

# Investigation of the use of Deep Learning and emotion detection for the improvement of Text-based medical conversational agent

Bing Yuan<sup>a</sup>, Haithem Afli<sup>b</sup>

<sup>a</sup>Munster Technological University of Ireland, Bishopstown Campus Cork, Ireland.T12 P928

<sup>b</sup>Munster Technological University of Ireland, Bishopstown Campus Cork, Ireland.T12 P928

## Abstract

In the last decade, there was an explosion in Natural Language Process applications led by the progress of Deep Learning architectures. Recently, Google researchers published BERT (Bidirectional Encoder Representations from Transformers) [1], a deep bidirectional language model based on the Transformer architecture, and advanced the state-of-the-art in many popular NLP tasks, including Machine Translation, Information Retrieval, and Sentiment Analysis. Emotion Detection is a sub-branch of Sentiment Analysis that focuses on detecting specific emotions. This project investigated the combination of emotion detection using Deep learning methods with rule-based Chatbot such as ELIZA. To improve the performance of conversation of agent focus on helping People with Obesity.

## Keywords

emotion detection, BERT(Bidirectional Encoder Representations from Transformer), NLP(Natural Language Process), DL(Deep Learning), PwO(Person with Obesity), STOP(STOP Obesity Platform), LSTM,

## 1. Introduction

Conversational interfaces provide a different (and possibly easier) way to do things, e.g., engage in a conversation to ask a query or provide a response instead of navigating menus, forms, and drop-down boxes a traditional graphical user interface. As a result, chatbots have emerged as a new type of interface to serve on the web. It has also become possible in the e-health medical area due to the significant advances in artificial intelligence and speech and language technologies.

This project aims to design an intelligent conversational agent Smart-ELIZA to enable the STOP platform interactive communication with People with Obesity(PwO). The second main feature is the integration of the emotion detection model. When the user's emotion is negative, the Smart-ELIZA Chatbot will respond differently from what ELIZA system's original reply; for example, the new reply can be: "I am sorry to hear that, how can I help you?".

This project use the new State-art-of Deep Learning model: the BERT model and a fine-grained training dataset such as GoEmotions containing 58000 data and twenty eight emotions.

---

CERC 2021: Collaborative European Research Conference, September 09–10, 2021, Cork, Ireland

✉ hbingyuan@gmail.com (B. Yuan); Haithem.Afli@cit.ie (H. Afli)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

## 2. Contribution

The primary contributions of this study are summarized below:

### From the research aspect

1. This project is feasibility research regarding adding the emotion detection function to the ELIZA Chatbot with emotion detection function using Deep Learning technologies. ELIZA Chatbot is best known for acting in the manner of a psychotherapist, and the therapist "reflects" on questions by turning the questions back at the patient. By adding emotion detection within ELIZA, which empowers the ELIZA's function.
2. For the emotion detection model development: This project demonstrates the effect of using the GoEmotions dataset to train the Bert pre-trained model for emotion detection. To my knowledge, this is the first study to use Bert Model with the GoEmotions dataset combination to detect twenty eight different emotions.

### From industry application aspect

1. The ELIZA Chatbot offer to the PwO is a novel approach that offers the ability to initiate a conversation any time and anywhere through a smartphone or laptop. It shows friendly companionship and also acts as an assistant to the PwO. By understanding emotions in textual dialogue to provide emotionally aware responses to e-health users enhances the healthcare-related platforms.
2. The first-hand chat information through the conversation with PwO provides insights into the common queries, concerns, and general psychological state of PwO. The captured chat data will be fused with further sensor data and knowledge resource database within the STOP platform and passed to an analysis pipelines service to provide insight for healthcare professionals.

## 3. Background

### 3.1. Introduction to Chatbot

Conversational agents are "systems that mimic human conversation" using communication channels such as speech, text, and facial expressions and gestures [2]. Conversational agents roughly consist of three main categories: Chatbots without embodiment, virtually embodied avatars, and physically embodied robots.

A Chatbot is a contraction of chat and robot, and it is a computer program that conducts conversations by using dialogue or text [3]. A Chatbot is an instant messaging application, an automated chat system; it is a web-based application that is easy to use and no need to install.

Chatbot application usage categories include commerce (e-commerce via chat), education, entertainment, finance, health, news, and productivity.[4]. With artificial intelligence (AI) technologies, including deep learning, machine learning (ML) algorithms, and natural language processing, Chatbot can proactively assist and predict the user or customer behaviors, requests, and needs that bring many domains.

### 3.1.1. Chatbot for HealthCare Overview

Chatbots have been used in many health-related practices, such as sports activities, psychological health, drug compliance, and depression and anxiety, as well as expressing sympathy and compassion [5]. In addition, some of the e-health Chatbot could accept consultation in the form of pictures, images, and even voice.

Semantic analysis technology allows Chatbots to autonomously converse with the patient to identify and respond to their needs. Thus Chatbots have become preferred by the HealthCare domain for communicating instantaneously with users in terms of efficiency and benefits. In addition, research shows that the deployment of conversational agents in service encounters is growing exponentially in many other sectors, including healthcare.[2]

### 3.2. Introduction to ELIZA

ELIZA was programmed in 1966 by Joseph Weizenbaum at MIT, and it was the first public known Chatbot. ELIZA was an early test case for the Turing Test. ELIZA does not understand the conversation; ELIZA uses pattern matching and returns the response by the selection scheme based on templates. ELIZA can only accept text format input; the input sentences are analyzed based on decomposition rules triggered by keywords appearing in the input text. Responses are generated by reassembly rules associated with selected decomposition rules.[6] ELIZA only contains 200 keywords and rules.[7] However, many users believe they were talking to a natural person when ELIZA was invented.

