Documentation

Depression Therapy using Chatbot

Programming Club - IIT Kanpur

Harish Rajagopal
160552

Vishwas Lathi
160808

Harshit Sharma
160283

Acknowledgements: Mohammad Daud

Week 1

Coursera - Intro to Machine Learning

Completed chapters from the course by Andrew Ng on Coursera from Stanford University, pertaining mostly to supervised learning, including Linear Regression, Polynomial Multivariate Regression, Logistic Regression, Neural Networks and Back-propagation, Support Vector Machines, etc.

Udacity - Intro to Machine Learning

Completed chapter on text learning from Introduction to Machine Learning course on Udacity, based on preprocessing of text by tokenisation of document, removal of stopwords, word stemming and term-frequency inverse-term-frequency (tf-idf) representation of documents using the python libraries scikit-learn and NLTK.

Week 2

Neural Networks and Deep Learning

Completed chapters of the online ebook *neuralnetworksanddeeplearning.com* on the basics of neural networks, neural network architectures, the back-propagation algorithm, optimizing hyper-parameters, regularization methods and improving performance of neural networks, gradient instability in deep neural networks, and basics of CNNs.

Woebot - Congnitive Behavioural Therapy Chatbot

Members chatting with woebot, a chatbot made by Stanford researchers which provides supplementary therapy using Cognitive Behavioural Therapy techniques, available on Facebook Messenger. Aim is to obtain maximum responses from Woebot to learn about its tree structure, which will aid in building a tree structure for the project chatbot (virtual human).

Week 3

Cognitive Behavioural Therapy

Obtained and studied several transcripts of conversations between patients and therapists practicing Cognitive Behavioural Therapy in therapy sessions, along with a book on the basics of Cognitive Behavioural Therapy, in order to study common methods employed by psychologists to finalise the tree structure.

USC ICT Virtual Human Toolkit

Downloaded and installed the USC ICT Virtual Human Toolkit. Learned from tutorials on the VH Toolkit including the VH Builder tutorial, Character Customization tutorial and reading documentation on the toolkit.

Week 4

Tree Structure

Created tree structure of questions to be asked and responses given by the chatbot, according to the result obtained after applying sentiment analysis to the user's answer.

Merged and integrated the concepts of Cognitive Behavioural Therapy, as learned from various online resources and books, to the tree structure.

Obtained specific psychological advice from online and offline sources for specific cases of depression. Integrated these cases as a separate branch of the tree for tackling slightly different cases of depression.

Week 5

Sentiment Analysis

Obtained datasets from Kaggle, Twitter, Stanford Sentiment Treebank and sampled them for identifying those with more relevant and accurate information.

Preprocessed the Twitter dataset, containing 1600000 tweets classified as positive and negative, initialized a vocabulary using Gensim Dictionary and replaced words with positions in the vocabulary. Limited each tweet to 20 words and split them if they exceed the limit.

Coded the model in python using the libraries Tensorflow, Keras, and Gensim. Initialised a word embedding and implemented an LSTM in the neural network model for identifying context in the tweet in accurately classifying into positive or negative.

Week 6

Depression Therapy Chatbot

Created the chatbot python file and implemented the tree structure in python. Integrated sentiment analysis in the chatbot functioning, with thresholds for classification into leaf nodes tweaked for high accuracy.

Experimented with the model architecture and two different methods of implementation. Fine-tuned the parameters for the sentiment analysis model for optimum balance between performance and accuracy.

Debugging, testing and documentation of the chatbot.