

# **Cover Page**

## **Multilingual Chatbot in Healthcare**

Project report submitted for  
**4<sup>th</sup> Semester Minor Project**

in  
**Department of ECE & DSAI**

By,  
**Parth Bhandakkar(211010241)**  
**Srijan Ratrey(211020454)**



**Dr. Shyama Prasad Mukherjee**  
**International Institute of Information Technology, Naya Raipur**  
**(A Joint Initiative of Govt. of Chhattisgarh and NTPC)**  
**Email: [iiitnr@iiitnr.ac.in](mailto:iiitnr@iiitnr.ac.in), Tel: (0771) 2474040, Web: [www.iiitnr.ac.in](http://www.iiitnr.ac.in)**

# **CERTIFICATE**

This is to certify that the project titled “Multilingual Chatbot in Healthcare” by  
“Parth Bhandakkar and Srijan Ratrey” has been carried out under my/our  
supervision and that this work has not been submitted elsewhere for a degree.

(Signature of Guide)

---

**Dr. Anurag Singh**  
**ASSISTANT PROFESSOR**  
**Department of ECE**  
**Dr. SPM IIIT-NR**  
**May,2023**

## **Declaration**

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

---

**Parth Bhandakkar**

**(211010241)**

---

**Srijan Ratrey**

**(211020454)**

**Date: 15/05/2023**

## Approval Sheet

This project report entitled “Multilingual Chatbot in Healthcare” by “Parth Bhandakkar and Srijan Ratrey ” is approved for 4<sup>th</sup> Semester Project.

(Signature of Examiner - I)

---

Name of Examiner -I

(Signature of Examiner - II)

---

Name of Examiner -II

---

Dr. Debanjan Das

Date: 15/05/2023 Place: Dr. SPM IIIT-NR

# Table of Contents

Title	Page No.
<b>TABLE OF CONTENTS.....</b>	<b>i</b>
<b>LIST OF FIGURES.....</b>	<b>ii</b>
<b>ABSTRACT.....</b>	<b>iii</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>1</b>
1. Research Area I .....	1
2. Research Area II.....	2
3. Research Area III.....	2
4. Challenges.....	2
<b>CHAPTER 3 PROPOSED SOLUTION</b>	<b>2</b>
1. Enhance Natural Language Processing(NLP).....	2
2. Incorporate Clinical Decision Support (CDS) .....	2
3. Ensure Data Privacy and Security.....	3
4. Use Patient Feedback to Improve the Chatbot.....	3
5. Integrate with Electronic Health Records (EHRs) .....	3
6. Provide Multilingual Support.....	4
7. Text to Speech feature.....	4
<b>CHAPTER 4 RESULTS</b>	<b>4</b>
1. Training Dataset.....	4
2. Model.....	5
3. Bot GUI .....	5
<b>CHAPTER 5 CONCLUSIONS</b>	<b>6</b>
<b>REFERENCES</b>	<b>6</b>

## List of All Figures

Figure No.	Figure Title	Page Number
3.1	Flowchart of NLP .....	4
4.1	Training Dataset.....	4
4.2	Preprocessing before training.....	5
4.3	Model used.....	5
4.4	Prototype GUI chatbot.....	5
4.5	Final GUI of chatbot.....	6
4.6	Query for treatment.....	6
4.7	Different interpretation.....	6
4.8	Hindi language support.....	6

# Multilingual Chatbot in Healthcare

## Abstract

This project uses natural language processing (NLP) techniques to develop a chatbot for the web. A chatbot is able to understand user queries and respond to them in a conversational manner, providing a more personalized and efficient experience for website visitors. The chatbot is trained on a dataset of frequently asked questions and uses techniques such as tokenization, sentiment analysis, and intent classification to understand and respond to user input. The ultimate goal of this project is to improve the user experience on the site by providing quick and accurate answers to common questions.

Index Terms—preprocessing, hyperparameters, *lemmatizer*

## I. INTRODUCTION

A chatbot is a software application used to conduct an online chat conversation via text or text-to-speech, instead of providing direct contact with a live human agent. Chatbots are computer programs that are able to maintain a conversation with a user in natural language, understand their intent and respond based on pre-set rules and data. Designed to convincingly simulate

the way a human would behave as a conversational partner, chatbot systems typically require constant tweaking and testing, with many people in production unable to converse adequately; in 2012, none of them passed the standard Turing test.

