

Group No. 7

**RECOMMENDATION ENGINE FOR CAMERA RENTAL SERVICES (KLACHAK)**

**Project Interim Report**

PGDBA-BI – Jan 2017

Abhishek Ramachandan

Jayaprakash Nallathambi

Santhosh Murali

Dr. Monica MIttal

Mentored By

****

Harish Ganesan

Interim on

**Recommendation Engine for Camera Rental Services (Klachak)**

**Mentor :** Dr.Monica Mittal

**Students**

|  |  |  |
| --- | --- | --- |
| **Jayaprakash Nallathambi** | BACJAN17022 | 9940170133 |
| **Abhishek Ramachandran** | BACJAN17002 | 9841411556 |
| **Harish Ganesan** | BACJAN17020 | 9790793927 |
| **Santhosh Murali** | BACJAN17062 | 9677291041 |

Contents

[Introduction 4](#_Toc495844972)

[Camera Rental Services Industry 4](#_Toc495844973)

[Klachak.com 5](#_Toc495844974)

[Equipment Rentals 5](#_Toc495844975)

[Photography Services 5](#_Toc495844976)

[Problem/Opportunity Statement 6](#_Toc495844977)

[Analysis Methodology 6](#_Toc495844978)

[Methodology 6](#_Toc495844979)

[Data Source 7](#_Toc495844980)

[Data Preparation 7](#_Toc495844981)

[Modeling 7](#_Toc495844982)

[Tools and Software 8](#_Toc495844983)

[Exploratory Data Analysis 9](#_Toc495844984)

[Sales Yearly Trend. 9](#_Toc495844985)

[Seasonal Trend 9](#_Toc495844986)

[Percentage of Revenue 10](#_Toc495844987)

[Cohort Analysis 11](#_Toc495844988)

[Data Gathering Techniques 12](#_Toc495844989)

[Web Scrapping – Using Scrapy 12](#_Toc495844990)

[Information Gathering Flow 13](#_Toc495844991)

[Scrapy Data Flow 13](#_Toc495844992)

[Recommendations & Applications 14](#_Toc495844993)

[Overall Architecture 14](#_Toc495844994)

[Status 15](#_Toc495844995)

# Introduction

## Camera Rental Services Industry

This is a budding industry with only very few active players. So, what is the business? Let’s assume that Someone called Sam is a great Photographer and would not mind crossing any limits to get best click. He wanders around the globe to capture the moments. Every photographer has their own gears which they used carry all around with them ever. {gear: Camera Body, Lenses, Filters, Tripod, Lights etc). Like that Sam also has his own set of gears, he has Canon 5D Mark II Body, Canon 70-200 IS2 and Canon 16-35 L lens, Tripod and Travel bag. His total gear value is provided below.

|  |  |
| --- | --- |
| Canon 5D M2 | ₹ 70,000.00 |
| Canon 70-200 IS2 | ₹ 120,000.00 |
| Canon 16-35 L | ₹ 110,000.00 |
| Tripod | ₹ 27,000.00 |
| Bag | ₹ 3,000.00 |
| Total | **₹ 330,000.00** |

With his existing set of gears limits his creativity. With Canon 70-200 IS2, he could capture long range objects but still not extensive telescopic range and fstops are not wide enough. On the other hand, he has 16-35mm wide angle lenses which will cover larger area and minimal depth. So, whatever photograph he takes it will either long distance and not too long with moderate depth information or a wide angle. Suppose if he wants to take a portrait his gear pack will not help him. He has to have either “Canon EF 35mm f/1.4L II USM” or “Canon EF 50mm F/1.2L USM”. It will cost him additional 200,000 or 100,000 respectively. If he wants to go for macro photography, he has to invest ~100,000 additionally. Sam can afford to buy these new lenses. But, there are Hobbyists, Cost sensitive photography professionals, who might not invest such a large amount. Also, not all lenses are used at the same frequency. Application of Camera and lenses are very dependent on time and situation(event). Not everyone can afford to buy all kids of cameras and lenses.

