Understanding Mental Health concepts derived from Cognitive Behavioural Therapy

CSD 422: Deep Learning
Final Project

Aditya Jain, Manav Prabhakar, Salil Saxena

May 17, 2021

Abstract

Mental Health comprises of a person's emotional, psychological and social well-being. It particularly relates to the way one handles stress, how one relates to others and makes choices and how one feels, thinks and reacts. The pandemic and consequent lockdowns have further deteriorated mental health conditions in individuals and this has been the source of motivation for doing this project. Cognitive Behavioural Therapy (CBT) is a type of treatment that helps people learn how to identify and change destructive or disturbing thought patterns that have a negative influence on behaviour and emotions. In this work we have tried to predict three classes, namely: Thinking errors, Emotions and Situations on an unseen mental health corpus, and compared the predictions produced by an LSTM model and an SVM.

Contents

1 Introduction

CBT focuses on changing the automatic negative thoughts that can contribute to and worsen emotional difficulties, like depression, anxiety, etc. These thoughts are are identified, challenged and replaced with more objective, realistic thoughts. The primary goal is to challenge unhelpful cognitive distortions like beliefs, attitudes and thoughts. Originally devised to treat depression, it has great potential in solving a wide variety of mental health problems. Conventionally, CBT is based on three factors: Thinking Error, Situations and Emotions. For whatever one says, the mental health personnel will identify these three that would best fit the statement and will accordingly take the conversation/therapy forward. In this work we consider understanding of mental health concepts of as a classification task. To facilitate this process, we use distributed representations in order to design a deep learning model that can identify the three factors when fed with statements from people undergoing therapy.

- **Thinking Error:** Thinking errors are faulty patterns of thinking that are self-defeating. They occur when the things you are thinking do not match up with reality. This is sometimes also referred to as cognitive distortions. Those who commit thinking errors often don't realise they are doing so. The thinking errors which we have taken in account are:

  1. **Black and White:** Only seeing things in absolutes
  2. **Blaming:** Holding others responsible for your pain. Not seeking to understand your own responsibility in situation
  3. **Catastrophizing:** Magnifying a (sometimes minor) negative event into potential disaster
  4. **Comparing:** Making dissatisfied comparison of self-versus others

5. **Disqualifying the positive:** Dismissing/discounting positive aspects of a situation or experience
6. **Emotional Reasoning:** Assuming feelings represent fact.
7. **Fortune Telling:** Predicting how things will be, unduly negatively
8. **Jumping to a negative conclusion:** Anticipating something will turn out badly, with little evidence to support it
9. **Labelling:** Using negative, sometimes highly coloured, language to describe self or other. Ignoring complexity of people
10. **Low frustration tolerance:** Assuming something is intolerable, rather than difficult to tolerate or a temporary discomfort
11. **Inflexibility:** Having rigid beliefs about how things or people must or ought to be
12. **Mental Filtering:** Focusing on the negative Filtering out all positive aspects of a situation
13. **Mind-Reading:** Assuming others think negative things or have negative motives and intentions
14. **Over-generalizing:** Generalising negatively, using words like always, nobody, never, etc
15. **Personalising:** Interpreting events as being related to you personally and overlooking other factors

- **Emotion:** Emotion is a subjective state of mind or a feeling which can be the reaction to the internal thoughts, memories, or can be due to the different mood, which a person tends to have. The emotions which we have taken for our work:

  1. **Anger:** A strong uncomfortable and non-cooperative response to a perceived provocation
  2. **Frustration:** Being upset or annoyed as a result of being unable to change or achieve somethings
  3. **Anxiety:** Any expression of fear, worry or anxiety
  4. **Depression:** Feeling down, hopeless, joyless, negative about self and/or life in general
  5. **Grief/Sadness:** Feeling sad, upset, bereft in relation to a major loss
  6. **Guilt:** Feeling blameworthy for a wrongdoing or something not done
  7. **Hurt:** Feeling wounded and/or badly treated
  8. **Jealousy:** Antagonistic feeling towards another either wish to be like or to have what they have
  9. **Loneliness:** Feeling of alone-ness, isolation, friendlessness, not understood by anyone

10. **Suicidal:** Deeply unhappy or depressed and likely to commit suicide.
11. **Shame:** Feeling distress, humiliation, disgrace in relation to own behaviour or feelings

- **Situation:** The situations under which these thinking errors are formed, and the emotions which the people are feeling.

  1. **Bereavement:** Being deprived or bereaved of something or someone
  2. **Existential:** Existence, Individuality
  3. **Relationships:** Having problems in a relationship
  4. **Work:** Facing problems in work including schools, colleges or at office
  5. **Life:**Facing problems in your day-to day lives like bad health, unable to make ends meet
  6. **Other:** Those that do not come under the scope of other given Situations

2 Literature review

2.1 Deep Learning for language understanding of mental health concepts derived from Cognitive Behavioural Therapy

In recent years, deep learning and distributed representations of words and sentences has impacted a number of natural language processing tasks, such as similarity, entailment and sentiment analysis. This paper [1] has introduced a task which is "Understanding of mental health concepts derived from CBT(Cognitive Behavioural Therapy)". A mental health ontology is defined based on the CBT principles, taken a large amount of data in forms of annotated texts, where this phenomena is exhibited. After which an understanding was made using the deep learning and distributed representations. The results of this paper shows that the performance of deep learning models combined with word embeddings for sentence embeddings significantly outperform non-deep-learning models in this difficult task. The module will be an essential component of a statistical dialogue system delivering therapy. This understanding module will be an essential component of a statistical dialogue system delivering therapy.

