Report

On

PRODUCT RECOMMENDATION SYSTEM

FOR LOCAL FASHION PRODUCT VENDORS &

SMALL SCALE E-COMMERCE WEBSITES

By TEAM DHANDHUKIYA

Date: 08-08-2022

Members:
- DHRUV KUMAR DHANDHUKIYA
- ANISH DESAI
- ANSH ASHOK SHRIWAS
- ADNAN HABIB
- ANIKET SINGH

Table of Contents

1. ABSTRACT

In this report, we have proposed the prototype of using Apriori algorithm, which suggests the top products based on customer reviews (from e-commerce websites), helping them to restock particular products which are in demand and trend in that particular area.

Product is the number one thing any business strives on, especially start-ups related to fashion. There are more than 17,000 fashion companies in India and more than 3,00,000 in the world.

Using this project, start-ups and many more small-scale fashion industries, will be helped to choose their products over millions of choices and compete globally with various e-commerce giants.

Market Basket Analysis has extensively been used by e-commerce companies to boost up their sales. The proposed service will analyse the customer purchase behaviour for fashion and identify the top products.

Association rule mining is a major factor and a method of data mining. Apriori algorithm is a tool in association rule mining used to implement it. Data mining helps using the huge amount of data in an efficient way, extract the required data and predict the future based on customer behaviour.

2. PROBLEM STATEMENT

The problem statement is, many shops/local vendors struggle to compete with the tech giants. Bewildered by tons of options, especially in fashion. My project will help them in suggesting and analysing the product, based on demand and reliability (in their particular locality). It goes through extensive and exhaustive dataset containing product reviews, the issues, frequently asked questions, etc. Then it is built using an ML algorithm called Apriori algorithm.

Size Of Southeast Asia's E-Commerce Market
In Billions, USD

Year	Value
2015E	$6
2016E	$7
2017E	$10
2018E	$13
2019E	$17
2020E	$22
2021E	$29
2022E	$38
2023E	$51
2024E	$67
2025E	$88

We see the rapid rise of e-shopping over the years, with increasing sales of e-commerce websites, the sales of local shops and vendors are likely to go down. Hence, we see a vital need for them to compete in this changing world.

3. MARKET/CUSTOMER/BUSINESS NEED ASSESSMENT

Product is the most imperative subject in any business. With this project, businesses in fashion will get an overview of which products and brands are in demand, along with their quality check.

The customer buying preferences have significantly changed due to the pandemic, with these changing trends and a constant generation gap, it is difficult for any business to keep up with the demands, more so in fashion. Therefore, by using this technique, we aim to provide small businesses with useful insights from the available data and ways to generate more revenue.

The fashion choices vary geographically and with my model, the product recommendations are to be made for that specific place.

If this system is successfully implemented at local shops and small-scale start-ups, the customers will start exploring these small-scale businesses with great selection of products rather than relying on the online giants.

Usually, customers get that feel of comfort and assurance while buying from physical stores. However due to a lack of variety and a less attractive product collection, they revert back to online sites. This project is going to eliminate that problem to a great extent.

4. TARGET SPECIFICATION

The proposed system/service will provide the shopkeepers and vendors with techniques so that their sales boost up and they no longer have to go through an economic crisis.

It will suggest them certain top demanding products (on major e-commerce websites) in their locality, based on the analysis performed by the Apriori algorithm. With much better collection of products in their shops, the sales are likely to go up.

Based on online prices in the dataset, and their demand, they can purchase these items' quantities in accordance to it, or can develop products with similar style and give further discounts to compete with the e-commerce giants.

5. EXTERNAL SEARCH

The sources I have used as reference for analysing the need of such a system for local businesses and how E-commerce giants have been using the technique to boost up online sales, have been mentioned below:

- Understanding the problems faced by local fashion products vendors.
- How E-commerce sites benefit from Market Basket Analysis.
- Apriori algorithm explained
- A study on Understanding Changing Trends of Customer Behaviour.
- A study on change of styles with geography.
- Customer preference analysis based on reviews.