That’s where the rental services industry identified the opportunity. These companies will have an inventory of all kinds of cameras, lenses and its accessories from almost all brands. They rent these gears to customers charged based on daily based rentals. This way it becomes win-win deal for both the company and Customers. All their inventory has now started minting money while Customers(Photographers) need not invest whooping money for lenses, instead spend fraction of that as rent and also gain access to wide range of options. Though it seems to be very attractive, this industry is not exceptions from challenges from market and competition. Top 5 challenges are listed below.

* Inventory cost
  + Average lens cost comes out to be 135000.
  + If the Rental company wants to keep all the lenses available in market, it will go beyond $100 M.
* Narrow customer segment
  + Unlike other rentals products like Car, garments etc, where anyone can be a prospective customer. In Camera rental services, it is mostly the Photography aspirants and professionals. Their population is negligible.
* Services and handling
  + Services and handling charges for these gears are very costly.
* Securing the devices
  + These are very costly devices

. It requires at most care and make sure the gears that are rented out are safe. Customer’s background verification test is also important to avoid gear loses.

* Pricing
  + Not all the camera gears go out for rent in same frequency and not everything is of same price. Some gears are less expensive with aggressive movement, some are very expensive and little movement. Rental pricing should be customized at the gear level to meet early break even.

Though we have such challenges, the industry is growing as more and more people are interested in Photography. Not only that, Social media and media services over the internet are complementing the camera industry for is growth.

## Klachak.com

Klachak.com is a company based out of Chennai, founded by a team of entrepreneurs, who are professionals in their own fields of expertise, yet have a great passion and skill in the art of photography. The company was formed with a vision to make photography accessible to all sincere enthusiasts who would like to create art with light.

It facilitates artists of every level to overcome their barriers to engage with photography, and help inspire creativity and ingenuity in this form of art. It's aim is to provide our discerning clientele services of every kind that is related to the field of photography including education, lens rental, and experiences on field both within and outside the country.

### Equipment Rentals

Klachak provides a top of the line service in equipment rentals to assist professionals and amateurs to hire equipment that is usually out of reach financially. Be it lenses, cameras or any other photography related product, Klachak aims to make it accessible at a fair price, while still maintaining the best quality of service. Whether you have a wedding to shoot, a friend’s birthday party to capture, or bring home memories of wild animals in the forest.

### Photography Services

Through both in-house photographers, as well as through selected professionals in our network, Klachak provides a range of photography services such as educational workshops, commercial shoots, portfolio development, for both individuals, corporates or advertising agencies. We also hold a high-quality stock photography collection, that can be licensed for a wide range of needs.

# Problem/Opportunity Statement

“*Primary Objective is to help users in finding the cameras, lenses or any other camera related gadget they would like to rent by predicting likeliness score or a list of top 5 or 10 recommended items for the given users, by using Recommender System Algorithms”.* To Reach that Goal, we have to perform very extensive Exploratory Data Analysis, which will produce Key Reports and Dashboards. This also gives opportunity to Perform Cohort Analysis and RFM Analysis as well, which will help Kalchaks in Segmenting and Targeting their loyal customers, there by it can recommend the Items to users through offline channels also.

