International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 5, Issue 06, June-2019

DEVELOPING A MATHEMETICAL MODEL USING MACHINE LEARNING IN FAULT DIAGNOSIS OF AUTOMOBILE SYSTEMS

Vinayak P B[1], Madhusudan K[2]

[1]Assistant Professor, Department of Mechanical Engineering, New Horizon college of Engineering, Bangalore.
[2]Assistant Professor, Department of Mechanical Engineering, New Horizon college of Engineering, Bangalore.

Abstract— ***Online fault diagnosis system level for a mechatronic system makes use of a number of sensors which provide number of data from various components that they are linked with the help of the reliability of any system. It increases reducing the run time error by providing accurate results from the previous data measurement. The current technological development in automobile industry demand Accuracy and Precision more than anything to meet such high expectations. One must incorporate the knowledge of various learning techniques which provide a gateway to overcome this reverberating situation. This is where machine learning plays vital role because of number of options it provides to solve a particular problem. Decision tree and regression techniques being one among them are the ones incorporated here in order to perform fault diagnosis on any system. The interrelations between all the sensor measurements must be timely monitored to collect data which can be put in any one of the machine learning technique to enhance the output.***

This is challenging and often demands for better understanding of the techniques so as to select the best one for our problem which in the end generates the required output in this project. We try to enhance the diagnosis process with the help of basic level machine learning techniques which meet the preliminary requirements of the data collected from automobile that is the Go Kart which we made use of. A sufficient amount of data is collected with the help of sensors and then later used for enhancing the output of the system.

Keywords— Fault Diagnosis, Go Kart, Logistic Regression

I. INTRODUCTION

Machine learning is one of the fastest growing interdisciplinary areas of Engineering, with wide range of applications. The project aims to introduce machine learning, and the algorithmic paradigms it offers, in a principled way. It is one of the technique from which computers can learn from input available to them. The algorithm learns by the input data we are giving which inturn represents an experience on a particular task. The learning depends majorly of the data we have collected. We are trying to study particular pattern a data is exhibiting. So the data collection has to be genuine.

EXISTING FAULT DETECTION PROBLEMS:

The correct diagnosis of fault is of utmost importance automotive systems both from safety and maintenance point of view. These systems consists of software and hardware components. High level of automation with software assistance is continuously being used. Up to recent driver assist was provided in the form of warnings, but increasingly such assistance will be provided by actively manipulating actuators that control vehicle longitudinal acceleration and deceleration, lateral position, and vertical displacement. The long term trend is towards partial or even fully autonomous operation of a single vehicle, or even of fleets of vehicles.

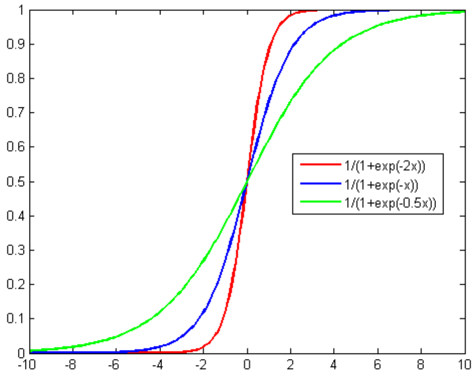
Usually the fault diagnosis is conducted offline in the present scenario i.e when the vehicle in not in its functioning mode. The challenge now is to do an online diagnosis so that preventive or corrective actions can be taken immediately to avoid major altercations. But for this the challenge is to correctly predict the fault with minimum error. Also the time taken for this correct diagnosis of the fault plays a key role.

Failure diagnosis goes beyond fault detection by providing extended information on the underlying cause of a system failure. Failure diagnosis is distinguished from fault detection in that detection aims mainly to determine that some fault occurred, while failure diagnosis might reveal what kind of fault occurred, what component(s) is/are responsible, and what caused the fault.

LOGISTIC REGRESSION IN MACHINE LEARNING:

Classification problems are an important category of problems in analytics in which the response variable (Y) takes a discrete value. The primary objective is to predict the class of a customer (or class probability) based on the values of explanatory variables or predictors.

$$\frac{e^{Z}}{1+e^{Z}}$$



logistic regression estimate probability of one class of event
Logistic Function (Sigmoidal function)

$$P(Y=1)=\pi(z)=\frac{1}{1+e^{-z}}=\frac{e^{z}}{1+e^{z}}$$
$$z=\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_nx_n$$

II. Experiment

Model Building:
The data is obtained from a GO-KART engine. The variables selected for the evaluation of fault diagnosis are Coolant temperature, Engine Load, Engine rpm and Intake manifold pressure. A total of 150 observations were recorded.

Coolant Temp	Eng Load	RPM	Intake Manifold Pressure	Air Intake Temp	Fault Detected	Fault
80	33	1009	49	59	NO	0
80	32	1003	52	59	NO	0
80	32	995	51	59	NO	0
80	32	1004	51	60	NO	0
80	32	1005	49	60	NO	0
81	32	1019	50	60	NO	0
81	33	1002	49	60	NO	0
81	33	1001	50	60	NO	0
81	33	1004	52	59	NO	0
81	32	1010	52	58	NO	0
81	34	1008	49	59	NO	0
91	34	1005	50	59	NO	0
81	34	1004	51	62	YES	1
91	32	998	53	60	NO	0
81	34	999	68	63	YES	1
81	37	1589	47	64	YES	1
81	26	2955	33	65	YES	1
82	36	1050	50	65	YES	1
77	29	1005	46	42	NO	0

