International Journal of Advanced Science and Technology
Vol. 29, No. 9s, (2020), pp. 3667-3676

Fault Diagnosis In Satellite Power Systems Using PCA And Deep Learning

Niranjana L., Sathana S., Arvindh M., Rashmikka K., Ganesan M., Lavanya R.
Department of Electronics and Communication Engineering,
Amrita School of Engineering, Coimbatore
niranjana.naras98@gmail.com,sathana.98@gmail.com,m_ganesan1@cb.amrita.edu

Abstract

Fault detection and diagnosis is one of the key technologies for monitoring the functions of power systems in satellites. Most of the machine learning approaches used currently in spacecraft fault diagnosis have limitations in representing complex functions within small samples and cells. In addition, the ability of these methods to generalise is limited. Therefore, in this paper, we use Principal Component Analysis to identify the features and Deep Neural Network to diagnose the faults. Usually a classifier trains faulty data to identify the causes of the anomalies and this aspect has generally limited the use of model-driven approaches to fault detection tasks. The paper suggests a technique for using methods of machine learning for purposes of fault detection and diagnosis. The proposed method is more suitable for characterizing complex features of equipment information, allowing for more accurate identification of equipment health status under complex monitoring tasks. Here, the dataset used is acquired from the Advanced Diagnostic and Prognostic Testbed (ADAPT). The accuracy obtained through this method is 80% and thus can prove to be efficient in healthcare management of systems.

Keywords: *power system, machine learning, fault diagnosis, principal component analysis, deep neural network.*

1. Introduction

In satellites, Electric Power Systems (EPS) are of fundamental importance. The EPS provides power for subsystems for the necessary functions of the satellite such as propulsion, communications, navigation and control functions. Faults in power systems may have a serious impact on the satellite or mission. The EPS consists typically of three main units: solar, battery and payload units. In all satellites, both the solar and the battery units are the same. In each satellite, the load unit varies according to the necessary functions. During real-time missions, the testing of faults will cause considerable costs and is dangerous for the human crew accompanying for the mission. Therefore, in a similar smaller platform, methods of failure diagnosis should be developed and tested[1].

The Advanced Diagnostics and Prognostics Testbed (ADAPT) is a fully working, machine-like device similar to the satellite power systems for testing, observing and validation developed by NASA Ames and Research Centre. The ADAPT dataset consists of information reflecting the health of the sensors in the power system, in which the proper device operation is controlled and a fault can be detected [2].

Machine learning is the ability to learn from experience automatically without being specifically programmed. The aim of machine learning is to create computer programs that can use and access data on their own. Methodologies that are powered by data are increasingly being more used for fault detection while monitoring the machinery conditions [3]. The program is equipped first with ordinaryandincorrect data to distinguish the faults in real-time applications from the original data. From fault detection to fault identification, machine learning capability can be enhanced with the limited data available from the ADAPT dataset. In this paper, Principal Component Analysis (PCA) is used for the process of feature extraction to compress the data to a smaller dimension with higher characteristic contribution. Deep learning is a type of algorithm in machine learning that employs several layers in order to gradually extract higher levels of raw input. It enables us to train a system to forecast results when a number of inputs are given [4]. Deep learning is a dynamic method of extracting features from a computational model perspective and classifying the data. We can effectively carry out a thorough analysis of satellite power system test data using profound learning techniques, like neural networks. Deep learning offers a range of methods, mainly DeepNeural Network (DNN). These deep learning models can transform low-level features of the data into high-level features. Due to this characteristic, deep learning models can be more effective than machine learning models in feature extraction and representation[5].

We evaluate complex data, i.e. multi-dimensional data, in real world data analysis activities. We compile the data and find different trends or use it to train some models. The growth in information dimensions also raises the complexity of processing and manipulating it. In order to decrease data dimensions, the redundant dimensions first are removed and only the most appropriate dimension is retained. PCA finds a new array of data so that all dimensions are orthogonal and linearlyindependent.

