EmoCare

Carla Mendes[a]

[a]Computer Science and Communications Research Centre, School of Technology and Management, Polytechnic of Leiria, 2411-901 Leiria, Portugal

ARTICLE INFO

Keywords:

ABSTRACT

Abstract goes here.

1. Introdution

2. Related work

3. Background and contextualization

This section aims to review existing literature regarding Conversational Agents (CA), Sentiment Analysis (SA) and mood assessment questionnaires.

3.1. Conversational agents

Chatbots, referred to as CA or Conversational Systems (CS), are software applications created to imitate human-computer interactions [1].

3.1.1. Classification methods

Chatbots can be categorized based on various factors, including their goal, interaction mode, knowledge domain, and response-generation method [2].

CS operate using different modes of interaction, namely based on text or speech. In the text-based mode, users communicate with the chatbot by typing their queries or statements, commonly through chat applications, messaging platforms, or web-based chat interfaces. The chatbot responds with text-based messages in return. On the other hand, speech-based chatbots enable users to interact with the chatbot using spoken language. These chatbots employ Speech Recognition (SR) technology to convert the user's voice input into text, which is then processed and analyzed to generate appropriate responses. Voice-based chatbots are typically found in voice assistants like Amazon Alexa, Google Assistant, or Apple Siri. They offer a convenient and hands-free method of interacting with the chatbot, enabling users to engage in natural conversations and perform tasks using voice commands. Additionally, chatbots can also adopt a multi-modal approach, allowing interaction through both text and speech [1].

When considering their objective, chatbots are categorized as either task-oriented or non-task-oriented. Task-oriented chatbots are designed with a specific purpose, focusing on handling particular tasks and engaging in brief conversations, typically within a limited domain. Conversely, non-task-oriented chatbots specialize in emulating conversations with individuals and participating in casual chitchat primarily for entertainment. As a result, they operate in open domains, facilitating more diverse and unrestricted conversations [2].

In terms of response generation methods, chatbots are also classified based on the techniques and algorithms used to generate suitable and meaningful responses to user queries or inputs. These methods can be categorized as follows:

- **Rule-based**: rely on a predefined set of rules and patterns to generate responses. Human experts typically create these rules and program them into the chatbot system. The CA then matches user inputs with specific patterns or keywords and retrieves corresponding responses. Rule-based systems work well for simple and structured conversations but may struggle with complex or unpredictable queries.

- **Retrieval-based**: use Machine Learning (ML) techniques to select the most appropriate response based on the user's input from a large dataset of predefined responses. Retrieval-based systems may struggle with creative responses, however provide contextually relevant responses.

- **Template-based**: use pre-built response templates that are filled with relevant information based on user inputs. These templates contain placeholders for dynamic content such as names, dates, or specific details. The chatbot captures the intents of the user's input and selects an appropriate template to generate a response. Template-based systems may lack flexibility and creativity in generating unique responses but are relatively simple to implement.

- **Generative-based**: rely on advanced Natural Language Processing (NLP) techniques and ML models to create responses from scratch, namely sequence-to-sequence models or transformers. These models are trained on large datasets and learn the patterns and structures of human language. Generative models are more versatile in handling diverse user inputs. However, they can be more computationally intensive and require significant computational resources for training and inference.

- **Hybrid Approach** - combine multiple methods mentioned above to leverage the strengths of each approach.

carla.c.mendes@ipleiria.pt (C. Mendes)
ORCID(s): 0000-0001-7138-7124 (C. Mendes)
https://www.linkedin.com/profile/view?id='carla-mendes-5b3586233' (C. Mendes)

Lastly, in the knowledge domain dimension, chatbots possess specialized knowledge and capabilities tailored to serve specific purposes within a limited scope. They possess knowledge and capabilities customized to fulfill specific objectives or offer assistance within a restricted topic. These chatbots undergo training and programming to comprehend and address inquiries regarding a specific domain. For instance, a chatbot tailored to a particular subject can aid in customer support for an e-commerce platform, provide medical advice in the healthcare sector, or offer travel recommendations within the tourism industry. On the other hand, open-domain chatbots engage in conversations spanning a broad spectrum of topics, free from confinement to any specific domain. They aim to emulate human-like interactions, delivering casual conversation, entertainment, or general information across various subjects [3].

