

# Machine Learning Models for Sneaker Price Forecasting

<sup>1</sup>Dr.Abdul Rasool Md, <sup>2</sup>Mohammed Mohiuddin, <sup>3</sup>Mohammad Fayaz, <sup>4</sup>Ibaad Ahmed Ajaz

<sup>1</sup>Associate professor, Dept of CSE-AI&ML, Lords Institute of Engineering and Technology, Hyd.

[rasool.501@gmail.com](mailto:rasool.501@gmail.com)

<sup>2,3,4</sup>BE Student, Dept of CSE-AI&ML, Lords Institute of Engineering and Technology, Hyd.

[anasmohiuddin56@gmail.com](mailto:anasmohiuddin56@gmail.com), [mfayaz7001@gmail.com](mailto:mfayaz7001@gmail.com), [ibaadahmed077@gmail.com](mailto:ibaadahmed077@gmail.com)

***Abstract:** Sneaker tradition is a cultural identity created by using young humans. In the Internet and digital generation, sports activities shoes were rewarded with interest and financial blessings from predictions, which has caused the concept of sports footwear with cooperation of everybody. According to human conduct, the damaged system behind considering footwear is worth thinking about. There are structures of conduct chains and logics inside the conduct of shoes. It is critical to apply the attitude to show the dynamic technique of convergence from micro to macro conduct. Therefore, gadget gaining knowledge of may be used to predict the price movement of shoes inside the sneaker marketplace. All the information for this have a study is from the stock. In research, tool mastering is used to look at the price of sports footwear and examine statistics. Four linear regression fashions were used for estimation. It is determined that the accuracy of OLS regression model is excessive. The OLS version is consequently the extremely good model for these facts.*

**Keywords-** Machine learning; linear regression; price prediction; sneakers price

## I. INTRODUCTION

It changed into located that amongst 2011 and 2019, the worldwide sports activities footwear market length persevered to develop and charges multiplied sharply. In 2018, the shoes market reached \$146.5 billion, an increase of \$seventy. Three billion in assessment to 2011. For the shoe enterprise, which has the developments of the organisation, there may be plenty less

research on the use of tool gaining knowledge of to anticipate the price of footwear. Therefore, in this have a look at, the real trouble to be investigated is how to analyze the accrued information, and what sort of call is most appropriate for the records primarily based mostly on all the assessment information. Shoe rate forecasting may be useful to shoe charge supervisor using shoe platform and

marketplace instructor evaluation using shoe sales platform, to get higher storage fees.

**II METHODOLOGY**

**A. Dataset Analysis and Data visualization**

All Off-White x Nike and Yeezy 350 budget are to be had from January nine, 2017 (the month Off-White first released the "Ten" series) till today. The record includes 99,956 total income, 27,794 Off-White sales, and seventy two, 162 Yeezy incomes. The pattern consists of simplest US profits. First, the author eliminated the information, i.e. Changed the product facts to the sort, removed the "\$" in the sales rate from the attribute and changed it with the statistics of the numeric kind. Renaming the hierarchical line of zones is likewise essential.

The end result of this huge quantity is the mistake of the perspective, and the denominator is the standard length minus the cost of the predicted fee. The error tolerance of these indicators is supplied as follows:

Sales by way of sneaker call, sneaker income use the time shown. Then, the distinction between the fee of the sneakers and the retail charge is analyzed. The retail fee here is the charge list for every brand of sneaker. Then, the author created a view

of the average income rate through the years [8].

Sales and dates are as follows: the sooner the time, the decrease the profits of footwear.