Chatbots are used in dialog systems for a variety of purposes, including customer service, routing requests, or gathering information. While some chatbot applications use extensive

others simply search for general keywords and generate answers using common phrases obtained from an associated library or database.

Most chatbots are accessed online through website pop-ups or via virtual assistants. They can be divided into usage categories that include: business (e-commerce via chat), education, entertainment, finance, health, news and productivity.

## II. LITERATURE REVIEW

Chatbots, also known as conversational agents or virtual assistants, have become increasingly popular in the medical field in recent years. They are designed to mimic human conversation and allow patients to communicate with healthcare providers in a more natural and efficient way. The following literature review provides an overview of recent studies on the use of chatbots in the medical field.

### A. Research Area I

One of the most promising areas for the use of chatbots is mental health. A study by Fitzpatrick et al. (2020) evaluated the effectiveness of a chatbot in individuals with anxiety and depression. The study found that the chatbot was also able to significantly reduce symptoms of anxiety and depression

to improve overall mental health and well-being.

### *B. Research Area II*

Another study by Yaghoubzadeh et al. (2019) evaluated the use of a chatbot in patients with chronic diseases such as diabetes and hypertension. The study found that the chatbot was effective in improving patient self-management, reducing hospitalizations and increasing patient satisfaction.

### *C. Research Area III*

Chatbots are also used in patient education and health promotion. A study by Sharma et al. (2020) evaluated the effectiveness of a chatbot for promoting healthy behavior in individuals at risk of cardiovascular disease. The study found that the chatbot was effective in increasing knowledge and awareness of cardiovascular risk factors as well as promoting healthy lifestyles.

In addition to patient care, chatbots have also been used for medical education and training. A study by Al-Ta'ar et al. (2020) evaluated the effectiveness of a chatbot for medical students in learning about clinical decision making. The study found that the chatbot was effective in improving students' clinical decision-making and knowledge retention skills.

### *D. Challenges*

However, there are also concerns about the use of chatbots in healthcare. One of the main concerns is patient privacy and confidentiality. A study by Abbasi et al. (2021) emphasized the importance of ensuring that chatbots comply with privacy regulations and that patients are aware of the risks and benefits of using them. Also, chatbots need to be trained

and built in different native languages so that users can easily access them.

Overall, the literature suggests that chatbots have the potential to improve patient care, education, and training in the medical field. However, further research is needed to fully understand their effectiveness and to address any privacy and confidentiality concerns.

## **.III. PROPOSED SOLUTION**

Some key troubleshooting steps are as follows  
—

### **A. Enhancing Natural Language Processing (NLP)**

Natural Language Processing (NLP) plays a crucial role in the development of chatbots. NLP enables chatbots to understand, interpret, and generate human-like responses in natural language, allowing them to engage in meaningful and effective conversations with users. Here are some key aspects of NLP for chatbot development:

1. **Text Preprocessing:** NLP techniques are applied to preprocess and clean the user input before it is fed into the chatbot. This typically involves tasks like tokenization, removing stopwords, stemming or lemmatization, and handling special characters or emojis. Preprocessing helps to standardize the input and extract relevant information for further processing.
2. **Intent Recognition:** Chatbots need to understand the intent behind user queries or statements. Intent recognition involves classifying user input into predefined categories or intents, such as "how to treat illness," "what is symptoms of," or "which disease is this"



This task is often accomplished using techniques like machine learning algorithms, rule-based approaches, or more advanced methods like deep learning and neural networks.

3. Named Entity Recognition (NER): NER is the task of identifying and extracting specific entities from user input, such as names of people, organizations, locations, dates, or product names. NER is crucial for chatbots to identify important information in user queries and generate accurate and contextually relevant responses.

4. Sentiment Analysis: Understanding the sentiment or emotion behind user messages is essential for chatbots to provide appropriate and empathetic responses. Sentiment analysis techniques are used to classify the sentiment of user input as positive, negative, or neutral. This helps chatbots adapt their tone, response style, and level of empathy based on user sentiment.