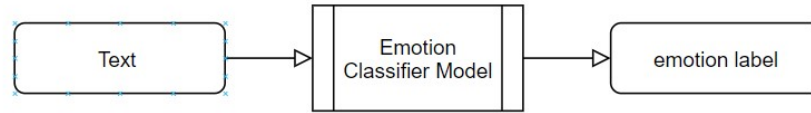
### 3.3. Introduction to Emotion Detection

Based on Merriam-Webster definition: Emotion which is a state of feeling, is the affective aspect of consciousness: Feeling is a conscious mental reaction (such as anger or fear) subjectively experienced as a strong feeling usually directed toward a specific object and typically accompanied by physiological and behavioral changes in the body. The prominent work in understanding and categorizing emotions include Ekman's six class categorization and Plutchik's "Wheel of Emotion," which suggested eight primary bipolar emotions.

Emotion detection is the state or fact of emotion being detected. There are several ways to recognize human emotion, such as source from the text message, voice tones, facial expressions, gestures, etc. Myriam Munezero's research shows that current emotions detection is from text has focused on capturing emotion words based on three emotion models, that is, categories of basic emotions, emotion dimensions, or cognitive appraisal categories. Suppose emotion detection application can recognize emotions such as joy, sadness, anger with possible high accuracy. In that case, it will bring a wide range of impact for many emotion-based services, such as medical applications and public emotion sense services.

### 3.4. Introduction to Sentiment Analysis

Sentiments are "socially constructed patterns of sensations, expressive gestures, and cultural meanings organized around a relationship to a social object, usually another person or group such as a family." Examples of sentiments include romantic love or nostalgic feeling. Both emotions and sentiments refer to "experiences that result from the combined influences of the biological, the cognitive, and the social". However, sentiments are differentiated from emotions by the duration in which they



**Figure 1:** A high-level overview of an emotion detection classifier shows input and output

are experienced. Although emotions are brief episodes of brain, autonomic, and behavioral changes, sentiments have been found to form and be held for a more extended period. Sentiments, in addition, are more stable and dispositional than emotions.

Sentiment Analysis is a natural language processing (NLP) task. It has been defined as analyzing the sentiment according to the opinion expressed about a given subject is positive, negative, (but sometimes also neutral or vague). It aims to classify the text-based content, but sometimes it also includes audio and video. Sentiment analysis is a suitcase research problem that requires tackling many natural language processing (NLP) tasks. Emotion Detection is a branch of sentiment analysis within the NLP research domain.

### 3.5. Introduction to Natural Language Processing

Natural Language Processing (NLP) started in 1950, is an intersection of artificial intelligence and linguistics domain. NLP is a theoretically motivated range of computational techniques for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis to achieve human-like language processing for a range of tasks or applications (Liddy, 2003). NLP is a large area of research and application, and the NLP tasks include automatic text summarization, machine translation, information search, question answering, and sentiment analysis, etc. Because NLP helps machines to "read" text by simulating the human ability to understand language; plus the availability of cloud computing, big data technologies, and deep learning algorithms, the NLP applications has been widely used in many industries such as linguistic analysis, speech recognition, and notes analyze in electronic health records, etc.

## 4. Using Deep Learning techniques to develop emotion detection models

Emotion detection and Sentiment Analysis attract diverse attention, including researchers from machine learning, natural language processing, and computational linguistics. The Emotion Detection classification within this project is a reference to a form of text classification in which text content will be classified into the pre-defined emotion classes. This section will present a state-of-the-art overview of various deep learning technologies and relevant datasets for text-based emotion detection.

### 4.1. Literature Review

From the dataset aspect, many pieces of research are done by using human physical reaction signals (Such as respiration signals, heartbeat and breathing signals) work together with a deep learn-

**Table 1**

Deep Learning Model and Datasets Summary for emotion detection

| Deep Learning Model                                                      | Dataset                                                              | Detection Target | Author                               |
|--------------------------------------------------------------------------|----------------------------------------------------------------------|------------------|--------------------------------------|
| LSTM+CNN                                                                 | Public Twitter datasets of SemEval Task-1, Affect in Tweets dataset. | Text             | Malak Abdullah [8]                   |
| LSTM with GloVe and SSWE as embedding layer                              | The Twitter dataset with 17.62 million tweet conversational pairs    | Text             | Ankush Chatterjee from Microsoft [9] |
| LSTM                                                                     | electroencephalography (EEG) signals data from DEAP Dataset          | Text             | Alhagry S. [10]                      |
| multiple-fusion-layer based ensemble classifier of a stacked autoencoder | multimodal physiological signals from DEAP Dataset                   | Text             | Zhong Yin [11]                       |

ing model. Alhagry S. who use electroencephalography (EEG) signals data together with the LSTM model for emotion prediction and classified category for emotions. They used the Dataset for Emotion Analysis using Physiological signals (DEAP) data set in the study, and a classification process was performed for valence, arousal, and liking classes. Zhong Yin uses multimodal physiological signals and a multiple-fusion-layer-based ensemble classifier of stacked autoencoder (MESAE) for recognizing emotions. He also uses DEAP multimodal database for the study.

From the Deep learning model aspect, there are quite several pieces of research done by using the Long-short term memory (LSTM) model. LSTM is a particular type of Recurrent Neural Network(RNN) with DL family, its capability well knows of modeling the semantic relationships between words in the text, or combine of LSTM with other neuro network method. Such as Malak Abdullah use Convolutional Neural Networks CNN combine with LSTM on public Twitter datasets of SemEval Task-1, Affect in Tweets dataset. Ankush Chatterjee from Microsoft uses LSTM with GloVe and SSWE as an embedding layer on an extensive dataset: the Twitter dataset with 17.62 million tweet conversational pairs. The table 1 shows the comparison details :

1.

With the development of neural language models such as word vectors , paragraph vectors , and GloVe , the transfer learning (pre-training and fine-tuning) revolution started in NLP. Recently, Google researchers published BERT (Bidirectional Encoder Representations from Transformers) , a deep bidirectional language model based on the Transformer architecture [1] and advanced the state-of-the-art in many popular NLP tasks.