5.1. BENCHMARKING

Most of the e-commerce giants like Flipkart, Amazon, Myntra, etc. uses AI based recommendation systems using Apriori Algorithm This identifies purchasing habits of customers and uses this information to cross-sell and up-sell relevant items. But this technique would also be beneficial

when applied to the small businesses since most of the people revert to physical stores every once in a while.

5.2. APPLICABLE PATENTS

▪ Patent 1 – Recommendation system based on the customers behaviour

A recommendation service is a computer-implemented service that recommends items from a database of items. The recommendations are customized to particular users based on information known about the users. One common application for recommendation services involves recommending products to online customers. For example, online merchants commonly provide services for recommending products (books, compact discs, videos, etc.) to customers based on profiles that have been developed for such customers. Recommendation services are also common for recommending Web sites, articles, and other types of informational content to users.

The types of items that can be recommended by the service include, without limitation, books, compact discs ("CDs"), videos, authors, artists, item categories, Web sites, and chat groups. The service may be implemented, for example, as part of a Web site, online services network, e-mail notification service, document filtering system, or other type of computer system that explicitly or implicitly recommends items to users. In a preferred embodiment described herein, the service is used to recommend works such as book titles and music titles to users of an online merchant's Web site.

https://patents.google.com/patent/US7908183B2/en

▪ Patent 2 - Enhanced Market Basket Analysis Model

An object of the present invention is to provide a method and system for data mining in which the premise and/or conclusion of an association rule can include an attribute of an entire basket (e.g., its total dollar value). The items contained in a single aggregate sale (e.g., all of the purchased items in a particular market basket, referred to herein as a "market basket grouping") are characterized according to predetermined attributes. Each attribute is identified and an "imaginary item" is included in the data for each market basket grouping which possesses an identified attribute. When the data is subjected to traditional association analysis, the imaginary items are included in the analysis and may be utilized to identify frequent item sets that are typically found in market basket groupings having the identified characteristics.

Other objects and advantages of the present invention will be set forth in part in the description and the drawings which follow, and, in part, will be obvious from the description or may be learned by practice of the invention.

https://patents.google.com/patent/US6976000

- <u>Patent 3 - Method and apparatus for recommendation engine using pair-wise co-occurrence consistency</u>

The invention, referred to herein as PeaCoCk, uses a unique blend of technologies from statistics, information theory, and graph theory to quantify and discover patterns in relationships between entities, such as products and customers, as evidenced by purchase behaviour. In contrast to traditional purchase-frequency based market basket analysis techniques, such as association rules which mostly generate obvious and spurious associations, PeaCoCk employs information-theoretic notions of consistency and similarity, which allows robust statistical analysis of the true, statistically significant, and logical associations between products. Therefore, PeaCoCk lends itself to reliable, robust predictive analytics based on purchase-behaviour.

https://patents.google.com/patent/US20070094067

There are many more related patents, but these 3 are the most related and important ones.

5.3. APPLICABLE CONSTRAINTS

- Data Collection from shopkeepers and vendors (stock availability and locations)
- Maintaining a dynamic dataset
- Reliability of the model in a subjective field like fashion.
- Convincing the shopkeepers to implement the system in their shops.

Hybrid Recommendation engines are essentially the combination of diverse rating and sorting algorithms. For instance, a hybrid recommendation engine could use collaborative filtering and product-based filtering in tandem to recommend a broader range of products to customers with accurate precision.

5.4. APPLICABLE REGULATIONS

- Data protection and privacy regulations (Customer reviews and asked questions)
- Govt Regulations for small businesses
- Employment Laws
- Antitrust Regulations
- Regulations against false advertising

6. BUSINESS OPPORTUNITY

Since the above technique has only been used by large companies, this can be

extended for small businesses, not only shopkeepers or vendors, but also any other businesses selling physical goods. Therefore, there is a fair chance of this service being a great business opportunity. Every small business that depends on sales can and would want to opt for using this service in order to always know what their customers want. The emergence of every small business is thus a fairly great business opportunity for the service provided by it.