The probability is mathematically represented by the below formula:

$$P(Y=1) = \frac{e^{Z}}{1+e^{Z}}$$

Where,
$Z=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4$
Where, β0,β1, β2, β3, β4 are constants
X1= Coolant temperature;
X2= Engine Load;
X3= Engine rpm;
X4= Intake manifold pressure

```
Summary of the Logistic Regression model (built using glm):

Call:
glm(formula = FAULT ~ ., family = binomial(link = "logit"), data = crs$dataset[crs$train,
    c(crs$input, crs$target)])

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7856  -0.4896  -0.4059   0.1302   2.2299

Coefficients:
                          Estimate  Std. Error z value Pr(>|z|)
(Intercept)            -21.0278717   5.6678377  -3.710 0.000207 ***
coolant.temp             0.0873624   0.0533350   1.638 0.101423
eng.load                 0.1648615   0.0482641   3.416 0.000636 ***
rpm                      0.0034861   0.0009683   3.600 0.000318 ***
intake.manifold.pressure 0.0740000   0.0265926   2.783 0.005390 **
---
```

Using Rattle for model building the equation obtained is as below
Z= -21.037 + 0.087*X1 + 0.164*X2 + 0.0034*X3+ 0.074*X4

III. Result and Discussion

Model Validation

Classification Table (Cut-Off Probability = 0.5)

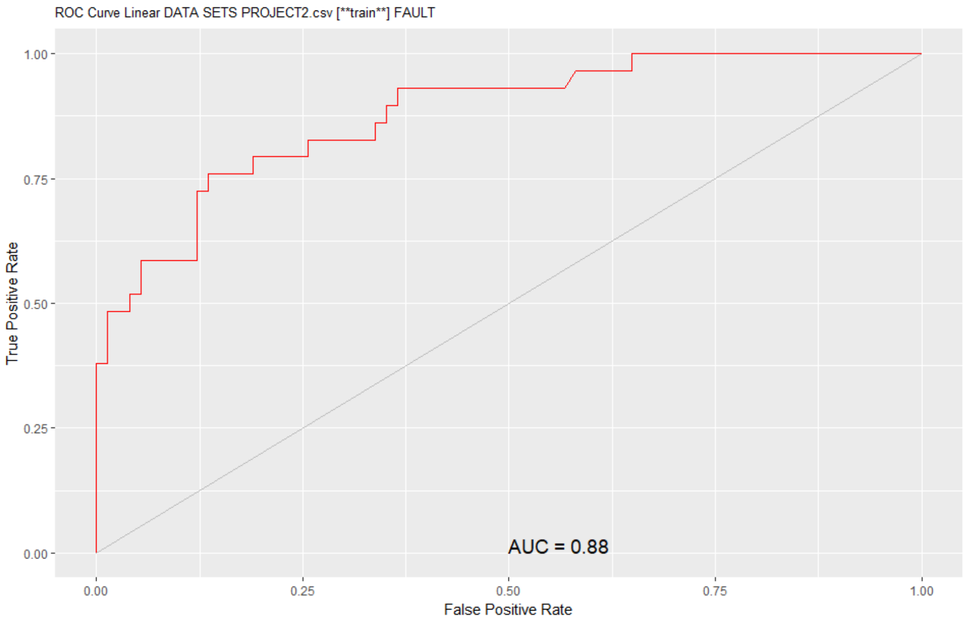
	0	**1**
0	98 (TN)	7(FP)
1	14(FN)	31 (TP)

SENSITIVITY = TP/ (FN+TP) = 0.688

SPECIFICITY = TN/ (TN+FP) = 0.9333

PRECISION = TP/ (TP+FP) = 0.815

ROC Curve



IV. Conclusion

The ROC Curve built for the Model is giving the Area Under the Curve value of 0.88 which can be considered satisfactory. The Specificity value of the model i.e the accuracy of the model to correctly predict no maintenance required when no actual maintenance required is 93.33 % which is acceptable from the Insurance application point of view.

REFERENCES

[1] Hong Guo, J.A.Crossman, Automotive signal diagnostics using wavelets and machine learning, IEEE Transactions on Vehicular Technology, vol 49, issue 5,September 2000,pp.1650-1662.

[2] Chaitanya Sankavaram, Data-driven fault diagnosis in a hybrid electric vehicle regenerative braking system, 2012 IEEE Aerospace Conference, vol 2, April 2012, pp.1225-1239.

[3] Yimin Shao, Fault Prognosis and Diagnosis of an Automotive Rear Axle Gear Using a RBF-BP Neural Network, Journal of Physics, vol 305, number 1, September 2015, pp.63-69.

[4] T. Praveen Kumar, Fault Diagnosis of Automobile Gearbox Based on Machine Learning Techniques, Elsevier, vol 97, 2014, pp.2092-2098.

[5] Lei Liu, Research on Logistic Regression Algorithm of Breast Cancer Diagnose Data by Machine Learning, International Conference of Robots and Intelligent systems, July 2018.

[6] Glenn. V. Ostir, "Logistic Regression: A Nontechnical Review", American Journal of Physical Medicine & Rehabilitation., vol. 6, 2000, pp. 565-572.

[7] C. Carter, J. Catlett, "Assessing credit card applications using machine learning", IEEE Expert, vol. 2, no. 3, 1987, pp. 71-79.

[8] G. S. Fang, "A note on optimal selection of independent observations", IEEE Trans. on Systems Man and Cybernetics, vol. 9, May 1979, pp. 309-311.