International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Innovations in the Sneakers Industry and Leveraging Machine Learning Models for Predictions

Tanmay Gautam

Mayoor School, Noida, India
Email: *tanmay.gautam456838[at]gmail.com*

Abstract: ***The shoe industry has been a vital part of global fashion, leading to a diverse range of footwear styles tailored to meet the ever - evolving needs and preferences of consumers worldwide. With the rapid integration of cutting - edge technology and AI, the shoe industry has witnessed groundbreaking innovations in materials, manufacturing and personalized design solutions, increasing comfort, performance and much more. This paper highlights the vast world of artificial intelligence in the footwear industry. Machine learning algorithms are used in this paper to predict the sale price of sneakers from a publicly available database. Among the algorithms considered, the XGBoost Regressor algorithm was found to be the most effective producing the maximum accuracy. This paper has delved into the fascinating world of algorithms and testing its effectiveness and reviewed their weaknesses and strengths. As technology advances, we can expect further innovations in algorithm development leading to more accurate and versatile models pertaining to the footwear industry***

Keywords: Machine learning, Sneakers, Futuristic shoes, ML algorithms

1. Introduction

From being on our phones to environmental monitoring, computer science has become the backbone of the human race. The world broadcasts the new technological developments in various fields like artificial intelligence, automobiles, electronics, textile and many more but something of daily utilization like shoes should not be written off. Sneakers touch the boundaries of fashion and technology, revolutionizing every activity of ours, be it run or walk with its advanced designs and materials. We have reached a stage where there are shoes with inbuilt sensors in the shoe to adapt to the size of the person's foot, these futuristic shoes are a testament to the limitless routes of today's technology. With their stylish looks and flush mechanisms, the footwear industry is becoming more and more intriguing with futuristic innovations to give an unparalleled experience.

This paper holds descriptive details about various kinds of futuristic shoes that involve cutting edge technology which have endless potential in the shoe world. The aim of the paper is to discuss the recent innovations in sneakers and use various machine learning models to predict the sale price using a publicly available dataset. This paper combines machine learning with smart fashion exploring different types of shoes and various machine learning models. The mechanism of the algorithms are also briefly discussed.

Innovations in the footwear industry

The Nike go fly ease [1], shown in Figure 1 (a), has an individualistic mechanism and design, it works on the principle of a hinge which allows us to wear and take off the shoe without the use of our hands. The shoe has a simple but effective design and is lace - less which proves to be helpful for the disabled. The concept of no laces can be seen in a lot of futuristic shoes after all the future is all about less effort. Li Ning Yun you [2] following the concept of no laces, is a sleeping bag like shoe, as it can be seen in the Figure 1 (c), which has a very soft sole and it incorporates multiple thickened quilted structures into the shoes which expresses the warmth and comfort of winter.

Figure 1: (a) Nike go fly ease [1], (b) Clogs ral7000 X Bershka [2], (c) Li Ning Yun you, (d) Yeezy foam runners

Adidas Yeezy foam runners [3] has a crazy, funky and unique design which can be seen in Figure 1 (d). Apart from that the shoe's manufacturing uses algae and less toxic fuels and more sustainable methods. The foam is responsible for the comfort within the shoe and the crazy design makes the shoe long lasting and easy to clean. Another shoe that uses environment friendly means for manufacturing are Clogs from ral7000 Studio in collaboration with [4] Bershka, the futuristic shoe in Figure 1 (b) uses an XL Extralight technology which makes the shoe's material 3 times lighter than other materials. This technology is composed of 51% pre - consumer industrial waste. This makes each single pair unique due to the uneven surface and the random distribution of the recycled parts across the product.

Nike has been behind the idea of a self - lacing shoe for a long time and finally made one known as the Nike air mags [5]. This shoe, shown in Figure 2 (a), has a technology which allows the shoe to adjust the laces according to the foot of the person. There are buttons present on the shoe for manual control on the laces which can also be done through an app. This idea can also be seen in other Nike shoes like Nike jordan 11, Nike adapt B 2.0. The brand Shooz [6] adds a zip to every shoe so that one can customize the shoe by just zipping a top with the sole, a new top gives you a new shoe [as shown in Figure 2 (b)].

There are shoes which are specifically designed for certain activities like [7] Roame zero is very popular within the motorbike community. It has certain materials which easily absorb various shocks without any damage to the human foot. The shoe is built with shock resistant leather, nylon and double stitching which can be seen in Figure 2 (c). Following the idea of advanced sole and shock resistance in shoes, [8] Balmain unicorn is a sneaker with different materials used in the upper portion of the shoe and the sole is crafted from an 8 sectional mold which is responsible for the shock absorbing cushion system of the sole [as shown in the Figure 2 (d)].

The scientists at the University of Wisconsin have created an energy harvesting technology using human energy to charge electronic devices. This technology is based on the concept of kinetic energy which is then converted to electrical energy. It can be inserted into the footwear and can store power inside a battery generated by walking [as shown in the Figure 2 (e)]. They also use a method known as the bubbler method which allows the shoe to harness maximum energy from a small area, it can create upto 10000 watts of power.