PCA may also be used to filter noisy datasets like compressed images. The first principal component displays the greatest variation. Each additional component represents less variation and more noise, thereby retaining the signal and removing the noise from the original signal by representing data with a smaller sub-set of main components [6]. By rotating the axes, PCA shows the maximum variability in the dataset. The PCA lists the key axes to define the base data set before classifying the underlying ones by the amount of variance collected. This data is then used to train the deep learning model to get accurate predictions.

2. Datasets And Adapt

The ADAPT is a testbed developed by the NASA Ames and research center to serve as a benchmark or a platform for testing and verification. All data samples in the datasets are presented as a list of faults and under normal conditions includingall sensor's data sampledwithdifferent frequencies, during a span of time, covering the time of particular fault injection.
As shown in Fig 1, the basic equipment of ADAPT basically Generates, Stores and distributes Power. The generation part consists of solar panels, the storage part consists of 3 batteries and it is connected to 2 load banks for the Distribution of the Power [7].
Further, the ADAPT system assigns 3 actors. They are the User who simulates and plays the role of a crew member, Antagonist who injects the faults into the system and the observer who observes and records the data. We have acquired datasets from the Advanced Diagnostic and Prognostic Testbed results to perform this experiment of fault detection. The datasets consist of multiple experiments in Multivariant Time Series with different types of sensors. There are different types of sensors and each sensor has different notation. For example, the Relay sensor is notified as EY, Voltage Sensor as E, Current sensor as I, and Position sensor as X [8].

This particular dataset has 21 sensor outputs taken across 4 minutes time duration. The fault is injected to the sensor and the fault maybe of two types – Manual fault and Software fault. The fault modes are stuck, offset, failed open, failed off, stuck open, underspeed, overspeed and degraded. These faults can occur in different locations. For example, the stuck open and stuck faults occur on the relay side. Faults like overspeed and underspeed occur on the load side. Hence, each fault occurs on the specific locations[9].

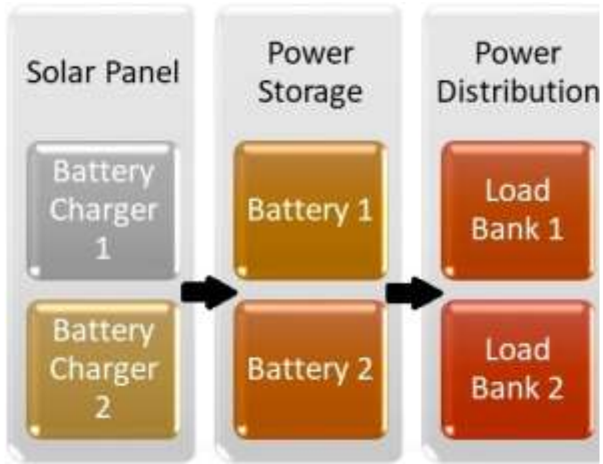

Fig 1. Physical structure of ADAPT

3. Methodology

The overall methodology of the experiment is demonstrated in the following block diagram Fig.2. The first stem is data preprocessing. This step is done to process the data because the telemetry data obtained from the dataset is of high dimension and of high complexity. Hence, to reduce the dimensions, PCA (Principal Component Analysis) and KPCA (Kernel PCA) is performed on the data. PCA is a statistical procedure to reduce the computational complexity of the data, and the Kernel PCA is another form of the PCA. Hence the procedure becomes easier after the PCA and KPCA is performed [10].

The next step is the application of DNN on the datasets. It is performed in the Python Programming Language using the library ‘Keras’. It is an open source neural network python library. The nodes or neurons are defined using the activation function. The activation function introduces nonlinear function into the network. 3 main layers used are Rectified Linear Unit Activation Function (ReLU) as 2 layers and the Sigmoid Function as the third layer. Finally, the accuracy of the system is performed to evaluate the execution of the model [11].