7. CONCEPT GENERATION

Looking at the common figures, how tech giants like Amazon, Flipkart, etc. are making billions. Their success is inversely proportional to the success of the local vendors/small scale online sites. Apriori algorithm helps in suggesting the products by using the user reviews. In fashion, product is what makes a brand successful. Hence, helping them with product selection is definitely going to buff up their sales.

8. CONCEPT DEVELOPMENT/PRODUCT DETAILS

The project will first use the dynamic data from one of the e-commerce giants. The algorithm is going to be developed on this dataset, and hence a set of top products based on user reviews and sales are going to be suggested. Shopkeepers/vendors/small websites will have to share their stock availability, capital and their locality. Based on this data, a set of products with particular quantities specific to their locality is going to be suggested by the model.

This project can be easily implemented by an individual or a group of 2. However, maintaining the dynamic dataset is the tedious part, which requires a team. The cost of this project is the cost for maintaining the dataset. This doesn't require any additional software, just a python compiler.

9. FINAL PRODUCT PROTOTYPE

GitHub Link: Adnan232/Fashion-Market-Recommendation-System (github.com)

The final service product is going to suggest the vendors or suppliers which products are to be ordered according to the demand and trend.

There are three major components of the Apriori algorithm:

1. Support: Support is basically telling how popular the item is. It is calculated by the no. of order or transactions having the product B out of total no of order.

```
Support(B) = (Transactions containing (B))/(Total Transactions)
```

2. Confidence: Confidence refers to the likelihood that an item B is also bought if item A is bought. It can be calculated by finding the number of transactions where A and B are bought together, divided by total number of transactions where A is bought. Mathematically, it can be represented as:

```
Confidence(A,B) = (Transactions containing both (A and
B))/(Transactions containing A)
```

3. Lift (A, B): It refers to the increase in the ratio of sale of B when A is sold. Lift (A, B) can be calculated by dividing Confidence (A, B) divided by Support(B). Mathematically it can be represented as:

```
Lift(A→B) = (Confidence (A,B))/(Support (B))
```

We can draw some conclusion by seeing lift value. Lift value 1 tells that there is no association between the products. Lift value greater than 1 means the products are likely to bought together. Let's say lift value is coming as 5 for two product A and B. This means product A and B are 5

times more likely to bought together than just buying product B alone. A lift value less than 1 suggest that products are unlikely to bought together.
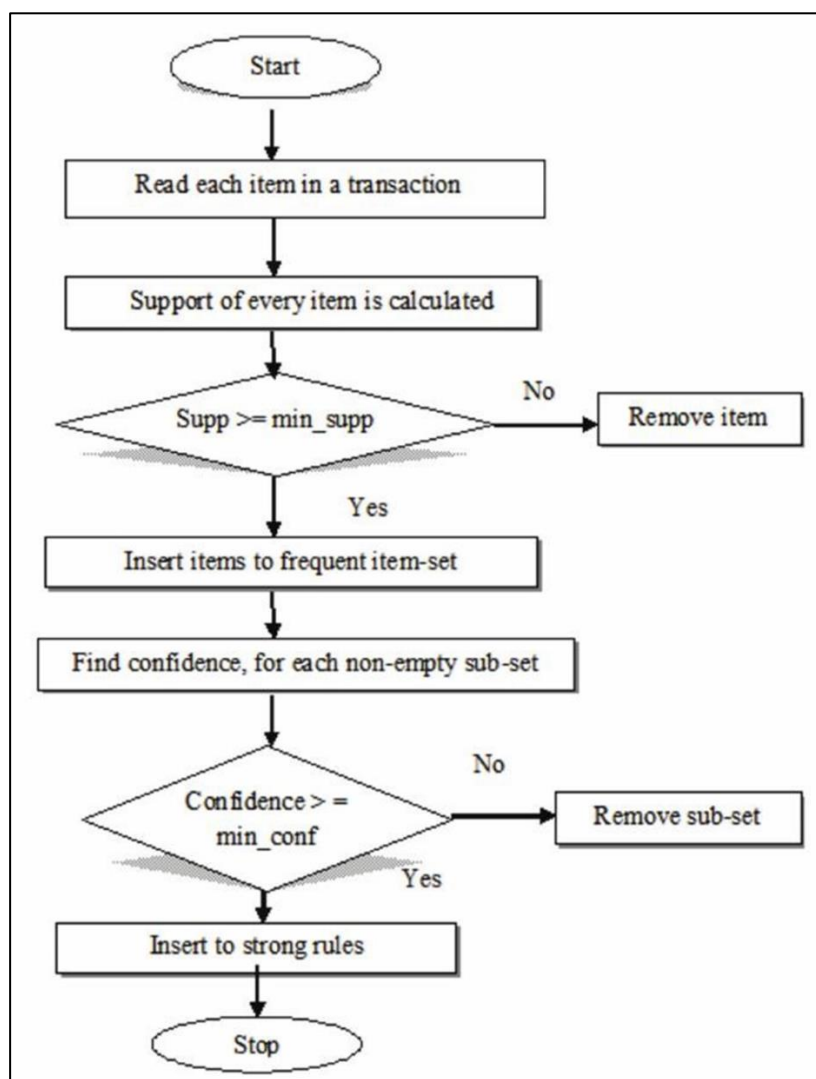
Step-1: Determine the support of item sets in the transactional database, and select the minimum support and confidence.

