DESIGN AND DEVELOPMENT OF AI BASED INTELLIGENT FAULT DETECTION SYSTEM FOR INDUSTRIAL HYDRAULIC SYSTEMS

A Project report submitted by students of
Department of Mechanical Engineering

By

ARITRA ROY (20JE0180)
ARUN KUMAR MISHRA (20JE0186)
AMAR SINGH (20JE0112)

Under the supervision of

Prof. AJIT KUMAR

(Department of MECHANICAL Engineering)

ACKNOWLEDGEMENT

It is our great fortune that we have got opportunity to carry out this project work under the supervision of **Prof. AJIT KUMAR**, in the Department of Mechanical Engineering, Indian Institute of Technology IIT (ISM) DHANBAD. We express our sincere thanks and deepest sense of gratitude to my guide for his constant support, unparalleled guidance and limitless encouragement.

We wish to convey our gratitude to **Prof. Ajit Kumar**, Department of Mechanical Engineering, IIT (ISM) DHANBAD and to the authority of IIT ISM Dhanbad for providing all kinds of infrastructural facility towards the research work.

We would also like to convey our gratitude to all the faculty members and staffs of the NVCTI, IIT ISM DHANBAD for their whole hearted cooperation to make this work turn into reality.

ABSTRACT

Maintenance of hydraulic circuits is very tedious and very costly, ranging from 15% to 70%, depending on the industry, a well-known fact confirmed by numerous studies. For the technology infrastructure, losses from the downtime can be quite significant, with hourly loss ranging from hundreds of thousands to millions of rupees. Maintenance best practices have been evolving from reactive maintenance to preventive maintenance to predictive maintenance,with Intelligent Fault Diagnosis (IFD), using recent advances in industrial Internet of Things, artificial intelligence, edge computing, etc. Many studies examine system monitoring and suggest various mathematical models and methods for fault detection. Literature Survey considers model- based and DSP-based fault prediction, while papers and more recent ones use data-driven approaches. Data-driven and deep learning-based methods show great results not only in computer vision applications, speech recognition, natural language processing, and medical imaging, but also as classifiers for induction motor fault classification, railway vehicle wheels diagnosis, industrial machinery, hydraulic system malfunction identification.

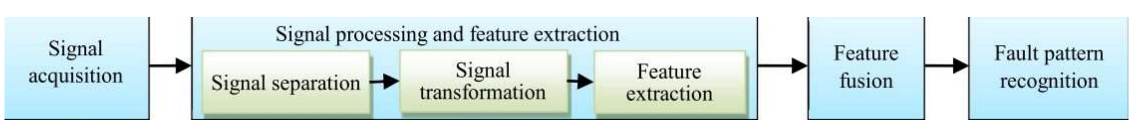
In this project an intelligent AI model is designed for fault diagnosis for industrial hydraulic systems. This model can diagnosis two types of faults: (a) Fluid leakage inside hydraulic cylinder (b) Spool blockage in directional control valve. A new classifier model will use both simulation and the physical system data. The Physical system is simulated in Automation Studio P-6.4 and MATLAB SIMULINK software. The model is trained on the simulated dataset mixed with a small portion of the data collected from the physical system. Including the real data, even a small amount of it compared to the simulated data, improved classification performance when the trained model was a physical system.

INTRODUCTION

Hydraulic systems are widely used in modern machinery owing to a multitude of advantages, such as a fast response, significant load stiffness, large power density, and superior stability. A hydraulic system is often the core component of engineering equipment, such as control and power transmission systems, which are typically operated in the field. A hydraulic system can be damaged by exposure to sunshine, rain forests, and dust particles, among other factors, and by unstable working conditions such as a high load or severe impact. Therefore, such systems areprone to faults, and if certain initial abnormalities are not located and eliminated in time, they may develop into a functional disability and even lead to a dangerous condition. Therefore, it is extremely important to diagnose and remove such problems in time. As indicated in the figure below, the common technology during each process of a signal-based intelligent fault diagnosis will be thoroughly discussed.

Traditional fault diagnosis approaches are generally based on mechanism, characteristic frequency or faulty feature extraction. The signal of faulty hydraulic pump is a union of machinery and hydraulic, so its model and feature are difficult to acquire. In addition, although the hydraulic pump is failure for different reasons, it may generate the same fault characteristic frequency. For the purpose of raising the fault diagnosis accuracy rate for hydraulic pump, a lot of scholars have done many fruitful researches. Some of the pioneers during literature survey that motivated us for this project are discussed here. *Wang et al.* [1] analyzed the statistics of the characteristic's frequency measured from hydraulic pump, and calculated the failure threshold by experience, in order to diagnose the multiple faults of hydraulic pump. *Gao Y. et al.* [2] Used wavelet packet coefficient and residual of the signal energy for fault recognition of hydraulic pump, and he also set threshold through the experience. *Dong M. et al.* [3] established the hidden semi- Markov models for hydraulic pump using vibration signals under different conditions, in order to recognize the failure. *Du jingyi et al.* [4] extracted the statistics of the vibration signal for fault recognition hydraulic pump.