Fig 2. Methodology

A. Standardisation

Different methods of preprocessing of data isstandardization and normalization.
Normalization is a process by which all the values ina dataset are rescaled into values between 0 and 1. But this technique is not used to avoid bounding problems in the data. Hence standardization ispreferred.

Standardization is the process of converting variables to the same scale of having their mean as 0 and the standard deviation as 1. It is important to standardize the data to make sure that the data is comparable and is internally consistent, which means that each data type has the same content and format. Here, in this particular dataset, each variable define different parameters (voltage, current, position, digital), which are not comparable in their normal state. Hence, for easier analysis, standardization is performed to bring all the variables to a common, comparabletype.

The formula for standardizationis

$X = \frac{x-\mu}{\sigma}$ (1)
where,

x= Input variable, μ = mean
σ = Standard Deviation, X = Standardised value

B. PCA and KPCA

The next step to be approached is the calculation of the Principle Component Analysis. It is an unsupervised technique for linear transformation most predominantly used for feature extraction and dimensionality reduction. In basic algebra, PCA has an algorithm that explains the relation between the data containing the variables as columns and samples as rows. The aim of the PCA algorithm is to reduce the large related variables into smaller number [11]. Its central idea is to reduce the dimensionality of the data consisting of unrelated variableswhilekeepingbackthevariationsthatispresent in the available datasets. In the analysis, first the orthogonal Eigen vectors of the covariance matrix of the variables is calculated[12].

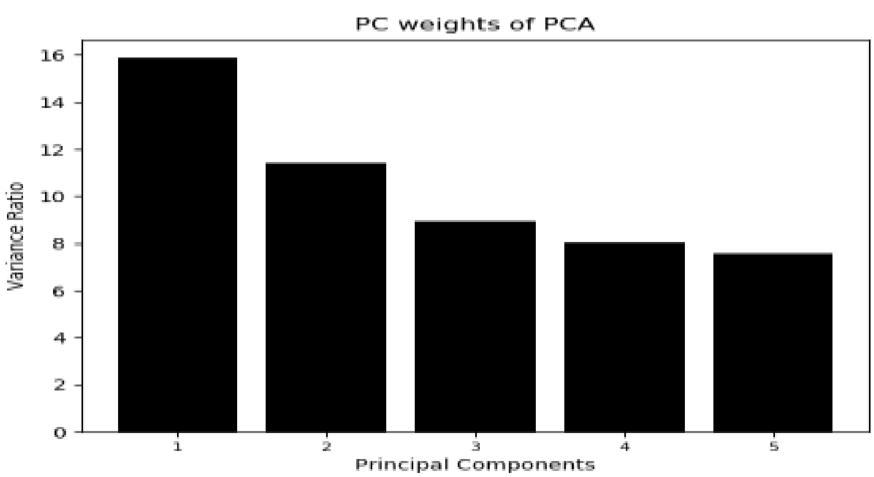

Fig 3. Variance ratio graph of PCA

Algorithm:

- Standardization of the dataset

- Computation of the covariance matrix in the standardized dataset
- Determining the Eigen values and Eigen vectors to determine the principal components.
- To determine the dominant principal component vectors.

The first principal component accounts for the maximum variation percentage and the second principal component accounts for the next largest variation. Finally, it explains the amount of variance with few numbers of components.
The mathematical formula for covariance matrix is given by:

$$C(x) = \frac{(X - X^I)(X - X^I)^T)}{n-1}$$

where,
$C\ (x)$ = Covariance Matrix
X = Sample Value of X
X^I = Sample mean of X
N= No. of variables
$(X - X^I)^T$= transpose matrix of $(X - X^I)$

Here, for this datasets, five principal components are taken to project the variance and is plotted accordingly.

