



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: III Month of publication: March 2023

DOI: <https://doi.org/10.22214/ijraset.2023.49809>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Laptop Performance Prediction

Adithya T¹, Dr. T. A. Albinaa²

¹PG Research Scholar, ²Assistant Professor, Department of Data Analytics (PG), PSGR Krishnammal College for Women, Coimbatore, Tamil Nadu

Abstract: The “LAPTOP PERFORMANCE PREDICTION” is an important metric used to predict the performance of laptops. Laptop manufacturers spend a lot of time, resources, and money designing new systems and newer configurations. Their ability to reduce costs, charge competitive prices, and gain market share depends on how well these systems perform. In this work, concentrate on the system and architectural design processes for parallel computers and develop methods to expedite them. The methodology relies on extracting the performance levels of a small fraction of the machines in the design space and using this information to develop machine learning models to predict the performance of any machine in the whole design space. Laptop performance prediction is useful to accelerate the design space exploration significantly and aid in reducing the corresponding research/development cost and time-to-market.

Keywords: Extra Tree Classifier, Laptop performance, Machine learning, Prediction

I. INTRODUCTION

Laptop performance prediction especially when the laptop is coming direct from the factory to Electronic Market/Stores is both a critical and important task. The mad rush for laptops to support remote work and learning is no longer there. In India, the demand for Laptops soared after the Nationwide lockdown, leading to 4.1 million unit shipments in the June quarter of 2021, the highest in years. Accurate Laptop performance prediction involves expert knowledge, because of distinctive features and factors. Typically, the most significant brands, models, processors, operating systems, storage, Graphics cards, displays, Warranty, Price Different methods, and techniques are used in order to achieve higher precision of the laptop performance.

II. RELATED WORK

Predicting the performance of laptops has been studied extensively in various kinds of research. Listian discussed, in her paper written for Master thesis, that a regression model that was built using Decision Tree & Random Forest Regressor can predict the price of a laptop that has been leased with better precision than multivariate regression or some simple multiple regression. This is on the grounds that Decision Tree Algorithm is better in dealing with datasets with more dimensions and it is less prone to overfitting and underfitting. The weakness of this research is that a change of simple regression with more advanced Decision Tree Algorithm regression was not shown in basic indicators like mean, variance or standard deviation.

III. PROBLEM DESCRIPTION

The problem statement is that if any user wants to buy a laptop then our application should be compatible to provide a tentative performance of the laptop according to the user configurations. To come up with the best method to select the laptop in the industry for customers to find the laptop suitable for their needs and the clients over the e-commerce industry to improve selling out of laptops and predict the requirements of the laptops. if any user wants to buy a laptop then our application should be compatible to provide a tentative performance of the laptop according to the user configurations.

IV. METHODOLOGY

A software system has intercommunicating components based on software forming part of a computer system (a combination of hardware and software). With the retail market getting more and more competitive by the day, there has never been anything more important than optimizing service business processes when trying to satisfy customers' expectations. Channelizing and managing data with the aim of working in favor of the customer as well as generating profits is very significant for survival. For big retail players all over the world, data analytics is applied more these days at all stages of the retail process – taking track of popular products that are emerging, doing forecasts of sales and future demand via predictive simulation, optimizing placements of laptops and offers through heat-mapping of customers and many others. By using machine learning algorithms such as Random forest classifier and logistic regression, The prediction rate of the algorithms will be calculated.

And the model is built for the prediction of the calculated results based on the testing and training data. The First step is Data gathering. This step is very important because the quality and quantity of the data gathered will directly affect the level of the prediction model. So this data is a list of over 332 records of laptops from laptop retailers. The dataset includes basic product information, price, ratings, and various hardware and software parts. The attributes such as Brand, Processor, Ram(GB), and Price plays a vital role in predicting the performance of the laptops. Brand – The different brands of laptops with versions, Processor – Processing unit of laptops i.e memory, RAM – Random Access Memory in GB, Operating System – OS (Windows or MAC), Storage – Capacity of the laptops in SSD (Flash Based Memory), Frequency in HZ i.e., 1,2,3,4,5, Graphics Card – Visions for categories i.e. gaming- advanced and normal versions, Display – Inches of the screen in Centimeters, Warranty – Number of months or years and the type of warranty, Price – market price of the laptop from the retailers, Rating – customer ratings of the laptops.

