Dynamic Pricing in Small Business using Machine Learning

Project report by

Zainab Fatima

ML intern,FeyNLabs

11.01.2022

Presented to

Sanjay

Project Mentor,FeyNNLabs

Abstract:

E-commerce has certainly changed the way we shop. As technology is evolving day by day, we can also witness the change in business models within e-commerce. Machine learning is undoubtedly the main reason for providing more and more opportunities for small and big businesses.

The project here will maximize the revenue for the e-commerce companies using Machine Learning. Here's how, the probability for an online buyer to buy something depends on many factors. Since even a small webshop can create millions of relevant data points daily, one thing we are sure of is that data about these factors can be easily collected.

The project can actually predict the right price of product for which a buyer can buy it. Here we try to maximize the price for which the customer will buy and not the maximum price the seller wants to sell it for, the seller only limits the minimum price for the product. That is how the product can be sold on uttermost profit in order to maximize overall revenue.

Price is, unsurprisingly, one of the most important factors when considering a purchase. In fact, price is one of the major drivers for at least 47% of customers. There are many other factors that may influence your prices – supply and demand, which promotions are running, what time of the day or year it is and so on.

Via machine learning, these factors could be assessed in a fraction of a second and your site will display dynamic pricing: showing the customer the most up-to-date price for them based on the aforementioned factors.

The goal is to find the optimal pricing of different products based on base price and discounts in order to create maximum sales/revenue on a particular day. Demand prediction on a particular day plays a key role. In addition, the price elasticity of demand plays a key role in solving the pricing optimization problem. Lower the price, higher the demand and vice versa.

Problem Statement:

The e-commerce website owners face a huge problem while setting the price of their products. And once they set it, another problem arises that it cannot be changed meanwhile it is required to change with time. Here's why, the possibility of a product being sold depends on many factors and these factors change with time. So the price of the product cannot be fixed and has to be changed dynamically.

The project here solves this problem by collecting customer data and using a Machine Learning model to dynamically set the price of a product changing from time to time.

Market/Customer/Business Need Assessment:

Here we thoroughly look at the customer needs which in this case are the e-commerce website owners.
The main need of the customer is to get the service which solves the problem of predicting the best price for every product and get rid of the headache of setting up prices.

- Prices that generate maximum profit.
- Prices are predicted every time there is a change in factors and dynamically set every time on the website.
- If there is more demand for a product the price should be increased by itself but if the demand is less i.e. lesser number of people who want to buy the product then the price should go low so that demand can be increased.
- The e-commerce owners also want the service to take care of the comparison factor. If the competitive websites are showing the lesser price for the exact same product then this will be a big turn-off for the company. So while setting up prices also include the competitor's prices as a factor.

Target Specifications and Characterization:

(1) Customers must be heterogeneous in their willingness to pay, that is they should be willing to pay different prices for the same products or services.

(2) The market must be segmentable, that is, it should be possible to identify different groups of buyers. The web has significantly improved a company's ability to profile their customers and track their behaviour. Two examples: (1) A grocery customer might sort products by price before choosing (price-sensitive customer) or might use non-price attributes like brand name or quality to shortlist and select (price-insensitive customer). (2) In airlines booking, a business customer is sensitive to the time and date of departure while a pricesensitive customer wishes to choose the minimum price schedule.

(3) Arbitrage should be limited, that is, a customer who bought a product at a lower price should not be able to resell it for profit to customers who have a higher willingness-to-pay. For example, a cheaper airline ticket has so many restrictions that to resell it at a higher price is almost impossible.

(4) The cost of segmenting and price differentiation must not exceed revenue due to price customization. For example, airlines have some of the most sophisticated price customization schemes which took millions of dollars to implement. However, these schemes have led to revenues that are far in excess of the setup costs, leading to the successful deployment of dynamic pricing schemes. Similarly, the presence of companies like Priceline.com has helped the airline industry by generating additional revenues on seats that would have otherwise gone unsold.

