International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307 (Online), Volume-12 Issue-3, July 2022

Empathetic Robot for the Elderly using Machine Learning

Varsha Ryali, Sristi Kumari, Syed Mahmood Nabeel Anwar, Sheela SV

Abstract: ***According to WHO estimates, there is a growing population of over 1 billion people aged above 60 years of age. There has been an increasing shortage of caregivers for the aging populations. This has opened a market of USD 7 billion dollars in senior care alone. In a struggle to care for our elderly and capitalize from this industry in efficiency, this paper presents an empathetic robot that can make day-to-day activities hassle free. It is a voice assistant that can detect the emotion of the speaker and reply with contextual awareness to bridge the superstitious gap between the elderly and technology. It achieves this with multi-modal classification of emotion using audio and text. i) Machine learning model using Librosa to engineer features and Support vector classifier (SVC) , ii) BERT based model using transfer learning to categorize text. The robot can also set reminders for appointments and medicine intake to help with forgetfulness. The robot is customizable by a caregiver or loved one through the web application where details of any messages, reminders, descriptions of prescriptions could be entered. The robot comes with a pill dispenser that can rotate and dispense pills at the correct time and also notify the senior through the speaker. A raspberry pi is used to convert the speech to text and vice versa. The design of this robot paves the path to providinga realistic, care-giving experience to the elderly.***

Keywords: ***The robot can also set reminders for appointments and medicine intake to help with forgetfulness.***

I. INTRODUCTION

The supply and demand issue in elder care is a major issue. The number of caregivers does not expand in lockstep with the growth of the elderly population. The growing rate of loneliness and the depression amongst elderly people calls for major attention. It has been seen that with age, people start to share less and suppress their emotions, leading to stress, anger, loneliness and sometimes major illness. The use of medical robots in the care of the elderly will drastically cut the existing enormous costs of elderly care. In order to understand their emotions, respond to it with appropriate answers, take care of them personally, a very smart and empathetic robotis needed.

Manuscript received on 16 June 2022 | Revised Manuscript received on 23 June 2022 | Manuscript Accepted on 15 July 2022 | Manuscript published on 30 July 2022.

***** Correspondence Author

Varsha Ryali*, Department of Information Science and Engineering, BMS College of Engineering, Bangalore (Karnataka), India. E-mail: varsharyali@gmail.com

Sristi Kumari, Department of Information Science and Engineering, BMS College of Engineering, Bangalore (Karnataka), India. E-mail: sristi.sonal8686@gmail.com

Syed Mahmood Nabeel Anwar, Department of Information Science and Engineering, BMS College of Engineering, Bangalore (Karnataka), India. E-mail:smnanwar18@gmail.com

Dr. Sheela S V, Professor, Department of Information Science and Engineering, BMS College of Engineering, Bangalore (Karnataka), India.

The robot being used in elder care are not smart enough to consider both the mental as well as physical health of the elderly. In order to reduce the stress and anxiety level, the elderly person must be looked after closely and must have someone to talk all the time. To find an empathetic robot that looks out for the emotional aspect and help with medicines, reminders and physical aspects is a task in itself. And this brings to the need of having an empathetic caregiver robot with the above-mentioned functionalities i.e., a voice assistant that recognises the emotions and respond accordingly along with taking care of medicines and reminders. Emotion detection is a subset of sentiment analysis that aims to extract fine- grained emotions from speech, image, or text data. Regardless of the amount of text data accessible, detecting emotions from texts has been difficult. This is largely owing to the lack of speech modulation, facial expressions, and other indications that could help with context and relation extraction. Another reason is the lack of a suitable method for extracting context from texts. Furthermore, because certain texts convey many emotional expressions, the need for emotion- conveying words disambiguation to verify categorised emotions as true emotions offers a substantial stumbling block in the sector. With all the emotional support the robot is providing, the need of help for physical work also increases. With each passing day, the memory weakens and the need of timely medications, exercise increases. People of all ages, whether young or old, are prone to forgetting their medications. For the treatment of any ailment, prompt medication is essential. An automatic pill dispenser with the ability to distribute many pills to a single patient or a private individual can help patients and the elderly remember when and how much of their medication they need to take. Not only a reminder to take medications but also other daily reminders, a caregiver is needed which the chatbot completes in our proposed System. Thus, the objective of this paper is the development of a robot that acts as a chatbot for normal daily conversations and provide empathetic responses according to the emotion detected in the sentence followed by the phrase “Let’s Talk”. To detect the emotion, we use two different model namely, BERT and SVC using librosa library, applied on text and audio recorded respectively. Along with the empathetic response, it also takes care of the reminders to be read out loud and pills to be given at the particular time slots.