There have been many Sentiment Analysis pieces of research that used the BERT model. For example, Quoc Thai Nguyen [12] used the Fine-Tuning BERT together with LSTM/TextCNN/RcNN got an average 90% F1 score for negative and positive classification of Vietnamese Reviews. Manish Munikar [13] uses the BERT to classify five-category sentiments: very negative, negative, neutral, positive, and very positive. Another fine-grained BERTBase and BERTLarge model was performed on sentiment classification with the Stanford Sentiment Treebank dataset, which got 94% accuracy. Yun Xiang Zhang [14] also combine BERT with Deep neural networks DNN to propose a classification model applied to aspect-level sentiments classification.

The detection or recognition of emotions means the extraction of finer-grained sentiments. Text-based emotion recognition is a sub-branch of emotion recognition (ER) that focuses on extracting fine-grained emotions from texts. [15] Although BERT has been widely used in sentence-level sentiment classification, as shown above, BERT has not found widespread application in sentence-level emotion classification. However, only several research pieces have been done by using BERT for emo-

**Table 2**

Deep Learning Model and Datasets Summary for Sentiment Analysis

| Deep Learning Model                                         | Dataset                                                                    | Emotions Categories                                                | Author                     |
|-------------------------------------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------|----------------------------|
| CNN, biLSTM, BERT                                           | Mass Shooting dataset (MS), the Terrorism dataset (TR)                     | 7 emotions: fear, anger, joy, sad, contempt, disgust, and surprise | Jonathas G.D. Harb         |
| BERT model+ bi-LSTM classifier                              | ISEAR dataset with 7666 sentences                                          | 7 emotions: joy, anger, sadness, shame, guilt, surprise, and fear. | Acheampong Francisca Adoma |
| BERT, RoBERTa, DistilBERT, and XLNet                        | ISEAR dataset with 7666 sentences                                          | 7 emotions: joy, anger, sadness, shame, guilt, surprise, and fear. | Acheampong Francisca Adoma |
| BERT Embeddings+GloVe word embeddings+BiLSTM neural network | Set of psycholinguistic features (e.g. from AffectiveTweets Weka-package.) | Four categories: Happy, Sad, Angry, and Other                      | Hani Al-Omari              |

tion detection; this is a very new research domain.

Jonathas G.D. Harb [16] design a framework using CNN, biLSTM, and BERT to perform the sentiment classification in terms of Ekman's seven basic emotions. The platform is to analyze the emotional reactions to mass violent events on Twitter. The research was done by applying three different DL technologies into different datasets Mass Shooting dataset (MS), the Terrorism dataset (TR), or MS+TR. CNN(MS+TR) produced the best results for three emotions and BERT(MS) for two emotions. biLSTM(MS) models produced statistically superior f-measure results for Anger only.

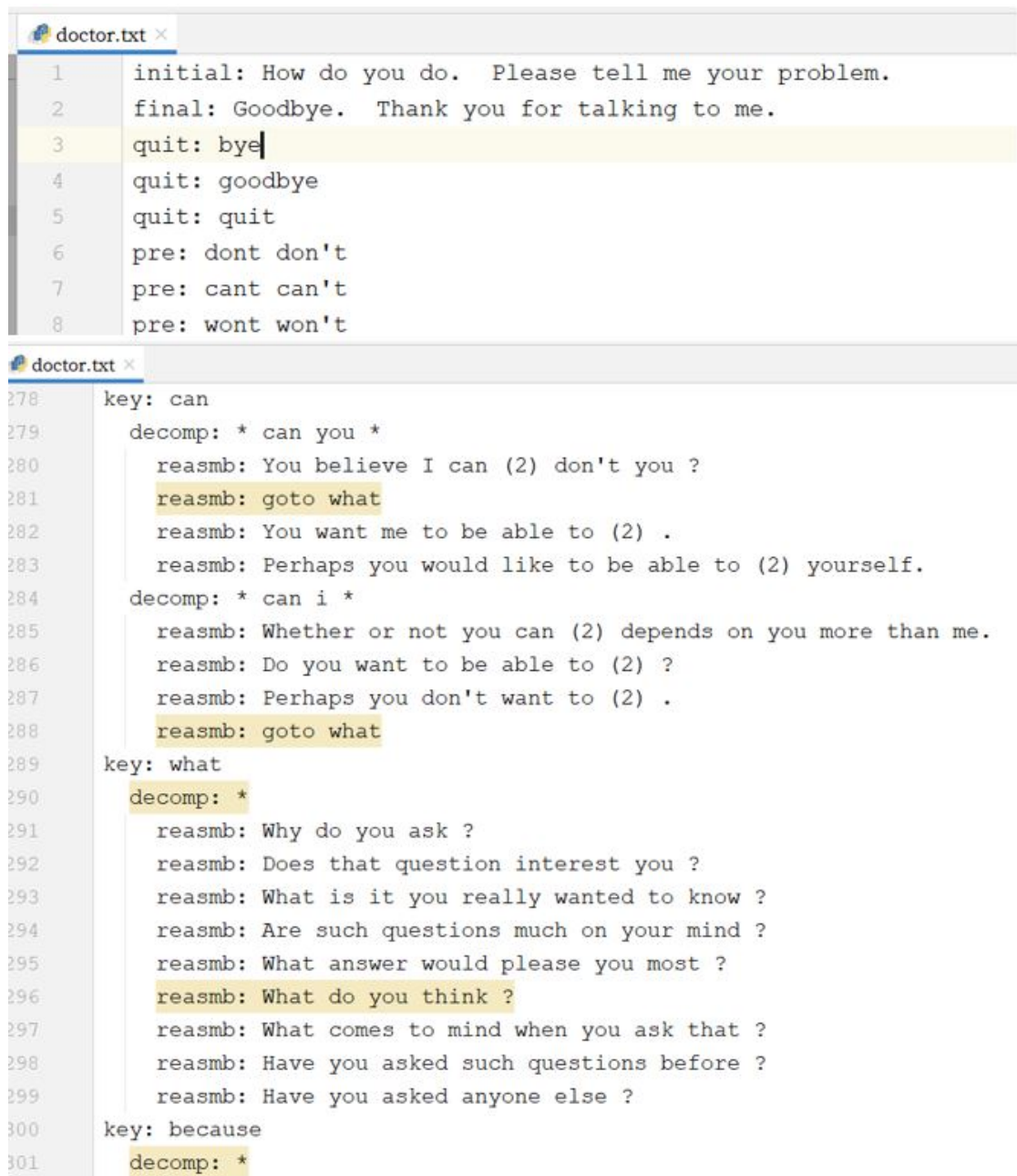
Acheampong Francisca Adoma [17] used the BERT model and bi-LSTM classifier to classify seven emotions from the text-based ISEAR dataset, which contains 7666 sentences. It attains an overall accuracy of 72.64%. That research shows that the BERT model performs an average much better performance than the machine learning-based classifier, including SVM, Naïve Bayer, and Randon Forest. They performed additional research to analyze the efficacy of BERT, RoBERTa, DistilBERT, and XLNet pre-trained transformer models in recognizing emotions from texts [15]; the recorded model accuracies in decreasing order are 0.7431, 0.7299, 0.7009, 0.6693 for RoBERTa, XLNet, BERT, and DistilBERT, respectively.