Over the last two decades, different approaches have been used to diagnose hydraulic system faults. Currently, Fault Diagnosis systems apply Artificial Neural Networks (ANNs) to diagnose faults of some system components such as valves, actuators, pumps, or sensors. *Le et al.* [5] presented a method using ANNs to detect single and multiple leakage types in a fluid power system, including a servo valve and a single rod actuator. *Karpenkoa et al.* [6] investigated a neural-network-based scheme to detect and identify actuator faults in a typical process control valve. The fault diagnosis process includes the following tasks: fault detection, which indicates that something is going wrong in the system; fault isolation which determines the exact location of a fault; and fault identification, which determines the magnitude of fault severity.



OBJECTIVE

Hydraulic systems play an important role in a wide variety of industrial applications, such asrobotics, manufacturing, aerospace, and engineering machinery. Monitoring the condition of hydraulic equipment can not only effectively improve productivity and reduce maintenance costs and downtime, but also improve the reliability and safety of this equipment in its application. Fluid power systems condition monitoring are earning more and more consideration to reduce the cost of maintenance and prevent the system fault from further deteriorating by detecting the fault in its incipient stage.

Leakage, one of the major faults will cause inefficient performance in a hydraulic system. With the increase in the leakage flow, the efficiency of the system will eventually decrease. A gradually increasing leakage could also trigger serious consequences when the system is working with load and is always a major threat to machine operators. Many studies on fault diagnosis of the hydraulic valve have been conducted by theoretical approaches and test measurements, and certain research results have been obtained. The pumps operate at high pressure differential and its speed is fast, so it is prone to failure. The objectives of the project were to predict the following faults occurring in an industrial hydraulics system: -

1) Fluid leakage inside hydraulic cylinder
2) Spool blockage in directional control valve.

In this project, a method is proposed to design and develop comprehensive intelligent fault diagnosis method for industrial hydraulic systems. A new classifier model will be built that uses both simulated and obtained from the physical system data. Our physical system was simulated with Automation Studio P-6.4 software. The model will be trained on the simulated dataset mixed with a small portion (less than 0.2%) of the data collected from the physical system. The inclusion of the real data, even a small amount of it compared to the simulated data, improved classification performance when the trainedmodel was a physical system.

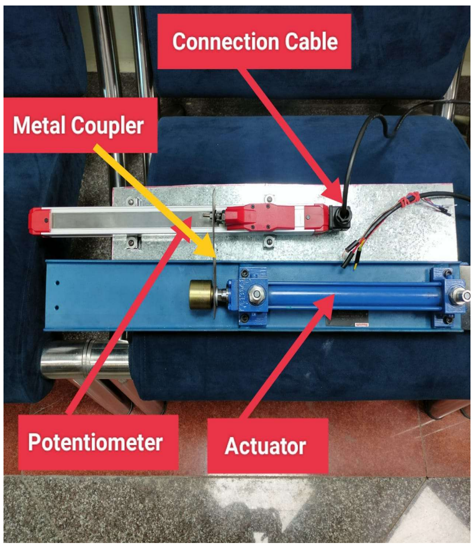
CHAPTER 2: PROPOSED SYSTEM HARDWARE AND ARCHITECTURE