II. PROPOSED SYSTEM

To achieve the said objectives of this paper, the proposed system is trisected into three important conceptual modules:

i) Emotion detection and empathetic response generator: Emotion recognition using meaning of the conversation with the robot Allows the system to detect emotions and categorize it into the five moods for further responses.

ii) Voice assistant: The question preparation and the answer retrieval, Reminder and Alert announcement

iii) Medicine Dispensing System: Scheduled Medicine dispensing, Alert/On Spot Medicine dispensing

In order to attain the empathetic aspect of the robot, it needs to be trained using two different models and applied on two different modes of input (Linguistic and textual input). SVC model with features of librosa is applied to detect emotion in voice message captured and BERT is used for textual emotion detection of the same voice captured. Final mood predicted must then be passed on to fetch the required response from the datasets. The next module proposed takes in the input from user as voice message. This message is then passed to the chatbot to get the appropriate reply. After having the appropriate replies from both the modules, it is read out through the connected speaker after conversion of text to speech. Alerts/reminders already stored in tables are to be read out loud through the same speaker after checking at specified time (thrice a day in this case). To cater the needs of timely medicine dispensing, a pill dispenser connected to the robot is proposed. Pills box rotation according to the slot selected and stored in the backend tablesis expected.

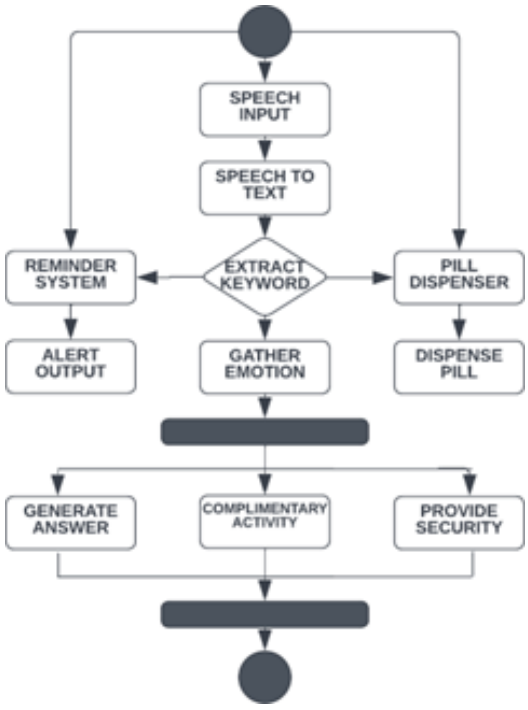

Fig. 1. (Activity diagram helps in modelling the coordination of the use cases to depict the end-to-end flow)

The flow begins with gathering the speech input and converting it to text. The system identifies the keywords that trigger either the reminder system, empathetic response generator or the pill dispenser. The alert system when called will read out the alert. The medicine dispenser is capable of rotating and dispensing required pill on command. The conversation system gathers emotion and replies accordingly.

III. IMPLEMENTATION

The workflow kicks off with the conversion of speech to text. A microphone is connected to the Raspberry Pi which uses the Speech Recognition package in Python. This package is not only known for its ease of use but it's popularly used asa wrapper for many speech APIs. The user has three routes to take from here that are triggered by the respective keywords which are: "Let's talk", "Medicine" and "Reminder".

A. Emotion Recognition

BERT based model: If the user opted to have a conversation with the robot (let's talk), this triggers the emotion detection models on the text that was extracted from the speech and the recorded audio of the speech as well. Holistic emotion detection can be implemented from the lexical, audio and video features. This system incorporates a combination of the first two to give a more heightened and refined accuracy score. A model run on BERT pre-training is used to predict the emotion from the converted text. BERT is a transformer-based machine learning technique for NLP pre-training[2]. A dataset called Daily Dialog is used that consists of written dialogs, messages and short stories. Each dialog utterance/message is labelled with one of the five emotion categories: joy, anger, sadness, fear, neutral. The model then executes multiclass text classification. Ktrain is additionally used to leverage a simple prediction API for saving and deploying both models and data pre-processing steps to make predictions on new raw data.

```
model = text.text_classifier('bert', train_data=(x_train, y_train), preproc=preproc)

Is Multi-Label? False
maxlen is 350
done.

learner = ktrain.get_learner(model, train_data=(x_train, y_train),
                             val_data=(x_test, y_test),
                             batch_size=6)

learner.fit_onecycle(2e-5, 3)
```