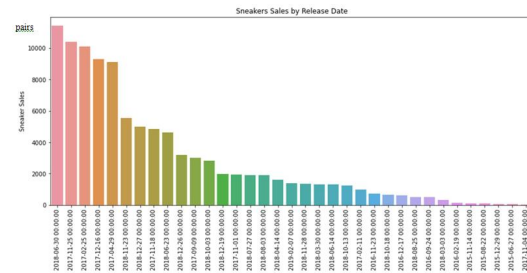


Figure1. Sneakers sales by release date

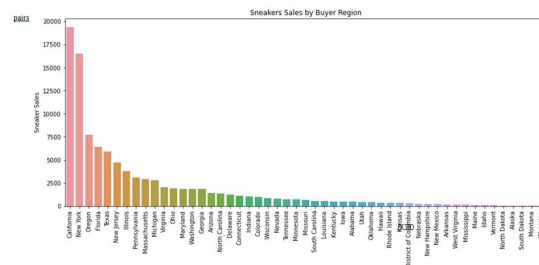


Figure2. Sneakers sales by buyer region

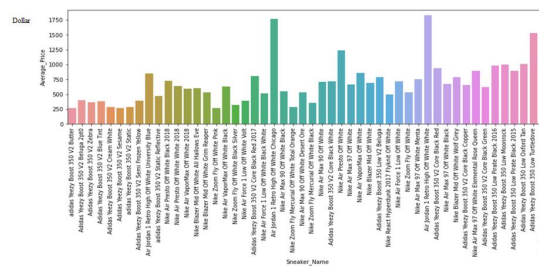


Figure3. Average price for each sneaker name

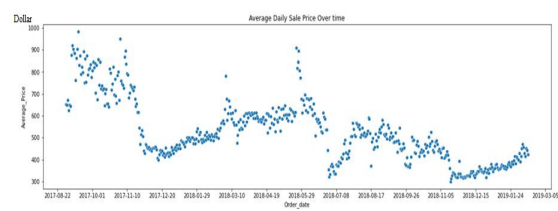


Figure4. Average daily sale price overtime

**Modelling**

Behaviour may be predicted just like the selling rate. First, the writer divided the training procedure and the trying out procedure. The writer set the dimensions: scale length = 0.2. The schooling method

has eighty% of the impact of the regression model. The end result is 1274056226783.0996. Fourth, the coefficient ratings of every variable must be mentioned.

	Coefficient
Sneaker_Name_Nike Blazer Mid Off White All Hallows Eve	1.054632e+09
Sneaker_Name_Nike Blazer Mid Off White Grim Reaper	1.054632e+09
Sneaker_Name_Nike Blazer Mid Off White Wolf Grey	9.870218e+08
Sneaker_Name_Nike Air Max 90 Off White Black	7.333264e+08
Sneaker_Name_Nike Air Max 90 Off White Desert Ore	7.333264e+08
...	...
Sneaker_Name_Adidas Yeezy Boost 350 Low Oxford Tan	-1.587586e+09
Sneaker_Name_Adidas Yeezy Boost 350 Low Moonrock	-1.665599e+09
Sneaker_Name_Nike Air VaporMax Off White	-1.785643e+09
Sneaker_Name_Adidas Yeezy Boost 350 Low Pirate Black 2015	-1.811222e+09
Sneaker_Name_Adidas Yeezy Boost 350 Low Turtledove	-1.908304e+09

107 rows x 1 columns

Figure5. Coefficient scores of each variable

The next is to keep predictions and run assessment metrics. To see the stop result of the model, the author uses MAE to test. The Mean absolute loss (MAE) inside the regression version is the common blunders. MAE is the sum of absolutely the values of the distinction most of the intention price and the expected rate. It first-rate can measure the common modulus of the prediction errors without considering the direction, and the price variety is zero to extraordinary infinity.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{9}$$

MAE value for these models: 60.3380912301444.

### III RESULT EVALUATION

The writer took the take a look at facts at the version [11]. The writer plotted the relationship among  $y_{train}$  and  $y_{train\_price}$ , and used the seaborn's feature despot () to attract the photograph. Infigure6, it's far may be seen that most of the sorting out samples of  $y$  just equals zero.01.

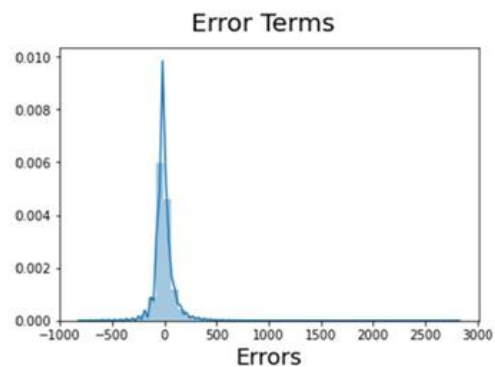


Figure6. Error terms

The author plotted  $y_{test}$  and  $y_{pred}$  to understand the spread. X axis is  $y_{test}$ , and the y axis is the  $y_{pred}$ .