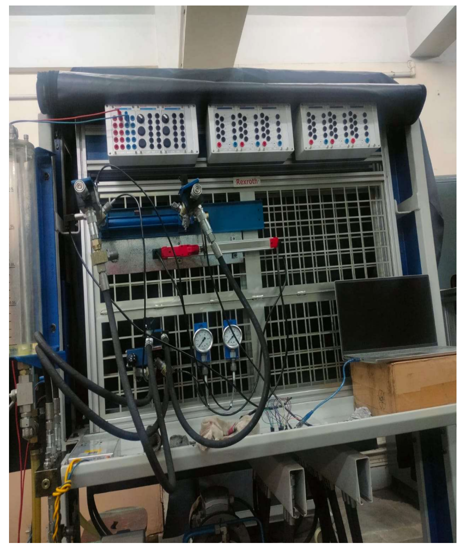
COMPONENT LIST

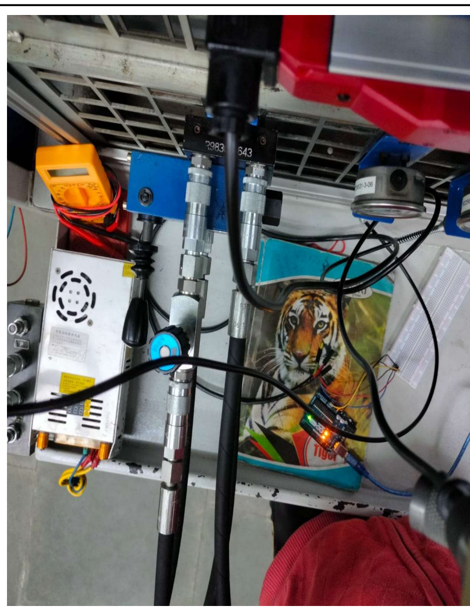
SL NO	NAME OF COMPONENTS	QUANTITY
1	Arduino UNO	1
2	Pressure Transducer	2
3	Linear Potentiometer	1
4	Double acting hydraulic actuator	1
5	Power Supply cord for Arduino	1
6	Male – Male Connecting wires	10
7	Female – Female Connecting wires	10
8	Male – Female Connecting wires	10
9	SMPS (0-24V)	1
10	4/3 Directional control valves manually actuated	1
11	12V – 2A AC Adapter for Arduino	1

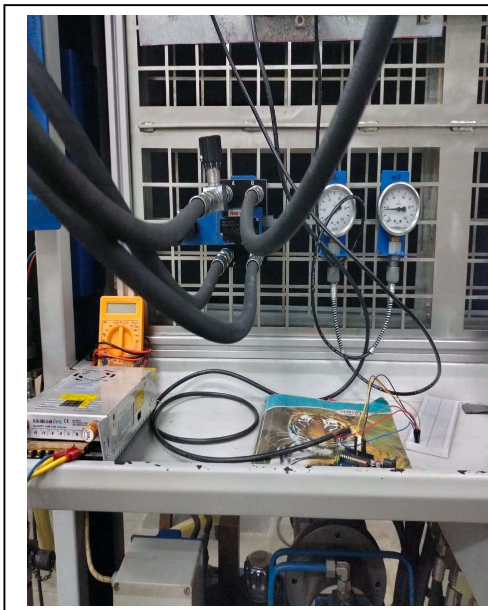
Table 1: Component list of proposed system

Designed Hydraulic Fault Diagnosis System









CHAPTER 3: MATHEMATICAL MODEL OF THE OIL LEAKAGE FAULT

Figure shows a schematic diagram of a typical hydraulic valve- actuator system. It is assumed that valve orifices are matched and symmetrical, and $\mathbf{P_1=P_S}$ and $\mathbf{P_2=P_E}$, where P_1 and P_2 are the pressures at the high-pressure side and the low-pressure side of the actuator respectively, P_S and P_E are the supply and exit pressures respectively. And it is assumed that: the fluid properties are not changed, *i.e.* the effective bulk modulus β_e and the density of the fluid are constant, the supply and exit pressures are constant, both valve and actuator parameters are constant and the effect of local variations of working temperature is negligible during the experiments.

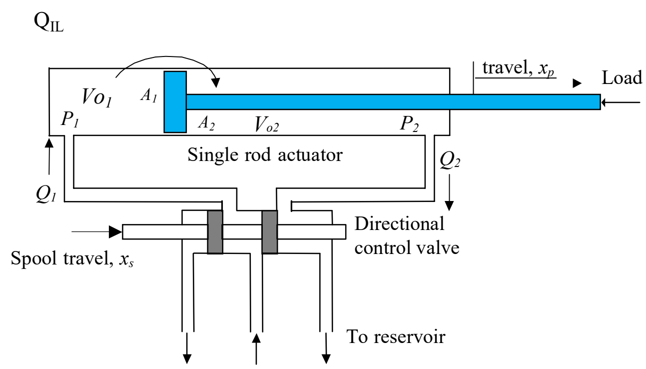


Fig. A Schematic diagram of Hydraulic Actuator System

For internal leakage derived fault indices $\mathbf{a_1}$:

$$\mathbf{a_1 = Q_{IL} / Q_1}$$

For valve spool blockage fault indices $\mathbf{a_2}$:

$$\mathbf{a_2 = x_s / s_s}$$

where a_1 and a_2 are the indices of the internal leakage and valve spool blockage faults respectively, Q_{IL} is the amount of the leakage flow rate between the two chambers of the actuator, s_s is the total stroke of the valve spool.

CHAPTER 4: ALGORITHM DEVELOPMENT FOR ANN

a. **Defining the ML problem statement**: - A task, is defined as a multiclass classification task for N types of faults. The class "zero" corresponds to the healthy system and other classes labelled from 1 to N correspond to specific faults. This reduces the problem of fault diagnosis to the problem of (N + 1)-class classification. In this proposal it is only considered mutually exclusive faults; multiple faults can be added as additional fault classes. The classification problem is solved using 'M' amount of data. Its assumed that all data has a uniform sampling in the time domain, where it is not the case, interpolation can always be performed. The data with the control action applied to the system is assumed to come from one of the sensors. Classification for each fault class c starts with a set of sensor measurements of the length W, the window size. A fixed-length window is being used, measured in seconds, andthe number of samples in the window depends on the sampling rate. In general, the state of the system can be described in the form of a fault class c, where zero class corresponds to a healthy system:

$$c = G(X, a)$$

where X is a $W \times M$ data matrix, each column represents the data from an individual sensor, each row represents a single measurement, a is a vector of the classifier parameters determined during the training process, and G is the decision rule of the classifier that determines the overall state of the system. A sliding window W is moved over the training data to create the training dataset X.