3.1.2. General Architecture

A robust CA system will possess several key components, represented in Figure 1.

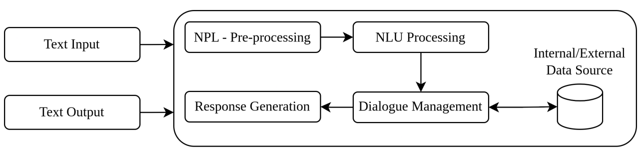

Figure 1: Conversational agents general architecture.

During the NLP phase, the user's request undergoes various techniques, including tokenization, lemmatization, and stemming. These techniques help extract structured data from the request, which is then passed on to the subsequent component, known as the Natural Language Understanding (NLU) module, responsible for analyzing each incoming user request using various strategies, namely, parsing the request to understand the user's intention and the associated details. The dialogue management module focuses on keeping track of the dialogue context and defining the following actions to perform by analyzing the input request that has been transformed into understandable structured data by the CA system. The data sources serve as repositories for information and data utilized by the dialogue manager. These sources can be either internal or external. Internally, chatbots can access data from templates or rules to understand user requests and generate appropriate responses. Moreover, CA can also build their databases from scratch or leverage existing databases that align with their domain and functionality. In contrast, external data sources can be accessed through third-party services like Web APIs, which provide the necessary information. The response generator module plays a crucial role in generating an appropriate response from a pool of potential options after executing an action. This component utilizes the approaches mentioned earlier to generate the most suitable response for the given context. [**? ?**]

3.1.3. Tools

Rasa [4] is an open-source dialogue framework for building conversational Artificial Inteligence (AI) applications. It uses NLP techniques and dialogue management to enable interactive and context-aware conversations. Rasa consists of two main components: the NLU module for processing user inputs and extracting intents and entities, and the Dialogue Management module for handling conversation flow and decision-making. It supports personalized dialogue policies, provides tools for training and evaluation, and integrates with different channels and platforms. Rasa supports both text-based and voice-based interactions, making it versatile for various applications.

Amazon Lex [5] is a service provided by Amazon Web Services that allows developers to build, test, and deploy CA powered by AI. It is designed to create interactive chatbots and virtual assistants that can understand natural language inputs and provide appropriate responses. Amazon Lex leverages advanced natural language models and ML algorithms to enable accurate understanding and interpretation of user inputs. It supports both text and speech inputs and outputs, making it suitable for various applications. With Amazon Lex, developers can easily integrate CA into their applications or platforms, enabling more intuitive and engaging user experiences.

Dialogflow [6] is a NLU platform developed by Google. It provides tools and capabilities for building CA, chatbots, and virtual assistants. With Dialogflow, developers can create, manage, and deploy CA across multiple platforms and systems. It supports both text and speech inputs and outputs, allowing users to interact with the CA through various channels such as messaging platforms, voice assistants, and websites. This platform utilizes advanced ML algorithms to understand and interpret user inputs, extracting important information such as intents (the user's intention) and entities (specific pieces of information). It offers a range of pre-built NLU components and features, including Named Entity Recognition (ER) and SA. Additionally, Dialogflow provides a visual interface for designing conversation flows, managing dialogues, and defining responses.

OpenDial [7] is an open-source Java-based toolkit used for building and evaluating speech-based CA. It provides a framework and set of tools that enable developers to create interactive dialogue systems capable of engaging in natural language conversations. Furthermore, the toolkit offers a range of features and functionalities for building CA. It provides modules for NLU, dialogue management, and speech synthesis. OpenDial allows developers to define dialogue policies and strategies to guide the system's behaviour and response generation. It also includes components for handling user input, managing context, and generating appropriate spoken responses. Overall, OpenDial emphasizes modularity and extensibility, enabling developers to customize and adapt the toolkit according to their specific requirements.

Botpress [8] is an open-source platform that enables developers to build, deploy, and manage chatbots and virtual assistants. It provides a visual interface for designing conversational flows and supports both text-based and voice-based interactions. Botpress is written in JavaScript and can be deployed on various platforms. One of the key features of Botpress is its visual flow builder, which allows developers to create complex conversational flows using a drag-and-drop interface. This makes it easy to design the dialogue flow of the chatbot and define the interactions between the user and the bot. Botpress also offers built-in NLU capabilities, allowing developers to train the chatbot to understand user intents and extract entities from user inputs.