3 Other related works

The aim of building an automated therapist has been around since the first time researchers attempted to build a dialogue system [3]. Automated health advice systems built to date typically rely on expert coded rules and have limited conversational capabilities [4], [5], [6], [7], [8]. One particular system that we would like to highlight is an affectively aware virtual therapist [8]. This system is based on Cognitive Behavioural Therapy and the system behaviour is scripted using VoiceXML. There is no language understanding: the agent simply asks questions and the user selects answers from a given list. The agent is however able to interpret hand gestures, posture shifts, and facial expressions. Another notable system [7] has a multi-modal perception unit which captures and analyses user behaviour for both behavioural understanding and interaction. The measurements contribute to the indicator analysis of affect, gesture, emotion and engagement. Again, no statistical language understanding takes place and the behaviour of the system is scripted. The system does not provide therapy to the user but is rather a tool that can support healthcare decisions (by human healthcare professionals). The Stanford Woebot chat-bot proposed by [9] is designed for delivering CBT to young adults with depression and anxiety. It has been shown that the interaction with this chat-bot can significantly reduce the symptoms of depression when compared to a group of people directed to a read a CBT manual. The conversational agent appears to be effective in engaging the users. However, the understanding component of Woebot has not been fully described. The dialogue decisions are based on decision trees. At each node, the user is expected to choose one of several predefined responses. Limited language understanding was introduced at specific points in the tree to determine routing to subsequent conversational nodes. Still, one of the main deficiencies reported by the trial participants in [9] was the inability to converse naturally. Here we address this problem by performing statistical natural language understanding.

4 Method Description

4.1 Preprocessing

The individual annotated statements were the training examples for the model. Tokenization of each statement and removal of punctuation marks and stop words were done, so that the redundant information can be removed, and only the important and meaningful words thus left are then used for the training. These words were then converted to a vector form with the use of Word2Vec, which is just the numerical representation of the word features, so that it can be understood by the computer. These words are then padded according to their length, so that each word which is going as an input to the model is of same length. Finally One-Hot Encoding is done on the transformation thus produced, so that it is beneficial for us in multi-class classification, of the labels.

4.2 Model Building

4.2.1 LSTM:

Long Short Term Memory networks – usually just called "LSTMs" – are a special kind of RNN, capable of learning long-term dependencies. They were introduced by [10], and were refined and popularized by many people in following work.1 They work tremendously well on a large variety of problems, and are now widely used. LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn! All recurrent neural networks have the form of a chain of repeating modules of neural network. In standard RNNs shown in figure 1, this repeating module will have a very simple structure, such as a single tanh layer.

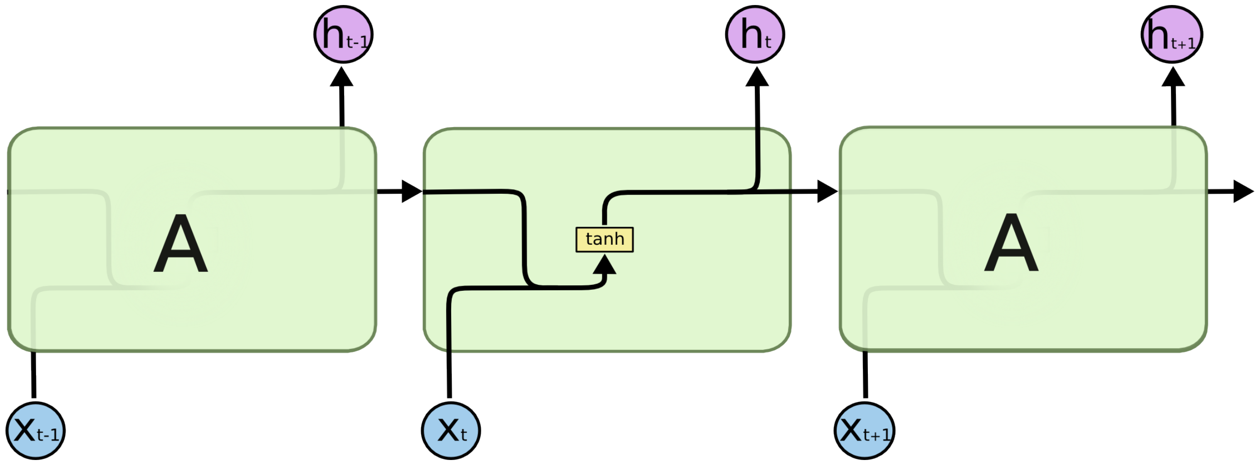

Figure 1: Depiction of layers in simple RNN

LSTMs also have this chain like structure shown in figure 2, but the repeating module has a different structure. Instead of having a single neural network layer, there are four, interacting in a very special way.

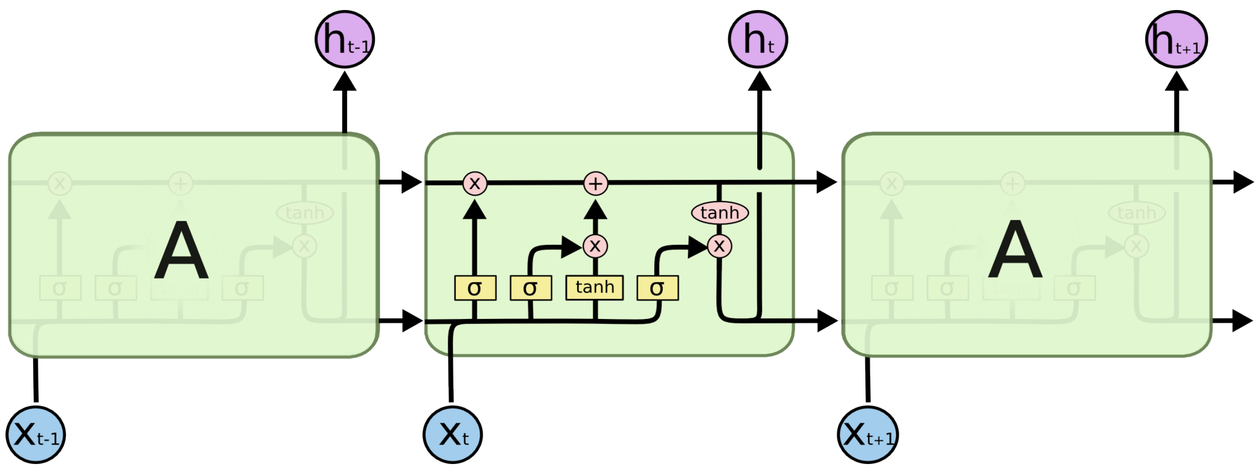

Figure 2: Depiction of layers in LSTM

The figure 3 shows the meaning of the symboles used in figures 1 and 2. Each line carries an entire vector, from the output of one node to the inputs of others. The pink circles represent pointwise operations, like vector addition, while the yellow boxes are learned neural network layers. Lines merging denote concatenation, while a line forking denote its content being copied and the copies going to different locations.