Another completely out of the box sneaker is the Moonwalker, this sneaker can provide one with a similar experience of walking on the moon. It has 2 layers of 45 magnets as seen in the Figure 2 (f) which creates the anti gravity force and decreases the effect of gravity.

Figure 2: (a) Nike air mags, (b) Shooz, © Roame zero, (d) Balmain unicorn, (e) Charging Sneaker, (f) Moonwalker

Introduction to Machine Learning

Similar to human evolution, Computers can also learn and get better at tasks through experience. Like humans, machines can also learn and become more reliable. Machine learning brings together statistics and computing science to allow computers to do a given task without a program; it uses statistics patterns within the data and tries to provide us with an accurate output. By feeding data to the machine, it will catch distinctive details within the data to categorize or make predictions. Just as humans get wise with experience, the algorithm in the computer gets accurate progressively with the amount of data it receives. There are various places these algorithms are used like facial recognition, text with speech recognition, spam filters on your mailbox, credit card fraud detection and so much more. If a machine is run through regularly with a satisfactory amount of data with combined statistics and computer science, a machine can solve complex problems more efficiently with less computing power. An example for a very simple algorithm is the K - Nearest neighbor algorithm: in this algorithm the output relies on the majority in this case to give us an output for a complex set of data.

Few of popular Machine Learning Algorithms are as follows:

- **Linear Regression**

Linear Regression is a widely understood algorithm used in statistics and machine learning. This model attempts to explain the relationship between any two variables using a best fitting straight line termed as the regression line. The independent variable predicts the value of the other dependent variable. To calculate this prediction there are various strategies, the most popular strategy is least square regression. This strategy involves a line drawn between the 2 variables of the data set and the measurement of the vertical distance from the data point to the line should be the least.

The smaller the distance the better the regression line, the better the prediction. The regression line allows us to read the relationship and make estimations between two linearly correlated variables.

- **Decision trees**

The clutch point about this algorithm is that it can be applied when two classes or two variables are not linearly separable. The decision tree recursively splits the data set until we are left with the data of only one type of class; this pure data is termed as leaf nodes this also helps us decide the class of a new data point. There is another type of node called the decision node. These contain a condition to split the data. The first step or the first classifying factor is called the root node and based on its condition relies on the outcome of the whole tree. There shall be no further new conditions applied to the data if the resultant data is a pure node. This may, however, turn out faulty for complex data therefore for simplicity we take the case of 2 classes. When we don't have a 100 percent pure leaf node a majority voting is done for the classification. The model needs to learn which features/conditions to take and the corresponding value to optimally split the data.

- **XBGRegressor**

XGBoost constructs a sequence of trees in an ensemble, where each successive tree is trained to rectify the errors made by its predecessors. It commences with a basic decision tree that forecasts the target variable based on a single feature. The discrepancies between these forecasts and the actual target values, known as residuals, are then computed. A new decision tree is constructed by XGBoost to predict these residuals. The objective of this tree is to minimize the residuals, effectively learning to anticipate the remaining errors from the preceding trees. The newly created tree is incorporated into the existing ensemble of trees, and their predictions are combined. This amalgamated prediction represents an updated estimation of the target variable. This iterative process continues, generating additional trees and fine - tuning the predictions. Each newly added tree focuses on rectifying the errors made by the previous ones. The iteration persists until a predetermined number of trees is achieved or until no further enhancement is observed. The final prediction is attained by aggregating the forecasts from all the trees in the ensemble.

2. Method

The dataset was taken from Kaggle [9] with the name StockX 2019 Data Contest. The dataset had sufficient data with 99956 rows and 8 columns. After importing it in the code, the classification of values was done to distinguish between data types. The columns considered for the dataset were:

1) Brand: Refers to whether the shoe belonged to Yeezy or off white
2) Sneaker Name: Refers to name of the sneaker
3) Buyer Region: Refers to the location/area from where the sneaker is bought
4) Retail Price: Refers to the price of the shoe it's being sold in retail.
5) Sale Price: Refers to the original price of the shoe.
6) Buyer Region: Referral to region of the purchase of the sneaker.

An issue was noticed that due to the dollar sign in Sale and Retail Price the system identified it as a categorical value. The dollar sign was then removed and the price was made an integer data type. We also dropped the order and release data as it didn't have much relevance with our output (as seen in figure 3). Then we proceeded with statistical analysis, creating count plots for categorical values and distribution plots for numerical values. For the computer to make predictions we needed to encode the data and transform categorical values to numerical values. Then we split the data into features and targets which are further split into training and testing dataset so that we can test and train the model with different algorithms and accuracy for each model was found.