5. Language Understanding and Generation: NLP techniques enable chatbots to understand and generate human-like responses. Techniques like word embeddings, such as Word2Vec or GloVe, capture semantic relationships between words and help the chatbot understand the meaning of user input. Language models like GPT or BERT leverage deep learning to generate contextually relevant and coherent responses.

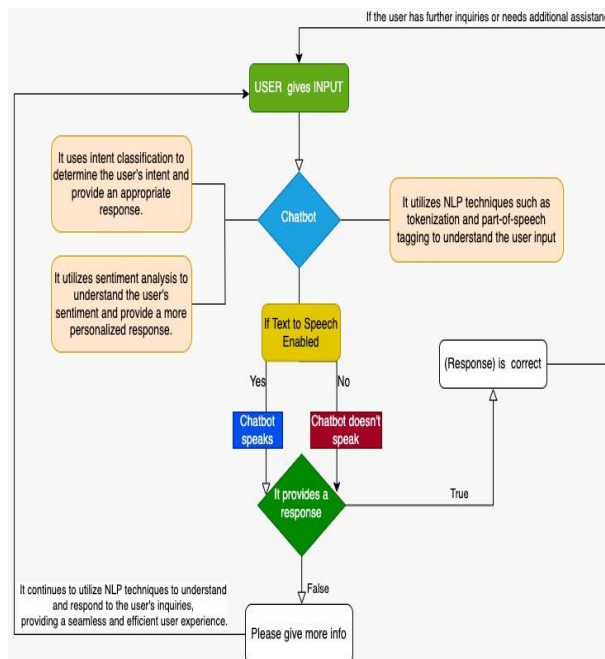
6. Contextual Understanding: Chatbots are often expected to handle multi-turn conversations, where the context of previous interactions is crucial. NLP techniques allow chatbots to maintain and utilize contextual information to provide more accurate and context-aware responses. Transformers models, such as GPT-3, have been particularly effective in capturing context and generating high-quality responses.

7. Knowledge Integration: Chatbots can be enhanced by integrating them with knowledge bases or external APIs to provide accurate and up-to-date information. NLP techniques help extract relevant information from these sources and incorporate it into the chatbot's knowledge to enrich its responses.

8. Error Handling and Misunderstanding: NLP techniques are also employed to handle errors, misunderstandings, and out-of-domain queries. Chatbots can be designed to detect and gracefully handle situations where user input is ambiguous, incomplete, or does not align with the chatbot's capabilities.

9. Continuous Learning and Improvement: NLP allows chatbots to continuously learn and improve over time. Feedback from user interactions can be used to update and refine the chatbot's models, enhancing its performance and making it more adept at handling a wide range of user queries and scenarios.

NLP is a rapidly evolving field, and ongoing research and advancements continue to drive improvements in chatbot capabilities. With the application of NLP techniques, chatbots can provide personalized, informative, and engaging interactions, creating more natural and effective communication channels between humans and machines.



## B. Incorporate Clinical Decision Support (CDS)

CDS can help chatbots provide more personalized recommendations to patients. For example, the chatbot can suggest different treatment options based on the patient's medical history and symptoms.

## C. Ensure Data Privacy and Security

Healthcare data is sensitive and should be treated with utmost care. The chatbot should comply with HIPAA regulations and ensure that patient data is protected.

## Training Dataset

We have managed to achieve training dataset from medical dictionary and trained our model based on the parameters. Also the GUI or website have also been made and is functional. The images of the same is inserted below-

We have used most used sentences in the field of healthcare which will help our chatbot to give better results.



Fig 4.1

The dataset was created as an intent.json file having numerous and vast medical vocabulary gathered from various trusted sources.

## Enhancements Done

### A. Use Patient Feedback to Improve the Chatbot

Collecting patient feedback can help identify areas of improvement for the chatbot. This feedback can be used to train the chatbot and enhance its effectiveness.

### B. Integrate with Electronic Health Records (EHRs)

Integrating the chatbot with EHRs can provide more accurate and personalized recommendations. The chatbot can access the patient's medical history and provide tailored advice based on that information.

### C. Provide Multilingual Support

Healthcare is a global issue, and chatbots should be able to communicate in different languages to be accessible to patients worldwide.

