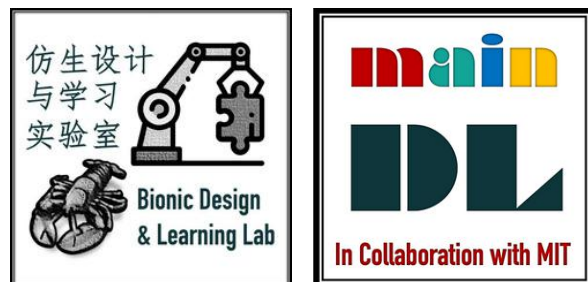


# Intention Recognition Robot Arm

Present Name 1: 12112816 张毅恒 | Present Name 2: 12112828 黄孜为  
Present Name 3: 12110521 苟国涛 | Present Name 4: 12110132 张宸翰  
Present Name 5: 12110128 封德凯 | Present Name 6: 12110303 张元卓

Group: 8

Date: 2024.3.23



AncoraSIR.com



SUSTech  
Southern University  
of Science and Technology

# Intention Recognition Robot Arm

1) Our goal is to design an assistive system for disabilities. It integrates multiple recognition means, including recognizing actions using IMU, recording keywords with speech recognition, and image recognition to identify gestures and object positions, for recognizing the disabled person's intention and performing the grasping of the corresponding target. We are interested in helping the disabled to improve their life.

2) The readings we depend are as follows:

IMU intent recognition: which illustrates how to determine grasp selection from arm trajectories via deep learning to enable functional hand movement in tetraplegia

Speech recognition: A review on speech recognition technique

Image recognition: Research and Application of Deep Learning in Image Recognition

# Intention Recognition Robot Arm

## 3) The data we will use:

For the voice data used for training, we will collect the voice commands of our team members in the way of recording. For the motion trajectory data, we plan to fix the IMU on the human wrist and use the posture information collected during the wrist movement as the feature data. For gesture data, we will collect the finger joint angle data through the mediapipe pack and send it to the machine learning model as feature data.

# Intention Recognition Robot Arm

4) The methods we will use including:

- a) IMU Intent Recognition: LSTM (long short-term memory networks)
- b) Camera Intent Recognition: Supervised Learning methods, such as Random Forest, SVC.....
- c) Camera Object Recognition: CNN (Convolutional Neural Network) such as VGG, ResNet, YOLO (You only look once)
- d) Voice Intent Recognition: LSTM

5) Finally, we will evaluate as follows:

- a) Confusion Matrix
- b) Accuracy

# What is the problem that you will be investigating?

---

- In our project, we will design an assistive system for people with disabilities. It integrates multiple recognition means, including recognizing actions using IMU, recording keywords with speech recognition, and image recognition to identify gestures and object positions, for recognizing the disabled person's intention and performing the grasping of the corresponding target.
- We want people with disabilities to be able to use our system in a natural and intuitive way to improve their quality of life.

# What reading will you examine to provide context and background?

---

IMU intent recognition:

(Determining grasp selection from arm trajectories via deep learning to enable functional hand movement in tetraplegia)

In this paper, the patient's wrist movement information is collected by inertial sensors and the motion trajectories are classified using the dynamic time warping (DTW) algorithm to identify the patient's intention.

# What reading will you examine to provide context and background?

---

Speech recognition:  
(A review on speech recognition technique)

This paper introduces the advantages and disadvantages of various speech recognition technologies, which helps us to choose the speech recognition technology suitable for our project.

# What reading will you examine to provide context and background?

---

Image recognition:  
(Research and Application of Deep Learning in Image Recognition)

This paper introduces three deep learning models in image recognition and compares their advantages and disadvantages, which is helpful for our project to choose a suitable image recognition algorithm.



# What data will you use?

*If you are collecting new data, how will you do it?*

- For the voice data used for training, we will collect the voice commands of our team members in the way of recording, and send the MP3 file to the machine learning model as feature information
- For the motion trajectory data, we plan to fix the IMU on the human wrist and use the posture information collected during the wrist movement as the feature data
- For gesture data, we will collect the finger joint angle data through the mediapipe pack and send it to the machine learning model as feature data

# What method or algorithm are you proposing?

*If there are existing implementations, will you use them, and how? How do you plan to improve or modify such implementations? You don't have to have an exact answer at this point, but you should have a general sense of how you will approach the problem you are working on.*

