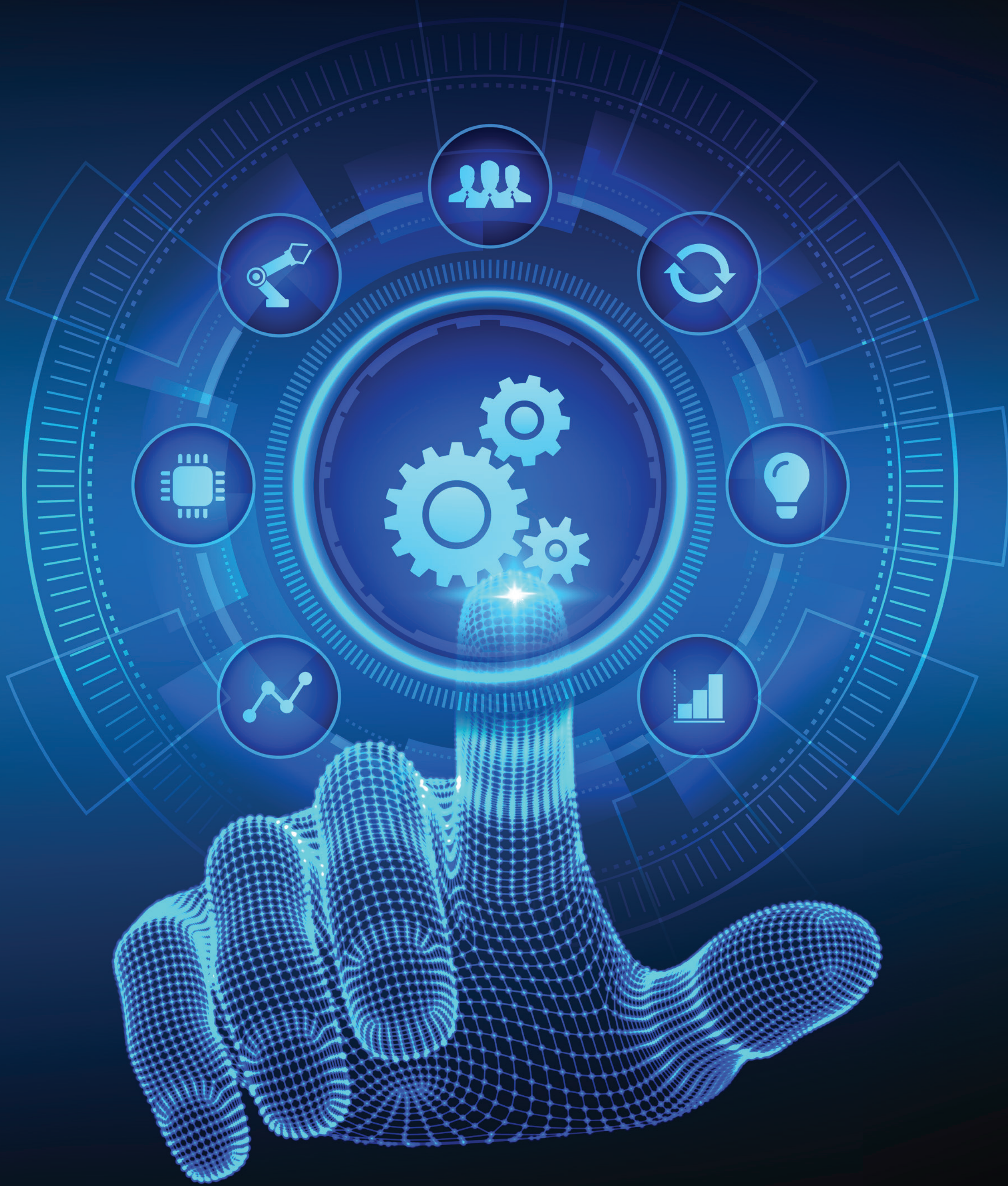
EXPRESSIVE MOTION: HARNESSING RANDOM FOREST FOR GESTURE-DRIVEN VOCALIZATION

Facilitating easy voice/text communication for the hearing- and speech-impaired



CONTENTS

ABSTRACT

Communication is an essential survival skill for human beings, but those with speech and hearing disabilities face difficulties in communicating with individuals without disabilities. Sign language, which uses hand gestures and motions, is widely used by speech- and hearing-impaired individuals to communicate, yet a huge communication barrier persists.