### D. Text to Speech (TTS) Feature

It is a technology that enables computers and digital devices to speak aloud text-based content in a natural-sounding voice. TTS technology is used in a wide variety of applications, including assistive technology for people with disabilities, language learning tools, and speech-enabled interfaces for smart home devices.

## IV. RESULTS

### Model

Using NLTK we have used lemmatizer to enhance model learning and understanding

Use of SGD as an optimizer have made model more efficient

```
17 for intent in intents[' intents ']:
18     for pattern in intents[ patterns ]:
19         #tokenize each word
20         w = nltk.word_tokenize(pattern)
21         words.extend(w)
22         #add documents in the corpus
23         documents.append((w, intent[' tag ']))
24
25 # add to our classes list
26 if intent[ tag ] not in classes:
27     classes.append(intent[ tag ])
28
29 # lemmatize and lower each word and remove duplicates
30 words = [lemmatizer.lemmatize(w.lower()) for w in words if w not in ignore_words]
31 words = sorted(list(set(words)))
32
33 # sort classes
34 classes = sorted(list(set(classes)))
35
36 # documents = combination between patterns and intents
37 print (len(documents), 'documents')
38
39 # classes = intents
40 print (len(classes), 'classes', classes)
41
42 # words = all words, vocabulary
43 print (len(words), 'unique lemmatized words', words)
44
45 pickle.dump(words, open('words.pkl', 'wb'))
46 pickle.dump(classes, open('classes.pkl', 'wb'))
47
48 # create our training data
```

Fig 4.2

```
17 random.shuffle(training)
18 training = np.array(training)
19 # create train and test lists, x = patterns, y = intents
20 train_x = list(training[ , 0])
21 train_y = list(training[ , 1])
22 print('training data created')
23
24 # Create model - 3 layers. First layer 128 neurons, second layer 64 neurons and 3rd output layer contains number of neurons
25 # equal to number of intents to predict output intent with softmax
26 model = Sequential()
27 model.add(Dense(128, input_shape=(len(train_x[0]),), activation='relu'))
28 model.add(Dropout(0.5))
29 model.add(Dense(64, activation='relu'))
30 model.add(Dropout(0.5))
31 model.add(Dense(len(train_y[0]), activation='softmax'))
32
33 # Compile model. Stochastic gradient descent with Nesterov accelerated gradient gives good results for this model
34 sgd = SGD(loss='categorical_crossentropy', optimizer=sgd, metrics=['accuracy'])
35 model.compile(loss='categorical_crossentropy', optimizer=sgd, metrics=['accuracy'])
36
37 # fitting and saving the model
38 hist = model.fit(np.array(train_x), np.array(train_y), epochs=200, batch_size=5, verbose=1)
39 model.save('chatbot_model.h5', hist)
40
41 print('model created')
```

Fig 4.3

### BOT GUI

The window for our chatbot appears as below which answers based on questions asked and it finds the answer via its database.

### Multilingual GUI

Healthcare is a global issue, and chatbots should be able to communicate in different languages to be accessible to patients worldwide.

We have multiple languages to translate the answers provided by the chatbot.