Fig. 2. (BERT model training)

SVC based model: A model run on the SVC classifier using the Librosa package on the recorded audio. The TESS (Toronto Emotional Speech Set) dataset is used for training and testing audio on the basis of 7 cardinal emotions thatis 2800 data points in total. These typically need to be ata minimum length of 5 seconds. The audio is split into five sections and each section is predicted to be one of the classes: [0- joy, 1-anger, 2- sadness, 3-fear, 4-neutral]. The SVC (Support Vector Classifier) is used to find the best-fit hyperplane for the audio where it divides and categorizes based on the emotions classified in the dataset. This performsvery well as the number of samples increases. The features are then fed to this classifier using Librosa. Librosa is a Python package for music and audio analysis. It is used to extract the following features from the audio to train and further predict: The zero-crossing rate of an audio time series, root-mean-square (RMS) value for each frame, either from the audio samples or from a spectrogram,

Retrieval Number: 100.1/ijsce.C35780712322
DOI: 10.35940/ijsce.C3578.0712322
Journal Website: www.ijsce.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Mel-frequency cepstral coefficients, spectral centroid, spectral roll-off frequency and p'th-order spectral bandwidth. The prediction is then extracted as the maximum occurring element in the predictions based onsections of the audio. The predictions from each of the models are pipelined into our final model that gauges the emotion based on the confidence score of each. The higher the score, the greater the weightage awarded in making the final prediction. Once the model predicts the emotion of user and finds the mood. The final mood is used to connect to one of the five response datasets prepared according to the respective moods. This further leads to random selection of one cell from the connected table to be passed as output response. If Sad is the final mood predicted, dataset specified for sadness is referredand one random selection of row is done. This output is then passed as input to the raspberry pi. After getting the input, Raspberry pi reads it out to user through the speaker connected to it using audio jack. To pass an audio message through speaker, the text line input to RPi is converted to audio using python code for conversion of text to Speech.

B. Reminder system

To satisfy the roles of a caregiver, the robot has pill dispenser system which is controlled by caretaker. Once the caretaker enters the correct login credentials in the web application, he/she gets the access to add medicine details with slot number and time to dispense it. These medicine details can also be viewed and their status for dispensingcan be controlled by the user of the app from the 'View Medicine' section. Along with this, there's another section "Add Reminder" to add reminders for the elderly person which needs to be read out at the specified time. And the last page "Add Contact" is to add phone number along with the relationto elderly person for emergency usages. At a specified time in Morning, Afternoon and Evening, the controller checks forany reminders being sent and stored in the reminder table. If any reminder is detected, the message specified is read out at the specified time by passing it. through the speaker after converting 'Text to speech'. In case 'emergency' word is spoken by the user of robot, the emergency contacts stored The results of each of these modules are presented in detail in the following section:

A. Emotion Recognition

1) BERT based model: To Classify the mood from the text (converted from speech), we used BERT.

```
[12] learner.validate(val_data=(x_test, y_test), class_names=class_names)
```

	precision	recall	f1-score	support
joy	0.86	0.82	0.84	1619
sadness	0.80	0.82	0.81	1641
fear	0.87	0.81	0.84	1492
anger	0.80	0.80	0.80	1566
neutral	0.78	0.85	0.81	1616
accuracy			0.82	7934
macro avg	0.82	0.82	0.82	7934
weighted avg	0.82	0.82	0.82	7934

```
array([[1321,   54,   34,   28,  182],
       [  52, 1338,   53,  118,   80],
       [  42,  101, 1204,  105,   40],
       [  38,  126,   69, 1248,   85],
       [  91,   54,   26,   69, 1376]])
```

Fig. 3. Result matrix for BERT model

2) SVC Model: To further determine the mood using the frequency of audio (captured simultaneously and stored in theform of .wav files), we used SVC model with Librosa library and its features.