Hani Al-Omari [18] then performed future research based on the above research with the combination of GloVe word embeddings, BERT Embeddings, and a set of psycholinguistic features (e.g., from AffectiveTweets Weka-package). The research combined a fully connected neural network architecture and BiLSTM neural, which achieve the result with F1-Score 0.748 for four emotions. The table 2 summerize the BERT model and the dataset.

## 5. Design

In order to design of the Smart-ELIZA Chatbot, this part reviews one of implemented ELIZA python version source code, summarizes how the ELIZA system was designed in detail with workflow. Also gives the overview of how the Smart-ELIZA will design based on the existing ELIZA system and narrows down the emotion detection working process within Smart-ELIZA then presents the detailed workflow design of Smart-ELIZA Chatbot. The final two parts are about the approaches for emotion modeling and the process for emotion detection.





```

doctor.txt x
1  initial: How do you do. Please tell me your problem.
2  final: Goodbye. Thank you for talking to me.
3  quit: bye
4  quit: goodbye
5  quit: quit
6  pre: dont don't
7  pre: cant can't
8  pre: wont won't

doctor.txt x
278 key: can
279   decomp: * can you *
280       reasmb: You believe I can (2) don't you ?
281       reasmb: goto what
282       reasmb: You want me to be able to (2) .
283       reasmb: Perhaps you would like to be able to (2) yourself.
284   decomp: * can i *
285       reasmb: Whether or not you can (2) depends on you more than me.
286       reasmb: Do you want to be able to (2) ?
287       reasmb: Perhaps you don't want to (2) .
288       reasmb: goto what
289   key: what
290       decomp: *
291       reasmb: Why do you ask ?
292       reasmb: Does that question interest you ?
293       reasmb: What is it you really wanted to know ?
294       reasmb: Are such questions much on your mind ?
295       reasmb: What answer would please you most ?
296       reasmb: What do you think ?
297       reasmb: What comes to mind when you ask that ?
298       reasmb: Have you asked such questions before ?
299       reasmb: Have you asked anyone else ?
300   key: because
301       decomp: *

```

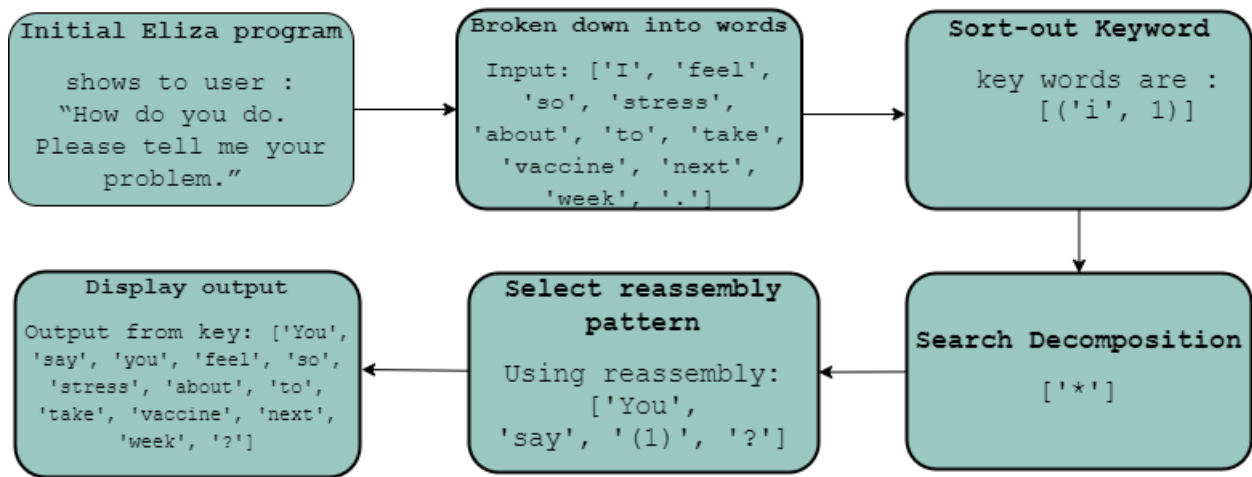
**Figure 2:** The screenshot of doctor.tx in ELIZA SourceCode

## 5.1. ELIZA Design Detail Review

With reference to Charles Hayden Source Code Readme document.[? ], and a step-by-step debug of the Source Code, below is the summary of how ELIZA was designed at the code level.