- IMU Intent Recognition
  - LSTM( long short-term memory networks)
- Camera Intent Recognition
  - Supervised Learning methods, such as Random Forest, SVC.....
- Camera Object Recognition
  - CNN (Convolutional Neural Network) such as VGG、 ResNet.
  - YOLO (You only look once)
- Voice Intent Recognition
  - LSTM

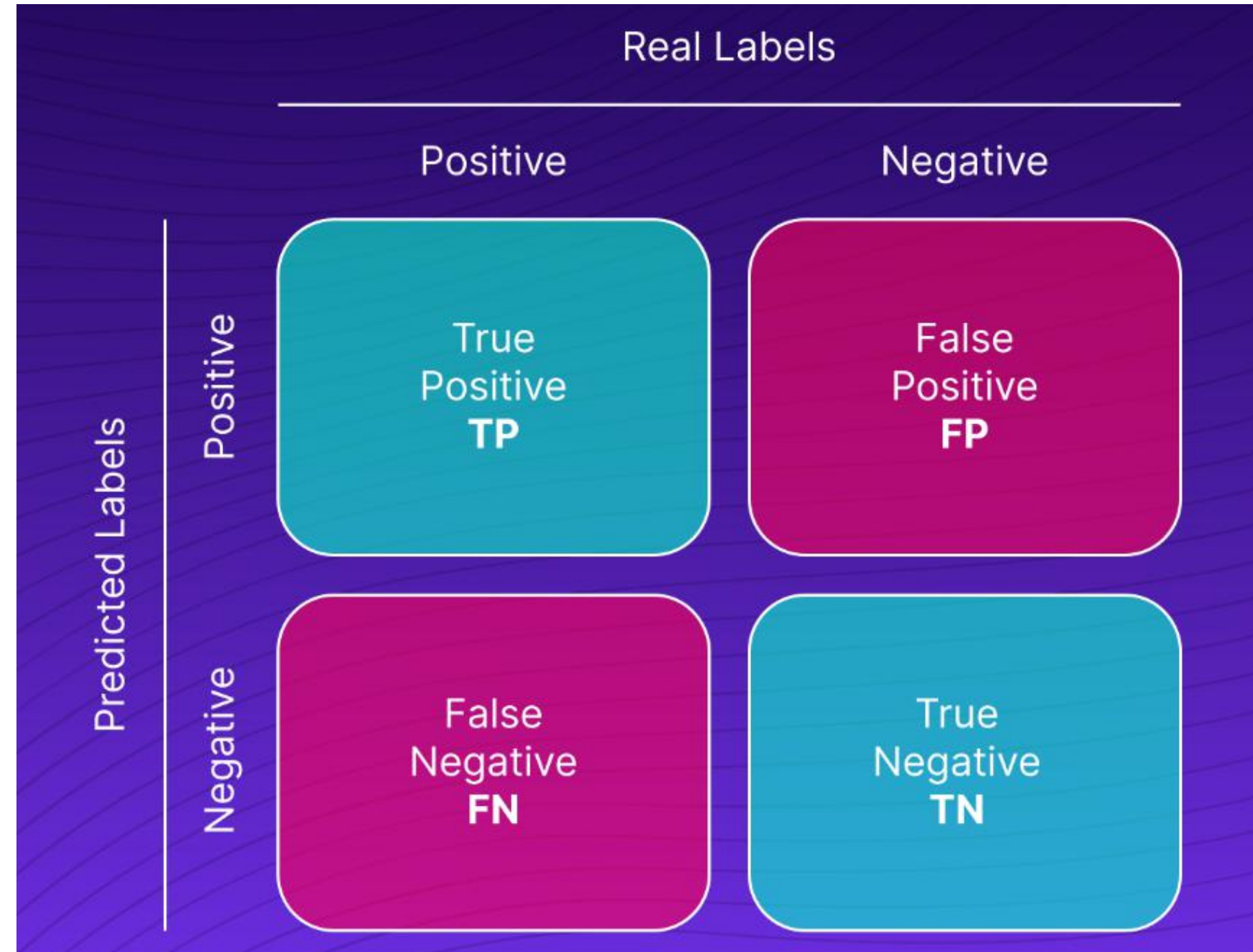
# How will you evaluate your results?

*Qualitatively, what kind of results do you expect (e.g., plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results (e.g., what performance metrics or statistical tests)?*

- Confusion Matrix
- Accuracy

# How will you evaluate your results?

- Confusion Matrix
- The confusion matrix is a square matrix where each row represents the instances in a predicted class, and each column represents the instances in an actual class. It helps us understand how well our model is performing in terms of correctly and incorrectly classifying instances.