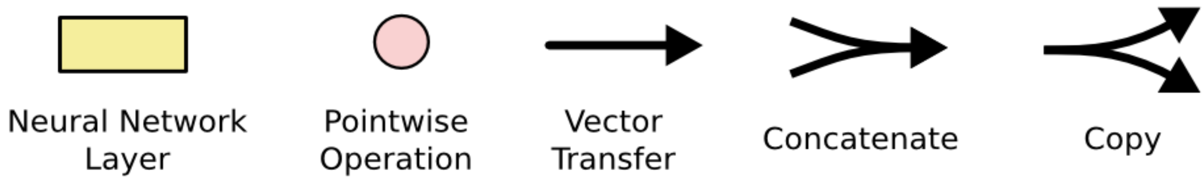

Figure 3: Notation used in RNN and LSTM diagrams

4.2.2 SVM:

The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N — the representing the number of features) that distinctly classifies the data points. As a hyperplane acts as a linear classifier, we also transform the input space to a new space such that new space is linearly classifiable, the process is usually carried out by kernels.
Hyperplanes are decision boundaries that help in classifying the data points. We can visualize the hyperplane of 2 feature model as a single line, and that of a 3 feature model as a plane. However, it is difficult to visualize if the number of features increases more than 3.

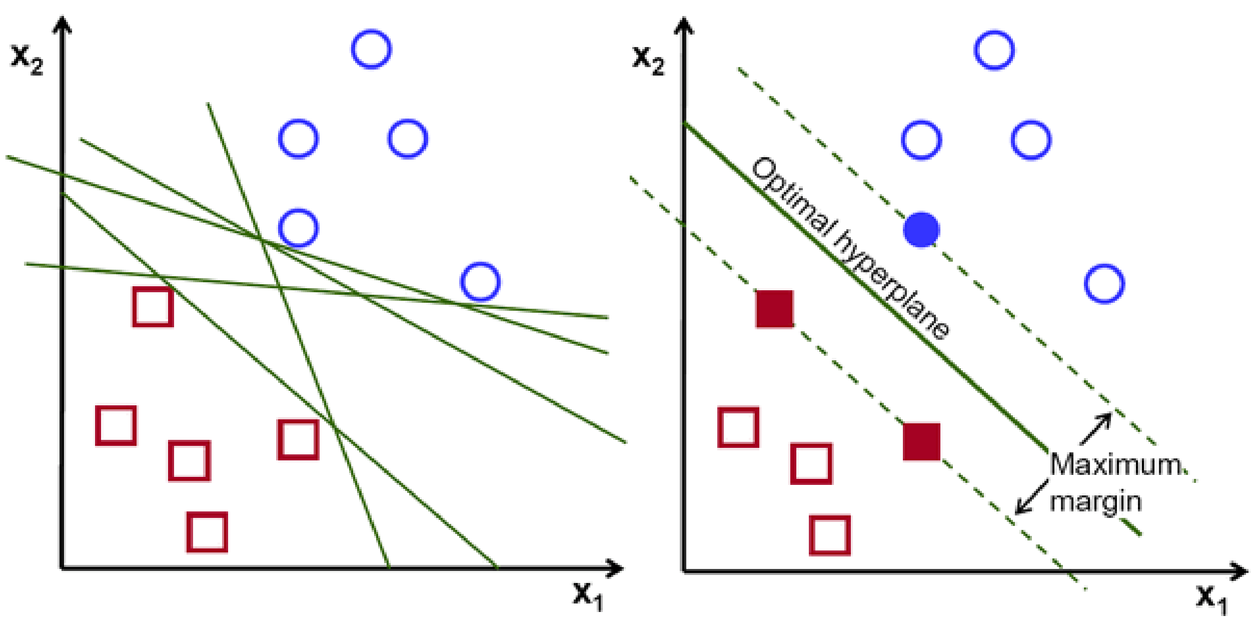

Figure 4: Possible Hyperplanes and Maximum margin hyperplane

As we can see from the figure 4, there are numerous hyperplanes from which we can choose from. The objective is to find the one which has the maximum margin, which can also be said as maximum distance between the data points of both the classes. This way we can ensure that the hyperplane thus formed is not biased to any class, and gives proper results for future data points.

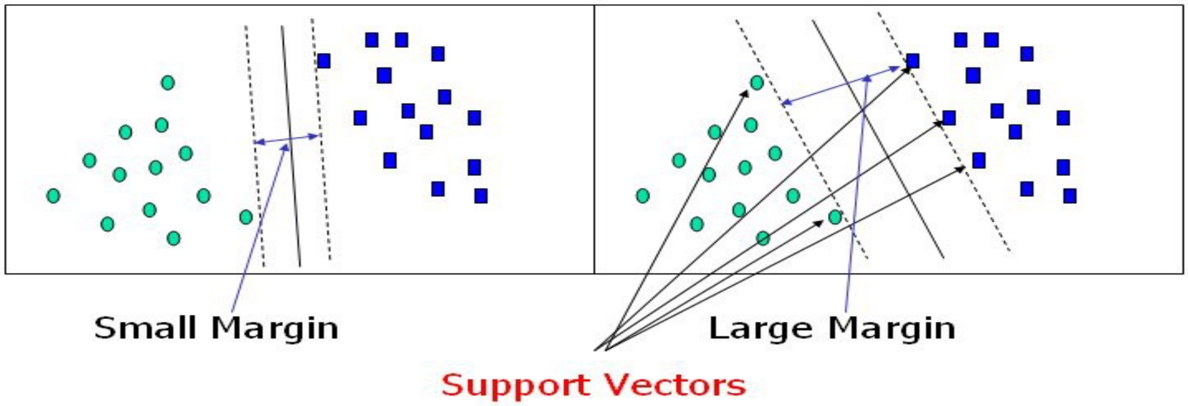

Figure 5: Support Vectors

Support vectors are the data points which are used to calculate the margin of the classifier. It can be taken as the data points which are closer to the hyperplane also shown in figure 5. If these data points are deleted, then a new hyperplane will be formed.