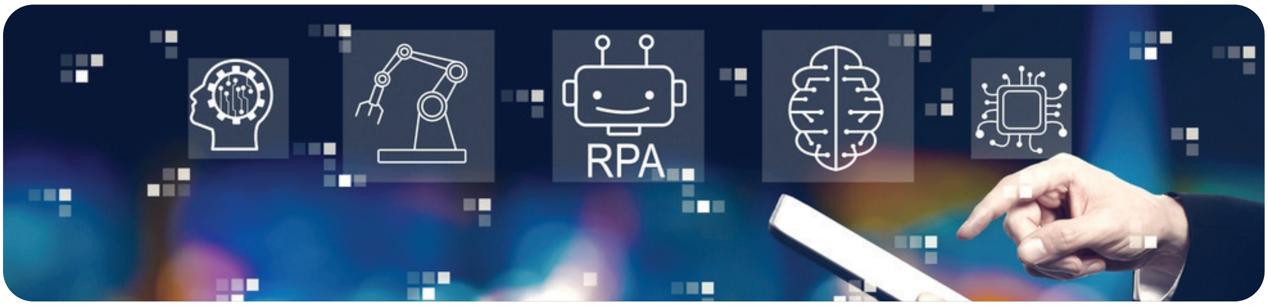
To address this issue, we propose a system that uses a hand glove gesture vocalizer that can convert hand gestures into specific sounds. The glove uses flex sensors to recognize gestures, which are then converted into speech/visuals on an LCD display. The project, discussed in this white paper, aims to present a prototype based on the random forest algorithm and flex sensors.

The project's objective is to create an affordable system that will enable the speech- and hearing-impaired to communicate without barriers. Gestures or signs is one of the basic forms of communication that humans have been using for centuries. Gestures and signs can convey emotions, thoughts, and ideas in a manner that transcends the limitations of spoken language. The best kinds of gesture communication used in history are emblems, illustrators, affect displays, regulators and adaptors.

INTRODUCTION

Speech and hand gestures are among the primary means humans use to communicate with each other. Communication, however, can become challenging for people with speech and hearing impairment. While the latter use sign language to communicate, it is difficult for common people to understand.

This white paper presents a project to simplify communication for the differently abled by translating gestures and signs into text or sounds on an LCD display. This is done by means of a data glove worn by an individual, who proceeds to use the hand to make gestures or signs to communicate. The data glove is fitted with flex sensors along each finger. These sensors detect the bending of the fingers, with each degree of bending resulting in a change in resistance that is converted into voltage. The system is designed to recognize these specific changes in voltage corresponding to different gestures and convert them into voice or text.



A gesture in sign language comprises movement of fingers and tilting of the hand in specific directions with specific angles. Sign language consists of different signs for different words. The proposed idea involves using these flex sensors to translate sign language into speech according to the bending of fingers, facilitating communication between the hearing- and speech-impaired and those who can hear and speak normally.

An ATmega328 controller interfaces this gesture-to-voice/text communication. For people who cannot hear, the converted text is displayed on an LCD screen. For those who cannot see, the system can be designed such that it provides speech output to communicate the message. This dual-mode output ensures that both the hearing-, speech- and visually impaired can benefit from the device. Additionally, the system can be dynamically reconfigured to work as a "smart device" that can adapt to all kinds of sign languages.

RANDOM FOREST ALGORITHM

Random forest is a powerful machine learning algorithm used for both classification and regression tasks. It is an ensemble learning method that involves combining multiple decision trees to create a more robust and accurate model.

The fundamental idea behind random forest is to generate multiple decision trees, each based on a different subset of training data, and then combine their predictions to make a final decision. Random forest is an extension of the decision tree algorithm, which involves creating a tree-like structure of nodes that split the data based on specific features until a prediction is made at the leaf nodes. While decision trees are effective at capturing complex relationships in the data, they are prone to overfitting-which means they may perform well on the training data but generalize poorly to new, unseen data. Random forest addresses this problem by generating multiple decision trees, each based on a random subset of the features and training data. This randomness helps to reduce overfitting and improve the model's ability to generalize to new data.