(5) Customers should perceive fairness while dealing with a vendor who practices dynamic pricing

External Search (information sources/references):

[1] Rahul Jadhav, Arpit Vais Hya, Mohammad Harris Nawarangee, Ajay Dhruv "DYNAMIC PRICING INTELLIGENCE APPROACH IN E-COMMERCE MARKET USING KNN"
[2] AI Multiple: https://research.aimultiple.com/dynamic-pricing-ecommerce/
[3] Pricing Optimization & Machine Learning Techniques https://vitalflux.com/pricing-optimization-machine-learning-techniques/
[4] "Dynamic pricing models for electronic business" Y NARAHARI1,CVL RAJU1, K RAVIKUMAR2 and SOURABH SHAH3
[5] https://www.coredna.com/blogs/vital-ecommerce-metrics
[6] https://www.coredna.com/blogs/dynamic-pricing-ecommerce

Benchmarking alternate products:

- **Amazon** is one of the largest global e-commerce platforms with ~300,000,000 active users. It's been estimated that Amazon changes its prices more than 2.5 million times a day to set their prices lower than their competitors. In 2016, it was claimed that dynamic pricing helped Amazon increase profits by 25%.
- **E-Bikeshop**, a UK bike supplier platform, leverages dynamic pricing to provide attractive prices in comparison to UK competitors such as ElectricRider and Electric Bike World.

Applicable Regulations (government and environmental):

As far as regulatory framework governing e-commerce activities is concerned, there are no dedicated e-commerce laws in India. Various ministries and department of the government of India deal with different aspect of e-commerce. For instance, the ministry of Electronics and Information Technology look after the technical aspects of e-commerce through the information technology Act, data privacy issues, etc. The Department of Consumer Affairs takes care of the consumer protection issues. The Department for promotion of Industry and Internal Trade deals with the foreign investment related matters on e-commerce. The Department of commerce deals with

the WTO discussions on e-commerce. Although there is no dedicated law governing e-commerce, such activities are governed by a number of laws & regulations applicable to various segments of the e-commerce value chain. Some of these laws come under the purview of central government whereas others fall within the jurisdiction of state governments.

Following are some regulations that are related to our business model:

- **REGISTRATION REGULATIONS**

Registration as an e-commerce companies Act, 2013. Registering an e-commerce business either as a company or a firm or a limited liability partnership [LLP] or sole proprietorship. At least two people are needed to register as a private limited company. A memorandum with the objective, Liability of member, capital clause of the company and article defines powers of the management.

- **TAX REGISTRATION**

Registration for goods and services tax [GST] if the sales or turnover crosses more than rupee 20 lakh annually or rupee 10 lakh in case the establishment is in the northeast states.

- **INFORMATION TECHNOLOGY REGULATIONS**

Information Technology Act, 2000. The IT Act 2000 is the sole cyber law in India which also governs, to some extent, the online issues of e-commerce in India. Although the IT Act focuses mainly on digital signature and related aspects, it mandates that the e-commerce entrepreneurs and owners must ensure cyber law due diligence in India.

- **DATA PROTECTION REGULATORY FRAMEWORK**

Section 79 of IT Amendment Act, 2008 sets out conditions under which an intermediary will not be liable for any third party information, data or communication link made available or hosted by him. Data protection has been made more explicit through the insertion of Clause 43A that provides for compensation to a person whose personal data may have been compromised by a company. Under section 72A, punishment for the disclosure of information in breach of a lawful contract is prescribed. Any person including an intermediary who has access to any material containing personal information about another person, as part of a lawful contract, discloses it without the consent of the person will attract punishment with imprisonment of up to 3 years and/or fine of 5 lakh. Section 69 on crimes against national security has been made stronger for interception and monitoring. Sections 66 & 67 on hacking and obscene material have been updated by dividing them into more crime-specific subsections, thereby making cybercrimes punishable.

- **PAYMENT AND SETTLEMENT REGULATIONS**

To allow online payments receipt and disbursements for e-commerce activities, one has to take a license from the reserve bank of India [RBI]. Payment & settlement systems are regulated by the payment & settlement systems Act, 2007 [PSS Act] and Settlement System Regulations, 2008. As per Section 4 of the PSS Act, no person other than the RBI can commence or operate a payment system in India unless authorized by the RBI. The RBI has since authorized payment system operators of prepaid payments instruments, card schemes, cross-border in-bound money transfers, automated teller machine [ATM] networks and centralized clearing arrangement.

- **THE DRAFT E-COMMERCE POLICY, 2018**

The government came out with a draft of much needed policy on e-commerce in 2018. both domestic and international concerns must have guided the framing of such policy. The draft policy has a number of facilitative elements for e-commerce players. For instance, the recommendation to establish a central registry for KYC will reduce the cost and burden of KYC compliance by the payment systems. Similarly the provisions to have controlling stake in spite of minority share by the promoters will allow the young entrepreneurs to grow and seek funding support without giving up control in their entity. There were criticisms of the draft e-commerce policy as it was considered as a regressive step by many. However, one must see this policy from developments at the international arena pertaining to e-commerce. The proposals submitted in the WTO on e-commerce tend to include almost everything under the sun, including physical trade, online trade, payment systems, consumer protections, telecommunication networks, spam mail and source code, to name a few. If all these things are being discussed in the WTO, and may be negotiated in future, the government must know its landing zones in international negotiations on various elements of e-commerce.

