EEE F343 - Power Electronics

Mini Project

Submitted in partial fulfillment of

Project of Power Electronics

By

Gargi Gupta—2016A3PS0288H

Akhilesh Aluvala—2016A3PS0281H

Jaju Vedant Vinod—2016A3PS0303H

Rishi Radia—2016A3PS0285H

Shreyas Kirtane—2016A3PS0314H

Tarunbir Singh—2016A3PS0301H

Supervisor

Dr. Sudha Radhika

Department of Electrical Engineering, BITS Pilani-Hyderabad Campus

*Abstract Fault detection of Induction motors using the current signals is a growing technology to detect the fault of an induction motor. Definite harmonic signals of the line current are located by a popular method known as Motor Current Signature Analysis. The actual fault detection by using the human involvement is widely replaced by the automated technology, like an on-line condition monitoring approach. Here we implement a real-time computer-aided advanced diagnosis system comprising of feature extraction, feature selection and feature classification.

Introduction Determination of health of an induction motor is a growing technology in todays world. Using Artificial neural networks and machine learning algorithms can automate the process of fault detection in an induction motor. The feature extraction of data which is being obtained in real time from machine can be used to determine the state of machine. This data is transmited from the motor to the ground station for the fault detection and give feedback to the motor controller if necessary

Methodology

I. Data Aquisition

Data acquisition (DAQ) is the process of measuring any phenomenon such as voltage, current etc. using computer. A Data Acquisition System includes sensors, DAQ measurement hardware, and a Personal Computer which has programmable software. PC-based DAQ systems use the connectivity capabilities, processing power of high-level computers which provides powerful and cost-effective measurement solution DAQs are used to transmit data from the sensors to raspberry using wired connection.

II. Data Processing

Based on sampling frequency ,the data is processed. After that, data features are extracted and transmit relevant features through wireless connection. Here DAQ sampling frequency is set to 10 kHz and three current measures at this frequency. Transmitting data at such high frequency through wireless medium is not economic so data processing is required. So the data is processed such that it is useful and sufficient enough for our machine learning model to identify the fault at the same time it is capable of transmission through Zigbee. Zigbee has its own practical limit to transmit data from one end to other.
To identify the required features feature selection is performed. Out of 10000 data points, features of every 1000 data points is formed. Following statistical features for training models were used to select features.

a) Mean($\overline{x}$) : The arithmetic mean (or simply mean) of a sample , is the sum of the sampled values divided by the number of items in the example.

b) Standard Deviation(σ) : The Standard Deviation is a measure of how spread out numbers are.Its symbol is (the greek letter sigma).

c) Skewness: Skewness essentially measures the relative size of the two tails.

d) Shape factor: It is the value that is affected by an object's Shape but is independent of its dimension.

e) Impulse factor: It is the maximum value divided by the mean of the absolute values of data entries.

f) Margin Factor: It is the maximum value divided by the square of the mean of the absolute values of data entries.

g) Current: current source of I1,I2,I3 are considered as 3 separate categorical features.

Apart from this **max value** ,**min value** and **pk-pk** value are also used as features.

Feature selection is done on basis of F score of Analysis of Variance test.(ANOVA test). It assigns score to every feature and best k features are selected on that basis.

Feature selection can also be done by selecting all possible subsets of features and then measuring the performance on basis of some metric such as accuracy, precision F1 score, recall, etc. For skewed data set we keep F1 measure is generally used while for balanced data set we use accuracy. This is computationally difficult due to exponential rise in complexity with increase in number of features. Feed forward or feed backward algorithms are also used which are greedy approaches and do not guarantee that best subset would be selected.Feed forward algorithm starts with one feature, measures the performance based on some fixed performance parameters and the one with the best performance is selected and then features are added one by one to that set. This is a greedy approach.