Step-2: Take all supports in the transaction with higher support value than the minimum or selected support value.

Step-3: Find all the rules of these subsets that have higher confidence value than the threshold or minimum confidence.

Step-4: Sort the rules as the decreasing order of lift.

Generating frequent itemset is the computationally expensive step, which can be improved using hash-based itemset counting, transaction reduction, partitioning, sampling and dynamic itemset counting.

DATASET:

- Dataset of 22,000 fashion products on Amazon
- This is a pre-crawled dataset, taken as subset of a bigger dataset (more than 7 million fashion products) that was created by extracting data from Amazon.
- This dataset was created by Prompt Cloud's in-house web-crawling service.

Product vs Count of Purchase
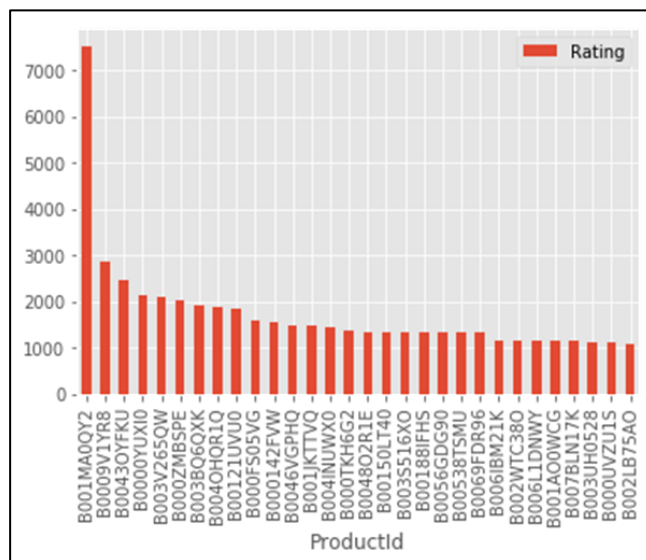
```
# Fitting K-Means to the dataset

X=X1

kmeans = KMeans(n_clusters = 10, init = 'k-means++')
y_kmeans = kmeans.fit_predict(X)
plt.plot(y_kmeans, ".")
plt.show()
```
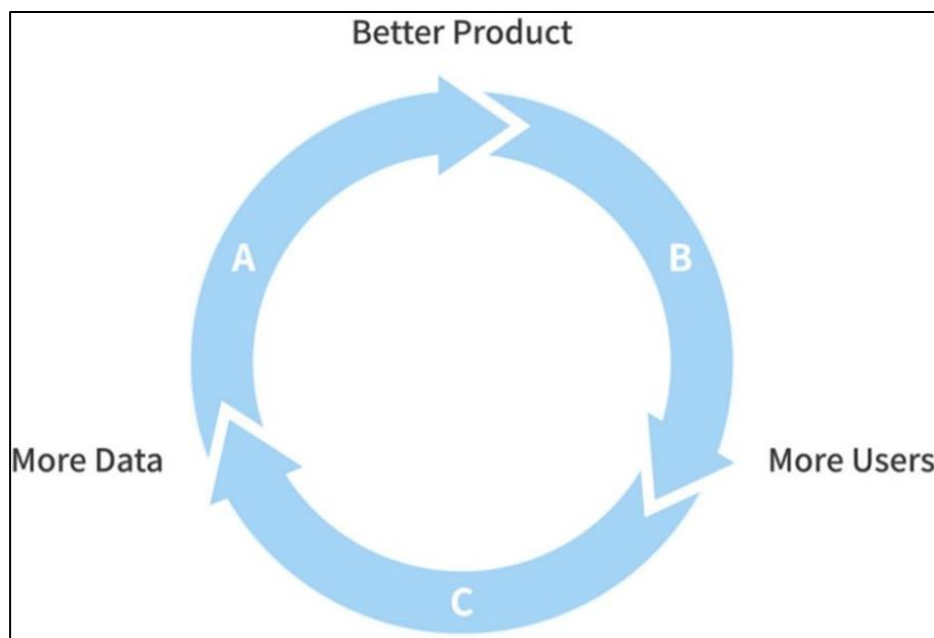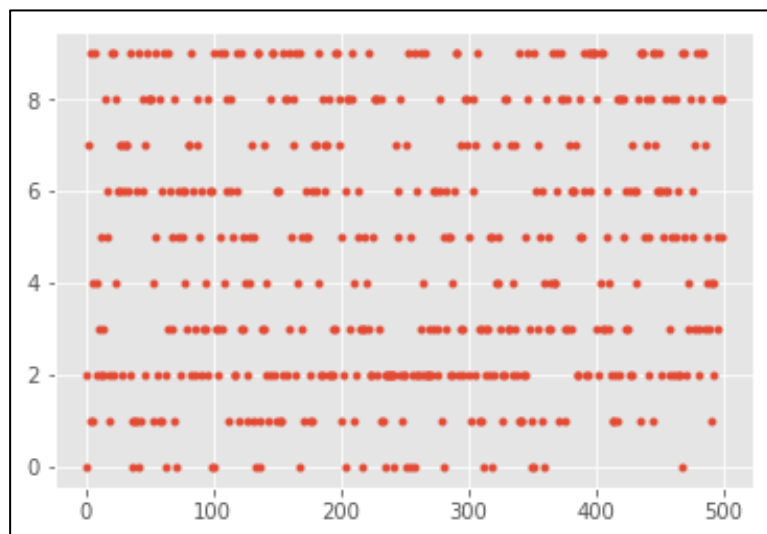

Figure 1: Overview of use of a recommendation system

The figure 2 below gives an overview of the architecture of the recommender system.