5 Experiments

5.1 Dataset Formation

There were three main steps which constituted the building up of the dataset:

- Collection: The dataset required for the task at hand is not publicly available. Thus, it had to be collected and annotated. The dataset comprises of statements that have been collected from a number of online community and discussion forums where people anonymously post about their emotions and feelings.

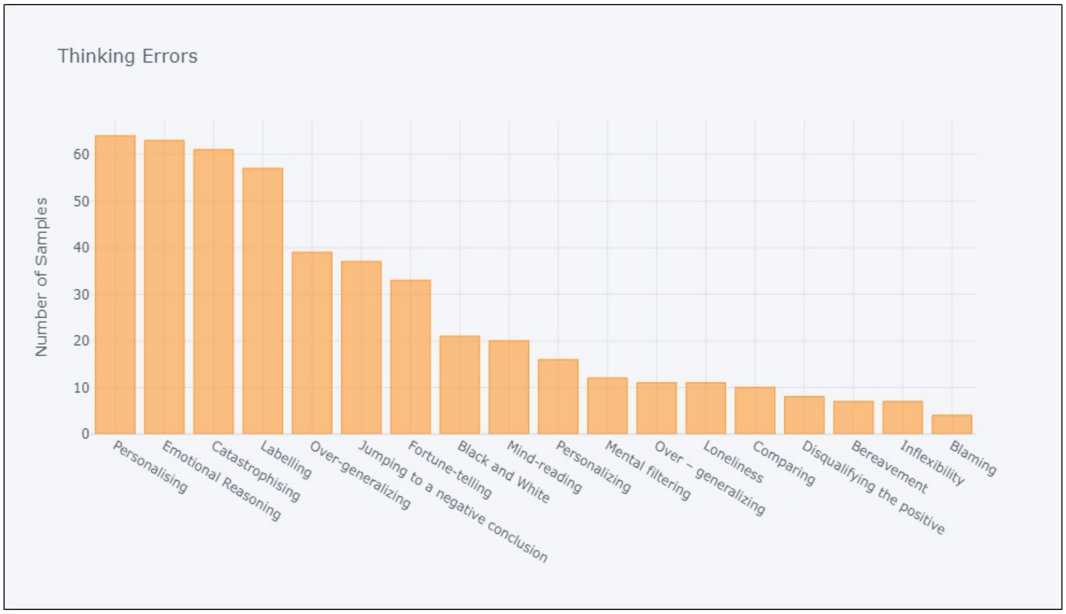

Figure 6: Thinking Errors

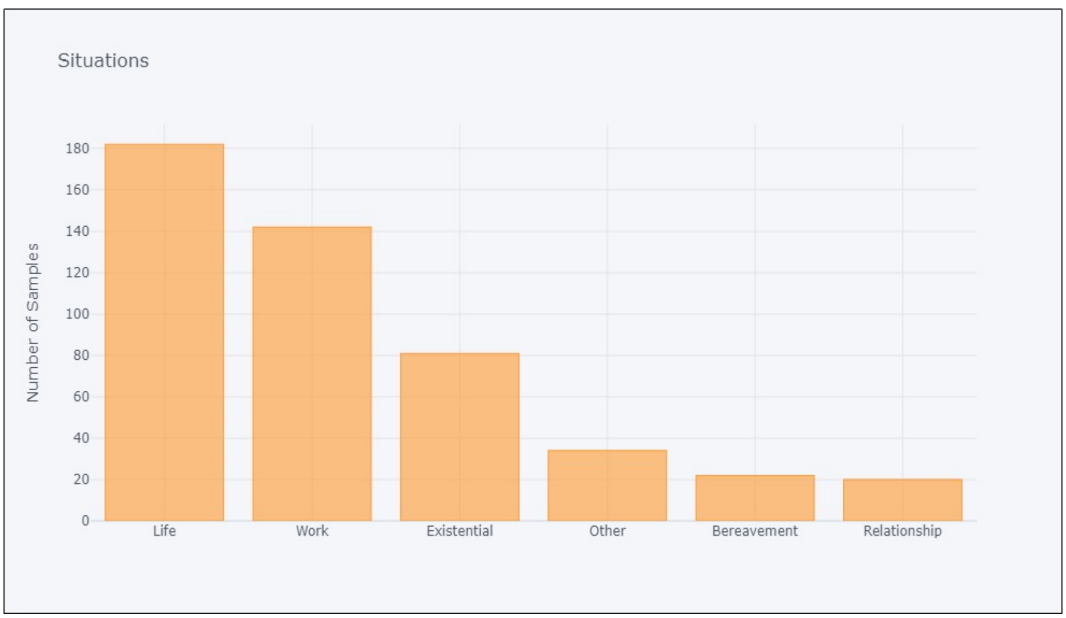

Figure 7: Situations

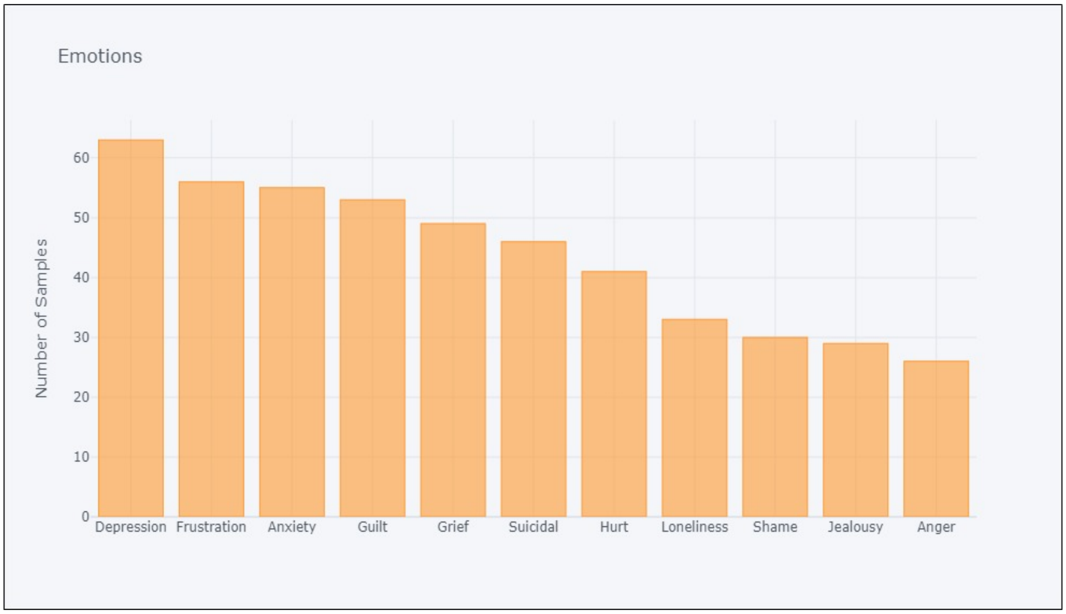

Figure 8: Emotions

- Annotation: The data collected was then annotated for their Thinking Errors, Emotions and Situations(see figure 9) which has been defined above in Section 1.

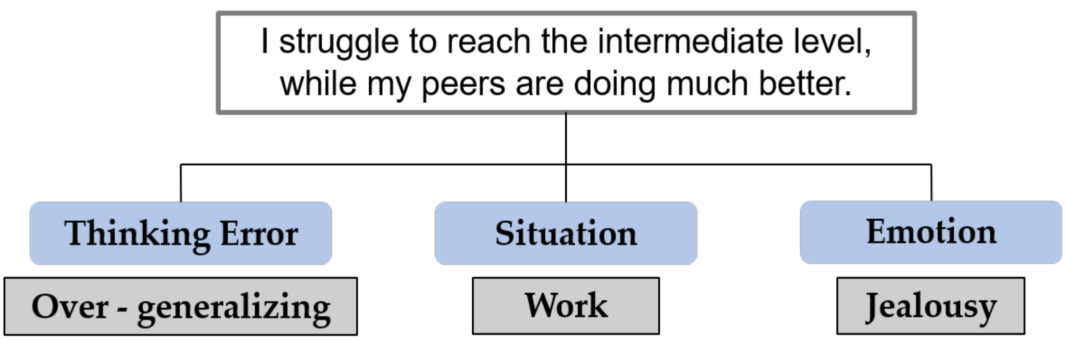

Figure 9: Annotation Example

- Expansion: After annotation was completed for the collected statements, back translation with Google Neural Machine Translation architecture [11] was used to expand the dataset by translating statements to and fro between English and other languages. This resulted in an expanded dataset with more words in the corpus.

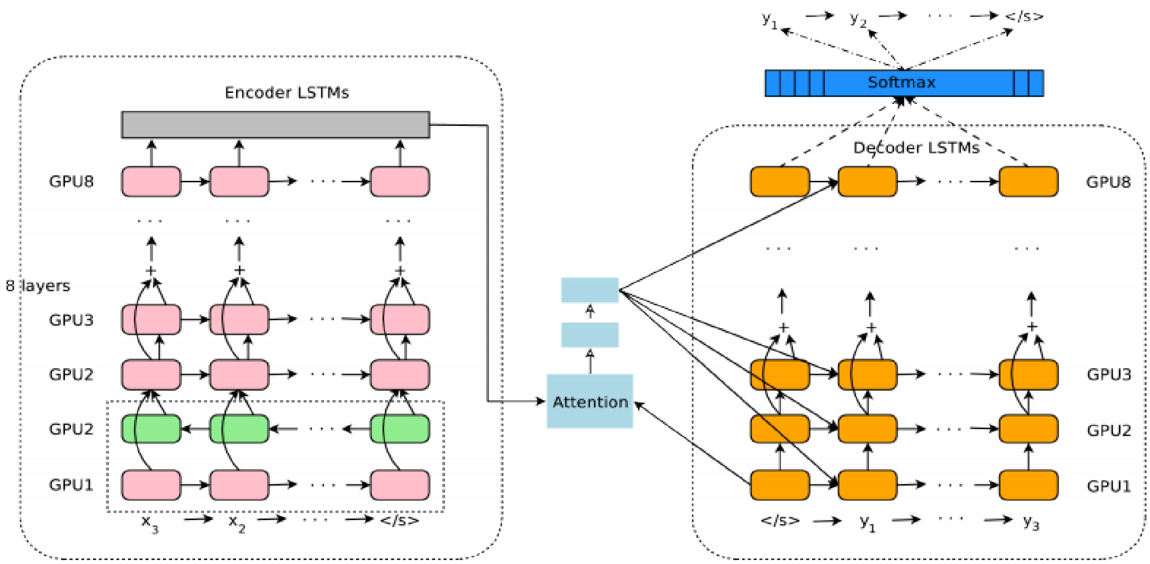

Figure 10: Architecture for Google NMT

5.2 Implementation details

The experiment was set up in three phases after the processed data was collected. This included sanity checks and making proper splits for each phase to make sure that we were preserving the data sanctity by separating testing and training data. Post preprocessing, we started training our models. The training data was built using 90-10 splits where 90% of the data was for training and rest was saved for testing. The models have been implemented using keras. An LSTM model was designed with an embedding layer, Dropout Layer, LSTM and a dense layer. Training was carried out in two different ways. One with Embedding Dimensions for Thinking Errors to be 120 while those of Situations and Emotions being 100 each and the other with all three dimensions to be the same. The higher number of Embedding dimensions were used so as to facilitate more number of classes (=15) when compared to those of Situations (=6) and Emotions (=11).

5.2.1 Training

We train using Adam, a batch size of 10, a dropout of 0.2, The model was trained for 10 epochs. The data was also trained using an SVM with a Gaussian Kernel. The model architecture with different embedding layers has been show in Figure 11 and Figure 12. Figure 13 shows the training loss and accuracy for the Deep learning model.