III. Classification Models

Artificial Neural network, Support vector Machine and random forest models are used for Fault Detection and Diagnosis.

a) Support Vector Machine: If a plot of two label classes on a graph is given and separator line for the classes is to be drawn , then this algorithm is useful. This is a learning model which uses learning algorithms to analyze data is used for regression analysis. It assigns new examples to the training data set which belongs to one category or other.This model represents examples to the points in space such that examples of different categories are separated by a clear and wide gap.And based on which side of the gap they fall, new examples are assigned their categories.

b) Random Forest Model: This is one of the most popular supervised machine learning algorithm in machine learning. It is basically used for doing both- it performs classification as well as regression part, which is the major part of current machine learning models.

Data Transmission

IV. ZigBee Modules

In the current world, we have many high data rate communication techniques available, but none of these are able to meet the communication standards of sensors and control devices. These communication standards require high data rate at low-latency and low-energy consumption even for smaller bandwidths. Zigbee technology is a wireless low power and low cost technology .It has excellent and superb characteristics which have made this communication most suitable for embedded applications ,home automation etc. It is especially built for sensor networks on IEEE 802.15.4 standard for wireless personal area networks (WPANs), . It is a product from Zigbee alliance. This communication technology defines physical and Media Access Control (MAC) layers for handling many devices at low-data rates. Zigbee WPANs work at 868 MHz, 902-928 MHz and 2.4 GHz frequencies. The best suited data rate is 250 kbps for periodic or intermediate two way transmission of data between controllers and sensors . Zigbee is a low-cost and power network mostly deployed to control and monitor areas where we need to cover only 10-100 meters within the range. This is a less expensive communication system and it is simpler than other proprietary short-range wireless sensor networks as Bluetooth and Wi-Fi.

Zigbee system structure mainly has three different types of devices

a) Zigbee coordinator: This forms the root of the network. The Mandatory node for all zigbee networks which has all the information of the network including the keys and acts as a trust centre playing a key role in the security. This device can never sleep.

b) Zigbee Router: This node can run an application function as well as act as a relay station for other zigbee devices in the network. The devices on this node can never sleep too.

c) End device: End devices have a very limited work which is to communicate with the parent nodes such that the battery power is saved. These device can sleep to save power.

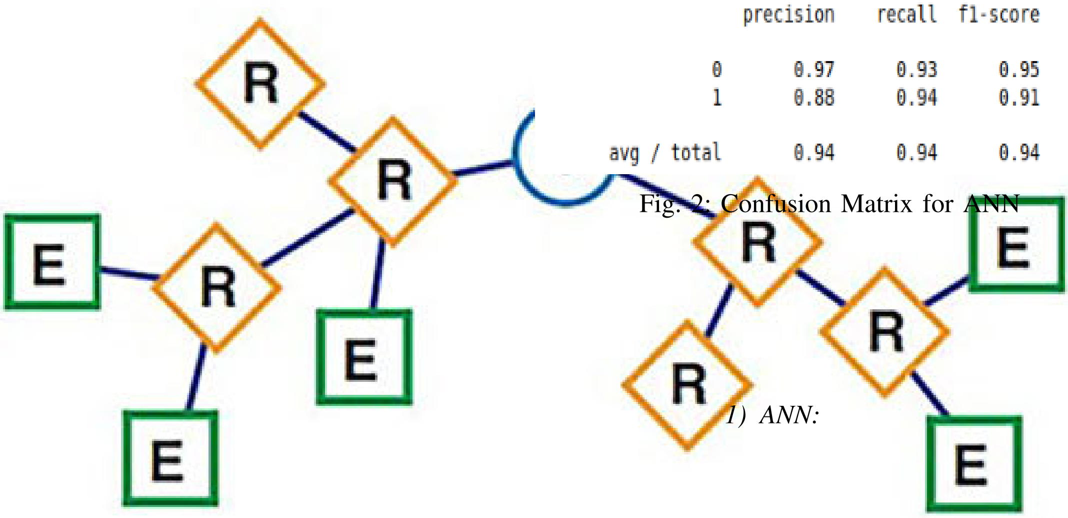

Fig. 1: A Mesh Network of Zigbee Devices

In the network currently implemented, there are only two nodes. One acting as the coordinator at the ground station and the other acting as an End point near the sensors. We used 802.15.4 protocol to achieve higher point to point transmission speed but this restricts the network to be only a point to point with no possibility of forming meshes of nodes.