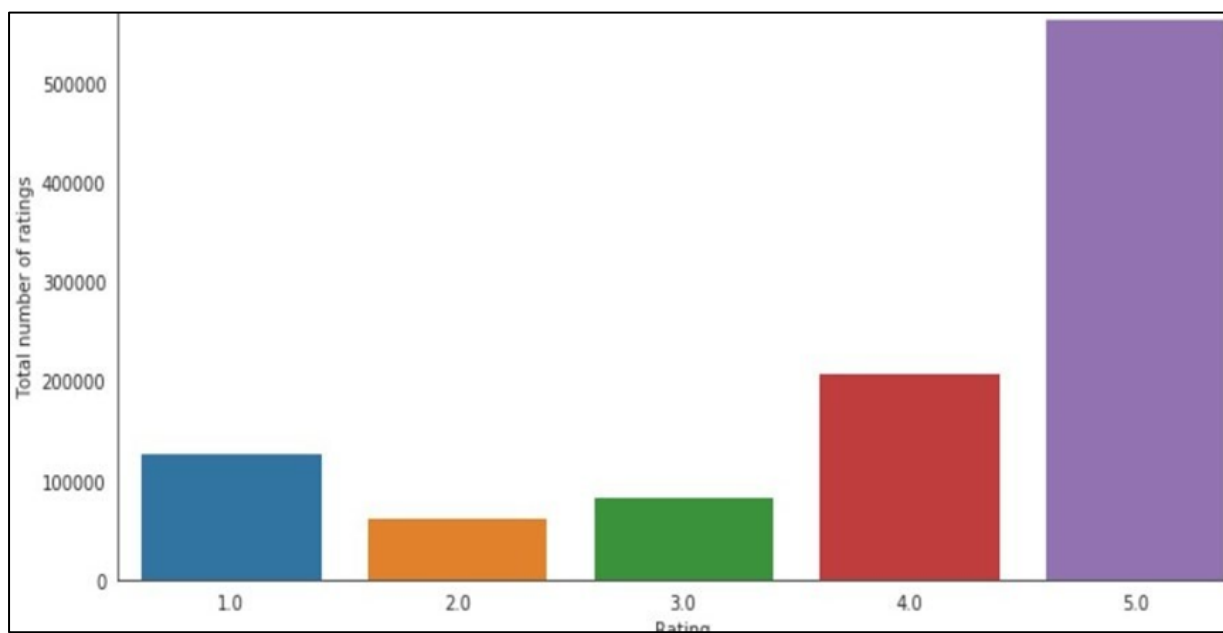Figure 2: Overall architecture of recommender system
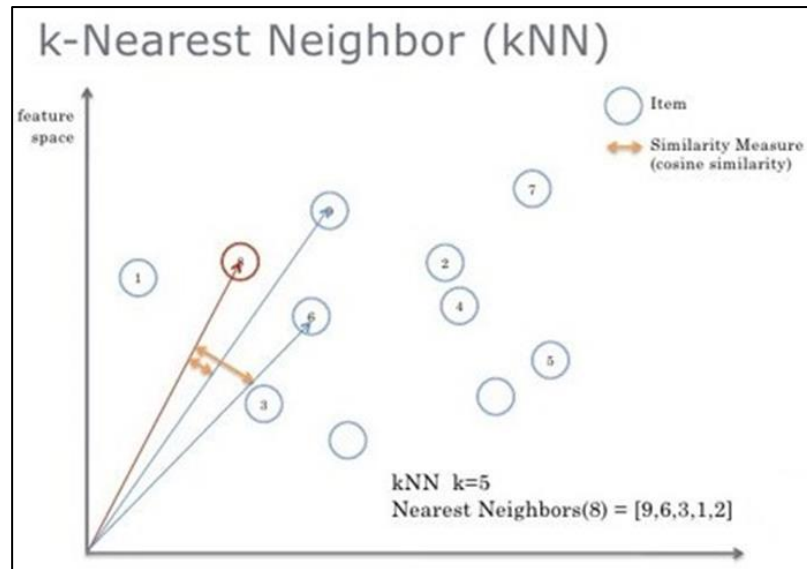

Figure 3: Distribution of Ratings

Figure 4: KNN Algorithm functioning

Once all this process is completed, we train our model on the dataset. The model will now be able to predict different products for the users. In the table 5 we can see the results which contains the Product id's that are finally recommended to the user by the Recommender System.

'1400501520'	'B0000010M5'
'1400532736'	'B00000J0D2'
'1400599997'	'B00000J0D8'
'1400698987'	'B00000J3UJ''
'6301977173'	'B00000J4EY'
'9573212919'	'B00000J6WY'
'9575871979'	'B00000JBAT'
'9625993428'	'B00000JBHP'
'9888002198'	'B00000JD34'
'9966694544'	'B00000JFIF'
'9984984354'	'B00000JMUG'
'9985511476'	'B00000JYLO'

14.

To evaluate the performance of the model we have used the Root Mean Square Error (RMSE) which gives the accuracy of the recommendations that are generated. The results of the RMSE can gives a value of 1. 3436.To eliminate the products that were already bought by the users, we decompose the outputs. This removes the products that were already bought by the user. Therefore, the final recommendation will not have any duplicates in it.