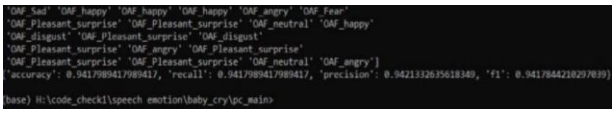

Fig. 4. Result matrix for SVC model

TABLE I. Model comparison

Model	Precision	Recall	F1-score	Support
BERT	0.82	0.82	0.82	7934
SVC	0.94	0.94	0.94	8132

C. Pill dispenser

At the specified time, controller looks for the slot numbers mentioned for pills and sends signal to motor driver. Once the Analog signals (sent using connections made from GPIO pins in the Raspberry Pi) are detected by the motor driver, the stepper motor rotates accordingly (1028*slot number – gives the rotation parameter since 2048/2 is equals to half step) and finally rotates back to original position to rotate again for anyfurther pills mentioned.

IV. RESULTS

This robot is designed to eliminate the feeling of loneliness in the growing population of the independent elderly today. It accomplishes this by machine learning models that can detect emotion and generate a response based on the contextual appropriacy. The dependency of the elderly for healthcare related needs are also brought to a minimum by the pill dispenser which dispenses based on the prescription and time and also alerts the user. It can also call the hospital/loved onesin case of an emergency where the elder person can merely utter a cry for help. Basic questions and conversation can also be delivered as it serves as a voice assistant. Lastly, it can set alerts and reminders that will be notified to the elderly on a speaker which can make day-to-day activities run smoothly.

B. Alert system

Alert input in Web application is stored in the table and searched for, after a fixed interval of time. If any alert is present, it is read out using the speaker connected. in contact details table is fetched and contacted.

C. Pill dispenser

Stepper motor used here, rotates every 45 degrees to reach next slot and stops at the desired slot (fetched from webapp)

V. CONCLUSION

To do the job of a good companion for an elderly, the robot has to take care of physical health as well as mental health i.e.,person's mood. The proposed system here takes care of their emotions by recognizing the emotion from the conversation and then it tries to improve the mood by playing some music,some trivia or some joke.

The robot does the job of taking careof the physical health by not only dispensing the medicines ontime scheduled by the caretaker but also keeping in check the person's conversation for any emergency word. If the word is being repeated by the elderly person, the robot catches the word and accordingly dispenses the emergency medicine, calls the caretaker, or calls the hospital/ambulance. Apart from this, all the reminders set by the caretaker like time for walk, exercise, etc are audio played. In all, this system servers as anempathetic caregiver.

FUTURE SCOPE

One of the three major aspect of emotion recognition, i.e., Facial expression can be integrated while training and used to get more accurate results/moods. Along with this the voice assistant could be modified further to send more contextually relevant replies. Using a better hardware model for pill dispenser (with more precise tablet dispensing) has future improvement scope.