Observations and Results

V. Context

Here we will be training three different models on the training dataset and testing the performance of the trained model on a test dataset. The feature selection for the dataset is varied from 3 features to 12 features out of maximum 15 features and their results are shown as below.

A. Best 3 features

Best 3 features as per our model are standard deviation, skewness and RMS value as per our anova test.

```
Confusion Matrix :
 [[1123   79]
 [  37  561]]
             precision    recall  f1-score   support

          0       0.97      0.93      0.95      1202
          1       0.88      0.94      0.91       598

avg / total       0.94      0.94      0.94      1800
```

Fig. 2: Confusion Matrix for ANN

1) ANN:

```
Test accuracy :  0.9322222222222222

Confusion Matrix :
 [[1133   69]
 [  53  545]]
             precision    recall  f1-score   support

          0       0.96      0.94      0.95      1202
          1       0.89      0.91      0.90       598

avg / total       0.93      0.93      0.93      1800
```

Fig. 3: Confusion Matrix for SVM

2) SVM:

```
Test accuracy :  0.9988888888888889

Confusion Matrix :
 [[1200    2]
 [   0  598]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      1.00      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 4: Confusion Matrix for Random Forest

3) Random Forest:

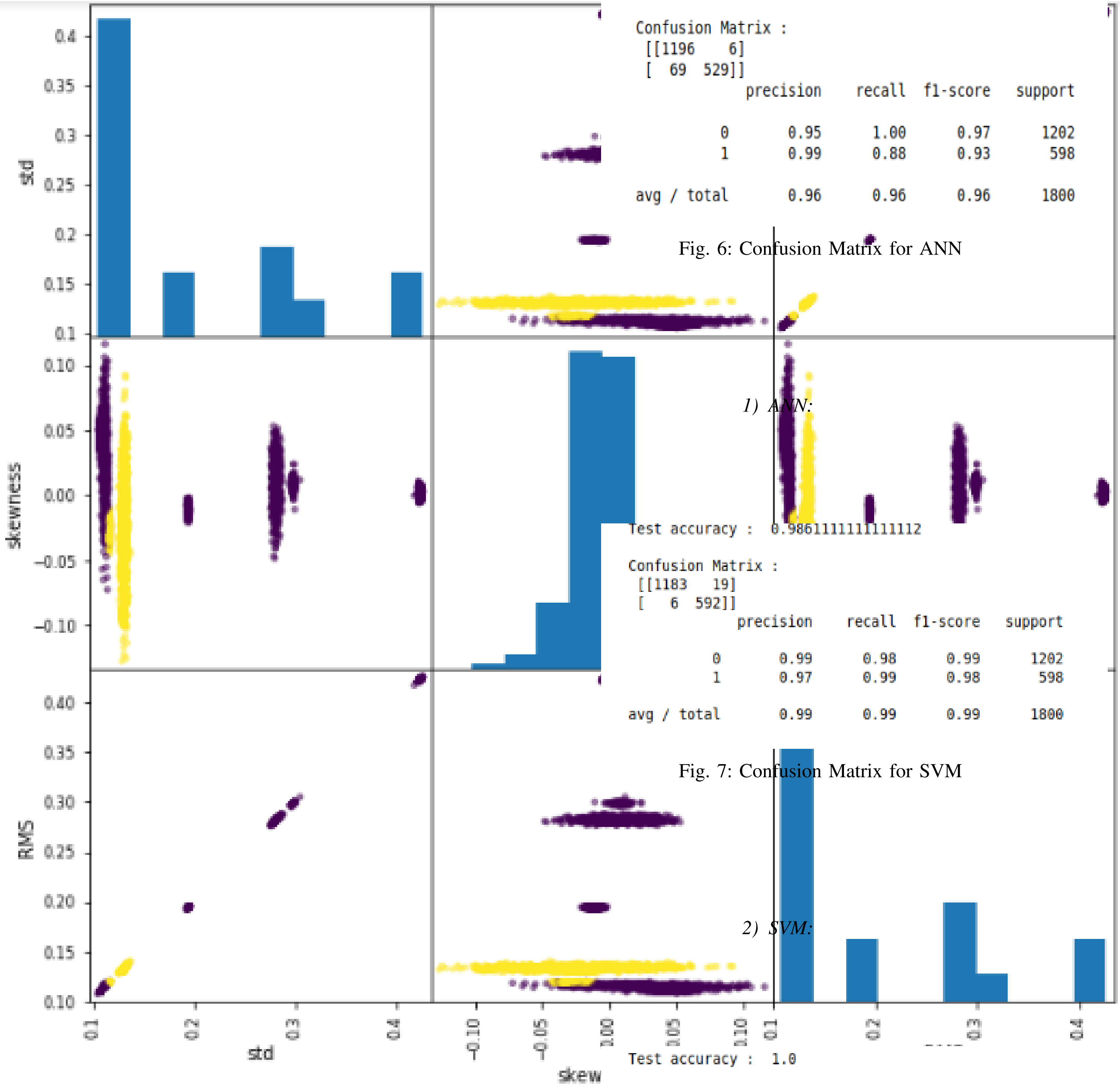

Fig. 5: Pair plots for 6 features

4) Pair Plots:

B. Best 6 features

Best 6 features as per our model are standard deviation, skewness, RMS value, max, pk-pk, and Margin factor as per our anova test.