KPCA or the kernel PCA is necessarily an extension of the PCA using kernel methods. Kernels are nonlinear functions that help to perform linear operations ina high-dimensional feature space, where it is much easier toseparatecomponentsinthedata[13].KernelPCAhas been employed as feature extraction method in fault diagnosis of satellite power systems combined with a Multi-layer Perceptron classifier [14]. PCA and KPCA have been obtained for the 60 experiments of ADAPT Industrial tier 1 Competition Dataset. The figures 3 and 4denotethevarianceratiosofthedatasetobtainedusing PCA and KPCAmethods.

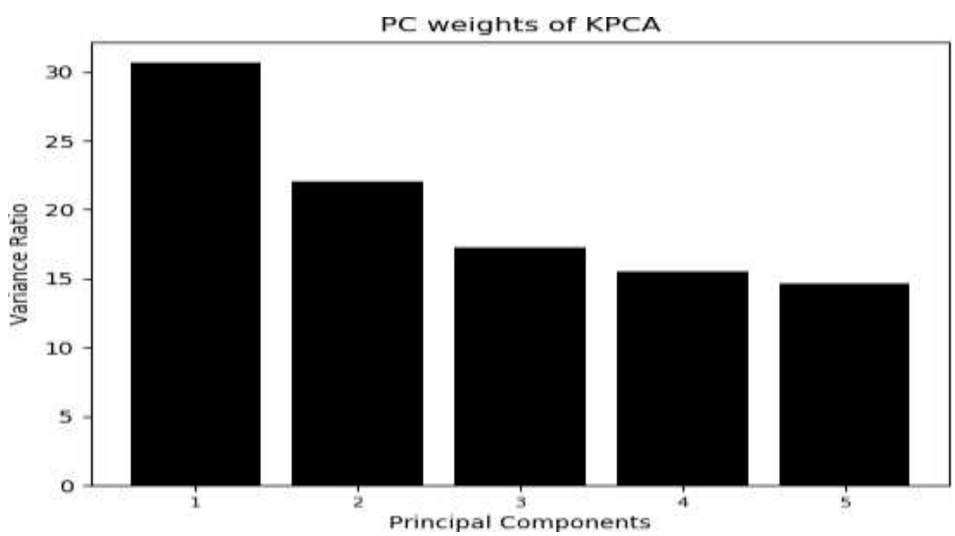

Fig 4. Variance ratio graph of KPCA

Method	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5
PCA	15.95	11.60	9.07	7.99	7.40
KPCA	30.00	22.75	17.57	16.09	15.70

Table 1: Variance Ratios Comparison of PCA and KPCA

From Table 1, we conclude that the variance ratios of the Eigen Vectors (Features) obtained through Kernal PCA is better than the that obtained through PCA.
KPCA has extracted more features in the first 5 Eigen-vectors than PCA. The first 5 Eigen Vectors

of KPCA gives almost 100% variance ratio when compared with PCA which gives 52%. Thus a more efficient feature extraction has been done by Kernel PCA.

4. Deep Learning Model Using Keras

Deep Learning Model is designed using Keras. Keras is a python library that uses Theano and Tensor Backends. Keras can be used as a classification model and as a regressionmodel. But for our need, a classification model will be used. It is a user friendly library that is designed to create consistency and provide clear feedback upon error. In this method, binary classification is done with the class variables 0 for no fault and 1 for fault. The activation functions used areReLuandSigmoid functions. Relurefers to normalization technique for removing negative values[15]. Finally, the validation set is used to determine if a trained system has accomplished the required accuracy with given trained set. Without the validation technique, the model can fall overfitting[16]. Figure 5 denotes the algorithm that is followed by Deep Neural Network.