A script file controls all the behavior of ELIZA with text format, which is only 359 lines. In this project, the file is named "doctor.txt." Every line of the script is prefaced by a tag that tells what list it is part of, the tags such as: initial, pre, post, key, etc. The script is used to construct the pre and post substitution lists, the keyword lists, and the decomposition and reassembly patterns, and a synonym matching facility.

The screenshot of the "doctor.txt" file shows in figure 2



**Figure 3:** ELIZA Chatbot working process

### 5.1.1. ELIZA Chatbot Working Process

The ELIZA Chatbot working process consists of the following six main steps as show in figure 3

## 5.2. Smart-ELIZA Design Overview

For Smart-ELIZA, the design analyzes the whole input sentence's emotion rather than word by word. So, the deep learning model will be engaged as long as the PwO returns a sentence to the Smart-ELIZA system. Reference the ELIZA original design explained in the above section. The Smart-ELIZA's emotion detection process should be the extra step between the "Initial ELIZA system" and "Broken down reply into words" steps.

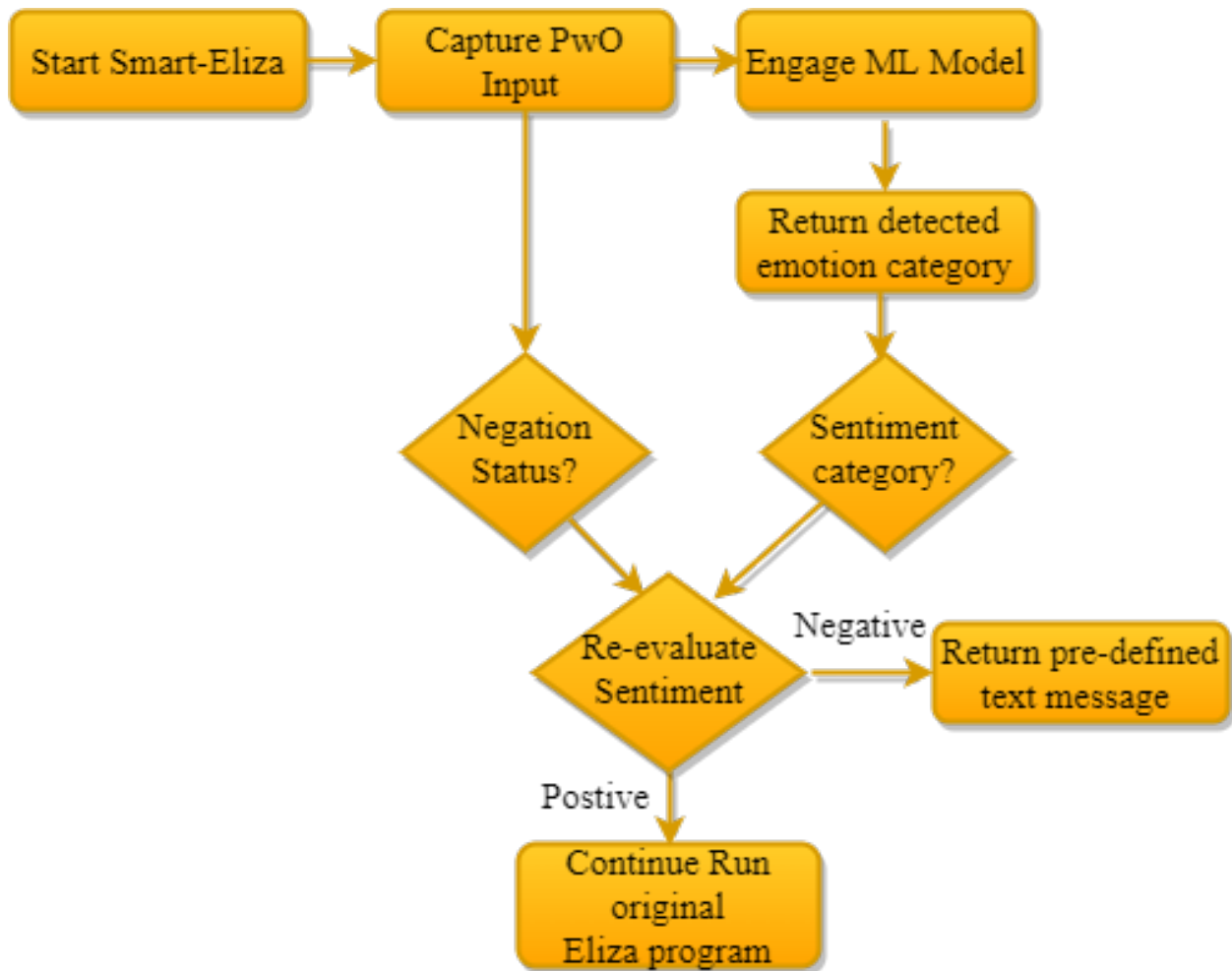
To perform the emotion detection for the user's input sentence, the DL model plays an essential role by analyzing the PwO's output and identifying the emotion category. Considering the sentence might contain negation (such as I am not feeling well these days during the lockdown.), it is essential to re-evaluate the sentiment in this situation. If the final emotion belongs to the negative list, then Smart-ELIZA will return a pre-defined text message to PwO; otherwise, continue the original ELIZA rest of the system's function. The workflow details show as 4

### 5.3. Process for emotion detection

Regarding text-based emotion detection, generally, it contains five steps to perform the emotion detection. The steps are Data Source selection, Text Pre-processing, Emotion Detection Model Development, Process text using model, and Model evaluation. See 5

1. **Data Source Selection:** The first step is to select the data source to be analyzed or used as a training dataset. The data source could be any text format of data, for example, tweeter tweets, amazon customer reviews, email messages, or service request tickets.
2. **Text Pre-processing:** Text pre-processing step is trying to sanitize the text/source data and transform the text with a specific format to train the model file. Depending on what ML model is selected, there are a few sub-processes that will be included. For example: stop words detection, tokenization, part-of-speech (pos) tagging, parsing (syntactic analysis), stemming, and lemmatization.