S.NO	Brand	Processor	RAM(GB)	OperatingSystem	Storage(SSD)	Frequency(Hz)	GraphicsCard	Display	Warranty	Price	rating
327	ASUS VivoBook S13	Intel Core i3 Processor (11th Gen)	8	64 bit Windows 10 Operating System	150	2	Radeon RX 6600 XT.	39.62 cm (15.6 inch) Display	1 Year Onsite Warranty	55000	4.8
328	HP Pavilion Core	Intel Core i5 Processor (10th Gen)	8	64 bit Windows 10 Operating System	512	1	RX 7900 XTX	33.78 cm (13.3 inch) Display	1 Year Onsite Warranty	35000	4.4
329	Lenovo IdeaPad Core	Intel Core i7 Processor (10th Gen)	8	64 bit Windows 10 Operating System	512	5	Ultra	39.62 cm (15.6 inch) Display	2 Year Warranty	65000	4.3
330	Lenovo Ideapad S145 Ryzen 3	Intel Core i7 Processor (10th Gen)	8	64 bit Windows 10 Operating System	512	6	FHD gaming	39.62 cm (15.6 inch) Display	1 Year International Travelers Warranty (ITW)	55000	5.0
331	HP 14s Core	Intel Core i3 Processor (10th Gen)	8	64 bit Windows 10 Operating System	256	4	Arc 450	40.64 cm (16 inch) Display	1 Year Onsite Warranty	35500	4.8

Fig.1 Dataset Description

V. DATA PREPROCESSING

The data are cleaned and processed before sending it into a model. The null values are checked and either removed or replaced. In this dataset, no null values were found and the data consist of categorical and numerical values. Both the numerical and categorical values will be converted into 0 and 1 for the modeling process. As laptop data does not contain any null value so the data present over the records consist of numerical and categorical values which will be evaluated in the modeling process.