b. **Training the ML model for fault diagnosis**: - Generating training samples in the form of arrays of the parameters values of the power supply subsystem obtained in its various states is an important step in the development of diagnostic systems since it directly affects the accuracy of classification. The formation of training samples can be represented in the form of an algorithm, the main stages of which are: virtual modelling, storage, processing and transmission of the received data to the hardware and software complex of the diagnostic system for the purpose of machine learning. An important component of the presented algorithm is the need to correct modelling conditions in order to improve the classification accuracy. To obtain the training samples as a result of calculation, arrays of pressure and flow rate values from time are written into a text file. The number of arrays depends on the number of faults intensity levels.

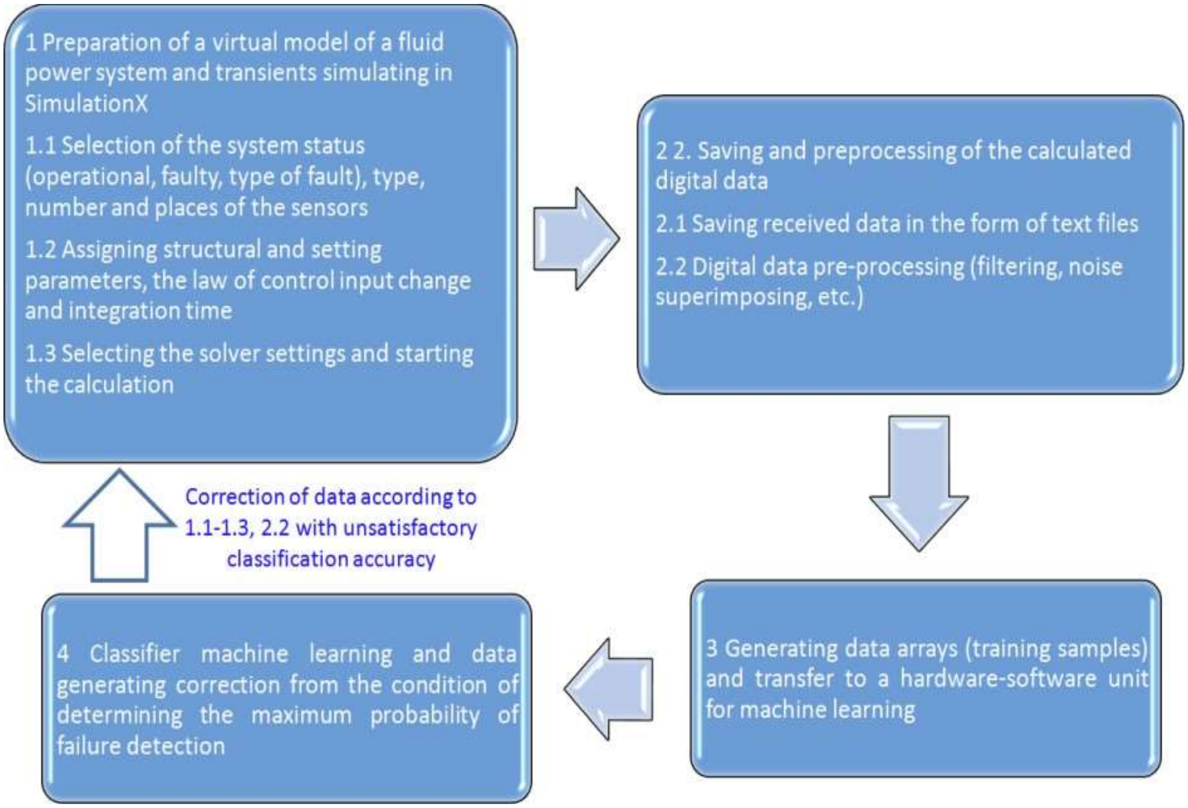
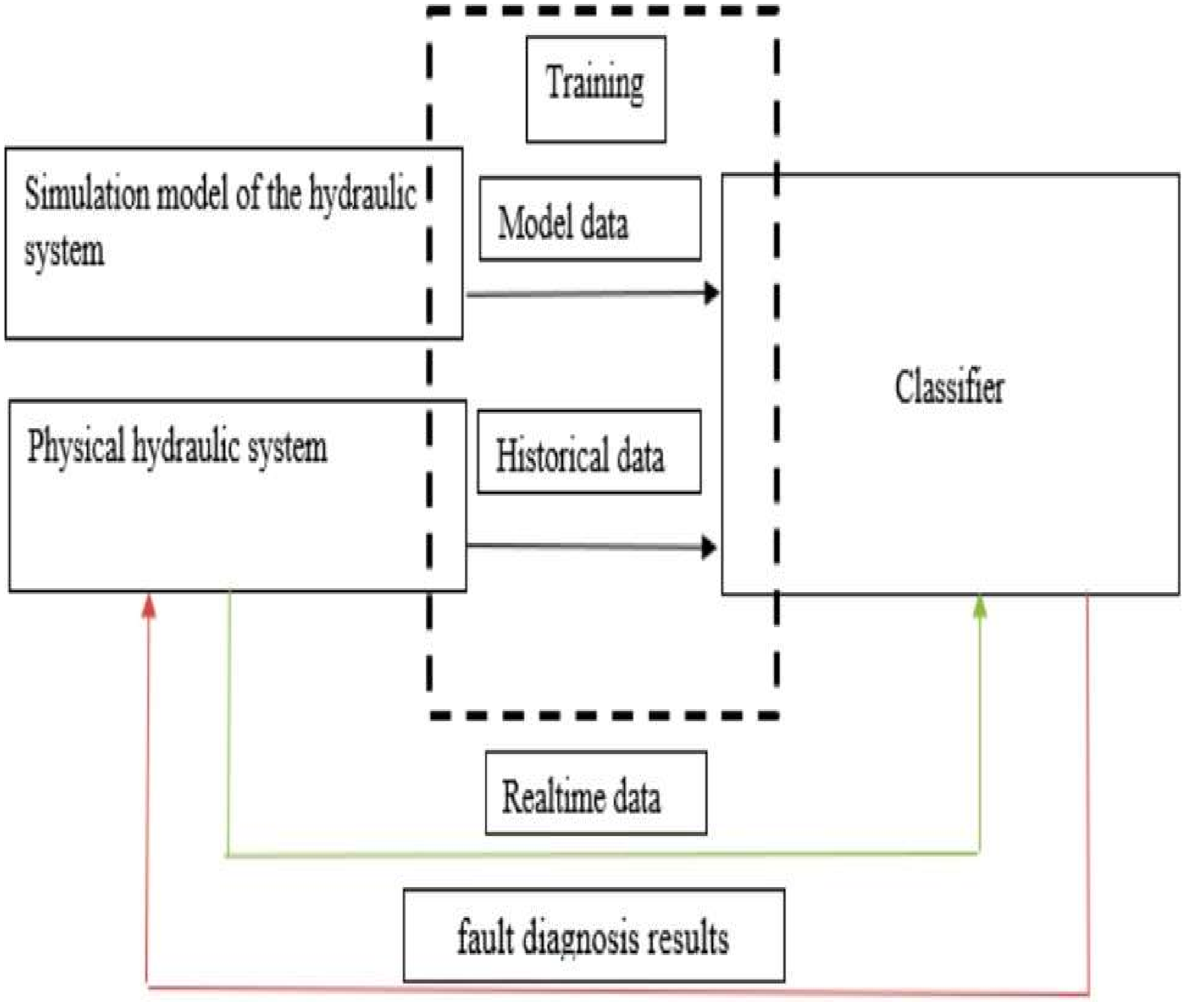