Applicable Constraints (need for space, budget, expertise):

The main components of dynamic pricing are:

- Manpower
- Software requirements
- Space required on cloud

1)Man Power
We will be needing a team of engineers to develop the project consisting of big data analysts, Machine learning engineers, and web developers.
2)Software requirements:

2.1)For e-commerce website development:

- Server: WAMP / LAMP / XAMPP
- Database: MySql, postgresql, Oracle, Big Data Databases
- Editor: Sublime / Eclipse PHP
- Web Browsers: Chrome, Mozilla, Safari (for testing purposes...also some browsers provide responsive or mobile view)
- Browser Add-ons: Firebug, Web Devloper etc
- POSTMAN: For REST API
- Javascript libraries: jQuery, Angular js

2.2)For Machine Learning:

- Hadoop
- Keras
- TensorFlow
- API or as a complete web service, or in a mobile device, Raspberry PI, etc.

3)Space:
The space required depends upon the number of products the website is willing to add to its inventory.
In general, we should plan with **at least 2 GB of disk space and 10 GB/month of bandwidth** as these are common, and will give you more than enough resources for an online store with thousands of products. This is a good starting point for your online store.

Business Opportunity:

The generation of maximum revenue by selling the most number of products at the right price is the main purpose of the model. This sums up the business opportunity by the model.It solely focuses on the need of e-commerce business owners and provides them with the efficiency of selling their product as per the market conditions.

Concept Generation:

The concept of marketing is the process of exploring, creating and distributing value in order to satisfy specific needs of consumers in a specific market while making a profit. Through marketing, all decisions and strategies adopted by a given company in achieving its objectives are defined.Marketing is governed by four fundamental levers, called the 4Ps: product, price, place and promotion. Product refers to the goods or services being sold. The variable price represents the monetary counterpart the consumer is willing to pay for a given product. It is determined by the company through specific policies called pricing. Place is the set of activities and structures or buildings necessary to bring a given product to the final consumer (e.g., stores and shops). Promotion refers to all communication activities between the company and the consumer in general, for example, customer care or advertising. The set of decisions concerning the 4Ps is also called a company's marketing mix. Applying this framework to an online context where multiple vendors are present for the same category of product, the marketing mix variables involved are reduced to three: price, place and promotion. There is no longer any competition in product characteristics, Forecasting 2021, 3 169 which are common to all competitors in the market. The variables of promotion and place are closely linked, and sometimes the boundary between the two is not clearly definable. Basically, one can consider the joint effect of place and promotion by interacting with the graphical interface of the website on which the product is sold. The e-platform not only replaces the traditional physical store where customers go to purchase products but also provides fundamental assistance to customers, such as effective management of returns, acceptance of complaints and general customer care throughout the customer purchase experience. With regard to price competition, there are various digital platforms specifically developed to allow consumers to compare prices of products online, such as Google Shopping. The very fact that multiple companies offer this service shows that although place and promotion have the same levels and characteristics, price plays a fundamental role in e-shoppers' purchase decision process. In this way, e-commerce companies have to deal very carefully with the pricing policies of their products, as even the slightest deviation from the competition can result in lost revenue or possible gains. Of course, the common goal of all companies is profit maximization. Given the high competitiveness in products that characterizes e-commerce, this maximization comes with a trade-off between sales profitability and market competitiveness. Although companies tend to seek the maximization of mark-up (i.e., profit, calculated as the difference between the price of a product and its cost of production), they cannot ignore that, for the same product, the price is inversely proportional to its competitiveness in the market. For example, a price that is too high would lead consumers to prefer the competitor's products. To adopt the highest