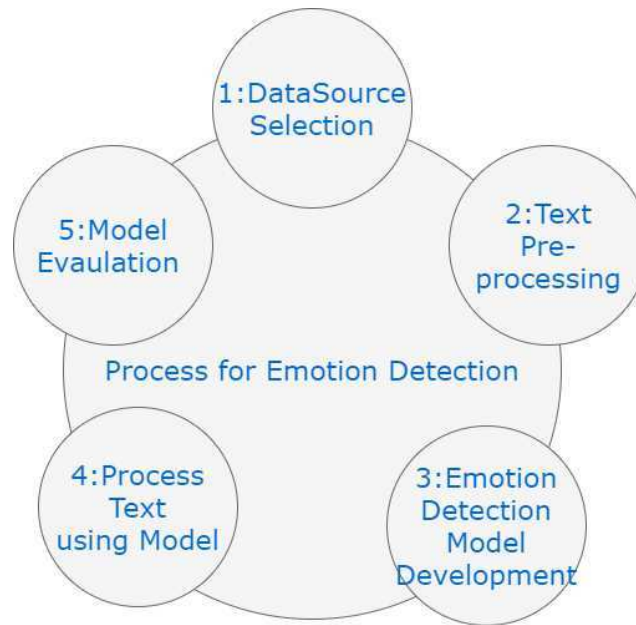


**Figure 4:** Smart-ELIZA Chatbot emotion detection working process

3. Emotion Detection Model Development: The emotion Detection Model is built using a method and works with several processes in which the emotions are detected in text source . They are corpus-based methods (e.g. lexical affinity) machine learning-based techniques (e.g. statistical methods), knowledge-based techniques (e.g. hand-crafted models, keyword spotting) and hybrid techniques. This project use state of the art deep learning model to implement emotion detection.
4. Process text using the model: Once the model file is generated, this step is to apply the text source with the model file and give the emotion detection result; the output emotion will be one of the emotion categories based on the source/training dataset's categories.
5. Model evaluation Three main evaluation metrics are mainly used in the research: Kappa Coefficient, Precision, and Recall, F-Score.

## 6. BERT Model implementation

BERT (Devlin et al., 2019) [1] Bidirectional Encoder Representations from Transformers is a method of pre-training language representations; it has been trained as a general-purpose "language understanding" model on a large text corpus (like Wikipedia). BERT model can be used for downstream NLP



**Figure 5:** Five steps to perform the emotion detection

tasks (such as question answering, emotion detection). BERT is the first unsupervised, deeply bidirectional system for pre-training NLP. BERT provides a practitioner with a model with a significant amount of general knowledge of the language.

Google releasing several pre-trained models from the paper are available to download. Google release BERT-Base and BERT-Large when the paper first published; in March 2020, Google release 24 smaller BERT models.

Google proposed two main BERT models: BERT-Base contains 12 encoder blocks, 12 head attention, and 110 million parameters. BERT-large contains 24 encoder blocks, 16 head attention, and 340 million parameters.

## 6.1. Dataset selection

To get better accuracy of the emotion detection, the DL model implement firstly focus also on to find out the best suit emotion datasets for the deep learning BERT model.

Google "GoEmotions have performed another research: A Dataset of Fine-Grained Emotions," Dorottya Demszky introduce "the largest manually annotated dataset of 58k English Reddit comments, labeled for 27 emotion categories or Neutral.", they also perform the "high quality of the annotations via Principal Preserved Component Analysis". The GoEmotions have also been tested on the BERTBase model and got the 46% across all the proposed taxonomy. This test shows the generalizability of the data across domains and taxonomies via transfer learning experiments.

### 6.1.1. GoEmotions dataset overview

GoEmotion GitHub website contains the entire dataset and Source Code that has been used for the research performed by Google. The complete datasets which are shows three CSV format files; the first few columns are text, id, author, subreddit, link\_id, parent\_id, created\_utc, rater\_id,

example\_very\_unclear. The rest of the columns are emotions, the value is either 0 or 1. By mapping emotions with sentiment, all the emotions are shown below:

- "positive": ["amusement", "excitement", "joy", "love", "desire", "optimism", "caring", "pride", "admiration", "gratitude", "relief", "approval"]
- "negative": ["fear", "nervousness", "remorse", "embarrassment", "disappointment", "sadness", "grief", "disgust", "anger", "annoyance", "disapproval"]
- "ambiguous": ["realization", "surprise", "curiosity", "confusion"]

## 6.2. Generate BERT Model

Feature extraction is one of the approaches to use BERT; this method uses BERT as a feature extraction model. The architecture of the BERT model is preserved, and its outputs are feature input vectors for subsequent classification models to solve the given problem.

This project uses BERT-Base, the pre-trained BERT model, and runs a further train on GoEmotions datasets without substantial task-specific architecture modifications.

Google's research mentioned that GoEmotions had been tested on BERTBase. However, they did not detail how the new Bert Model generated by the GoEmotions dataset and GitHub does not include the new BERT model for download.

Future research has been done regarding how to train BERTBase using GoEmotions to generate a new BERT model. The research is based on reading Google published paper [19] and debug the source code step by step. This part of the research is re-validate and investigating the procedure that has been performed in the Google GoEmotions research.

### 6.2.1. Detail Experiment Procedure

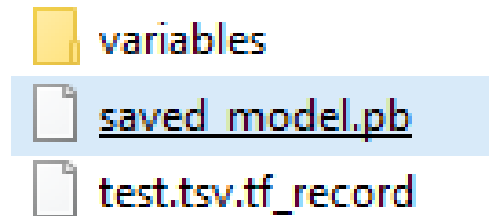
Below is the summarized the whole procedure based on the experiments locally:

1. Clone the GoEmotions code from [19] locally.
2. Clone the Bert code [19] from GitHub locally and put it under the GoEmotions folder.
3. Download the BERTBase model file from google storage, the folder contains three files:
  - A TensorFlow checkpoint (bert\_model.ckpt) containing the pre-trained weights
  - A vocab file (vocab.txt) to map WordPiece to word id.
  - A config file (bert\_config.json) specifies the hyperparameters of the model.
4. Install local running environment; the python libraries are listed in their requirements.txt file. There are more than 30 libraries. The TensorFlow version is straightly mentioned within the file.
5. bert\_classifier.py is the main file to perform fine-tuning on the GoEmotions dataset on top of BERTBase,
  - Many parameters need to configure to make it work. To centralize the configuration, to enhance the extra code was added into configuration\_constant.py, which specifies the variable: DATA\_DIR, SENTIMENT\_FILE, EMOTION\_FILE, etc.