Fig 8. Generating strategy for the algorithm and training the model.

c. <u>Analysis of the output datasets and plots from the model</u>: - Verification of the simulation model is carried out by comparing the calculated and experimental data obtained under identical conditions for the power supply subsystem in fault-free and faulty conditions: temperature of the working fluid; electric motor rotation frequency. In a power supply subsystem with a discharged HPA, more abrupt (spasmodic) changes in pressure occur as a result of the change in the flow area of the distributor compared to a fault-free state which is explained by a decrease in the system compliance capacity. Analysis of the graph when the valve is opened by more than 50%, there is a decrease in the flow rate of the working fluid in a system with a discharged HPA in relation to a fault-free working system due to the lack of an additional fluid flow from the HPA. The analysis of the results presented above shows that the considered fault has a significant effect on the characteristics of the heterostructure: the amplitude of the pulsations of the parameters grows and their gradients increase.
Finally, we can classify the fault and diagnose it into a healthy hydraulic system.

ARCHITECTURE OF AI WORKING MODEL



SOFTWARE DESIGNING – ANN ALGORITHM DEVELOPMENT

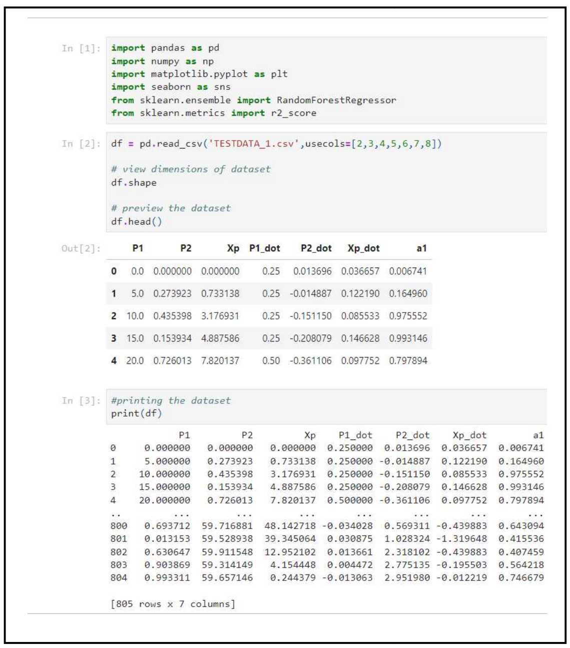
#code developed for MPDI project by IIT DHANBAD students

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import r2_score
df = pd.read_csv('TESTDATA_1.csv',usecols=[2,3,4,5,6,7,8])

# view dimensions of dataset
df.shape

# preview the dataset
df.head()
#printing the dataset
print(df)
#summary of the dataset
df.info()
# check missing values in variables

df.isnull().sum()
df['a1'].value_counts()
sns.pairplot(df)
#declare feature vector and target variable
X = df.drop(['a1'], axis=1)

y = df['a1']
# split data into training and testing sets

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.33, random_state = 42)
# check the shape of X_train and X_test

X_train.shape, X_test.shape
# check data types in X_train
X_train.dtypes
X_train.head()
X_test.head()
# instantiate the classifier with n_estimators = 100
rfr_100 = RandomForestRegressor(n_estimators=100, random_state=0)

# fit the model to the training set
rfr_100.fit(X_train, y_train)

# Predict on the test set results
y_pred_100 = rfr_100.predict(X_test)
# create the classifier with n_estimators = 100
clf = RandomForestRegressor(n_estimators=100, random_state=0)

# fit the model to the training set
clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)
```

 # view the feature scores
feature_scores = pd.Series(clf.feature_importances_, index=X_train.columns).sort_values(ascending=False)
feature_scores
#graphical representation of data using color-encoded matrix to give better visualization if data
plt.figure(figsize=(15, 15))
hm = sns.heatmap(data = df.corr(),annot=True,square=True,cmap=sns.color_palette("Spectral", as_cmap=True))

plt.show()
y_pred
df.corr()

<u>SYSTEM EFFICIENCY CHECK USING PLOTS</u>



In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import r2_score
```

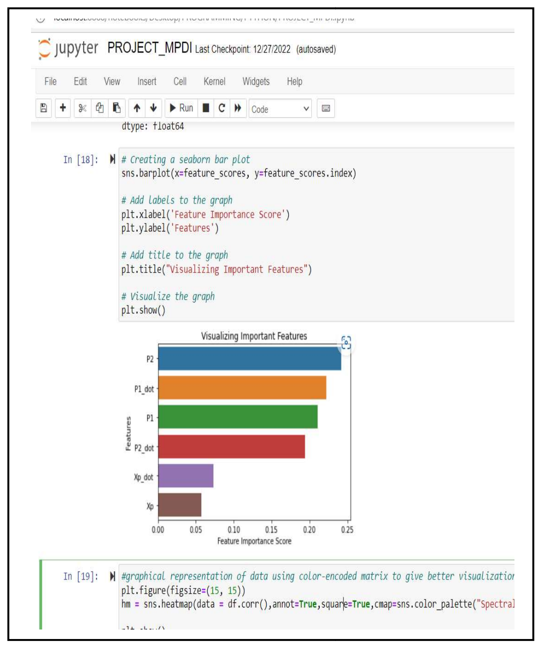
In [2]:
```
df = pd.read_csv('TESTDATA_1.csv',usecols=[2,3,4,5,6,7,8])

# view dimensions of dataset
df.shape

# preview the dataset
df.head()
```