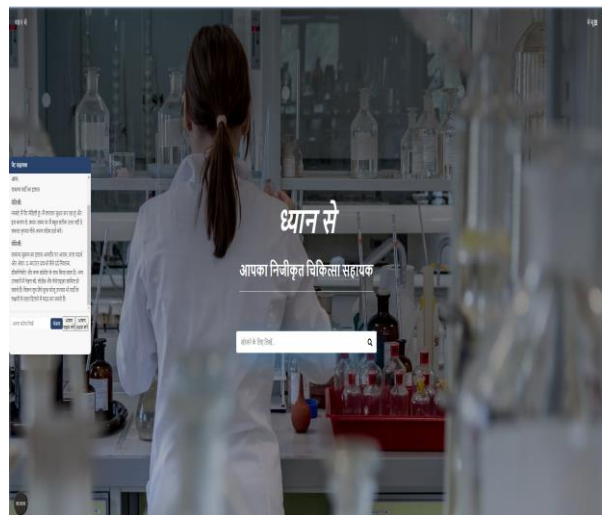


Fig 4.4

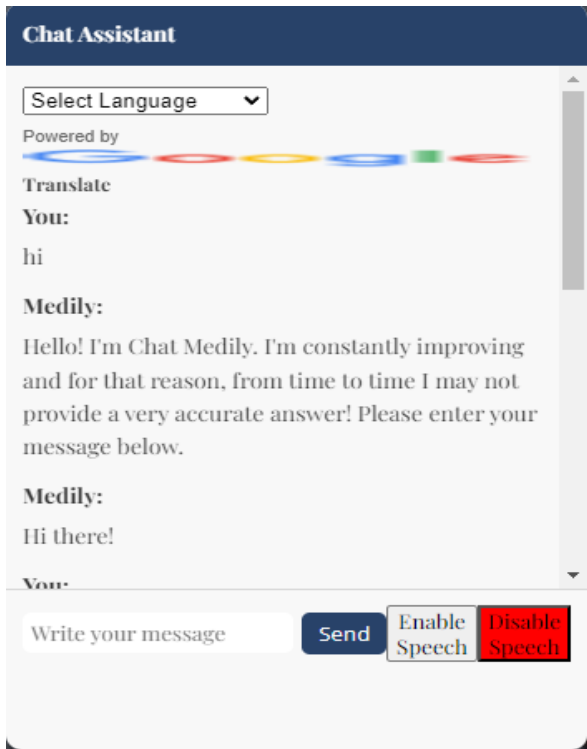


Fig 4.5

Responses given for specific query:

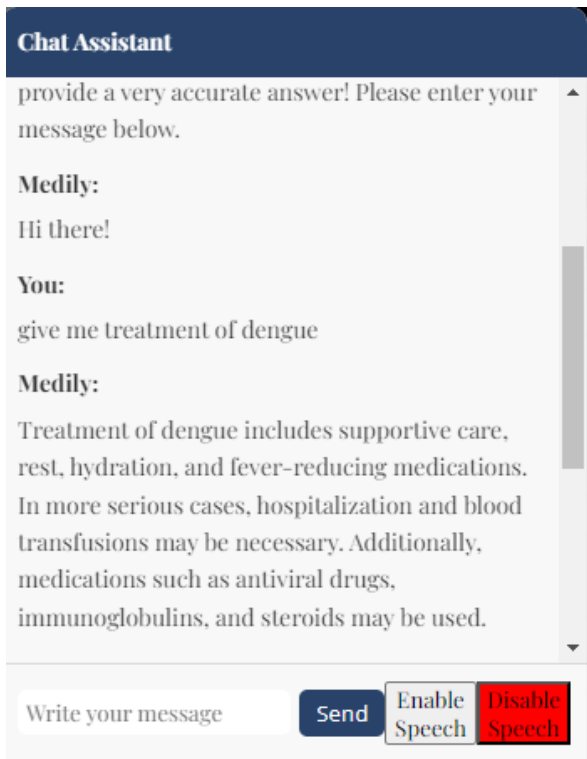


Fig 4.6

Can even answer if the query is given in different ways.

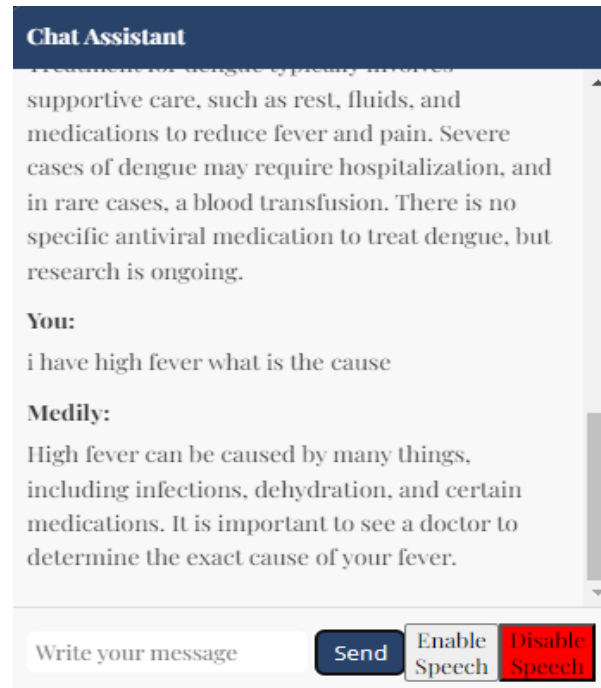


Fig 4.7

Can understand hindi language also.