Algorithm:

- the loaded data is split into input and output variables. The input variables of the dataset are the sensor outputs from the ADAPT model. This dataset consists of readings from 21 sensors. Hence 21 (0-20) variables. The output variable is the class variables (0 or 1). This is the 22nd column of the dataset.
- Keras modelconsistsofasequenceoflayers.The number of nodes in the hidden layers is done using trial and error method. In this method, fully connected network structure with 3 layers is used. The layers are defined using the Dense Class. It defines the fully connected neural network layer.
- The numberofneutronsornodesinthenetwork layer is defined using the activation function.
- The 3 layers used are: rectified linear unit activation function (2 layers) and a sigmoidfunction for the third layer.
- Now when the number of nodes are given and the activation function is defined, the model is compiled using Theano or Tensorflow. These are called backends. These backends choses the best way for the GPU to process the code and the input. They can compute multi-dimensional numpyorSciPy arrays efficiently.
- Training a network requires a loss function to evaluate the best set of weights to map the input and output. For this,a loss function is required. In this case, “binary crossentropy” is used as the loss function. It is the default loss function for binary classification problem. The other loss functionsthat can be used for this type of datasets are Hinge Loss and Squared HingeLoss.
- An optimizer is used to search throughdifferentweights and collect optimal data needed during training. In this case, “Adam” optimizer function is used. Adam is used to update the network weights iteratively using the training data.

- After this is done, the Keras model is fit and executed. Training occurs in epochs and each epochs consists of batches. For n epochs and m batch size, there will be (n/m) updates to the weights of themodel.
- This execution the neural network is evaluated by determining the accuracy of themodel.
- Finally, to understand the work graphically, the plots of loss function versus epoch and accuracy versus epoch isobtained.

We have attached the algorithm of the Deep Learning Model herewith.

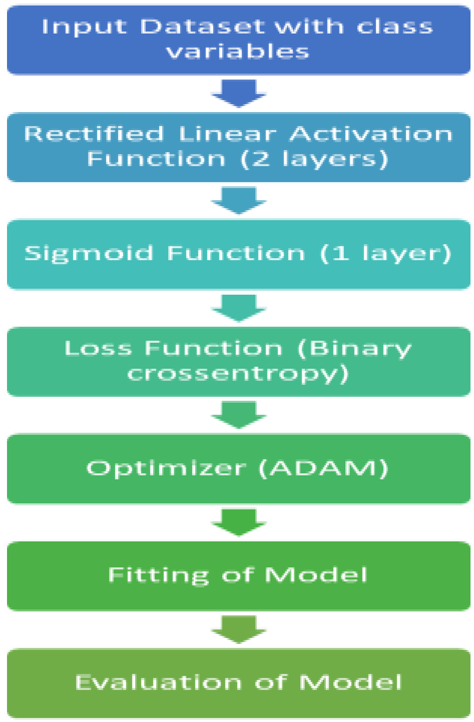

Fig 5. Deep Learning Algorithm

5. Results and Discussions

In this method, we have used 3 layers of neural network with the first 2 having ReLu activation function having 12 nodes in the first layer and 8 nodes in the second layer. The output layer has the activation function Sigmoid with a single node. The model has a batch size of 10 running 150 epochs. The figures 6,7 and 8 denote the accuracy curve, loss function curve and the receiver operating characteristic curve.