ChatterBot [9] is an open-source Python library that facilitates the development of chatbots. The primary focus of ChatterBot is to generate responses based on pre-defined conversational patterns. It uses a machine learning algorithm called Latent Semantic Analysis (LSA) to train a language model on a given corpus of text data and then generate appropriate responses based on the patterns it has learned. ChatterBot supports the use of multiple languages and provides various pre-trained language models that can be used out of the box. Additionally, it enables developers to customize the chatbot's behaviour by defining rules, selecting appropriate responses, and handling specific cases. One of the notable features of ChatterBot is its ability to learn and improve over time. It employs a technique called "conversational context" to maintain the history of the conversation and generate contextually relevant responses.

3.2. Sentiment Analysis

SA, a subfield of NLP, aims to derive the sentiments expressed in a piece of text based on its content, which can be conducted at different levels: Document Level, Sentence Level, Phrase Level, and Aspect Level [10, 11].

Document-level SA involves assessing the sentiment of an entire document and assigning a single polarity to it. Classification methods, both supervised and unsupervised, can be employed at this level to determine the sentiment conveyed in the document [11].

At the Sentence Level, each sentence is analyzed and classified into a polarity sentiment. This approach more valuable when a document contains a diverse range of sentiments. By independently determining the polarity of each sentence, either using methodologies similar to document-level analysis or with more extensive training data and processing resources, the overall sentiment of the document can be aggregated or analyzed sentence by sentence [11].

The Phrase Level of SA focuses on mining opinion words and sentiments at the phrase level. While document-level analysis broadly categorizes the entire document as either positive or negative, sentence-level analysis proves more advantageous because documents often contain both positive and negative statements. At this level, individual words become the fundamental units of language, and their polarity is intrinsically tied to the subjectivity of the sentences or documents in which they appear [11].

Lastly, Aspect Level delves even deeper, as it is performed on specific aspects within a sentence. Since a sentence may contain multiple aspects, this approach pays close attention to all aspects present and assigns polarity to each one. An aggregate sentiment is then calculated for the entire sentence, considering the sentiments of all its aspects.

To conduct SA, several main steps are followed, including data collection, feature selection, feature extraction, and the use of word embeddings. These steps aid in gathering relevant data, choosing essential features, extracting meaningful patterns, and representing words in a numerical format suitable for analysis [10].

3.2.1. Approaches

There are three main approaches commonly used for SA: the Lexicon Based Approach, the ML Approach, and the Hybrid Approach [11, 10].

Lexicons consist of tokens, with each token having a predefined score that indicates the neutral, positive, or negative nature of the text. The lexicon-based method is highly suitable for conducting SA at both the sentence and feature levels. Initially, the document is divided into tokens of individual words, and then the polarity of each token is calculated and aggregated. There are primarily two approaches used in Lexicon Based Approaches: the Corpus Based Approach and the Dictionary based approach [11].

The Corpus-based approach utilizes semantic and syntactic patterns to determine the emotion of a sentence. It begins with a predefined set of sentiment terms and their orientations, then explores syntactic or similar patterns within a vast corpus to identify sentiment tokens and their orientations. This method is specific to the situation and requires a substantial amount of labeled data for training. Within the corpus-based approach, there are two types of approaches: the Statistical Approach and the Semantic Approach [11, 10].

On the other hand, the Dictionary-Based Approach commences by manually collecting a set of opinion words to form a seed list. Next, dictionaries and thesauruses are consulted to find synonyms and antonyms of these words, which are then added to the seed list. This process continues until no new words are discovered. However, a drawback of this approach is the challenge of finding context or domain-oriented opinion words [11, 10].

There are two primary in ML approaches: Supervised ML and Lexicon-based unsupervised learning [11].

In Lexicon-based unsupervised strategies, knowledge bases, ontologies, databases, and lexicons containing specific and detailed information for SA are utilized. On the other hand, supervised learning methods are more widely used due to their high accuracy. These algorithms require training on a labeled dataset before they can be applied to real data. Features are extracted from the text data during the training process [10].