**Table 3**

BERTBase Model runs on GoEmotions Dataset

|                           |                            |
|---------------------------|----------------------------|
| eval_accuracy = 0.4354155 | eval_accuracy = 0.43209878 |
| eval_loss = 2.545235      | eval_loss = 2.5506651      |
| global_step = 10852       | global_step = 10852        |
| loss = 2.5417128          | loss = 2.5470078           |

**Figure 6:** New BERT Model file folder structure

ant &gt; output\_dir &gt; 1618244107 &gt; variables

| Name                          | Type                    | Size       | Date modified       |
|-------------------------------|-------------------------|------------|---------------------|
| variables.data-00000-of-00001 | DATA-00000-OF-00001 ... | 427,750 KB | 2021-04-12 05:15 PM |
| vocab.txt                     | Text Document           | 227 KB     | 2020-02-20 06:57 PM |
| variables.index               | INDEX File              | 9 KB       | 2021-04-12 05:15 PM |
| config.json                   | JSON File               | 1 KB       | 2021-04-11 10:32 PM |

**Figure 7:** New BERT Model file variables folder

- Parameter “do\_export” within the “flags.DEFINE\_bool” need to set as True to save the model file locally.
  - Other “train\_batch\_size,” learning\_rate,” num\_train\_epochs,” and the rest of the parameters are kept the same as the paper’s suggestion.
6. The finally working parameters are showing as below:
- ```
python bert_classifier.py -vocab_file=C:\BERT-Base-model\vocab.txt
-bert_config_file=C:\BERT-Base-model\bert_config.json -output_dir=C:\output_dir
```

### 6.2.2. The New BERT Model Running Result

The training model steps have been running twice; on average, it took more than 20 hours to complete the training on a local DELL laptop (CORE i7 8th Gen without GPU) with an accuracy of about 43.4%. This accuracy is very similar to Google Research paper’s result: 46%. The table 3 shows the running accuracy result.

The Saved Model contains a complete TensorFlow program, including trained parameters (i.e, tf.Variables) and computation.

The model files generated are under output\_dir with two files and one variables folder from the implementation. The screenshot of all files are showing in FIGURE 7 and 6.

Major files' function are :

- The saved\_model.pb file stores the actual TensorFlow program and a set of named signatures.
- The index file is a string-string immutable table.
- variables.data-00000-of-00001 is TensorBundle collection which stores the actual values of all the variables is about 427MB.

### 6.3. Apply New BERT model and Negation into ELIZA system

From the previous step, the New Bert model is generated. This section focuses on applying this new Bert model to make emotion detection for ELIZA chat input. A new "emotion\_prediction\_class" was created delicate to predict the chat context.

Several experiments have been performed to figure out how to use the above TensorFlow model file.

The first experiment uses "-do\_predict=true" to run the model to predict classifier base on the BERT website. However, this parameter will engage rerunning training and evaluation once again, which is about 20 hours long, which would not apply to the Chatbot system's requirement.

Second experiments is to investigation from GoEmotions bert\_classifier.py and try to how the Model file is saved.

The final working prediction detailed experiment steps are:

1. Load Graph (tf.graph) into TensorFlow session
2. Load the exported model to a TensorFlow session with tag is ['serve'] so it is available in the current context.
3. Pre-process the Chat text by adding extra "data\_type == 'chat\_prediction'" into GoEmotions "DataProcessor" class so it will create examples base on chat content.
4. Perform from single example to tf.train.Features format with Feature Tensors structure required for next step.
5. Adjust the tf.train.Features into tf.train.Example and feed into prediction.
6. Figure out the prediction session run "feed\_dic" structure from GoEmotions TensorFlow structure settings.

GoEmotions exist class like "DataProcessor" for this prediction; the extra step is to transfer the input text into CSV format and save it into local disk. The negation been applied to this sentiment result by reading the negative and positive list been defined from GoEmotions (see "GoEmotions dataset overview").

## 7. Conclusions and Future Work

### 7.1. Evaluate BERT version of Smart-ELIZA function

For the the function evaluation,when the user inputs every sentence, the system will perform the emotion negation detection. Once the final sentiment is negative, the return will be using the pre-defined sentence "I am sorry to hear that, how can I help you?". The system runs shows as in FIGURE ??

```
(base) C:\0.Data_Science_and_Data_analytic\0.CIT\AI_Master\4.FourthTerm\Code\Eliza\eliza_ml_goemotions_bert>python -W ignore eliza_with_ml_bert.py
How do you do. Please tell me your problem.
> I am doing great.
2021-05-08 17:07:21.285198: I tensorflow/core/platform/cpu_feature_guard.cc:142] Your CPU supports instructions that this TensorFlow binary was not compiled to use: AVX AVX2
*****DEBUG***** emotion detection result is : admiration
*****DEBUG***** negation_status is : False
*****DEBUG***** User is positive sentiment
Is it because you are doing great that you came to me ?
> I am worry about the exam.
*****DEBUG***** emotion detection result is : approval
*****DEBUG***** negation_status is : False
*****DEBUG***** User is positive sentiment
How long have you been worry about the exam ?
> I am not feeling well about it.
*****DEBUG***** emotion detection result is : neutral
*****DEBUG***** negation_status is : True
*****DEBUG***** User is positive sentiment
Do you believe it is normal to be not feeling well about it ?
> I am so worried about exam.
*****DEBUG***** emotion detection result is : sadness
*****DEBUG***** negation_status is : False
*****DEBUG***** User is negative sentiment
I am sorry to hear that,how can I help you?
> I am worry about my exam
*****DEBUG***** emotion detection result is : optimism
*****DEBUG***** negation_status is : False
*****DEBUG***** User is positive sentiment
Your exam ?
> I am worried about my exam
*****DEBUG***** emotion detection result is : sadness
*****DEBUG***** negation_status is : False
*****DEBUG***** User is negative sentiment
I am sorry to hear that,how can I help you?
```

