Available Online at www.ijcsmc.com/journal

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 11, Issue. 1, January 2022, pg.164 – 168

Laptop Price Prediction using Machine Learning

Prof. Vaishali Surjuse; Sankalp Lohakare; Aayush Barapatre; Abhishek Chapke
Department of Computer Technology, KDKCE & RTMNU University, India
sankalplohakare553@gmal.com; barapatre22aayush@gmail.com; abhichapke264@gmail.com
DOI: 10.47760/ijcsmc.2022.v11i01.021

Abstract— This paper presents a Laptop price prediction system by using the supervised machine learning technique. The research uses multiple linear regression as the machine learning prediction method which offered 81% prediction precision. Using multiple linear regression, there are multiple independent variables but one and only one dependent variable whose actual and predicted values are compared to find precision of results. This paper proposes a system where price is dependent variable which is predicted, and this price is derived from factors like Laptop's model, RAM, ROM (HDD/SSD), GPU, CPU, IPS Display, and Touch Screen.
Keywords— Multiple Linear regression, Laptop Price, Regression model, Machine Learning.

I. INTRODUCTION

Laptop price 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 that we saw in 2020 for laptops to support remote work and learning is no longer there. In India, demand of Laptops soared after the Nationwide lockdown, leading to 4.1-Million-unit shipments in the June quarter of 2021, the highest in the five years. Accurate Laptop price prediction involves expert knowledge, because price usually depends on many distinctive features and factors. Typically, most significant ones are brand and model, RAM, ROM, GPU, CPU, etc. In this paper, we applied different methods and techniques in order to achieve higher precision of the used laptop price prediction.

II. RELATED WORK

Predicting price of laptops has been studied extensively in various researches. Listian discussed, in her paper written for Master thesis, that 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. METHODOLOGY

To support the application of machine learning using the Decision Tree algorithm, of course the sample data is needed. Table below contains data about various laptops and their prices depending on their configuration. Sample data were obtained from Kaggle.com

Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37kg	71378.6832
Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34kg	47895.5232
HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	No OS	1.86kg	30636.0000
Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	512GB SSD	AMD Radeon Pro 455	macOS	1.83kg	135195.3360
Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37kg	96095.8080
...	...	...	...	...	...	...	...	...	...	...
Lenovo	2 in 1 Convertible	14.0	IPS Panel Full HD / Touchscreen 1920x1080	Intel Core i7 6500U 2.5GHz	4GB	128GB SSD	Intel HD Graphics 520	Windows 10	1.8kg	33992.6400
Lenovo	2 in 1 Convertible	13.3	IPS Panel Quad HD+ / Touchscreen 3200x1800	Intel Core i7 6500U 2.5GHz	16GB	512GB SSD	Intel HD Graphics 520	Windows 10	1.3kg	79866.7200
Lenovo	Notebook	14.0	1366x768	Intel Celeron Dual Core N3050 1.6GHz	2GB	64GB Flash Storage	Intel HD Graphics	Windows 10	1.5kg	12201.1200
HP	Notebook	15.6	1366x768	Intel Core i7 6500U 2.5GHz	6GB	1TB HDD	AMD Radeon R5 M330	Windows 10	2.19kg	40705.9200
Asus	Notebook	15.6	1366x768	Intel Celeron Dual Core N3050 1.6GHz	4GB	500GB HDD	Intel HD Graphics	Windows 10	2.2kg	19660.3200

The C4.5 algorithm (used as a Decision Tree Classifier which can be employed to generate a decision, based on sample dataset) starts with the process of selecting the highest gain attribute as the root of the tree, then creates a branch for each value, then divides the case in branches, then repeats the process for each branch until all cases in the branch have the same class.

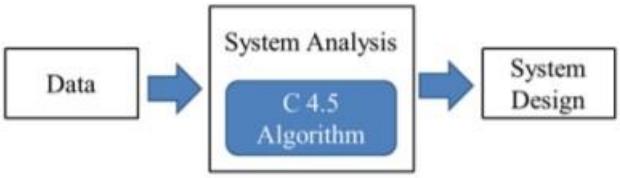

Figure: Flow of design and analysis

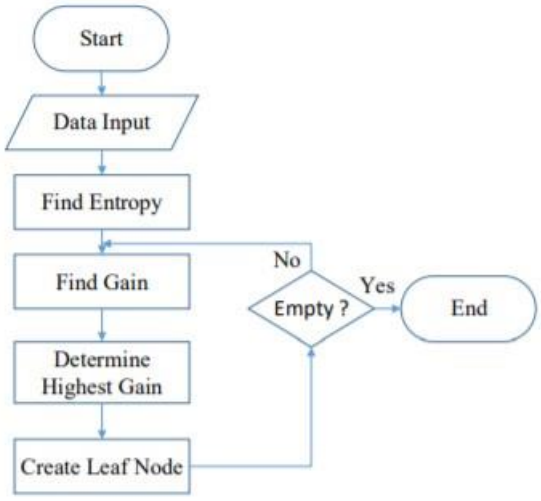

Figure: C 4.5 algorithms flowchart

To get accurate calculation results, calculations used Entropy and Gain for each variable. Entropy measures uncertainty between random variables in a data. The high Entropy value will affect the classification process.