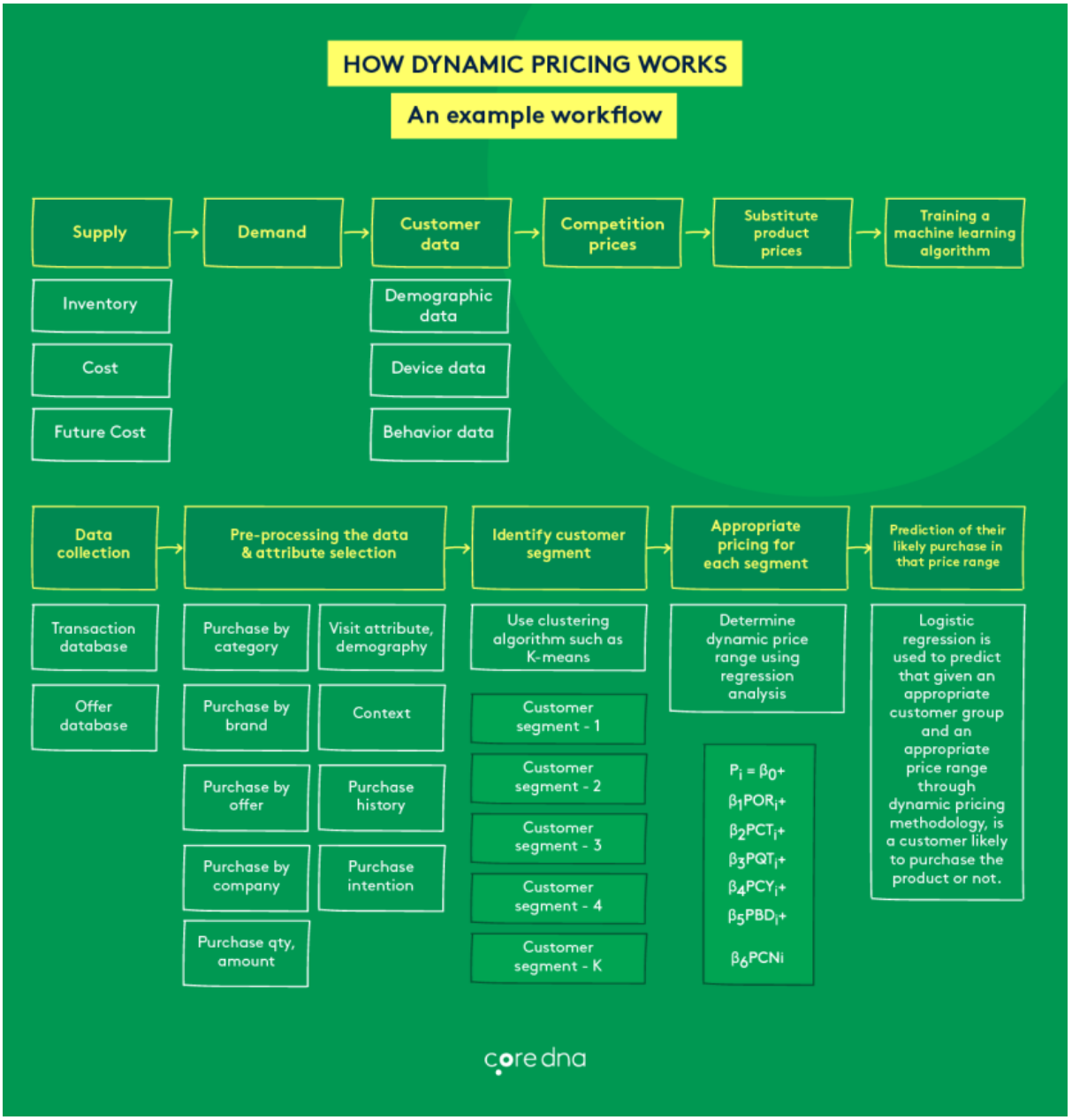
possible price for products, companies resort to various pricing strategies. The first strategy, called price skimming, aims to make customers more willing to spend on a given good or service. In this way, as much of the consumer's possible revenue is absorbed as possible. In practice, this strategy initially consists of setting the highest possible price, sometimes referred to as the launch price, and then gradually reducing the price to appeal to customers willing to pay less. The second strategy that e-commerce companies adopt is what is known as competitive pricing. This pricing strategy is most common for companies in the same sector and is an important characteristic for markets with free competition. If, for example, for the same product, one company's price is higher than the prices of all its competitors, the probability of making many sales will be very low, if at all. The third strategy, similar but in some ways the opposite of competitive pricing, is dynamic pricing. This strategy does not consist, as with competitive pricing, of simply taking competitors' prices into account, when setting one's own price, but of setting a price that is exclusively based on competitors' price levels. Sometimes, companies decide to monitor their competitors' prices in real-time, through specifically developed algorithms that process competitors' prices or consumers' data, and adjust their prices accordingly. In situations where a dynamic pricing strategy is adopted, in the event that a company raises its price, its competitors have several options for action. The first is to adjust price levels with a consequent price increase, given the decrease in the general price competition in the market.