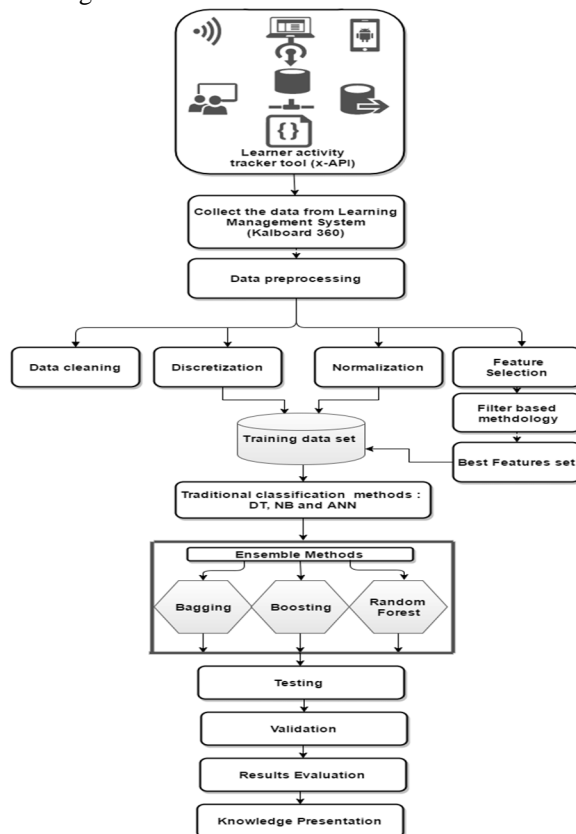


Fig.2 System flow diagram

VI. PROCESS FLOW

As laptop data does not contain any null value so the data present over the records consist of numerical and categorical values which will be evaluated in the modeling process. As data consists of more categorical values than numerical values, Which is the biggest drawback of building a model. So by finding the correlation of the data set will give the skewness of the data. By using statistical function `mode()`. Those categorical values get replaced with the mode of the particular attribute. This statistical function is followed for three attributes such as brand, Processor, and display. The Operating system is of two types MAC and Windows OS. And the ratings and price of the laptops are categorized into a mean of two. They are turned into 0 and 1 during the manipulation process. The specifications of the users such as Students, Developers, Basic users, and Gamers are mentioned in the data that to be trained. All the categorical values are turned into numerical values.

VII. MODEL EVALUATION

Logistic Regression, Random Forest Classifier, and Extra tree classifier are implemented for training and testing the model For choosing a model we split our dataset into training and testing. Here data are split into a 3:1 ratio which means training data has 70 percent and testing data has 30 percent. In this split process performing based on the `train_test_split` model. After splitting we get `x_train`, `x_test` and `y_train`, `y_test`[3,7].

A. Logistic Regression

Logistic regression is a predictive analysis. Logistic regression describes data and explains the relationship between one dependent binary variable and one or more nominal, ordinal, interval, or ratio-level independent variables. Another important consideration is the model fit when selecting the model for the logistic regression analysis. Split our dataset to train and test set and fit the dataset to the Logistic regression model. The assumptions made by logistic regression about the distribution and relationships in your data are much the same as the assumptions made in linear regression.. The trained logistic regression model and applying to a testing data set. The dependent variable is binary (Boolean). For each sample in the testing data set, while applying the logistic regression model to generate an accuracy level of 80 percent.

```
LogisticRegression()

#Training the model
logistic_model_2=LogisticRegression()

#Testing
predicted_value=logistic_model.predict(test_X)

#accuracy
metrics.accuracy_score(test_Y,predicted_value)+0.5
0.6764705882352942
```

Fig.3 Accuracy level of prediction using Logistic Regression

B. Random Forest Classifier

Random Forest is a flexible, easy-to-use machine learning algorithm that produces, even without hyper-parameter tuning, a great result most of the time. It is also one of the most used algorithms, because of its simplicity and diversity (it can be used for both classification and regression tasks). In random forest classification, multiple decision trees are created using different random subsets of the data and features. Each decision tree is like an expert, providing its opinion on how to classify the data. Predictions are made by calculating the prediction for each decision tree, then taking the most popular result.

```
from sklearn.ensemble import RandomForestClassifier
rf_classifier = RandomForestClassifier
rf_classifier = RandomForestClassifier(n_estimators = 20,criterion = 'entropy',random_state= 51)
rf_classifier.fit(X_train,y_train)
y_pred = rf_classifier.predict(X_test)
accuracy_score(y_test,y_pred)

0.02
```

Fig.4 Accuracy level of prediction using Random Forest Classifier

C. Extra Tree Classifier

An extra tree is an algorithm used for classification and regression tasks. It works by randomly selecting a subset of features and then training a Decision Tree. The tree is then pruned only to contain the most important features for making predictions. The Extra tree algorithm is considered an efficient and accurate machine learning method. It is similar to other methods such as decision trees and random forests, but it uses extra information about the data to improve predictive accuracy. Additionally, the extra tree algorithm is faster and easier to implement than other methods. As a result, it is a powerful tool for data mining and predictive modeling.

```

model = ExtraTreesClassifier(n_estimators=100,)
model.fit(test_X, test_Z)
pd.Series(model.feature_importances_).sort_values(ascending=False)

0    0.440656
1    0.260313
4    0.136753
3    0.099206
2    0.063072
dtype: float64

```

Fig.5 Accuracy level of prediction using Extra Tree Classifier

Table I

Algorithm	Accuracy
Logistic Regression	0.67
Random Forest Classifier	0.02
Extra Tree Classifier	0.85