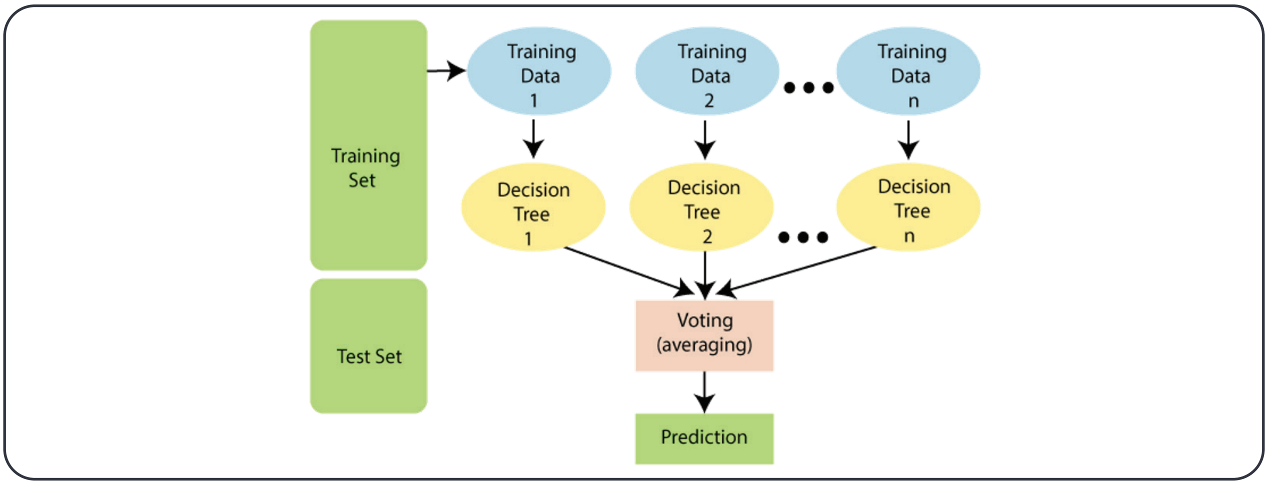


Fig.1 Random forest algorithm

SIMULATION AND RESULTS

The circuit for the data glove is connected using the Tinker Cad Software. The code is programmed into an Arduino, and based on this code, outputs are displayed on an LCD screen.

The data glove is designed with flex sensors attached along each finger. These sensors detect the bending of the fingers by changing resistance when the fingers are bent. When a person moves their middle finger to a certain extent, the flex sensor attached to the finger detects this movement. The bending of the finger alters the resistance in the flex sensor, producing a corresponding voltage. This change in voltage is then interpreted by the system, which translates it into a specific command. This command is displayed on the LCD screen and simultaneously vocalized, enabling both the hearing- and visually impaired to receive the information.