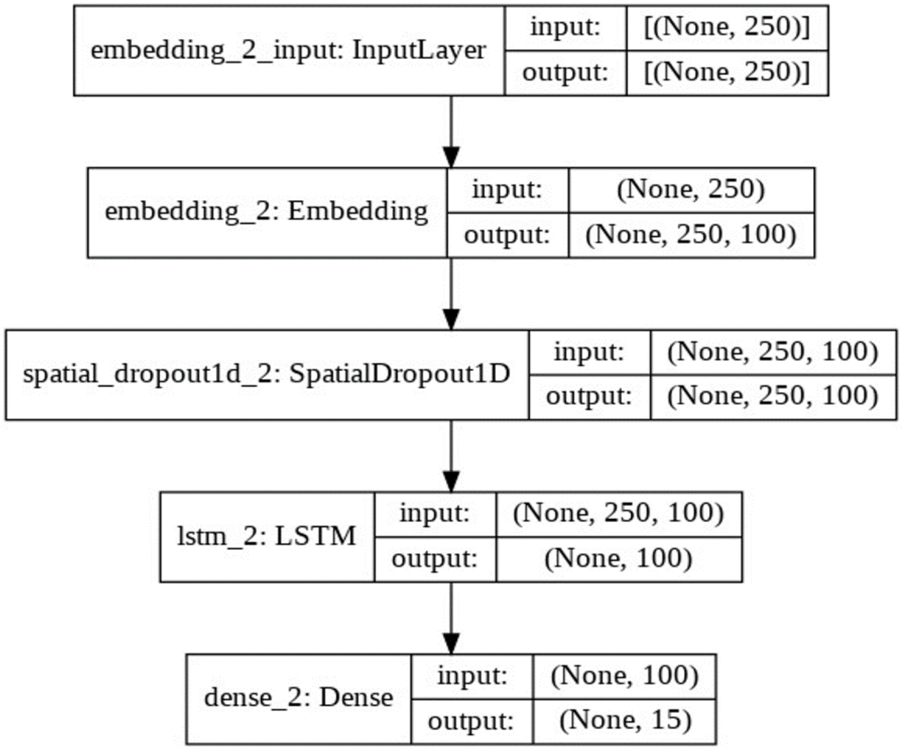

Figure 11: LSTM Model Architecture (Embedding Dimensions = 100)

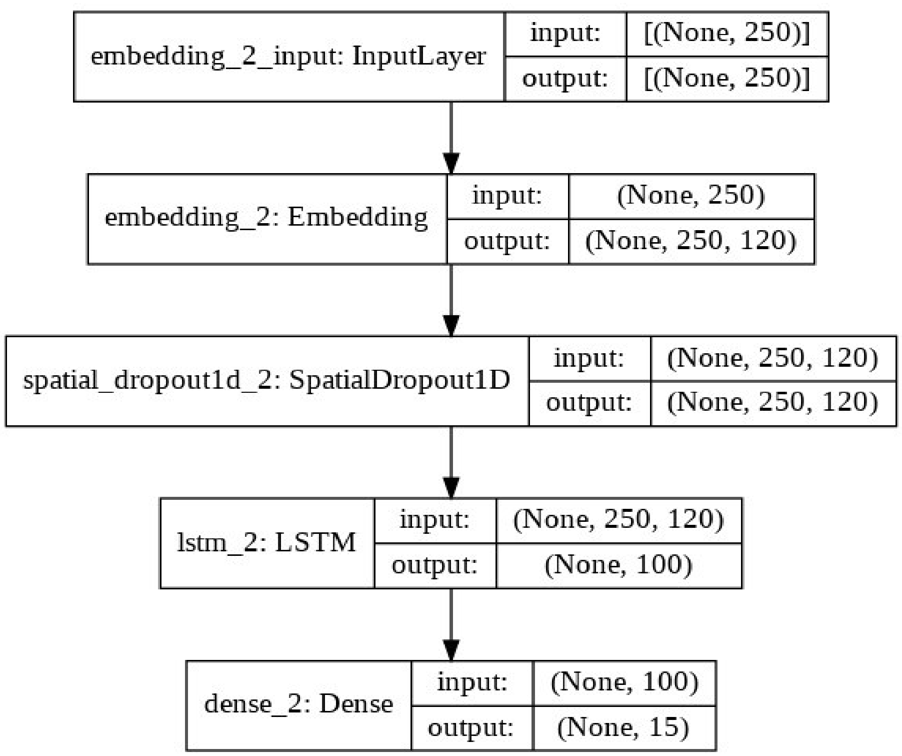

Figure 12: LSTM Model Architecture (Embedding Dimensions = 120)