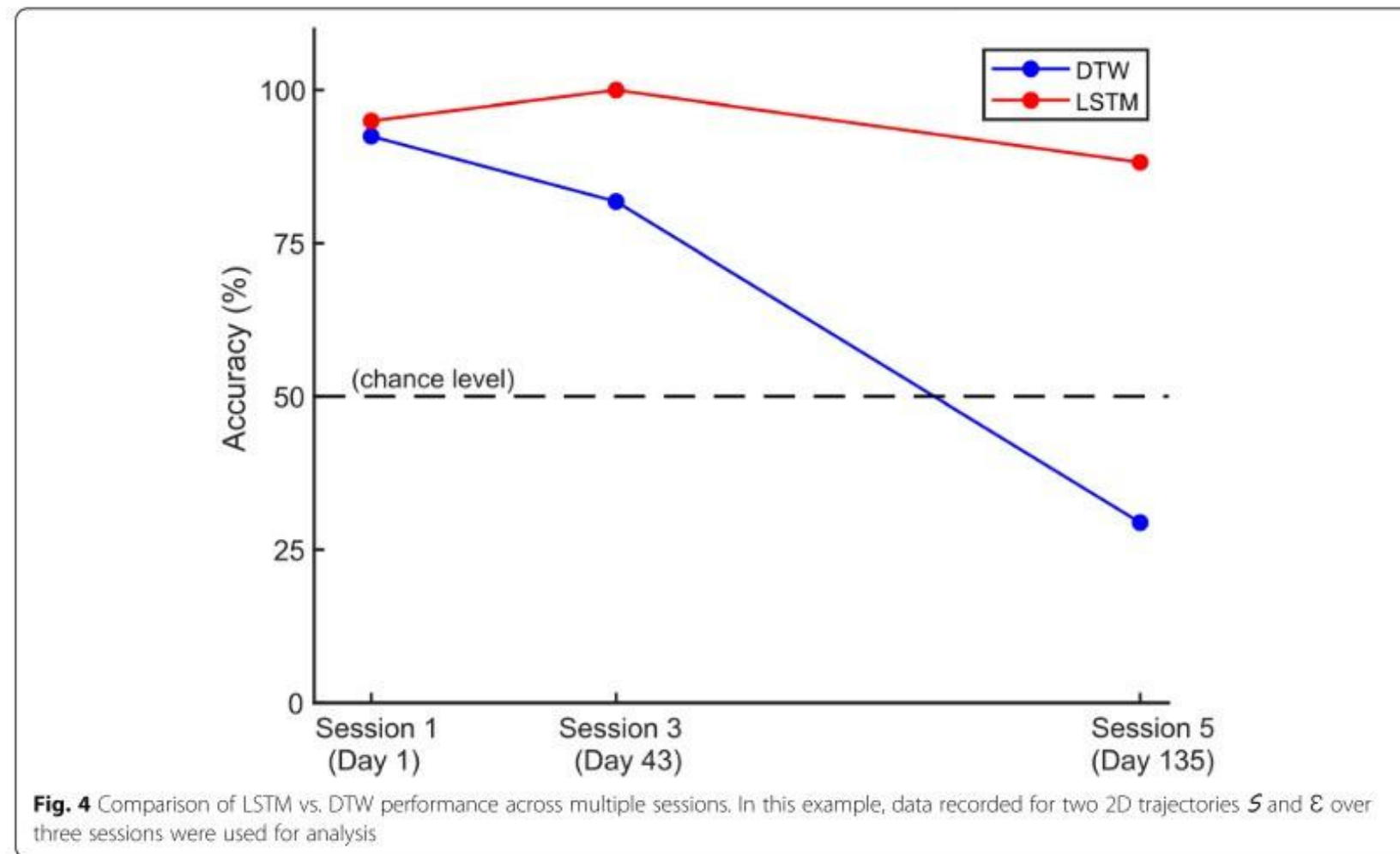
# How will you evaluate your results?

- Accuracy

Accuracy=Total Number of Predictions/Number of Correct Predictions

In the context of a confusion matrix:

$$\text{Accuracy}=(TP+TN)/(TP+TN+FP+FN)$$



# Thanks for your listening!

Group 8: Intention Recognition Robot Arm



AncoraSIR.com