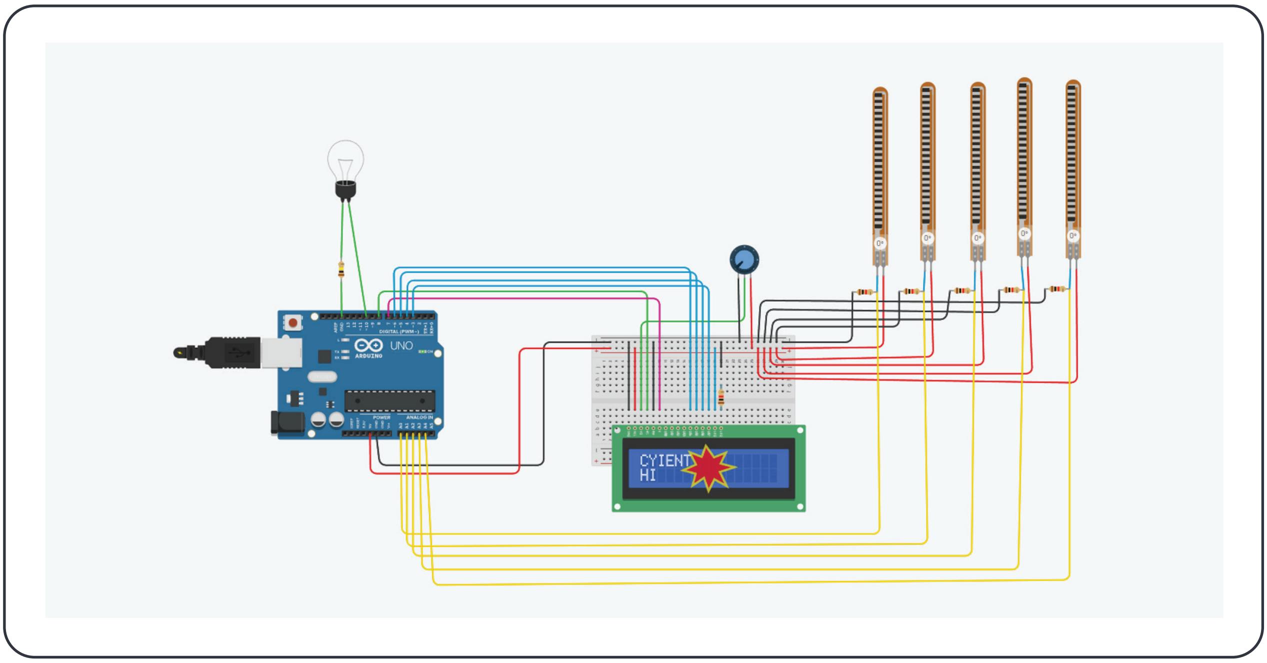
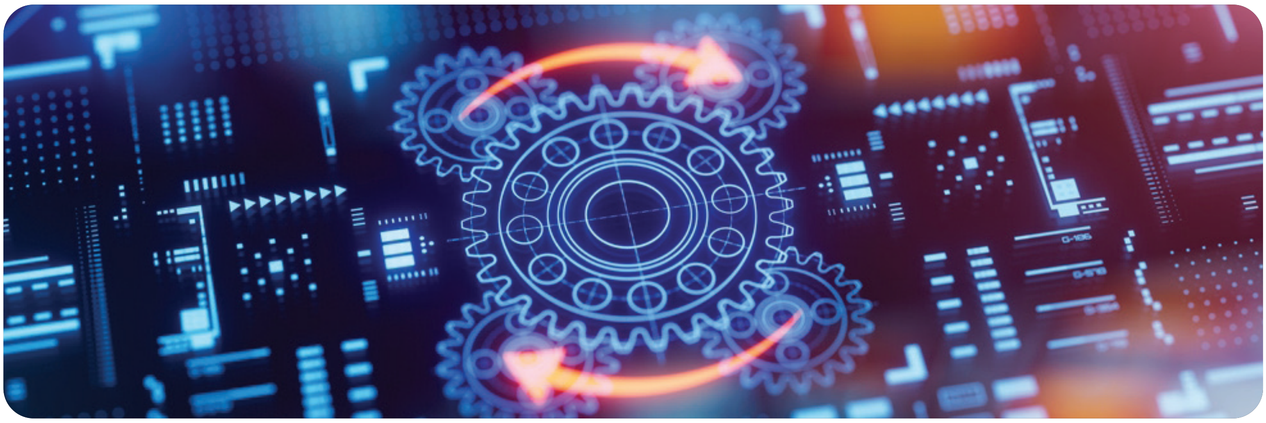


Fig.2 Circuit diagram



Implementation

Create a training set by using any programming methodology. For example: In this case Python

Install the required software (Jupyter) into your system

Import the required modules, libraries on to the file

Train the model using the data from the data table and form the training set

The algorithm will construct various decision trees for every training data set

Take some random data and test that against the trained model.

The simulation process of the prototype is done using programming language. Now using the results obtained we create a data table. Using the data table, we train the Arduino using the random forest algorithm.

Training Process

1. Import the libraries
2. Import the dataset
3. Put Feature variable to X and Target variable to Y
4. Import random forest classifier and fit the data
5. Do hyperparameter tuning for random forest using GridSearch CV and fit the data
6. Visualize
7. Sort the data based on feature importance

Graphical Representation

The main elements that can be shown in the result section of the software model of the Gesture Vocalizer using the random forest algorithm are:

- 01 **Feature importance**
- 02 **Confusion matrix**
- 03 **ROC curve**
- 04 **Precision-recall curve**
- 05 **Cross-validation results**

FEATURE IMPORTANCE

Random forest models can provide information about the importance of different features in predicting the target variable. A graph showing the relative importance of different features is useful in understanding which features are most important for the model's performance.