Concept Development:

A new phase of business "E-commerce" has arrived, and yet, there are certain crude old problems we still face — finding primary competitors, Keeping track of the market, acquisition of profitability. Traditionally manual track of prices, stocks and market trends of its competitors. Since with the advancement in the technology and higher demands of E-commerce transactions staying always at the edge of the curve is critical. One needs to be periodically updated with the market trends. With the increased competition and reliance on instant shopping by the customers , one needs to track its own products with the products sold by its primary competitors targeting its own primary audience .This brings in a proposed system as a solution for increasing the efficiency of business decisions. By developing a crawler for e-commerce website(primary competitors) which crawls and scrapes the required parameters for comparison of products and compares each of the client-side's product with same product sold in the market using KNN algorithm for reducing the human effort to analyse each of its product sold in the market. With the growing market trends and digitization in every sector, ecommerce websites have created a tremendous growth in terms of shopping which further leads to competition within

the markets. Main objective is to automate the entire system where the crawling and scraping of data happens in real time. The program of matching the product with the crawled data of market will run from the instance created on the cloud in the real time so as to reduce the infrastructure constraints. This will render and will display the complete analysis study of bulk data crawled from it's competitors E-commerce website and will help in improve business decisions

Final Product Prototype:



1. Live model having plug-in with the website:
   The model will be live as it needs to dynamically change the price which it will predict from time to time.

2. Model will also have connectors with competitors websites:
   As the price prediction of the model also depends on the factor at what price the competitor websites are selling the same product. It needs to feed to keep the record in the shifts of their prices.

3. Model will also keep collecting data about the increase or decrease in demand for the products.
4. In addition the model will also be fed with predefined factors like festivals which will affect the demand of the product or how users are willing to buy something specifically for the occasion.
5. Elasticity module of the model:
   The model uses time-series methods and big data analytics to calculate how a product's price affects demand, accounting for a wide variety of factors, including seasonality, cannibalization, and competitive moves.

Product details:

The model follows a data pipelining process which starts with acceptance of category or brand from the business client and finding their competitors in market with the help of automated crawler bot. This bot also scrapes entire product data from all websites which is further used to match the products and analyze them for various metrics.

Step 1: Acceptance of category from business client. Here, the client makes decisions about which category they wish to get detailed analytics so as, not only to improve their stand in market but also to improve their business.

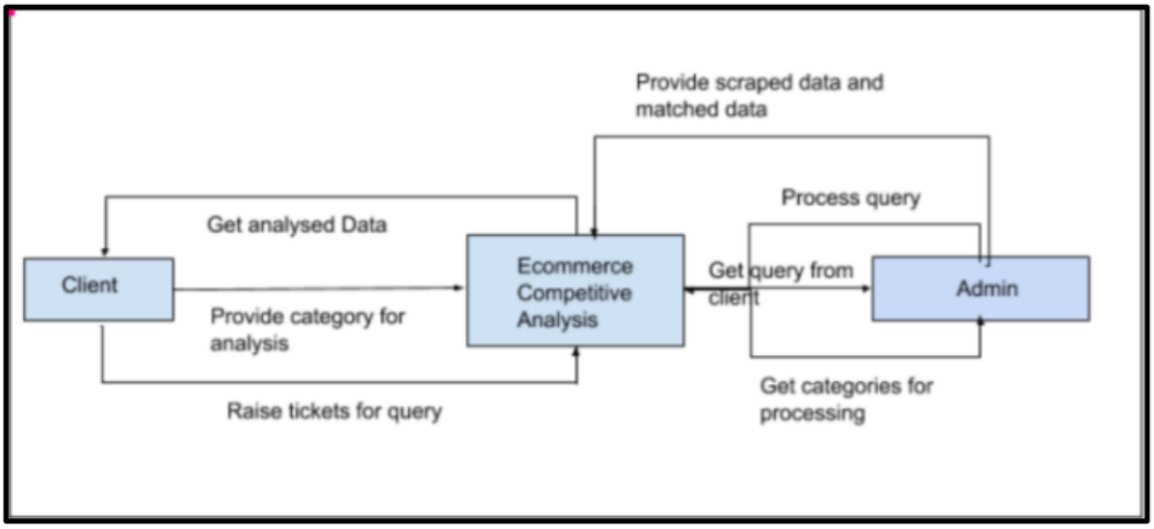


Fig. 1

Step 2: Crawling of client's data and their competitor's In this step, build an automatic crawler to scrape all the data related to the category provided by client. Then by mapping brand and category, effective competitors for each (category + brand) products is retrieved based on google search rankings. "Scrapy" framework built on python is used for web scraping. Scrapy is a fast high-level web crawling and web scraping framework, used to crawl websites and extract structured data from their pages. Below diagram shows the architecture of scrapy.