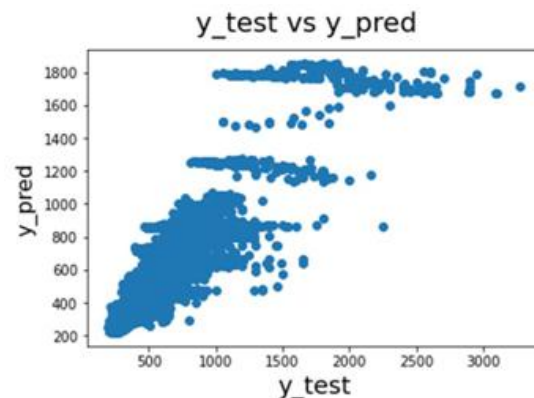


Figure7.  $y_{test}$ - $y_{pred}$

### IV. DISCUSSION

Linear regression is broadly utilized in debt and income forecasting fashions as it

is simple to recognize and provide a reason for. However, this approach can be too hard when coping with over fitting, outliers, and multicollinearity, no longer to say that it cannot capture statistical parameters. After the outcomes of information analysis and device getting to know, the accuracy of the OLS model has reached an awesome stage after changing the strolling conditions and considering the splendid goals. Of route, for the reason that test and the separate education facts for each run are proprietary, there is probably a difference in accuracy. Since there are certainly several dimensions, handiest a linear regression version is used to estimate the rate of shoes. After drawing the temperature graph, its miles clean that OLS is the first-rate suit for this facts set. It isn't constantly authentic to anticipate specific enterprise and style of sneaker business as a reading generation device from a macro view. Because there are many factors affecting the fashion of shoe enterprise price. For example, because of the retirement of a few basketball stars, the charge of the corresponding signature shoes will boom. Fashion tendencies are pretty unpredictable. This study only analyzes and predicts the facts that use the product from the attitude of the contemporary day generation and device

studying. Therefore, the precision will inevitably be insufficient.

## V CONCLUSION

From discovering and analyzing the dataset of recorded assault and machine observe model to amendment records, this article realizes the sneaker fee prediction that's furnished the usage of the complete of statistics. Shoe charge forecasting can be very powerful in some shoe profits packages. They can supply customers a prediction of the rate of a few shoes without the impact of an emergency. Overall, that is also an exquisite aspect for humans buying shoes. Predicting the rate of shoes can also help them purchase multiple footwear with protection rate. For the prediction version set up from the records set, even though linear regression is a easy and easy technique to recognize and put into effect, this kind may be sensitive even as studying with adjustment, outliers and multicollinearity. Therefore, the usage of specific techniques consisting of the random forest version may be greater powerful. However, it's far regrettable that the current set is simply too lengthy to set up a random wooded vicinity estimation model. There also are some gaps in this research. For instance, the dataset has fewer dimensions. For top

greenback, there is quality sale price and quality sale rate. If the rate behaviour of the dataset can be multiplied to multiple dimensions, which include recording the amazing charges of footwear over time, the rate one month after the sale, the fee during the last three days, and so forth. . With the boom within the size of the dataset, the tougher the model may be however the better the prediction accuracy can be.

## REFERENCES

1. Leigh Steinberg 2018, The Profitable Hidden Sneaker Market, <https://www.forbes.com/sites/leighsteinberg/2018/09/17/the-profitable-hidden-sneaker-market/?sh=7a39e7aa5925>
2. Altman, N., Krzywinski, M. Simple linear regression. *Nat Methods* 12,999–1000 (2015). <https://doi.org/10.1038/nmeth.3627>
3. Cohen, J., Cohen P., West, S.G., & Aiken, L.S. Applied multiple regression/correlation analysis for the behavioural sciences. Hillsdale, NJ: Lawrence Erlbaum Associates. 2003.
4. Cohen, J., Cohen P., West, S.G., & Aiken, L.S. Applied multiple regression/correlation analysis for the behavioural sciences. Hillsdale, NJ: Lawrence Erlbaum Associates. 2003.
5. Prasadi Peddi and Dr. Akash Saxena (2014), "EXPLORING THE IMPACT OF DATA MINING AND MACHINE LEARNING ON STUDENT PERFORMANCE", *International Journal of Emerging Technologies and Innovative Research* ([www.jetir.org](http://www.jetir.org)), ISSN:2349-5162, Vol.1, Issue 6, page no.314-318, November-2014, Available: <http://www.jetir.org/papers/JETIR1701B47.pdf>
6. Prasadu Peddi (2017) "Design of Simulators for Job Group Resource Allocation Scheduling In Grid and Cloud Computing Environments", ISSN: 2319- 8753 volume 6 issue 8 pp: 17805-17811.
7. N. Jain and P. Peddi, "Gender Classification Model based on the Resnet 152 Architecture," 2023 IEEE International Carnahan Conference on Security Technology (ICCST), Pune, India, 2023, pp. 1-7, doi: 10.1109/ICCST59048.2023.10474266.
8. G. J. Jagtap, P. Peddi and Y. D. Sinkar, "Application of YoloV5 for Brain Tumor Detection from MRI," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024,

pp. 1-5, doi:  
10.1109/ESCI59607.2024.10497232.