# Predictive Sales Analytics: A Machine Learning Revolution

Romika Yadav[1], Mamta Bhardwaj[2], Puja Rani[3]

[1]Assistant Professor, National Institute of Fashion Technology, Kangra

[2]Assistant Professor, Government College for Girls, Nacholi Faridabad

[3]Assistant Professor, National Institute of Fashion Technology, Patna

**Abstract -** In a modern world where nearly all manual tasks are converted into automated, the definition of manual is changing now. Machine and Deep Learning algorithms can help computers to play chess, perform surgeries, getting smarter gadgets.We are living in a modern era of constant technological advancements, and looking at how computing has progressed over the years, we can predict what's to come in the days ahead in the future. The objective of this case study is to analyse the relationship between minimum price, maximum price and the price on sale of women's shoes of different brands depending upon the dataset provided.

Keywords: Machine learning, Data, Analytics, Multiple regression, Personal computer

## I. Introduction

Machine learning is the science of getting computers to act upon without being explicitly programmed for automation. In the past decade, machine learning has given us self-driving cars, voice recognition, speech recognition, effective web search, and a vastly improved understanding of the human genome of the understanding. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many researchers also think that it is the best way to make progress towards human-level artificial intelligence and automation. **Techniques of machine learning are discussed in the coming sections of the machine and deep learning approach.** We are probably living in the most defining period in technological advancements.

The period when computing moved from large mainframes to personal computers to self-driving cars, gadgets and robots. But what makes it defining is not what has happened in the data driven approach of the methods, but what has gone into getting here. What makes this period exciting is the democratisation of the resources and techniques of data and machine learning. Data crunching which once took days, today takes mere minutes, all thanks to Machine and deep Learning Algorithms. So, Machine and deep Learning Algorithms can be categorised as supervised and unsupervised learning.

## II. Literature Review

Gowri et al. states about the importance of machine learning in the modern era of technological advancements. In the application of the various fields like 3D printing, speech recognition, augmented reality, virtual reality and various technological advancements [1]. Alzubi et al. stating about machine learning adapting human behaviour in the present context. Also presenting the various applications of machine learning that are based on various supervised and unsupervised algorithms [2]. Ayushi and preeting talking about the machine and deep learning that is going to be used in various applications of the present era. Deep learning fetching the data directly from the raw data then features extraction finally

classification for the desired output [3]. Issam et al, states that comparison of various algorithms to predict the wall of robotics. Providing the various algorithms to provide the solution of the problem which is defined in the paper [4].

## III. Problem Objective

The objective of this case study is to analyse the relationship between minimum price, maximum price and the price on sale of women's shoes of different brands depending upon the dataset provided**.**

## IV. Research Methodology

Multiple regression is an extension of linear regression models that allow predictions of systems with multiple independent variables. Multiple regression is specifically designed to create regressions on models with a single dependent variable and multiple independent variables.

$$y = B\_1 * x\_1 + B\_2 * x\_2 + … + B\_n * x\_n + A$$

**DATASET –** Women's shoes dataset contains maximum, minimum and sale prices of women's shoes of different brands of different specifications.

**Attribute Information :**

**BRAND –** The brand of the shoes.

**CATEGORIES –** Detail of the shoes.

**COLOURS –** Different colour of the shoes.

**NAME –** Style name given to the shoes by brand.

**PRICES AMOUNT MINIMUM –** Minimum Amount set by the brand.

**PRICES AMOUNT MAXIMUM –** Maximum Amount set by the brand.

**PRICES IN SALE –** price of the shoes when in sale.

**Data Preparation and building the dataset :**

To perform a cluster analysis in R, generally, the data should be prepared as follows:

1. Rows are observations (individuals) and columns are variables

2. Any missing value in the data must be removed or estimated.

3. The data must be standardised (i.e., scaled) to make variables comparable.

**Computing multiple regression in R :**

Creating a subset of response variable prices. inSale and predictor variables amountMin and prices.amountMax to analyse and establish relationships.

**INPUT**

```
model<-lm(prices.isSale~prices.amountMin+prices.amountMax, data = newdata)
summary(model)
```

**OUTPUT**

```
> model<-lm(prices.isSale~prices.amountMin+prices.amountMax, data = newdata)
> summary(model)

Call:
lm(formula = prices.isSale ~ prices.amountMin + prices.amountMax,
    data = newdata)

Residuals:
    Min      1Q  Median      3Q     Max
-4.0152 -2.0776  0.4627  1.6184  5.2964

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       74.2882     0.5180 143.407   <2e-16 ***
prices.amountMin  -0.4600     0.2068  -2.224   0.0311 *
prices.amountMax   0.4704     0.2052   2.293   0.0265 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.562 on 46 degrees of freedom
Multiple R-squared:  0.1831,    Adjusted R-squared:  0.1476
F-statistic: 5.155 on 2 and 46 DF,  p-value: 0.009549
```

The P value for the model is 0.009549 which suggests that the null hypothesis can be rejected and the model is built good.

**PREDICTING THE VALUES:**

**INPUT**

```
a<-data.frame(prices.amountMin=70,prices.amountMax=8)
a
result<-predict(model,a)
result
print(result)
```

**OUTPUT**

```
> a<-data.frame(prices.amountMin=70,prices.amountMax=8)
> a
  prices.amountMin prices.amountMax
1               70                8
> result<-predict(model,a)
> result
       1
45.85126
```

Based on the Established relationship, value for response variable prices.inSale is calculated as 45.85126 when prices.amountMin = 70 and prices.amountMax = 8

## V. Results and discussion

The objective of this case study was to analyse the relationship between minimum price, maximum price and the price on sale of women's shoes of different brands depending upon the dataset provided**.** The algorithm used is multiple regression.The P value for the model came out to be 0.009549 which suggests that null hypothesis can be rejected and the model is built good. Based on the Established relationship, value for response variable prices.inSale is calculated as 45.85126 when prices.amountMin = 70 and prices.amountMax = 8.

## VI Conclusion

This paper surveys various machine and deep learning algorithms and also application of the fashion industry. Today each and every person is using machine and deep learning approaches knowingly or unknowingly of data prediction. This paper gives an introduction to most of the popular machine learning algorithms on the fashion industry application.

## References

1. Gowri geetha et al, Machine Learning, international journal of research and analytics reviews, pp.197-208, 2019.
2. Alzubi et al, Machine Learning from Theory to Algorithms: An Overview, Second National Conference on Computational Intelligence, 2018.
3. Ayushi and Preeti, Machine learning and Deep learning, International Journal of Innovative Technology and Exploring Engineering, 2019.
4. Issam et al, A comparative study on machine learning algorithms for the control of a wall following Robot, IEEE, 2019.