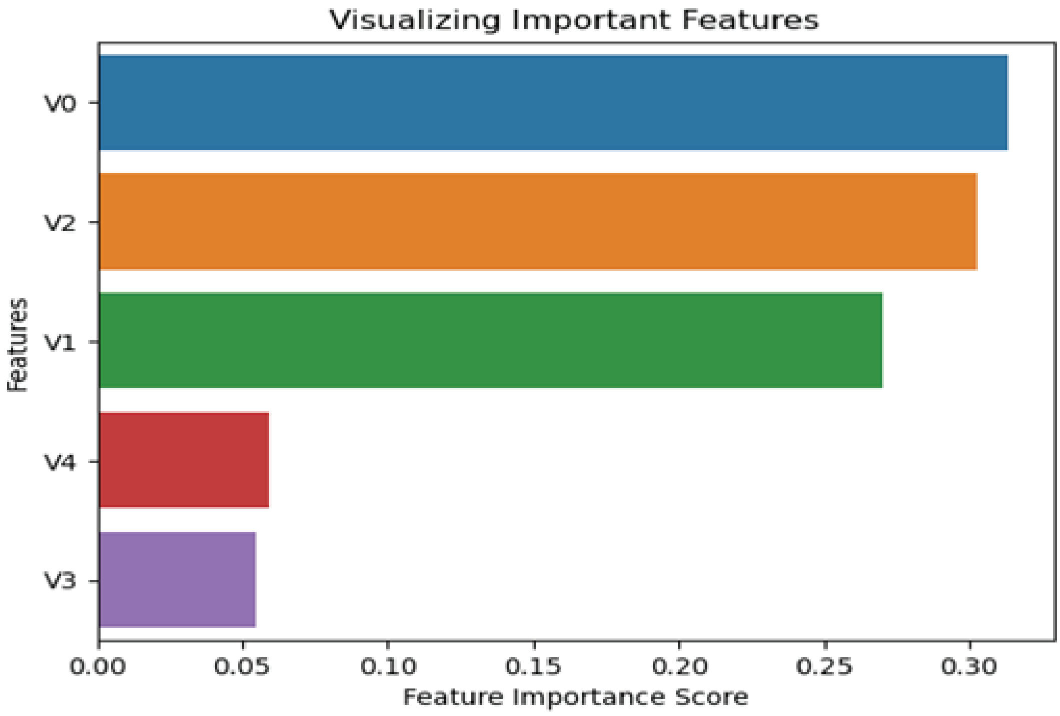


Fig.3 Feature importance

CONFUSION MATRIX

A confusion matrix is used to evaluate the performance of the model in predicting the correct gesture based on the vocal input. This matrix shows the number of true positives, true negatives, false positives, and false negatives, and can help to evaluate the accuracy, precision, recall, and F1-score.

```
array([[2, 0, 0, 0, 0, 0, 0, 0],
       [0, 4, 0, 0, 0, 0, 0, 0],
       [0, 0, 7, 0, 0, 0, 1, 0],
       [0, 0, 0, 5, 0, 0, 0, 0],
       [0, 0, 0, 0, 1, 0, 0, 0],
       [0, 0, 0, 0, 0, 3, 0, 0],
       [0, 0, 0, 0, 0, 0, 1, 0],
       [0, 1, 0, 0, 0, 0, 0, 2]], dtype=int64)
```

Fig.4 Confusion matrix

ROC CURVE

The ROC (receiver operating characteristic) curve can be used to evaluate the performance of the model in differentiating between various classes. This graph plots the true positive rate against the false positive rate at different threshold levels, and can help to evaluate the overall performance.

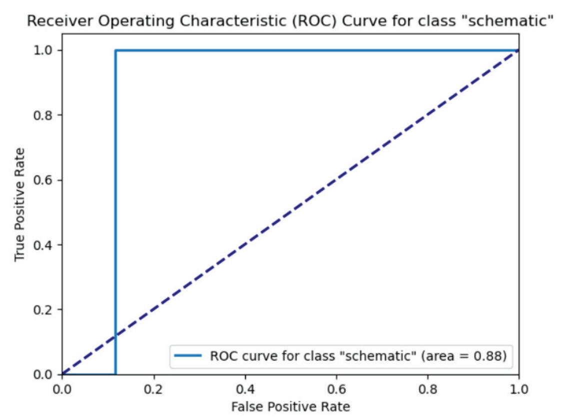


Fig.5 ROC curve

True Positive Rate (TPR), also known as Sensitivity or Recall, is calculated using the formula

$$\mathrm{TRP} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

Where:

- TP is the number of true positives (instances correctly predicted as positive)
- FN is the number of false negatives (instances incorrectly predicted as negative)

False Positive Rate (FPR) is calculated using the formula:

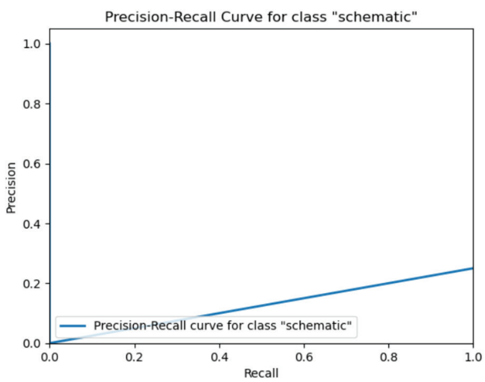
$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

Where:

- FP is the number of false positives (instances incorrectly predicted as positive)
- TN is the number of true negatives (instances correctly predicted as negative)

Precision-recall curve

The precision-recall curve is used to evaluate the performance of the model in predicting positive cases. This graph plots precision (the fraction of true positive predictions among all positive predictions) against recall (the fraction of true positive predictions among all actual positive cases) at different threshold levels.



CROSS-VALIDATION RESULTS

Cross-validation is used to evaluate and report the performance of the model. This could include the mean accuracy, precision, recall, or F1-score across different folds of the data.

```
[[2 0 0 0 0 0 0 0]
 [0 4 0 0 0 0 0 0]
 [0 0 7 0 0 0 1 0]
 [0 0 0 5 0 0 0 0]
 [0 0 0 0 1 0 0 0]
 [0 0 0 0 0 3 0 0]
 [0 0 0 0 0 0 1 0]
 [0 1 0 0 0 0 0 2]]
              precision    recall  f1-score   support

      cyient       1.00      1.00      1.00         2
  electrical       0.80      1.00      0.89         4
    engineer       1.00      0.88      0.93         8
   hyderabad       1.00      1.00      1.00         5
   manikonda       1.00      1.00      1.00         1
        sana       1.00      1.00      1.00         3
   schematic       0.50      1.00      0.67         1
      subash       1.00      0.67      0.80         3

    accuracy                           0.93        27
   macro avg       0.91      0.94      0.91        27
weighted avg       0.95      0.93      0.93        27

0.9259259259259259
```

Random Forest Algorithm

```
[[2 0 0 0 0 0 0 0]
 [0 2 0 0 0 0 2 0]
 [0 0 7 0 0 0 1 0]
 [0 0 0 5 0 0 0 0]
 [0 0 0 0 1 0 0 0]
 [0 0 0 0 0 3 0 0]
 [0 0 0 0 0 0 1 0]
 [0 0 0 0 0 0 0 3]]
              precision    recall  f1-score   support

      cyient       1.00      1.00      1.00         2
  electrical       1.00      0.50      0.67         4
    engineer       1.00      0.88      0.93         8
   hyderabad       1.00      1.00      1.00         5
   manikonda       1.00      1.00      1.00         1
        sana       1.00      1.00      1.00         3
   schematic       0.25      1.00      0.40         1
      subash       1.00      1.00      1.00         3

    accuracy                           0.89        27
   macro avg       0.91      0.92      0.88        27
weighted avg       0.97      0.89      0.91        27
```

Decision Tree Algorithm

Comparison of the random forest and decision tree algorithms results reveals that the accuracy is higher in random forest for the same data set.

This white paper therefore holds that **the random forest algorithm is more efficient than previous algorithms such as decision tree, etc.**

CONCLUSION

The gesture vocalizer is an easy-to-use hand glove device designed to help the hearing impaired and non-verbal individuals communicate with the rest of the world using acoustics or sounds. The vocalizer receives a gesture as input and provides text output on an LCD display. Hand gestures are more crucial than other types of gestures (arm, face, head, and body) as they swiftly reflect the user's perception. People without disabilities rarely learn sign language, making communication with the differently abled community difficult, further isolating them as a result.