The equation used to calculate Entropy and Gain, as follows:

$$Entropi\ (S) = \sum_{j=1}^{k} -p_j\ \log_2\ p_j$$

S: Case set
k: Number of S partition
Pj: Probability obtained from the total (Yes / No) divided by the total case

$$\text{Gain (S,A)} = \text{Entrophy (s)} - \sum_{i=1}^{n} \frac{|si|}{|S|} * \text{Entrophy (si)}$$

S: Case set
A: Attribute
n: Number of A attribute partition
|Si|: Number of cases on the i partition
|S|: Number of S partition

Explanatory Data Analysis (EDA)

Using our feature-engineered dataset, we can now plot graphs and compute tables to visualize how each feature relates to the variability of laptop prices. By using the barplot method imported from Matplotlib, we can test and verify our hypothesis or initial opinions on how some features will affect the pricing of laptops. Here's an illustration of plotting a barplot for the feature TypeName (type of laptop)

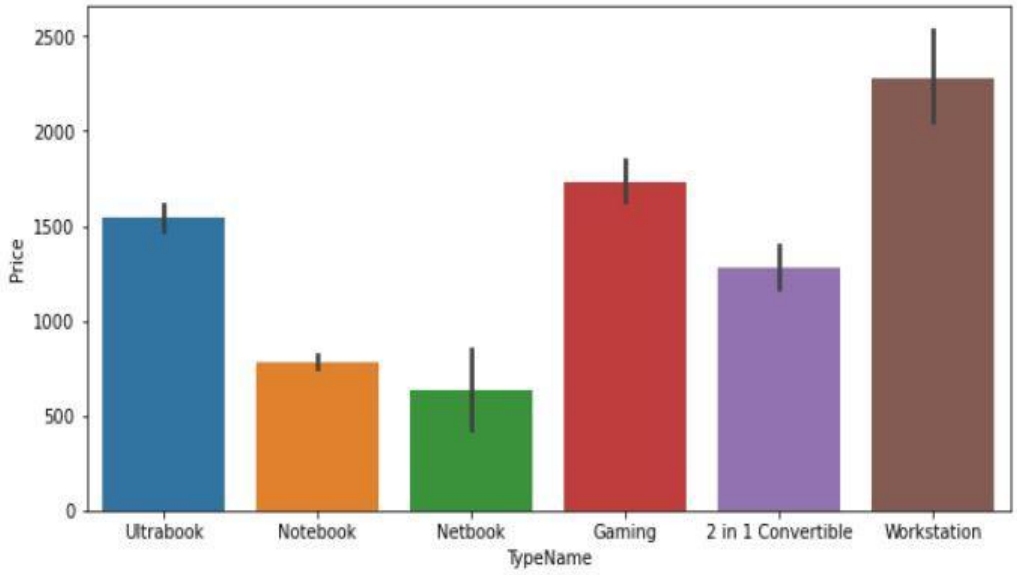

Figure: Data Visualisation using barplot

Result

Streamlit library is used to build this WebApp UI. Streamlit is an (open-source Python library) that makes it easy to create and share, custom web apps for machine learning and data science. Result with backend code is shown in following figures.


```python
import streamlit as st
import pickle
import numpy as np

# import the model
pipe = pickle.load(open('pipe.pkl', 'rb'))
df = pickle.load(open('df.pkl', 'rb'))

st.title("Laptop Predictor")

# Asking the user which brand laptop to choose
company = st.selectbox('Brand', df['Company'].unique())

# type of laptop
type = st.selectbox('Type', df['TypeName'].unique())


# Ram
ram = st.selectbox('RAM(in GB)', [2, 4, 6, 8, 12, 16, 24, 32, 64])

# Weight
weight = st.number_input('Weight of Laptop')

# Touchscreen
touchscreen = st.selectbox('Touchscreen', ['No', 'Yes'])

# IPS Display
ips = st.selectbox('IPS', ['No', 'Yes'])

```

Figure: Integrating ML model with Web Application

Conclusions

Predicting something through the application of machine learning using the Decision Tree algorithm makes it easy for students, especially in determining the choice of laptop specifications that are most desirable for students to meet student needs and in accordance with the purchasing power of students. Students no longer need to look for various sources to find laptop specifications that are needed by students in 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.

ACKNOWLEDGEMENT

Researcher is grateful to the team of entrepreneurs who provide insight and guidance that are very helpful in completing this research. This research is expected to be useful for students in determining laptop specifications along with their prices.

REFERENCES

[1]. Sorower MS. A literature survey on algorithms for multi-label learning. Oregon State University, Corvallis. 2010 Dec;18.

[2]. Pandey M, Sharma VK. A decision tree algorithm pertaining to the student performance analysis and prediction. International Journal of Computer Applications. 2013 Jan 1;61(13).

[3]. Priyama A, Abhijeeta RG, Ratheeb A, Srivastavab S. Comparative analysis of decision tree classification algorithms. International Journal of Current Engineering and Technology. 2013 Jun;3(2):334-7

[4]. Streamlit.io, Kaggle.com, Wikipedia.com

[5]. Ho, T. K. (1995, August). Random decision forests. In Document analysis and recognition, 1995., proceedings of the third international conference on (Vol. 1, pp. 278-282).

[6]. Weka 3 - Data Mining with Open Source Machine Learning Software in Java. (n.d.), Retrieved from: https://www.cs.waikato.ac.nz/ml/weka/. [August 04, 2018].

[7]. Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications, 167(9), 27-31.

[8]. Pudaruth, S. (2014). Predicting the price of used cars using machine learning techniques. Int. J. Inf. Comput. Technol, 4(7), 753-764.

[9]. Listiani, M. (2009). Support vector regression analysis for price prediction in a car leasing application (Doctoral dissertation, Master thesis, TU Hamburg-Harburg).

[10]. Agencija za statistiku BiH. (n.d.), retrieved from: http://www.bhas.ba . [accessed July 18, 2018.]

[11]. Utku A, Hacer (Uke) Karacan, Yildiz O, Akcayol MA. Implementation of a New Recommendation System Based on Decision Tree Using Implicit Relevance Feedback. JSW. 2015 Dec.