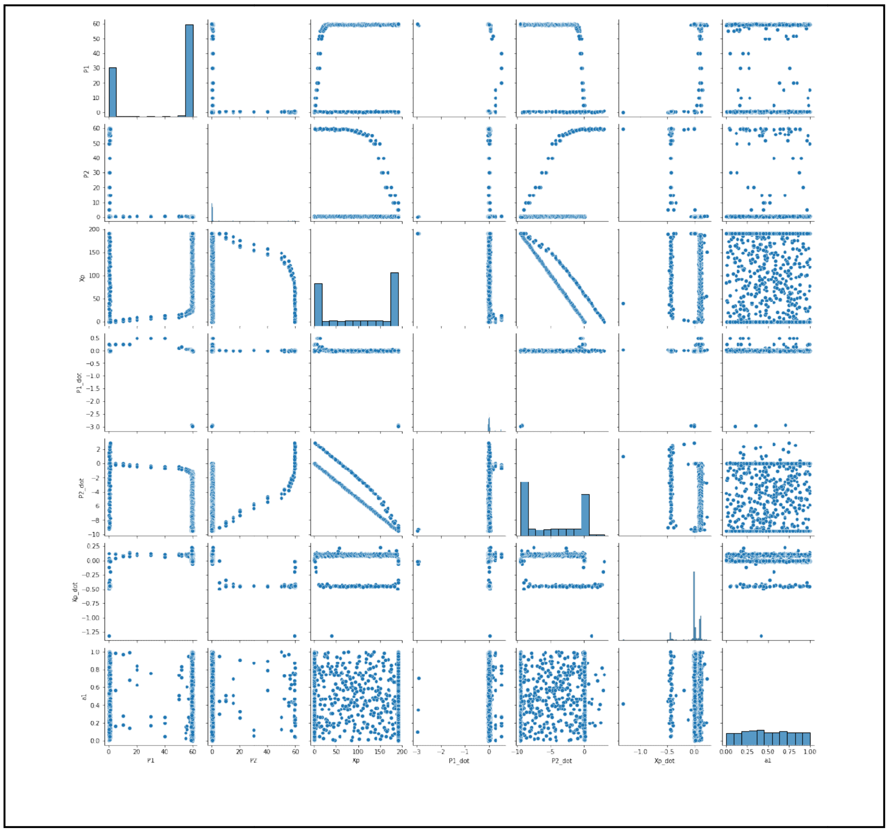
Out[2]:

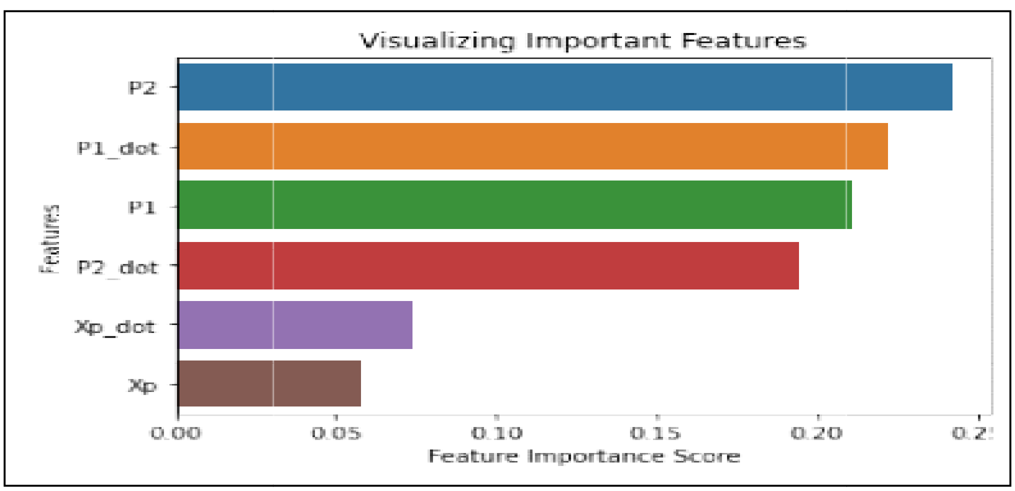
	P1	P2	Xp	P1_dot	P2_dot	Xp_dot	a1
0	0.0	0.000000	0.000000	0.25	0.013696	0.036657	0.006741
1	5.0	0.273923	0.733138	0.25	-0.014887	0.122190	0.164960
2	10.0	0.435398	3.176931	0.25	-0.151150	0.085533	0.975552
3	15.0	0.153934	4.887586	0.25	-0.208079	0.146628	0.993146
4	20.0	0.726013	7.820137	0.50	-0.361106	0.097752	0.797894