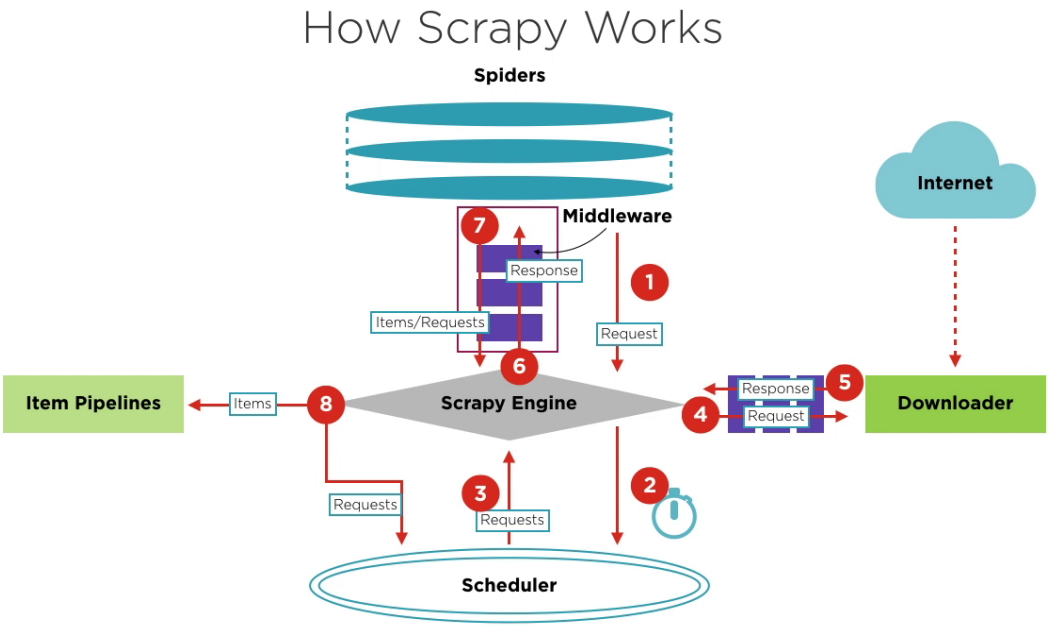


Fig.2

Crawlers built using scrapy processes the pages asynchronously which makes the speed of processing pages faster. Apart from building crawler for e-commerce websites one more crawler is built which finds the competitors of our client based on particular brand or category. The processing for finding competitors is done using google search ranking and these results are sorted in an array based on number of times they appear in search results. Scrapy makes it easier to solve the problem that occurred while crawling a webpage such as HttpAuth, HttpCache, RFC2616 policy, by passing robots.txt file etc. by using Downloader Middleware like DefaultHeadersMiddleware, DownloadTimeoutMiddleware, HttpAuthMiddleware etc. For maintaining the consistency of crawlers, hash value is used. A hash value of each page is pre-calculated using SHA-1. This precalculated value is compared with the value which is calculated before crawling starts.

Note: First comparison of hash value is done and then crawling begins because the structure of webpage remains same,only the data changes but over the time the structure may also change.

Step 3: Matching similar products among competitors. First most step is matching of e-commerce products from different websites, So primary matching is done with the title of each product from different sites like, Amazon, Ajio, LimeRoard “Levenshtein” distance matching algorithm compares the string by converting each character of the title into lowercase and the matches the word based on occurrence of word in each of the comparing titles.

Levenshtein distance (LD) is a measure of the similarity between two strings, which is referred to as the source string (s) and the target string (t). The distance is the number of deletions, insertions, or substitutions required to transform s into t.

For example,

• If s is "test" and t is "test", then LD(s,t) = 0, because no transformations are needed. The strings are already identical.

• If s is "test" and t is "tent", then LD(s,t) = 1, because one substitution (change "s" to "n") is sufficient to transforms into t. The greater the Levenshtein distance, the more different the strings are.