Cyient's gesture vocalizer will help to better connect the differently abled community to the world and make their lives easier with the ability to "talk" easily, effectively, and accurately.

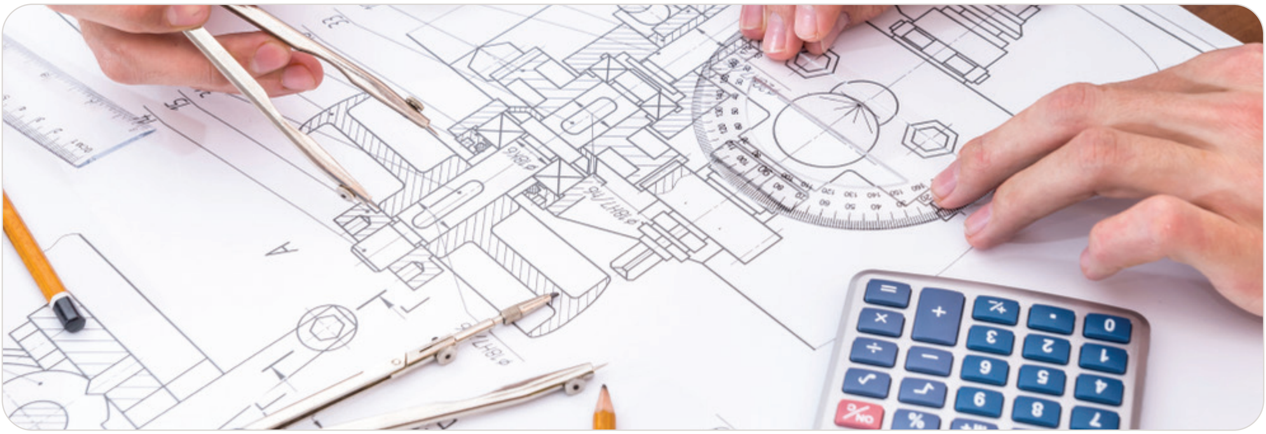
ABOUT THE AUTHORS

Madhavi Sanapala specializes in electrical schematics and system design. After six months of training as a schematic engineer, she now works as a reliability engineer in the aviation industry at Cyient. With a curious mind and drive for excellence, continuously contributing to advancements in electrical engineering.

Midatana Jahnavi excels in creating detailed electrical diagrams. After six months of training as a schematic engineer, she now works as a technical author in the aviation industry at Cyient. With a curious mind and drive for excellence, continuously contributing to advancements in electrical engineering.

Dake Anvesh brings expertise in system design and analysis. After six months of training as a schematic engineer, he now works as a technical author in the aviation industry at Cyient. With an inquisitive mind and a dedication to achieving greatness significantly contributing to industry advancements with his innovative solutions.

Sanikommu Srinivasulu excels in creating detailed electrical diagrams. After six months of training as a schematic engineer, he now works as a technical author in the Healthcare&Life Sciences (India) industry at Cyient. With a keen mind and a relentless drive for excellence significantly contributing to industry advancements with his innovative solutions.



ABOUT CYIENT

Cyient (Estd: 1991, NSE: CYIENT) partners with over 300 customers, including 40% of the top 100 global innovators of 2023, to deliver intelligent engineering and technology solutions for creating a digital, autonomous, and sustainable future. As a company, Cyient is committed to designing a culturally inclusive, socially responsible, and environmentally sustainable Tomorrow Together with our stakeholders.

For more information, please visit www.cyient.com

Follow us on:

CONTACT US

North America Headquarters
Cyient, Inc.
99 East River Drive
5th Floor
East Hartford, CT 06108
USA
T: +1 860 528 5430
F: +1 860 528 5873

Europe, Middle East, and Africa Headquarters
Cyient Europe Limited
Apex, Forbury Road,
Reading
RG1 1AX
UK
T: +44 118 3043720

Asia Pacific Headquarters
Cyient Limited
Level 1, 350 Collins Street
Melbourne, Victoria, 3000
Australia
T: +61 3 8605 4815
F: +61 3 8601 1180

Global Headquarters
Cyient Limited
Plot No. 11
Software Units Layout
Infocity, Madhapur
Hyderabad - 500081
India
T: +91 40 6764 1000
F: +91 40 2311 0352