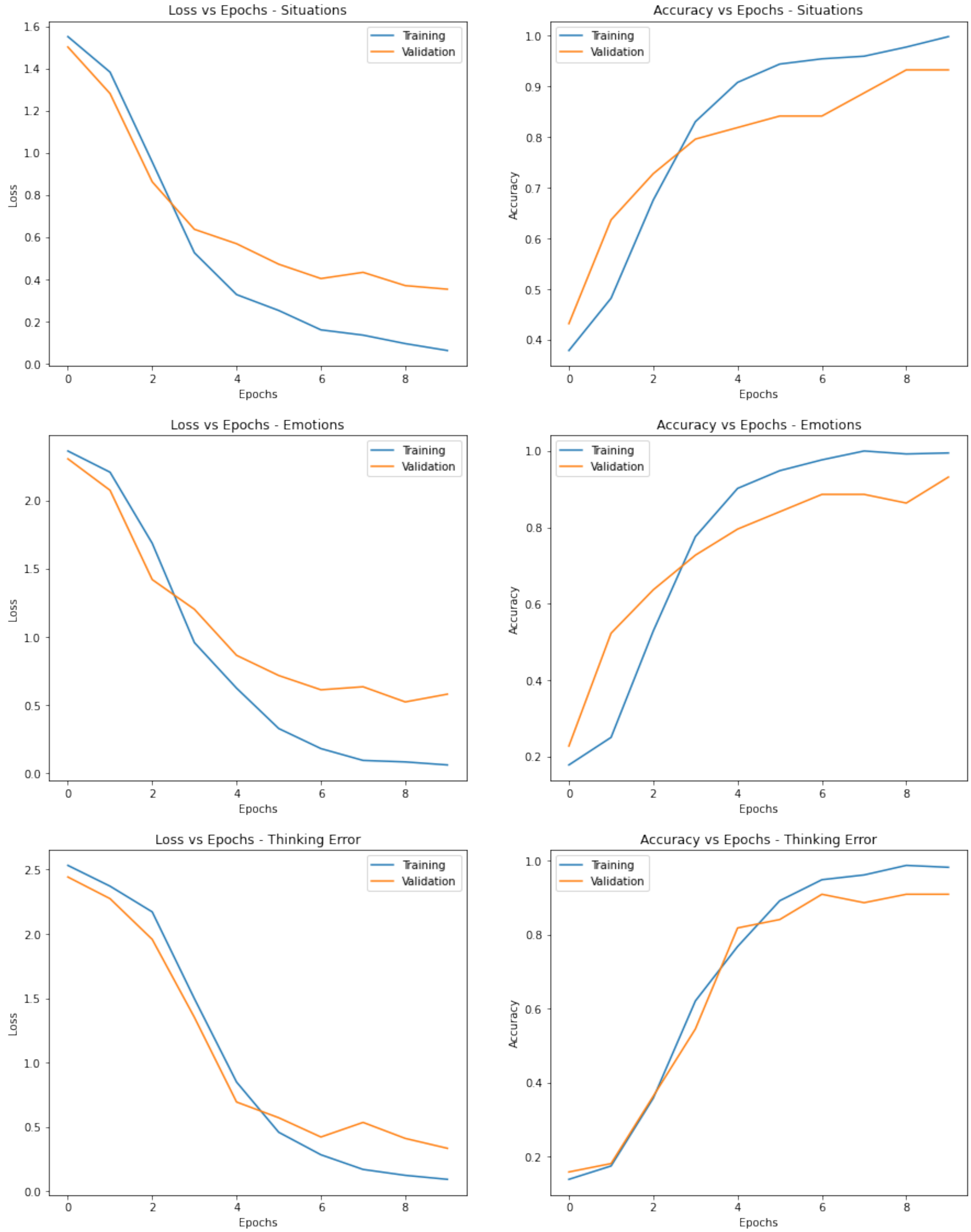

Figure 13: Training Loss and Accuracy

6 Result and Analysis

6.1 Testing

The models were tested on an unseen dataset of 49 statements. Figure 14 shows the testing accuracy for the three predicted variables.

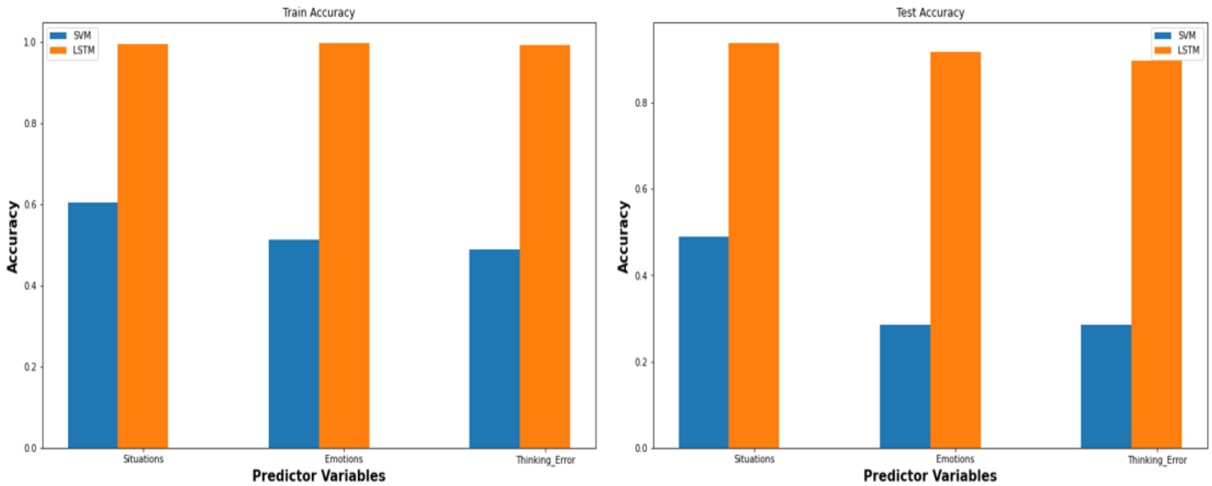

Figure 14: SVM vs LSTM on unseen data

6.2 Inference

Figure 15 shows the sentences in the unseen dataset and the corresponding expected and predicted Thinking Errors, Situations and Emotions.

0	Situations (actual)	Situations (predicted)	Emotions (actual)	Emotions (predicted)	Thinking Error (actual)	Thinking Error (predicted)
i just want to ignore the instinct to live	Bereavement	Bereavement	Suicidal	Suicidal	Jumping to a negative conclusion	Jumping to a negative conclusion
faliure to get into the college earned me the ...	Work	Work	Guilt	Guilt	Emotional Reasoning	Jumping to a negative conclusion
i struggle so hard to be barely average when m...	Work	Work	Jealousy	Jealousy	Over-generalizing	Over-generalizing
the only reason i didn't kill myself was that ...	Other	Existential	Suicidal	Suicidal	Emotional Reasoning	Emotional Reasoning
there is constant sadness in my life	Life	Life	Grief	Grief	Black and White	Black and White
my friend gets more points than me	Work	Work	Jealousy	Jealousy	Over-generalizing	Over-generalizing
i don't think i'll just love myself by having ...	Relationship	Existential	Hurt	Suicidal	Fortune-telling	Fortune-telling
everything is in place bad defective imperfect...	Life	Life	Grief	Grief	Catastrophising	Catastrophising
i am an unimportant anti social suicidal bitch	Bereavement	Bereavement	Suicidal	Suicidal	Catastrophising	Catastrophising
all my near ones have a better vehicle than me	Bereavement	Work	Jealousy	Jealousy	Personalizing	Over-generalizing

Figure 15: Prediction result on Unseen Data. The highlighted words show the incorrect predictions.