This algorithm is efficient enough to deal with the fuzziness of the word including the arrangement of words. But the problem arises when different e-commerce websites list their products with entirely different titles.

Scale-invariant feature transform (SIFT) is a feature detection algorithm using computer vision technology is used to detect and describe local features in images. This algorithm extract key-points from set of reference images and compare this key-points from the matching image using Euclidean distance algorithm. This algorithm will only work when the Euclidean distance of all the images from different sites are same. But as different products are captured from different angles this algorithm fails to deliver the desired accuracy. KNN (K-Nearest Neighbours), is a clustering algorithm which clusters data into groups based on certain common features. In this research, pixel of images is used as a feature for clustering images and calculate the distance between clusters using Euclidean distance between them. This helps us to rank the most matchings results based on the nearest neighbours. This can be used to implement this step with efficient accuracy. As a result, famous clustering algorithm “K-Nearest Neighbours”(KNN) was used along with tensor-flow framework of Machine Learning in python for effective GPU support and faster processing in training and testing of images in large quantity. Given a set of query images and database images, Image retrieval on database images is performed to get the top-k most similar database images using kNN on the image embeddings with cosine similarity as the distance metric. Libraries used were tensor-flow, skimage, sklearn, multiprocessing, numpy, matplotlib.

Step 4: Aggregating Data. Here matched results from title matching and image matching process is being aggregated into single json file for easy retrieval, display and effective analysis.

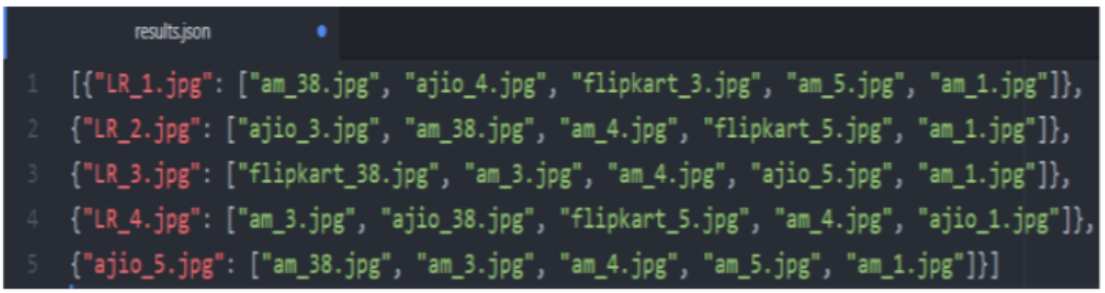


Fig 4. Aggregated Json Data

Step 5: Analyzing various metrics: Different metrics used to get intelligence in business processes are as follows:

(a) Total quantity of in stock and out of stock products categorized by brand and category as shown in fig 4.

(b) Timeline history of most selling products and least selling products.

(c) Profit comparison between competitors based on overall marked price, selling price, discounts.

(d) Comparing products among competitors for effective and dynamic pricing strategy.