REFERENCES

1. R. Yamazaki, S. Nishio, K. Ogawa and H. Ishigur,"Teleoperated android as an embodied communicationmedium: A case study with demented elderlies in a care facility," 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. [CrossRef]
2. J. Jiaetal., "Transformer-based Alarm Context- Vectorization Representation for Reliable Alarm Root Cause Identification in Optical Networks," 2021 European Conference on Optical Communication (ECOC), 2021, pp.1-4, doi: 10.1109/ECOC52684.2021.9606141. [CrossRef]
3. E. Kasthuri and S. Balaji, "A Chatbot for Changing Lifestyle in Education," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021 Rosario Sorbello, Antonio Chella, Carmelo Calí, Marcello Giardina, Shuichi Nishio, Hiroshi Ishigur [CrossRef]
4. Pereira, André Leite, Iolanda Mascarenhas, Samuel Martinho, Carlos Paiva, Ana. (2010). Using Empathy to Improve Human-Robot Relationships.Personal Assistant Vehicle Using Raspberry Pi and Zero UI Technology," 2021 6th International Conference for Convergence in Technology (I2CT)
5. Softić, J. B. Husić, A. Softić and S. Baraković, "Health Chatbot: Design, Implementation, Acceptance and Usage Motivation," 2021 20th International Symposium INFOTEH-JAHORINA (INFOTEH) [CrossRef]
6. Mohammad Monirujjaman Khan, Shahnoor Chowdhury Eshan (2020). Development of an e-commerce Sales Chatbot.
7. Shimada, M., Kanda, T. (2012). What is the appropriate speech rate for a communication robot? Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems [CrossRef]
8. Rosario Sorbello, Antonio Chella, Carmelo Calí, Marcello Giardina, Shuichi Nishioc, Hiroshi Ishiguro (2014).Telenoid android robot as an embodied perceptual social regulation medium engaging natural human–humanoidinteraction. [CrossRef]
9. Yamazaki, Ryuji Nishio, Shuichi Ogawa, Kohei Ishiguro, Hiroshi Matsumura, Kohei Koda, Kensuke Fujinami, Tsutomu. (2012). How Does Telenoid Affect the Communication between Children in Classroom Setting? Conference on Human Factors in Computing Systems. [CrossRef]
10. D. Sunehra, B. Jhansi and R. Sneha, "Smart Robotic Personal Assistant Vehicle Using Raspberry Pi and Zero UI Technology," 2021 6th International Conference for Convergence in Technology (I2CT) [CrossRef]
11. Mohanty, Subhadarshini; Mohapatra, Subasish; and Sahoo, Amlan (2022) "Speech Emotion Recognition System using Librosa for Better Customer Experience," Graduate Research in Engineering and Technology (GRET): Vol. 1: Iss.6, Article 7. [CrossRef]
12. Luca Iocchi, Emanuele Menegatti, Andrea Bonarini, Matteo Matteucci, (2013).Development of intelligent service robots. Intelligenza Artificiale [CrossRef]
13. Powers, A. D. I. Kramer, S. Lim, J. Kuo, Sau-lai Lee and S. Kiesler, "Eliciting information from people with a gendered humanoid robot," ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2020
14. Ishiguro, Hiroshi Nishio, Shuichi Chella, Antonio Sorbello, Rosario Balistreri, Giuseppe Giardina, Marcello calì, Carmelo. (2012). Perceptual Social Dimensions of Human-Humanoid Robot Interaction. [CrossRef]
15. Sorbello, Rosario Chella, Antonio Giardina, Marcello Nishio, Shuichi Ishiguro, Hiroshi. (2015). An Architecture for Telenoid Robot as Empathic Conversational Android Companion for Elderly People. Advances in Intelligent Systems and Computing. [CrossRef]
16. Ishiguro, Hiroshi Nishio, Shuichi Chella, Antonio Sorbello, Rosario Balistreri, Giuseppe Giardina, Marcello calì, Carmelo. (2013). Investigating Perceptual Features for a Natural Human-Humanoid Robot Interaction Inside a Spontaneous Setting. Biologically Inspired Cognitive Architectures. [CrossRef]

AUTHOR PROFILE

Varsha Ryali, is an undergraduate on the verge of completing her BE in Information Science Engineering in BMS College Of Engineering, Bangalore. Within the course of her degree, she has participated and won in hackathons and design thinking competitions. Her areas of interest lie majorly in Machine learning, Big data analytics, Natural language processing, Signal Processing and Data structures. She has presented her paper specializing in BDA at a conference. She hopes to contribute a lot more to research in this field and will pursue higher education in the same.

Sristi Kumari, was born 2001 and raised in Ranchi, Jharkhand. She was intrigued by technological breakthroughs as a teenager, and this fascination led to some early reading experience because she was drawn to publications about technology. Sristi is completing her BE in information science engineering at Bangalore's BMS College of Engineering. She is particularly interested in machine learning and deep learning, data structures, and algorithms as well as computer networks. She actively participated in UPAGRAHA, the BMSCE Satellite team. Over the course of her four years in engineering, She has worked on numerous mini projects. She aims to advance her technology knowledge by working for several businesses and eventually leading a tech team.

Syed Mahmood Nabeel Anwar, undergraduate completing final year from B.M.S College of Engineering, Bangalore. Department of Information Science and Engineering. Consistently striving to improve in various fields of study in my field. Areas of interest include cloud computing, computer networks, Robotics and web development. Have a publication . Experience of Full stack development. Actively participated in robotics from Carnegie mellon University, Qatar. Over the course of four years, have worked on various projects using numerous Technology stacks. deeply invested in achieving step-by-step in the field which have not succeeded yet. Aims to consistently strive and advance technological knowledge by operating in technical and business fields.

Dr. Sheela S V, holds B.E Computer Science & Engineering degree from Bangalore University in 1998 and M.Tech Computer Science degree from Visvesvaraya Technological University in 2004. She completed her Ph.D in Efficient and Novel Techniques for Eye Gaze Tracking and Recognition in 2013 from Visvesvaraya Technological University. Currently, she is Professor in the department of Information Science and Engineering, BMS College of Engineering. She has an experience of 21 years. Her areas of research is Human Computer Interaction, Machine learning, Biometrics and Pattern recognition. She has published about 20+ publications and 170+ Google scholar citations. She is a member of IEEE.