10. BUSINESS MODEL

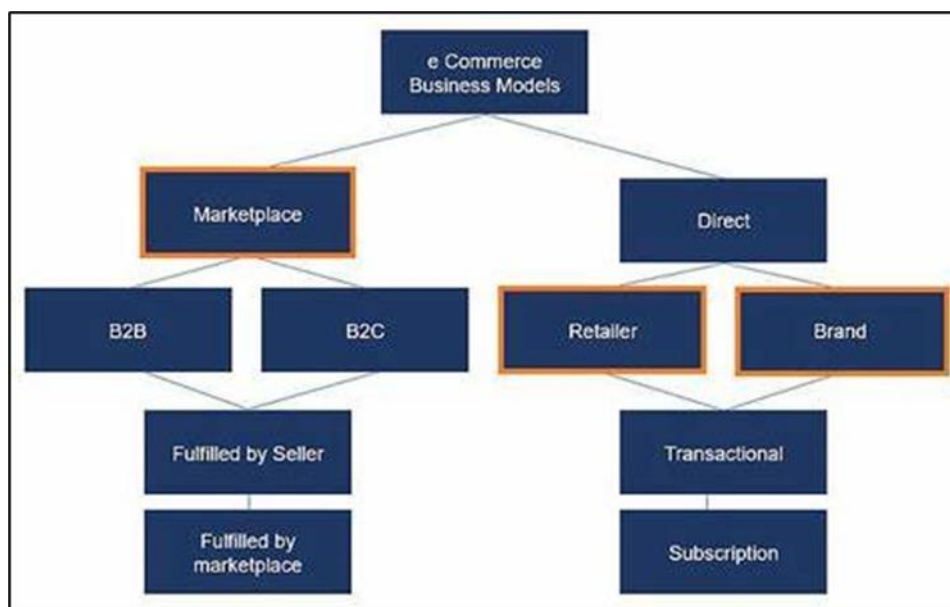What is Business Model?

A business model is a logical structure that supports the feasibility of the business. Designing a business model requires deep thought and analysis.

A business model provides a reason for the customer to choose the offering provided by your company over others. Business models focus on core strategy, customer interface, value network, and resources.

E-commerce Business Model:

E-COMMERCE BUSINESS MODEL

The eCommerce model focuses on selling products by creating a web-store on the internet (online shop). Amazon, Alibaba, eBay, olx, and Walmart are some of the big companies that have adopted an e-commerce business model.

E-COMMERCE MODEL
E-Commerce Business Model - Market Based

Step 4 a: Identify which Market the product/service will be launched into

The potential market for the Product Recommendation System is the Start-ups and Small-scale Fashion Industries. This Textile, Clothing and Fashion Industry is one of the oldest industries in India and across the world, adapting itself to the varying situations and is expected to grow exponentially by virtue of modernization and urbanization, changing lifestyles, fashion trends and especially the fashion-conscious youth of today.

The proposed Product Recommendation System uses Market Basket Analysis to gauge the sentiments of the consumers and analyse their buying patterns for fashion products. This is in use by many top-brands resulting in increased retention rate and high sales. This prototype is beneficial especially to the new fashion start-ups and small to mid-scale fashion businesses which face stiff competition from among themselves and from the big brands with immense stable financial backing.

<u>Step 4 b: Collect some data / Statistics regarding Market</u>

The fashion industries can be classified into many segments based on various criteria such as, product type, location, environment-compatibility factors, products for specific age-groups and genders, fashion type such as ethnic, sports, casual, price ranges of the products, the material of the products, etc.

The Market Segmentation Analysis can thus be beneficial to analyse such industries.

Brands like Benetton, Zodiac, Z3, Zara, Vero Moda, Calvin Klein, Diesel and Tommy Hilfiger have experienced good sales growth in the country, while names like Zara, Armani, Forever21 or Uniqlo appeal to the Indian audiences, attracting higher per square foot sales compared to the departmental or hypermarket stores.

On the other hand, Indian companies like the Arvind group, Madura Fashion and Lifestyle, Raymond Apparel, Trent Retail, Reliance Retail and Future Group has launched their own fashion labels.

Apparels command the largest share from amongst the various segments of the Indian fashion industry, accounting for 81 percent in 2017. Indian fashion practices which have traditionally been influenced by culture and region, had witnessed a gradual transition and has begun to align itself in line with the more western and fusion trends.