(e) Track of dynamic pricing strategy of competitors shown in fig 5

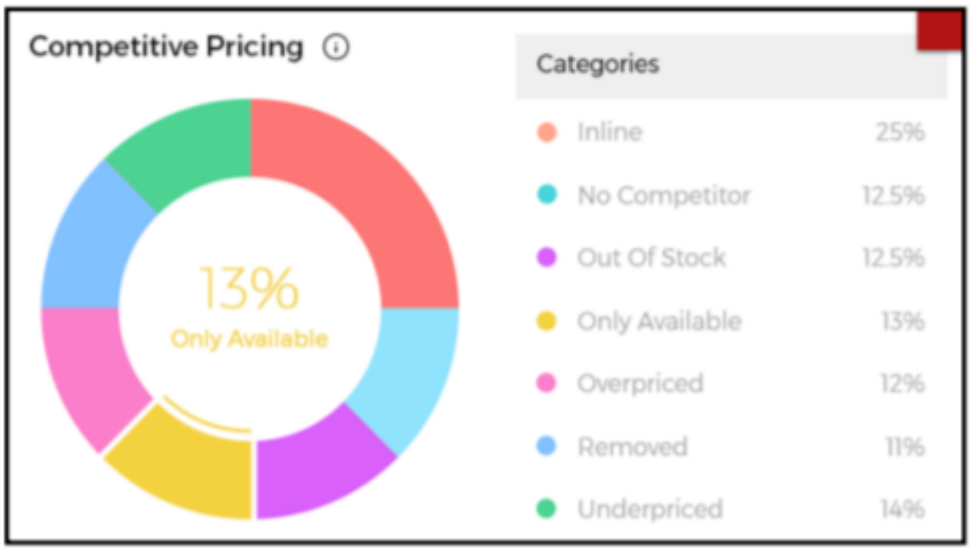


Fig. 4

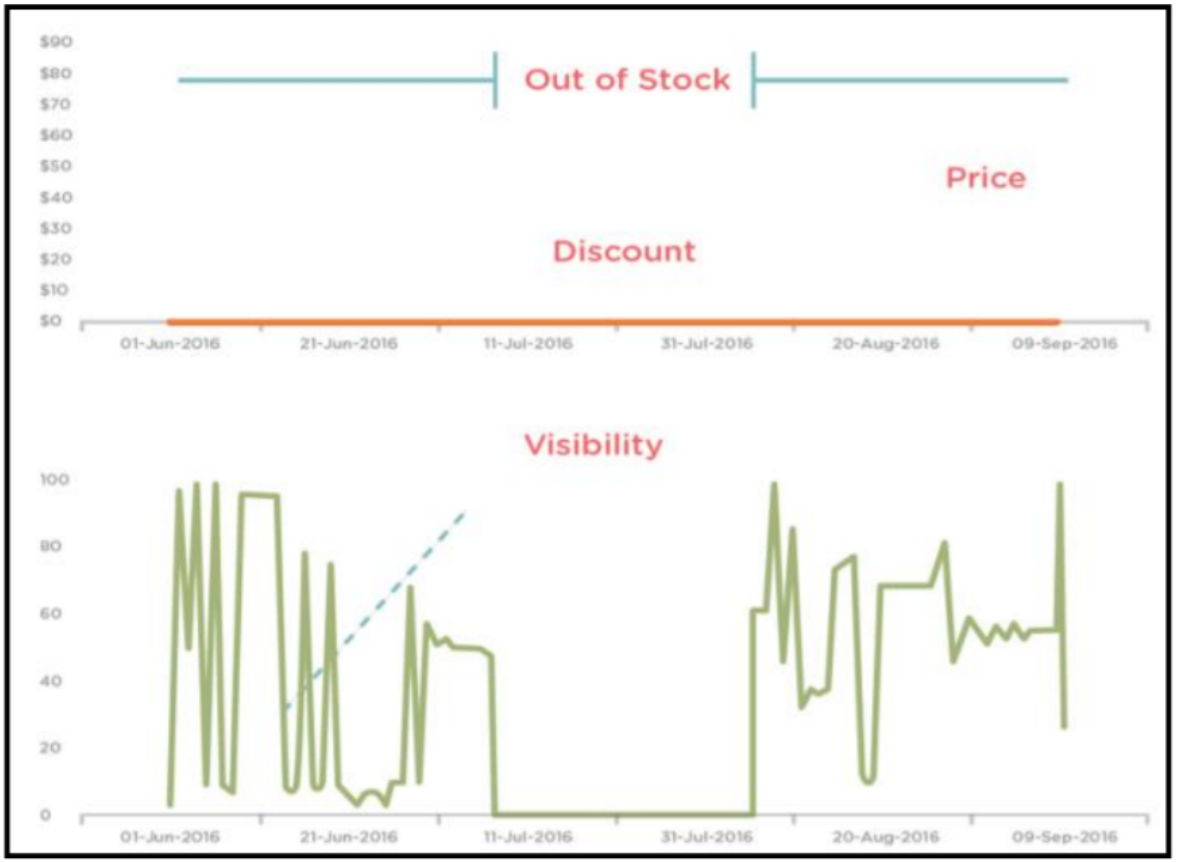


Fig. 5

Conclusion:

With the rising era of E-commerce transactions which brought growing competition, The proposed system has eradicated the computational, storage and human efforts of gathering and matching the client products with it 's primary competitors using KNN algorithm providing an accuracy of 66% in the worst case scenario. The client side is able to have a graphical and pictorial representation of the analysis of its own products with respect to its competitor. This has helped the client to overcome the problem of periodically updating itself with the growing dynamic environment of E-commerce. It helps by reducing the human efforts of manually keeping a track of their respective selected products with it's competitor's price, stocks and offers periodically and updating manually .Thereby providing a complete automated system where the computational ETL process is clustered in one system.