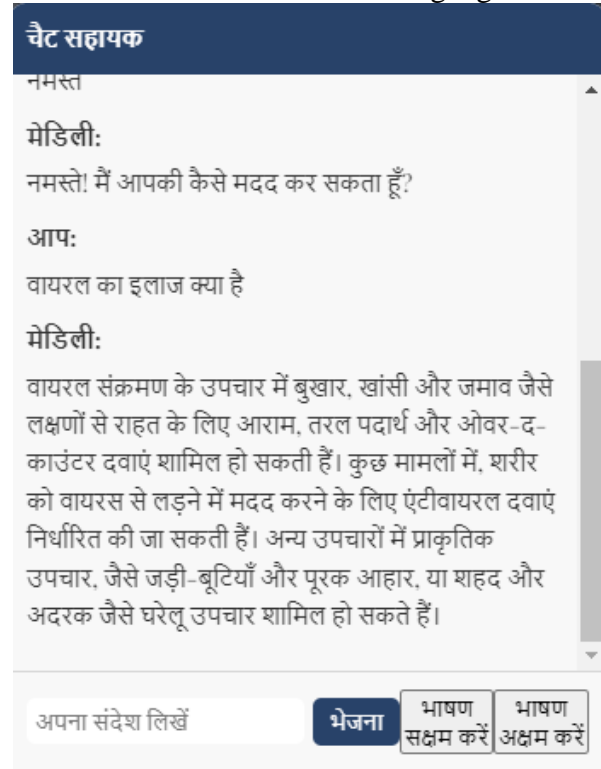


Fig 4.8

## V .CONCLUSION

In conclusion, chatbots have emerged as a game-changing technology in the healthcare industry. They have the potential to revolutionize the way patients interact with healthcare providers and access medical information. By leveraging natural language processing and machine learning, chatbots can provide personalized and timely assistance to patients, offer medical advice, schedule appointments, and even monitor patients remotely.

Moreover, chatbots can significantly reduce the workload of healthcare providers, allowing them to focus on more complex tasks and providing quality care to patients. With the increasing demand for healthcare services and shortage of healthcare professionals, chatbots can bridge the gap and provide reliable and consistent care to patients.

However, as with any emerging technology, there are concerns about the accuracy and privacy of medical information and the need for ethical guidelines for the development and deployment of healthcare chatbots. Therefore, it is essential to continue to monitor and evaluate the effectiveness of chatbots in

healthcare and establish proper regulations to ensure their safe and ethical use.

Overall, chatbots have immense potential to transform the healthcare industry and improve patient outcomes. As the technology continues to evolve, it is likely that we will see more sophisticated and advanced chatbots being developed, leading to even greater improvements in healthcare delivery.

## REFERENCES

1. Lu, Y., Chiang, T., & Zhang, G. (2019). Multilingual chatbot for healthcare based on deep learning. *IEEE Access*, 7, 120703-120711.
2. Abdullah, A., Zhang, W., Liu, Y., & Chen, H. (2020). Towards multilingual conversational agents in healthcare: An evaluation framework. *Journal of biomedical informatics*, 109, 103558.
3. Ribeiro, A., Gonçalves, R., Couto, F. M., & Lamurias, A. (2019). Multilingual chatbot for answering health-related questions. In *International Conference on Practical Applications of Agents and Multi-Agent Systems* (pp. 375-380). Springer, Cham.
4. Nwankpa, C., Jahan, A., & Kunapuli, A. (2018). Multilingual medical chatbot system for rural healthcare. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 321-326). IEEE.
5. Štajdohar, M., & Košmelj, K. (2021). Developing a multilingual chatbot for health-related purposes. *Journal of Healthcare Informatics Research*, 5(2), 118-132.
6. Krishna, A. (2021). A Review on Multilingual Healthcare Chatbots: A Case Study of 15 Indian Languages. *Journal of Advanced Research in Dynamical and Control Systems*, 13(6), 358-364.