Recently, a report stated that the Indian fashion industry can increase from its net worth of Rs 200 crore to Rs 1,000 crore in the next five to ten years. Currently, the worldwide designer wear market is amounted at $35 billion, with a 9 per cent growth rate, with the Indian fashion industry creating hardly 0.1 per cent of the international industry's net worth.

According to approximations, the total apparel market in India is calculated to be about Rs 20,000 crore. The branded apparel market's size is nearly one fourth of this or Rs 5,000 crore. Designer wear, in turn, covers nearly about 0.2 per cent of the branded apparel market.

At present, the largest sales turnover within the designer wear segment is about Rs25 crore, with other well-known names having less turnovers of Rs10-15 crore. In view of the prospects of the Indian fashion industry for growth, the figures are not very hopeful.

Many fashion designers and management experts foresee an average growth of about 10-12 per cent for the Indian fashion industry in the coming years. Though, the growth rate could be more than 15 per cent, if infrastructural and other logistical bottlenecks and drawbacks are overcome.

11. FINANCIAL MODELLING

Financial modelling, as a process, aims to evaluate the company's current financial expenses and profits in order to predict the impact of any future event and thereby helps in facilitating better financial decisions. It does this by creating a summary of the costs and revenues incurred in the functioning of a business enterprise. Financial modelling is essential for ensuring appropriate financial management in any corporate organisation since it provides an overview of what future costs and revenues will look like.

Developing Financial Equation:

Let,

       Total Profit generated in Fashion Sale = Y

       Price of each Fashion Product = F

       Total Sale generated as a function of time = X(t)

       Total Maintenance and Production Cost = C

Financial Equation:

$$\Rightarrow \ Y = F*X(t) - C$$

Constraints:

- $\Rightarrow$ X(t): This function is proposed to get the sale produced based on various factors such as type of clothing, local brands or foreign brands, material used for manufacturing, etc.

- $\Rightarrow$ Production Cost (C): This includes the manufacturing cost, price of materials used, types of clothing, production for local sale or for export.

12. ANALYSIS AND CONCLUSION

India is not a remarkable player in the global market with reference to brands because of its inability to add value to products. This is observed by the fact that nearly 50 per cent of its exports are apparel and made-ups where value addition is essential. Likewise, 75 per cent of domestic apparel market is commoditized and unbranded and very few Indian brands do survive in the foreign markets. Evidently, the Indian market has not made a strong stand and hence it is difficult to make Indian brands that can compete with global brands in India.

Another reason for the fashion industry's inadequate growth is the limited experience of the designers and the platform they are offered. The insignificance stalks from the reality that most of the young talent is hired by the bigger names to work in their studios, thus imprinting their work with the label of the big designers.

These challenges can be overcome by proper analysis of the market assisted by appropriate measures. Our prototype can solve such challenges by recommending products to the small

businesses so that they emerge themselves into the proposed product types and expand their business gradually earning them greater profits. They can compete with big brands and also on a global level. This is going to be the need of the hour for small scale fashion businesses in this age of online shopping and changing trends.

13. REFERENCES

1. Practitioner papers: Customer segmentation in fashion retailing: A conjoint study | Emerald Insight

2. [Birtwistle, Grete, Ian Clarke, and Paul Freathy. "Customer decision making in fashion retailing: a segmentation analysis." International Journal of Retail & Distribution Management (1998).]

3. Fan, S., Lau, R. Y. and Zhao, J. L. (2015), 'Demystifying big data analytics for Business Intelligence Through the Lens of Marketing Mix', Big Data Research 2015.

4. [Birtwistle, Grete, Ian Clarke, and Paul Freathy. "Practitioner papers: Customer segmentation in fashion retailing: A conjoint study." Journal of Fashion Marketing and Management: An International Journal (1999).]

5. In-Depth Guide: How Recommender Systems Work | Built In, https://builtin.com/data-science/recommender-systems

6. Fashion products on Amazon.com - dataset by promptcloud | data.world https://data.world/promptcloud/fashion-products-on-amazon-com