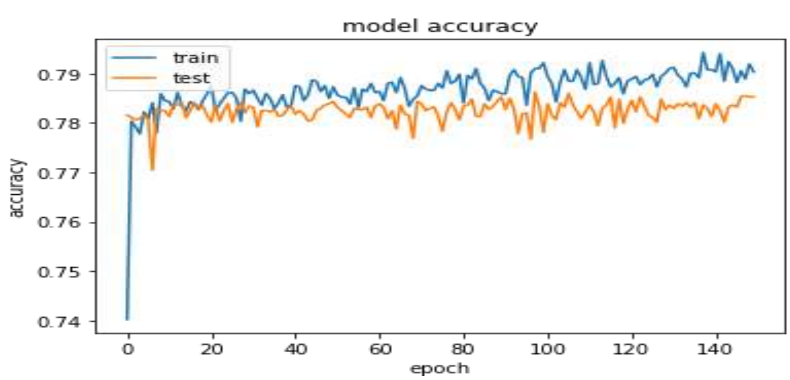

Fig 6. Model Accuracy Curve

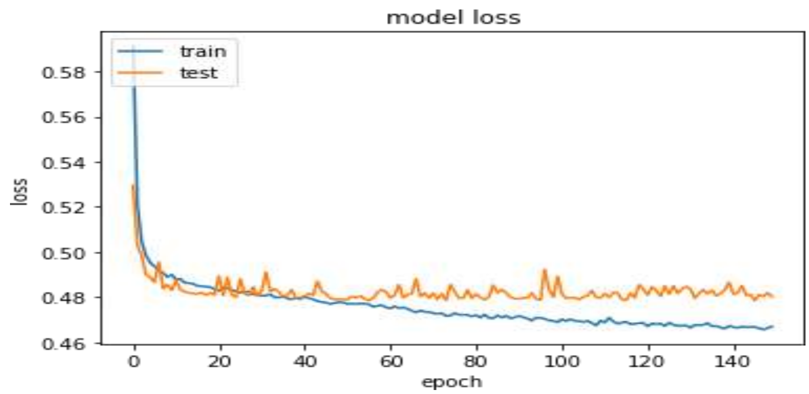

Fig 7. Model Loss Curve

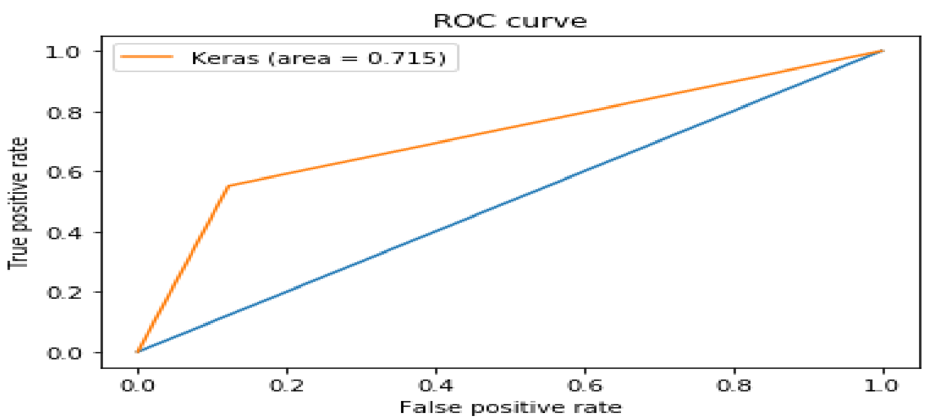

Fig 8. Receiver Operating Characteristic Curve

Table 2: Evaluation Parameters of Deep Learning Model

Accuracy	0.79834
Precision	0.665022
Recall	0.551865
Cohens Kappa	0.451264
ROC AUC	0.714869

From Table 2, we can see that the accuracy obtained is 80%. This method provides a good accurate model for prediction of faults that can occur in the power system of a satellite. The precision value denotes the percentage of results that are relevant. A precision score of 67% is obtained and the algorithm can classify more than 55% of the values correctly. Cohen's Kappa value helps in determining the performance of our classifier over a random classifier and the value 0.45 denotes a balanced, moderate character. The Receiver Operating Characteristic Curve gives an area of 0.71 under the curve. Thus, this classifier can be used to classify and detect faults in a power system of a satellite. Compared to [14], as we considered all fault conditions in all the 60 experiments in ADAPT Industrial Tier 1 Competition Dataset, a trade-off in accuracy was noted due to the huge complexity of the dataset used in ourexperiment.