6.3 Thinking Errors

Embedding Dimensions = 100			
Metrics	Training Set	Validation Set	Test Set(Unseen Data)
Loss	0.1031	0.3332	0.4263
Accuracy	0.9803	0.909	0.898

Embedding Dimensions = 120			
Metrics	Training Set	Validation Set	Test Set(Unseen Data)
Loss	0.0595	0.4191	0.390
Accuracy	0.9948	0.9091	0.918

The training and testing accuracy for the Thinking Errors is low as compared to Situations and Emotions. A primary reason for that is the higher number of classes in Thinking Errors and a highly unbalanced dataset. Another plausible reason could be that the data has been annotated incorrectly. Since, CBT is a complex medical subject, it would require the expertise of medical professionals for more accurate annotations.
As we increased the dimensions for representation of each sentence in the embedding layer from 100 to 120 the results seems to look better. By this decision we to hope to transform our input space to easily classifiable space.

6.4 Situations

Embedding Dimensions = 100			
Metrics	Training Set	Validation Set	Test Set(Unseen Data)
Loss	0.0520	0.3532	0.3098
Accuracy	0.9982	0.9318	0.9388

It shows the highest accuracy of all the three primarily because lower number of classes. However, there is a considerable gap between the training and testing accuracy. This could be attributed to the fact that the model is not immune to errors when it comes to predicting situations for statements which have words that are not there in the training corpus. One way of handling it could be a reinforcement learning based approach where new sentences can be added to the corpus and the model is trained with the new samples.

6.5 Emotions

Embedding Dimensions = 100			
Metrics	Training Set	Validation Set	Test Set(Unseen Data)
Loss	0.0620	0.5797	0.3852
Accuracy	0.9949	0.9318	0.9184

The training and testing accuracy do not show a huge difference for Emotions. This could be possible as the distribution of annotations for Emotions is relatively more balanced when compared to Situations and Thinking Errors.

7 Takeaways

In a nutshell, the following takeaways can be obtained after a careful study of the results.

- The data is self-annotated, however if it was done by a certified, professional psychiatrist, the quality of the data-set would increase and the model will give better results.
- Some labels may have biased prediction, as the division of the statements belonging to certain labels is not uniform. For example, there are far more statements for particular labels, and not for others. So the training of these might favour some of the labels.
- This model won't be able to predict for the words which are not present in its training data, and may give faulty outcomes.
- The data-set used by us is itself a small one. So there are large variety of texts, which are missed out on training

8 Conclusions and Future Scope

It takes a lot of knowledge, experience and hard work for psychiatrists to perform Cognitive Behavioural Therapy on their patients. Therefore the problem targeted in this work is challenging and requires sincere efforts in order to gain positive outcomes out of this. It can be used for helping in automating the process, so it could reach the masses. As the model cannot correctly predict on the statements, which contains words not present in the training set, reinforcement learning algorithms may be used which keep on updating the model based on new inputs. Also, the data-set can be expanded for the algorithm to perform better. This model after building it in a professional manner can also serve as a first level guidance to the patients who are wanting to seek care from psychologists. Efforts can be made for developing a chat-bot which when integrated with model predictions can act as a medical aid gaining insights about people from their conversations and helping them out if they are in distress.

9 Acknowledgement

We would like to thank our professor, Dr. Snehasis Mukherjee for providing us with the opportunity to learn about a challenging problem in the field of Deep Learning. We learnt a number of new concepts and framework tools to implement this work successfully. His focus on concepts and constant guidance and support throughout the course helped us a lot. The course, its assignments, evaluation and the work, all have contributed to a great learning experience for us.

References

[1] Lina Rojas-Barahona, Bo-Hsiang Tseng, Yinpei Dai, Clare Mansfield, Osman Ramadan, Stefan Ultes, Michael Crawford, and Milica Gasic. 2018. Deep learning for language

understanding of mental health concepts derived from cognitive behavioural therapy. arXiv preprint arXiv:1809.00640.

[2] Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. 2016. Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.

[3] Joseph Weizenbaum. 1966. Eliza, a computer program for the study of natural language communication between man and machine. ACM, 9(1):36–45.

[4] Lina Maria Rojas-Barahona and Toni Giorgino. 2009. Adaptable dialog architecture and runtime engine (adarte): A framework for rapid prototyping of health dialog systems. I. J. Medical Informatics, 78(Supplement-1):S56–S68.

[5] Lazlo Ring, Barbara Barry, Kathleen Totzke, and Timothy Bickmore. 2013. Addressing loneliness and isolation in older adults: Proactive affective agents provide better support. In Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII '13, pages 61–66, Washington, DC, USA. IEEE Computer Society.

[6] Giuseppe Riccardi. 2014. Towards healthcare personal agents. In Proceedings of the 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research Including Business Opportunities and Challenges, RFMIR '14, pages 53–56, New York, NY, USA. ACM.

[7] D DeVault, R Artstein, G Ben, T Dey, E Fast, A Gainer, K Georgila, J Gratch, A Hartholt, M Lhommet, G Lucas, S Marsella, F Morbini, A Nazarian, S Scherer, G Stratou, A Suri, D Traum, R Wood, Y Xu, A Rizzo, and L-P Morency. 2014. Simsensei kiosk: A virtual human interviewer for healthcare decision support. In International Conference on Autonomous Agents and Multiagent Systems.

[8] Lazlo Ring, Timothy Bickmore, and Paola Pedrelli. 2016. An affectively aware virtual therapist for depression counseling. In CHI 2016 Computing and Mental Health Workshop.

[9] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. JMIR mental health, 4(2).

[10] Sepp Hochreiter, Jürgen Schmidhuber; Long Short-Term Memory. Neural Comput 1997; 9 (8): 1735–1780.

[11] Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. 2016. Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.