The ML technique uses syntactic or linguistic factors to classify sentiment and constitute a text classification problem. Therefore, the model associates features of the underlying record with class labels and predicts the label for unknown instances [10]. Commonly used ML algorithms include:

- **Decision Tree (DT) classifier**
- **Linear classifier** - e.g. Support Vector Machine (SVM) and Neural Networks such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BILSTM), Transformers
- **Rule-based classifier**
- **Probabilistic classifiers** - e.g. Multinomial Naive Bayes (NB), Maximum Entropy (ME) and Bayesian Network (BN)
- **K-Nearest Neighbor (KNN)**

The hybrid approach combines ML and lexicon-based techniques. It refers to the integration of both methods to enhance the accuracy and effectiveness of the systems. By leveraging the strengths of both approaches, the hybrid method can provide more comprehensive insights into the sentiment expressed in the text [11].

3.2.2. *Datasets*

There are diverse datasets used to train SA algorithms, some of which include: SemEval 2007 task 14 corpus, ISEAR, SentiWordNet, IMDB, SST and NRC.

The SemEval 2007 task 14 corpus [12] consists of emotional newspaper headlines from reputable sources like the New York Times, CNN, BBC, or Google News. This corpus allows training models for emotions such as anger, disgust, fear, joy, sadness, and surprise.

ISEAR [13] is a dataset comprising psychological data from a survey conducted in 1990. In this dataset, 3,000 subjects described situations where they felt various emotions, including joy, sadness, fear, anger, disgust, shame, or guilt.

The SentiWordNet dataset [14] is a publicly available dataset that automatically labels synsets in the WordNet dataset based on their degree of positivity, negativity, and neutrality.

The IMDB Movie Reviews dataset [15] is publicly available and contains around 50,000 movie reviews, labeled as either positive or negative.

SST [16] is a sentiment analysis dataset from Stanford, containing 10,662 movie sentences from Rotten Tomatoes, labeled with corresponding emotions, ranging from "very negative" to "very positive."

NRC [17] is a comprehensive emotion Lexicon dataset, containing more than 14,000 words annotated with sentiment (positive or negative) and various emotions, such as anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.

3.3. Mood Assessment Questionnaires

4. Architecture

5. Proposed solution and implementation

6. Conclusion

CRediT authorship contribution statement

: Conceptualization of this study, Methodology, Software.

References

[1] J.L.Z. Montenegro, C.A. da Costa, and R. da Rosa Righi. Survey of conversational agents in health. *Expert Systems with Applications*, 129:56–67, 2019.

[2] Shafquat Hussain, Omid Ameri Sianaki, and Nedal Ababneh. A Survey on Conversational Agents/Chatbots Classification and Design Techniques. In Leonard Barolli, Makoto Takizawa, Fatos Xhafa, and Tomoya Enokido, editors, *Web, Artificial Intelligence and Network Applications*, Advances in Intelligent Systems and Computing, pages 946–956, Cham, 2019. Springer International Publishing.

[3] Sinarwati Mohamad Suhaili, Naomie Salim, and Mohamad Nazim Jambli. Service chatbots: A systematic review. *Expert Systems with Applications*, 184:115461, December 2021.

[4] Conversational AI Platform | Superior Customer Experiences Start Here, December 2020.

[5] AWS Lex - Amazon Web Services.

[6] Documentação do Dialogflow.

[7] Pierre Lison. plison/opendial, April 2023. original-date: 2015-03-21T18:11:36Z.

[8] Botpress | a plataforma Generative AI para ChatGPT Chatbots.

[9] About ChatterBot — ChatterBot 1.0.8 documentation.

[10] Harpreet Kaur, Veenu Mangat, and Nidhi. A survey of sentiment analysis techniques. In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pages 921–925, February 2017.

[11] Mayur Wankhade, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780, October 2022.

[12] Carlo Strapparava and Rada Mihalcea. Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing*, SAC '08, pages 1556–1560, New York, NY, USA, March 2008. Association for Computing Machinery.

[13] Papers with Code - ISEAR Dataset.

[14] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010. European Language Resources Association (ELRA).

[15] IMDb Non-Commercial Datasets.

[16] Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank.

[17] NRC Emotion Lexicon.