```
Confusion Matrix :
 [[1196    6]
 [  69  529]]
             precision    recall  f1-score   support

          0       0.95      1.00      0.97      1202
          1       0.99      0.88      0.93       598

avg / total       0.96      0.96      0.96      1800
```

Fig. 6: Confusion Matrix for ANN

1) ANN:

```
Test accuracy :  0.9861111111111112

Confusion Matrix :
 [[1183   19]
 [   6  592]]
             precision    recall  f1-score   support

          0       0.99      0.98      0.99      1202
          1       0.97      0.99      0.98       598

avg / total       0.99      0.99      0.99      1800
```

Fig. 7: Confusion Matrix for SVM

2) SVM:

```
Test accuracy :  1.0

Confusion Matrix :
 [[1202    0]
 [   0  598]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      1.00      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 8: Confusion Matrix for Random Forest

3) Random Forest:

C. Best 9 features

1) ANN:

```
Confusion Matrix :
 [[1201    1]
 [  30  568]]
             precision    recall  f1-score   support

          0       0.98      1.00      0.99      1202
          1       1.00      0.95      0.97       598

avg / total       0.98      0.98      0.98      1800
```

Fig. 10: Confusion Matrix for ANN

2) SVM:

```
Test accuracy :  0.9988888888888889

Confusion Matrix :
 [[1201    1]
 [   1  597]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      1.00      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 11: Confusion Matrix for SVM

3) Random Forest:

```
Test accuracy :  0.9994444444444445

Confusion Matrix :
 [[1201    1]
 [   0  598]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      1.00      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 12: Confusion Matrix for Random Forest

4) Pair Plots:

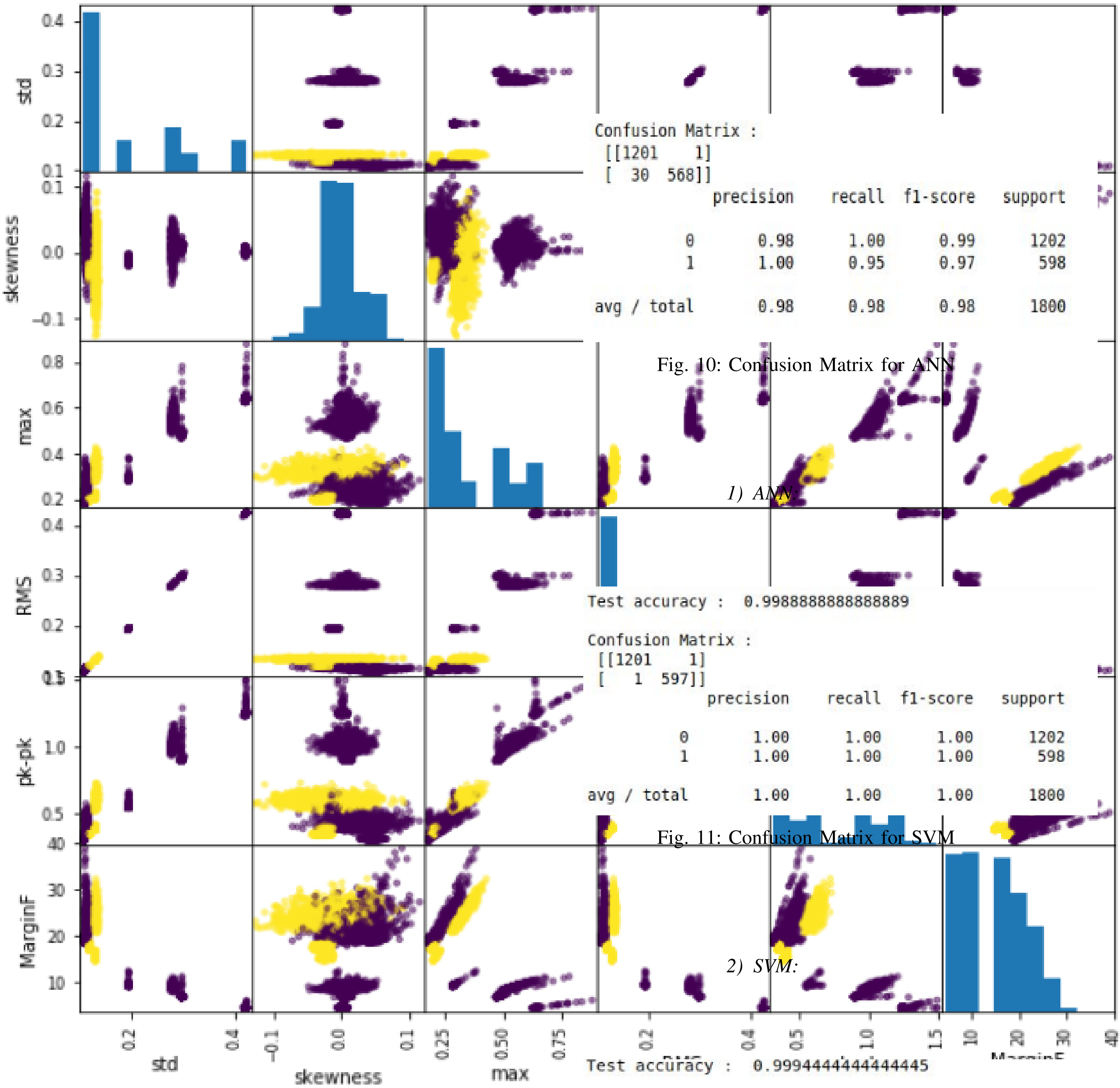

Fig. 9: Pair plots for 6 features

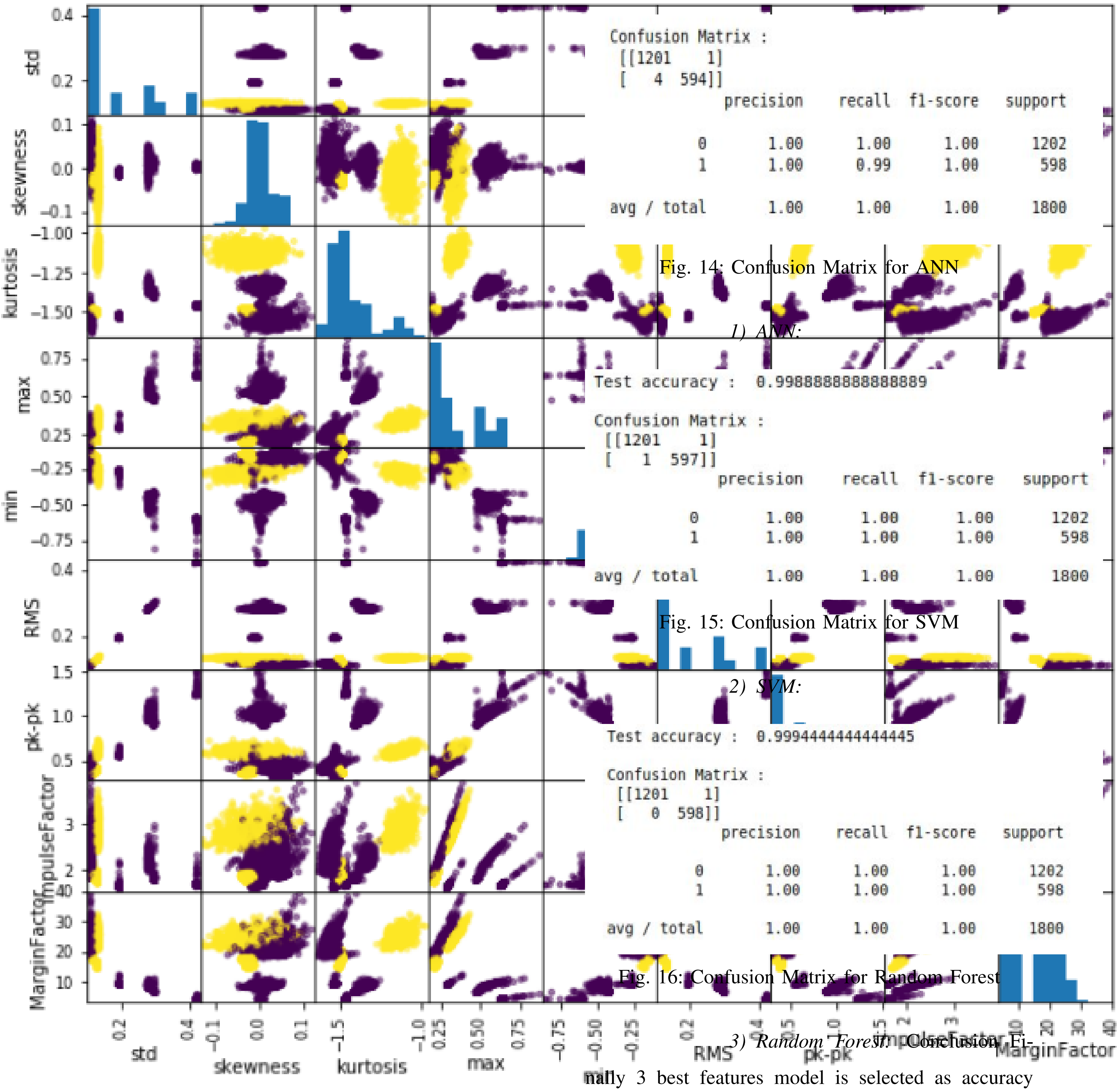

Fig. 13: Pair plots for 9 features

4) Pair Plots:

D. Best 12 features

```
Confusion Matrix :
 [[1201    1]
 [   4  594]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      0.99      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 14: Confusion Matrix for ANN

1) ANN:

```
Test accuracy :  0.9988888888888889

Confusion Matrix :
 [[1201    1]
 [   1  597]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      1.00      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 15: Confusion Matrix for SVM

2) SVM:

```
Test accuracy :  0.9994444444444445

Confusion Matrix :
 [[1201    1]
 [   0  598]]
             precision    recall  f1-score   support

          0       1.00      1.00      1.00      1202
          1       1.00      1.00      1.00       598

avg / total       1.00      1.00      1.00      1800
```

Fig. 16: Confusion Matrix for Random Forest

3) Random Forest:

E. Conclusion Finally 3 best features model is selected as accuracy obtained is high and less data needs to be transferred via Zigbee at a transmission rate of approximately 180 Bps. This can further be interfaced with the ANN to to get the status of the machine and perform any further action.