Table 1.Comparison Of Prediction Accuracy of algorithms

VIII. RESULT

In classification, accuracy is an important evaluation parameter. Accuracy is the proportion of the total number of predictions that were correct. It can be obtained by the sum of true positive and true negative instances divided by 100. And Precision is fraction of true positive and predicted yes instances. The accuracy score of the three algorithms is evaluated and the prediction model is built using the CSV file by defining the object function and reading the file. The categories of the predictions were defined for Student, Developer, Gamer, and Basic. Based on the features and specifications of the laptops, the categories are trained into the model. By using the python modules, those values will get imported. The dataset consists of 332 records, Where the functions and the performance of each laptop will get evaluated. The input and output values are imported and when the specified categories are given in the input function then the laptops which will be performed well over the other laptops with different functions are shown. The laptops for the categories based on the prediction and performance value will be visualized as the output.

IX. CONCLUSION

As Extra Tree classifier gives greater accuracy at 85%. Extra tree classifier groups the data into clusters and the data are trained and tested in five different groups and classified. Based on the training and testing those data will get five parts of accuracy and predictions of the laptops are made by a grouping of provided data. Predicting something through the application of machine learning using the Extra Tree algorithm makes it easy for students and retail stores, especially in determining the choice of laptop specifications that are most desirable for students, and IT professionals to meet needs and in accordance with the purchasing power of people. Students, Gamers, Developers, and Basic users no longer need to look for various sources to find laptop specifications that are needed by students meeting the needs of students, because the laptop specifications from the results of the machine learning application have provided the most desirable specifications with their prices of laptops.

X. FUTURE WORK

The data which consists of eleven attributes could be increased and the number of records over the data could be more. The prediction of the data will be improved by implementing many other algorithms.

- 1) Test more attributes and many categories of laptop users
- 2) Find a solvent feature extraction technique to improve results.



REFERENCES

- [1] Akay MF, Aci CI, Abut F (2015) Predicting the performance measures of a 2-dimensional message passing multiprocessor architecture by using machine learning methods. *Neural Netw World* 25:241–265
- [2] Bace R, *Intrusion Detection*, Macmillan Technical Publishing, 2000.. Agnar Aamodt, Enric Plaza. "Foundational Issues, Methodological Variations, System approaches." *AICom – Artificial Intelligence Communications*, IOS Press Vol. 7: 1, pp. 39-59.
- [3] B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", *International Journal of Computer Science and Information Security (IJCSIS)*, Vol. 9, No. 4, pp. 136-140, 2011.
- [4] Erkan Er. "Identifying At-Risk Students Using Machine Learning Techniques", *International Journal of Machine Learning and Computing*, Vol. 2, No. 4, pp. August 2012.
- [5] Jayakumar A, Murali P, Vadhiyar S (2015) Matching application signatures for performance predictions using a single execution. In: 2015 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp 1161–1170
- [6] S. Kotsiantis, C. Pierrakeas, and P. Pintelas, Preventing student dropout in distance learning systems using machine learning techniques, *AI Techniques in Web-Based Educational Systems at Seventh International Conference on Knowledge-Based Intelligent Information & Engineering Systems*, pp. 3-5, September 2003.
- [7] Lobachev O, Guthe M, Loogen R (2013) Estimating parallel performance. *J Parallel Distrib Comput* 73(6):876–887
- [8] Smith W (2007) Prediction services for distributed computing. In: *IEEE International Parallel and Distributed Processing Symposium, 2007. IPDPS 2007*, pp 1–10.
- [9] Springer, Berlin, pp 226–236. Prem H, Raghavan NRS (2006) A support vector machine based approach for forecasting of network weather services. *J Grid*
- [10] Sun J, Sun G, Zhan S, Zhang J, Chen Y (2020) Automated performance modeling of HPC applications using machine learning. *IEEE Trans Comput* 69(5):749–763 126. Zhang W, Hao M, Snir M (2016) Predicting HPC parallel program performance based on LLVM compiler. *Clust Comput* 20:1179–1192.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)