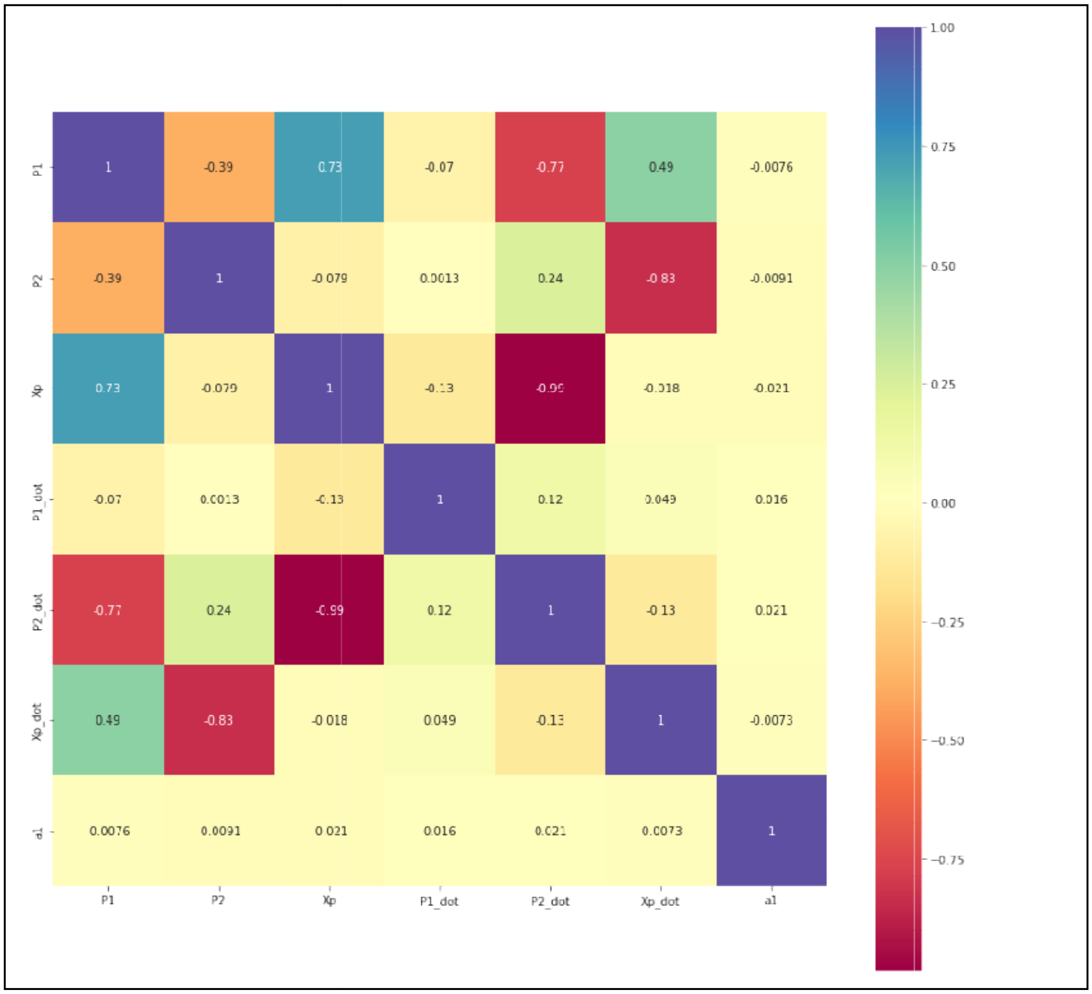
In [3]:
```
#printing the dataset
print(df)
```

```
            P1         P2         Xp    P1_dot    P2_dot    Xp_dot        a1
0     0.000000   0.000000   0.000000  0.250000  0.013696  0.036657  0.006741
1     5.000000   0.273923   0.733138  0.250000 -0.014887  0.122190  0.164960
2    10.000000   0.435398   3.176931  0.250000 -0.151150  0.085533  0.975552
3    15.000000   0.153934   4.887586  0.250000 -0.208079  0.146628  0.993146
4    20.000000   0.726013   7.820137  0.500000 -0.361106  0.097752  0.797894
..         ...        ...        ...       ...       ...       ...       ...
800   0.693712  59.716881  48.142718 -0.034028  0.569311 -0.439883  0.643094
801   0.013153  59.528938  39.345064  0.030875  1.028324 -1.319648  0.415536
802   0.630647  59.911548  12.952102  0.013661  2.318102 -0.439883  0.407459
803   0.903869  59.314149   4.154448  0.004472  2.775135 -0.195503  0.564218
804   0.993311  59.657146   0.244379 -0.013063  2.951980 -0.012219  0.746679

[805 rows x 7 columns]
```






Visualizing Important Features
P2
P1_dot
P1
P2_dot
Xp_dot
Xp
Features
0.00
0.05
0.10
0.15
0.20
Feature Importance Score


P1
P2
Xp
P1_dot
P2_dot
Xp_dot
a1
1
-0.39
0.73
-0.07
-0.77
0.49
-0.0076
1
-0.079
0.0013
0.24
-0.83
-0.0091
1
-0.13
-0.99
-0.018
-0.021
1
0.12
0.049
0.016
1
-0.13
0.021
1
-0.0073
1
0.0076
0.0091
0.021
0.016
0.021
0.0073
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

CHAPTER 5: CONCLUSION AND FUTURE WORK

CONCLUSION :-

The Artificial Intelligence Code for fault prediction using Random Forest Classifier is running successfully and giving expected output on the demo data after the training, with efficiency. It would be improved infuture for better efficiency.

FUTURE WORK: -

This system would be expanded in future from Arduino uno to raspberry pi4 for better speed, control, storage and communication facilities. The recorded data from raspberry will be stored in data logger where theintegrated AI Algorithm will be used to predict the suitable fault prevailing in the industrial hydraulics system especially in the heavy machinery used for mining purposes.