**Figure 8:** Smart-ELIZA function test using new BERT Model

Another finding is using a different way of expressing "worry": "I am worried about my exam," "I am so worried about my exam," "I am worried about my exam," it returns emotion detection with "approval," "sadness" and "optimism." The above test result leads the space to improve the model file. However, the answer from the original ELIZA system is still acceptable to the user. For example, the answers are "How long have you been worry about the exam?," "You exam?".

## 8. Conclusion

The project achieved a successful result with the state-of-art BERT model successfully detected ELIZA Chatbot input and returned a pretty high accuracy score considering this project using much more advanced emotion categories in the training datasets. In the meantime, the BERT version Smart\_ELIZA returns the answer simultaneously just after the user input the sentence. The emotion detection steps works as seamless with ELIZA system. This also indicates BERT Model can be used in the STOP system for real users.

## 9. Future Work for this project

From the technical point of view, the future work that could be done is to improve BERT model accuracy; for example, another experiment could be done by combining the BERT model with the biLSTM model and testing the accuracy because these two models' combination has not been used in a few other sentiment analyses research so far. However, no relevant research has been done for



emotion detection.

Another potential work could be using this project's emotion detection and the user's chat history to predict the new message input and generate a personalized message to PwO.

## 10. Acknowledgments

This research was conducted with the financial support of the Horizon 2020 project STOP Obesity Platform under Grant Agreement No. 823978 and at the ADAPT SFI Research Centre at Cork Institute Of Technology. The ADAPT SFI Centre for Digital Media Technology is funded by Science Foundation Ireland through the SFI Research Centres Programme and is co-funded under the European Regional Development Fund (ERDF) through Grant 13/RC/2106.

The project was partially supported by the Horizon 2020 project ITFLOWS under Grant Agreement No. 882986 and Dell-EMC (Ireland).

## References

- [1] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).
- [2] M. M. Van Pinxteren, M. Pluymaekers, J. G. Lemmink, Human-like communication in conversational agents: a literature review and research agenda, *Journal of Service Management* (2020).
- [3] A. Deshpande, A. Shahane, D. Gadre, M. Deshpande, P. M. Joshi, A survey of various chatbot implementation techniques, 2017.
- [4] J. BARON, 2017 messenger bot landscape, a public spreadsheet gathering 1000+ messenger bots (2017).
- [5] B. Liu, S. S. Sundar, Should machines express sympathy and empathy experiments with a health advice chatbot, *Cyberpsychology, Behavior, and Social Networking* 21 (2018) 625–636.
- [6] J. Weizenbaum, Eliza—a computer program for the study of natural language communication between man and machine, *Communications of the ACM* 9 (1966) 36–45.
- [7] B. Heller, M. Proctor, D. Mah, L. Jewell, B. Cheung, Freudbot: An investigation of chatbot technology in distance education, in: *EdMedia+ Innovate Learning*, Association for the Advancement of Computing in Education (AACE), 2005, pp. 3913–3918.
- [8] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781 (2013).
- [9] A. Chatterjee, U. Gupta, M. K. Chinnakotla, R. Srikanth, M. Galley, P. Agrawal, Understanding emotions in text using deep learning and big data, *Computers in Human Behavior* 93 (2019) 309–317.
- [10] S. Alhagry, A. A. Fahmy, R. A. El-Khoribi, Emotion recognition based on eeg using lstm recurrent neural network, *Emotion* 8 (2017) 355–358.
- [11] Z. Yin, M. Zhao, Y. Wang, J. Yang, J. Zhang, Recognition of emotions using multimodal physiological signals and an ensemble deep learning model, *Computer methods and programs in biomedicine* 140 (2017) 93–110.
- [12] Q. T. Nguyen, T. L. Nguyen, N. H. Luong, Q. H. Ngo, Fine-tuning bert for sentiment analysis of vietnamese reviews, arXiv preprint arXiv:2011.10426 (2020).
- [13] M. Munikar, S. Shakya, A. Shrestha, Fine-grained sentiment classification using bert, in: 2019

- Artificial Intelligence for Transforming Business and Society (AITB), volume 1, IEEE, 2019, pp. 1–5.
- [14] Y. Zhang, Z. Rao, Deep neural networks with pre-train model bert for aspect-level sentiments classification, in: 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), IEEE, 2020, pp. 923–927.
  - [15] A. F. Adoma, N.-M. Henry, W. Chen, Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition, in: 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), IEEE, 2020, pp. 117–121.
  - [16] J. G. Harb, R. Ebeling, K. Becker, A framework to analyze the emotional reactions to mass violent events on twitter and influential factors, *Information Processing & Management* 57 (2020) 102372.
  - [17] A. F. Adoma, N.-M. Henry, W. Chen, N. R. Andre, Recognizing emotions from texts using a bert-based approach, in: 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), IEEE, 2020, pp. 62–66.
  - [18] H. Al-Omari, M. A. Abdullah, S. Shaikh, Emotet2: Emotion detection in english textual dialogue using bert and bilstm models, in: 2020 11th International Conference on Information and Communication Systems (ICICS), IEEE, 2020, pp. 226–232.
  - [19] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, S. Ravi, Goemotions: A dataset of fine-grained emotions, *arXiv preprint arXiv:2005.00547* (2020).