Figure 3: Flowchart showing the process followed using the dataset

3. Results and Discussion

The exploratory data analysis revealed that the obvious preference of purchasing Yeezy shoes; there is approximately a difference of 40000 units. This might be because of the comfort the foam the shoe provides. From Figure 4, it can be observed that the prices are concentrated between approximately 200 to 1000 with more than 50% being less than 500. Majority of the people have a shoe size around 10 and the most of the shoe sizes are around 8 - 12 (as shown in Figure 5). People even buy shoes in different sizes as an investment as the value of these shoes appreciates. For example the shoe Red October Nike air Yeezy 2 was around $5000 in 2019, in 2022 its value skyrocketed to $22000.

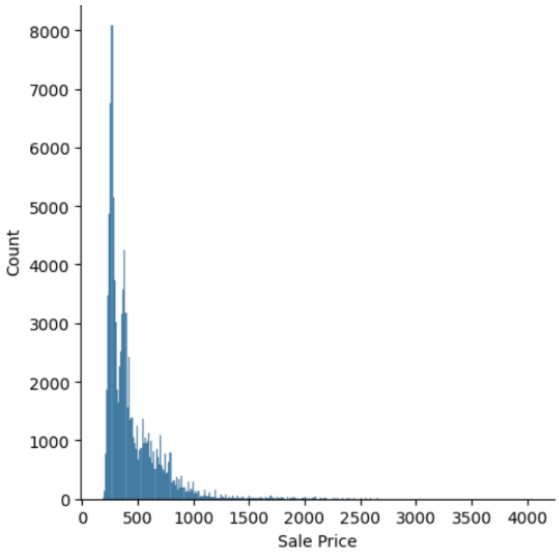

Figure 4: Sale price distribution in dataset

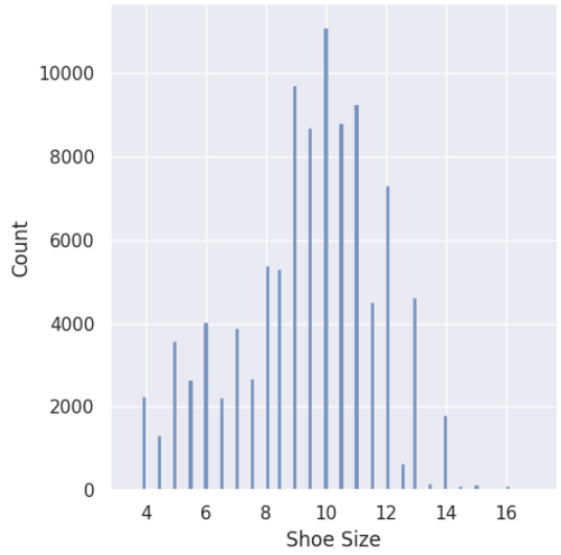

Figure 5: Shoe size distribution plot

The accuracies of the three machine learning models are presented in Table 1. The reason XGBoost Regressor has the maximum accuracy in testing data prediction is because it focuses on minimising the errors as it progresses unlike decision trees.

Table 1: Accuracies of various machine learning algorithms

Algorithms	Accuracy
Linear Regression	0.8269618975322777
Decision Tree	0.7381208851597119
XGBoost Regressor	0.8422524223394624

4. Conclusion

This paper covers the latest technology involved in the innovation and manufacturing of shoes. Shoes form an important part of the modern world's lifestyle and the growing field of artificial intelligence and machine learning can be a part of the development of the footwear Industry. This paper uses three machine learning models to predict the sale price of sneakers using a publicly available dataset. Among the three models, the XGBoost Regressor algorithm had the maximum accuracy of 84.22% and can be used as an effective model for sale price predictions with similar datasets.

References

[1] *Nike's first hands - free shoe: Go FlyeaseNike. com*. Available at: https: //www.nike. com/in/flyease/go - flyease (Accessed: 06 August 2023).
[2] *Nike's first hands - free shoe: Go FlyeaseNike. com*. Available at: https: //www.nike. com/in/flyease/go - flyease (Accessed: 06 August 2023).
[3] *Adidas official website India: Sportswear* (no date) *adidas India*. Available at: https: //www.adidas. co. in/ (Accessed: 06 August 2023).
[4] *Online fashion: Official storeBERSHKA*. Available at: https: //www.bershka. com/ (Accessed: 06 August 2023).
[5] *Sneakers, streetwear, trading cards, handbags, watchesStockX*. Available at: https: //stockx. com/ (Accessed: 06 August 2023).
[6] Milling, M. (2015) *Shooz: So easy to be stylish, The Manual*. Available at: https: //www.themanual. com/fashion/shooz/ (Accessed: 06 August 2023).
[7] *Roame* (no date) *ROAME*. Available at: https: //www.roame. com/ (Accessed: 06 August 2023).
[8] *BalmainBalmain Official Website India | Designer clothing & bags*. Available at: https: //in. balmain. com/en (Accessed: 06 August 2023).
[9] Stuck, H. (2020) *Stockx Sneaker Data contest, Kaggle*. Available at: https: //www.kaggle. com/datasets/hudsonstuck/stockx - data - contest (Accessed: 06 August 2023).