6. Conclusion

In this work, Deep Learning is used to detect and classify faults in a satellite power system. The dataset used consists of sensor output values obtained over a period of time. This method first extracts therequired features using dimensionality reduction methods of standardization and PCA. The extracted features is used to train and test the deep learning model. The feature extraction methods used are PCA and KPCA. From these methods, we could infer that for a non-linear dataset such as ADAPT, KPCA gives more accurate features. The main advantage of the deep learning model is the deep neural network which is very efficient as a classifier and are suitable for modelling complex non-linear relationships. The 3 layer model of DNN is built with the ReLu and Sigmoid activation function and this combination increases the accuracy and the precision of the system. The prediction model helps in predicting the outcomes while considering the sensor data and helps in identifying the faults. The accuracy obtained of the prediction model is 80%. This method can be very efficientusedforalargenumberofsensorinputsandcanbe very useful in determining faults in any kind of fault systems and can be an important part in the healthcare structure of systems. Different types of pre-processing methods can be pursued as a future scope to increase the accuracy and the different kinds of faults can be classified by using multi-class classification method.

References

[1] B. W. Ricks and O. J. Mengshoel, “Methods for probabilistic fault diagnosis: An electrical power system case study,” Annu. Conf. Progn. Heal. Manag. Soc. PHM 2009, pp. 1–12,2009.

[2] S. Poll, Advanced Diagnostics And Prognostics Testbed.2007.

[3] G. Nicchiotti, L. Fromaigeat, and L. Etienne, “Machine Learning Strategy for Fault Classification Using Only Nominal Data,” pp. 1–9,2013.

[4] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” vol. 9, pp. 249– 256, 2010.

[5] X. Du, Y, Cai, S, Wang and L. Zhang, "Overview of deep learning”, 31st Youth Academic Annual Conference of ChineseAssociation of Automation”, pp. 159-164,2016.

[6] B. Lee and X. Wang, “Fault detection and reconstruction for micro-satellite power subsystem based on PCA,” ISSCAA2010 - 3rd Int. Symp. Syst. Control Aeronaut. Astronaut., pp. 1169–1173,2010.

[7] T. M. Lim, A. M. Cramer, S. Member, S.A. Rawashdeh, and J.

[8] E. Lumpp, “A Modular Electrical Power System Architecture for Small Spacecraft,” vol. 9251, no. c,2018.

[9] M. Cho, “Failure mechanisms and protection methods of spacecraft power system,” Proc. Int. Symp. Electr. Insul. Mater., vol. 1, pp. 45–48,2005.

[10] D. Gorinevsky, S. Boyd, and S. Poll, “Estimation of faults inDC electrical power system,” Proc. Am. Control Conf., pp. 4334– 4339, 2009.

[11] S. Rajarajeswari and K. Somasundaram, “An Empirical Study on feature selection for Data Classification,” no. 3, pp. 3–7, 2012.

[12] Sruthi S., Harpreet S., Mohit A., Bhasker V. And Shantanu,“Data Analysis using principal component analysis”, International Conference on Medical Imaging, m-health and Emerging Communication Systems, pp. 45-49, Nov.2014.

[13] H. Fang, Y. Dong, H. Shi, H. Fan, and S. Ren, “Spacecraft power system fault diagnosis based on DNN,”2017.

[14] Syed Z. Rezvi, JavadMohammadpour, Roland Tóth and Nader Meskin, “A Kernel-Based PCA Approach to Model Reduction of Linear Parameter Varying Systems”, IEEE Transactions on Control Systems Technology, vol. 4, pp. 1883-1891, Sept.2016.

[15] Dheepadarshini S., Anandh. S., Bhavinaya K B. and LavanyaR., “Multivariate Time-series Classification for Automated Fault Detection in Satellite Power Systems”, International Conference on Communication and Signal Processing, pp. 814-817,2019.
[16] Emani, Sarma, K. P. Soman, VV Sajith Variyar, and S. Adarsh. "Obstacle Detection and Distance Estimation for Autonomous Electric Vehicle Using Stereo Vision and DNN." In Soft Computing and Signal Processing, pp. 639-648. Springer, Singapore,2019.
[17] Rajkumar, A., M. Ganesan, and R. Lavanya. "Arrhythmia classification on ECG using Deep Learning." In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), pp. 365-369. IEEE,2019.