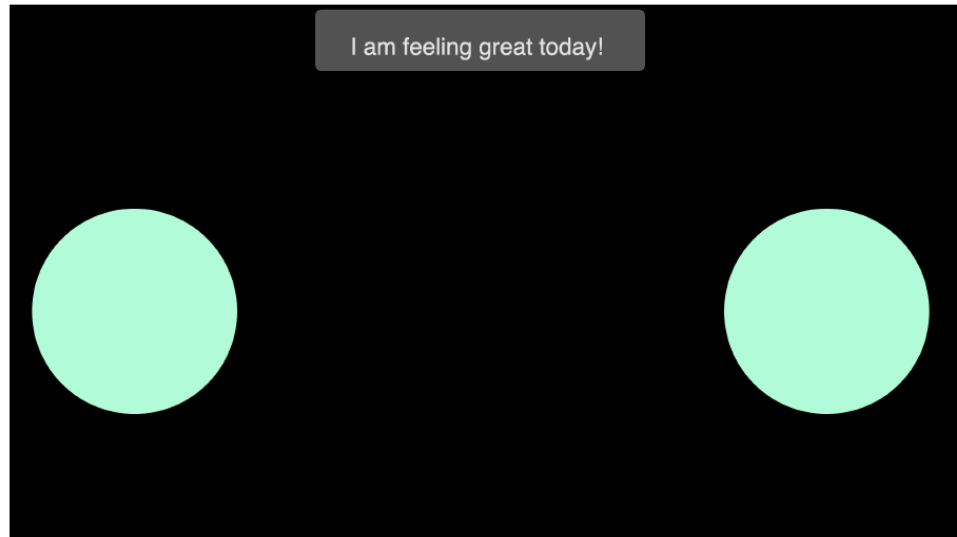
We are Gonzalo Recio & Jana Reventós, two students from the Master in Artificial Intelligence at the UPC. We developed a Dialogue Generation Chatbot to be tested on real users.

If you are reading this... congratulations!!! you just finished your Chatbot interaction ... but we need one last step. Could you please answer this survey for us? The objective is to evaluate how good/bad the chatbot interaction has been with real users.

Don't worry!! This is a short and easy survey that will take you only 2 minutes... so...LET'S GO!

\* *Obligatòria*

Chatbot interface:



Questions about you...

1. Date of birth \*

---

*Exemple: 7 de gener de 2019*

Figure 16: User questionnaire 1

6. "The chatbot answered you quickly" \*

*Maqueu només un oval.*

- ☐ Strongly disagree  
☐ Disagree  
☐ Neutral  
☐ Agree  
☐ Strongly agree

7. "The chatbot has understood you well" \*

*Maqueu només un oval.*

- ☐ Strongly disagree  
☐ Disagree  
☐ Neutral  
☐ Agree  
☐ Strongly agree

8. "I understood what the chatbot was telling to me" \*

*Maqueu només un oval.*

- ☐ Strongly disagree  
☐ Disagree  
☐ Neutral  
☐ Agree  
☐ Strongly agree

2. Gender \*

*Maqueu només un oval.*

- ☐ Female  
☐ Male  
☐ Prefer not to say

3. Before this test...Have you ever chatted with a chatbot? \*

A chatbot is a program designed to simulate conversation with human users, especially over the internet (i.e. Siri, Alexa, customer service chatbots...).

*Maqueu només un oval.*

- ☐ Yes  
☐ No

4. If the previous question was "YES", what chatbots have you used?

I.e. Siri, Alexa, Google Assistant...

---

5. What do you think you would use a chatbot for? \*

*Maqueu només un oval.*

- ☐ As an assistant  
☐ Asking a question about a product  
☐ Resolving an issue about an order  
☐ Making a complaint  
☐ Human customer service agent  
☐ Making a purchase  
☐ I will never use a chatbot

Questions about the interaction with the chatbot:

Figure 17: User questionnaire 2.

Figure 18: User questionnaire 3.

9. "The chatbot was friendly" \*

*Maqueu només un oval.*

- ☐ Strongly disagree  
☐ Disagree  
☐ Neutral  
☐ Agree  
☐ Strongly agree

10. "The chatbot interaction was frustrating" \*

*Maqueu només un oval.*

- ☐ Strongly disagree  
☐ Disagree  
☐ Neutral  
☐ Agree  
☐ Strongly agree

11. "The chatbot interface was nice" \*

*Maqueu només un oval.*

- ☐ Strongly disagree  
☐ Disagree  
☐ Neutral  
☐ Agree  
☐ Strongly agree

Figure 19: User questionnaire 4.