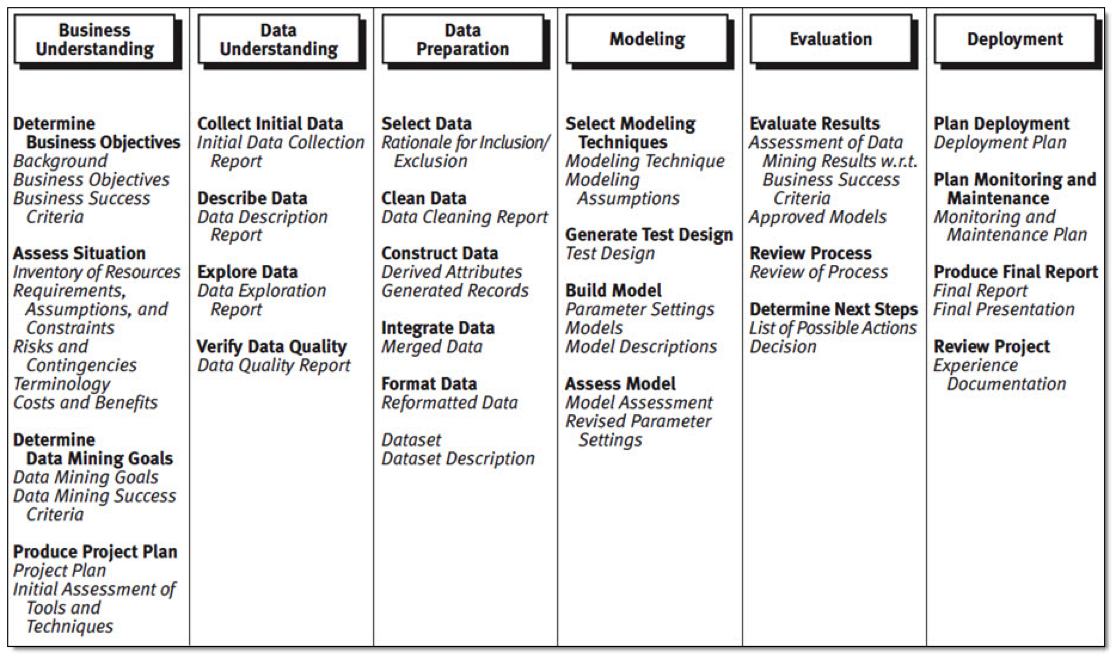
These report will further help Klachaks in,

1. Identifying all the factors contributing for more sales (in our case rentals)
2. How can Klachak take advantage of seasonality in rental activities.

# Analysis Methodology

## Methodology

This project will follow CRISP-DM model as a framework and methodology.



Understanding the business and data are paramount for this project. Once we have data, We will proceed with preparing the data for further Statistical and descriptive analysis. This data preparation will include data cleaning, masking, missing value treatment, removing bias by applying various sampling techniques.,etc., Before even proceeding with the project, initial study of data has been conducted as proof of concept and measure the feasibility.

Data preparation will be very critical. Wherever necessary we will apply transformation logics, like converting categorical variable to set of dummy variables, clubbing to variables to form a new variable, segmenting continuous variable to convert them to categorical variable,., etc., As we have data from 3 data sources, it is important to establish a relationship between them through some key variables.

Once the data set is prepared, Descriptive Statistics will be applied again on the prepared data to understand the data again and ensure the data is normal. Descriptive Stats includes studies like, central tendency, frequency distribution, variance study etc.,

## Data Source

We are using both Primary Data and Secondary data for this Project.

Primary Data is provided in an Excel format by klachak itself. This holds all information about transactions over the period of last 4 years.

We will access <https://www.dpreview.com> website to scrap, customers review and expert review of all lenses and cameras.

Data is also pulled from “flickr” database, for capturing the application of the camera and lens. For example, a camera “Canon 580ex ii” is used mostly for wildlife and Journalism.

|  |  |
| --- | --- |
| Primary Data Source | Secondary Data Source |
| Sales/Transaction Data from Klachaks | Detailed Product Information from  <https://www.dpreview.com>  Detailed Application of Product is obtained from  <https://www.flickr.com/> |

This data will help us in improving the accuracy of recommendation based on Situation.

## Data Preparation

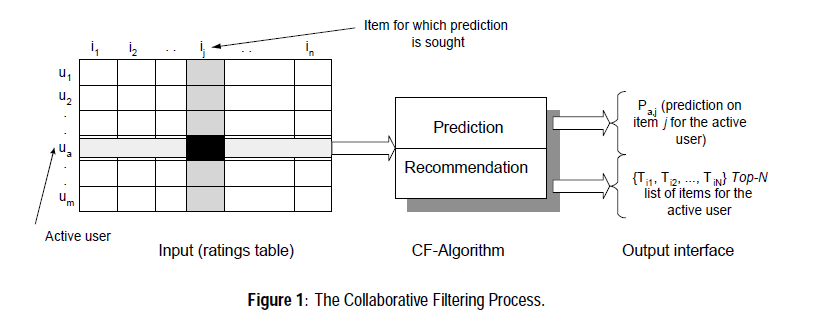
Once we have collected all the information about each variable including business implications, next we have to assess the significance of each variable, check for the existence of multi collinearity.

In case of existence of multi-collinearity, we may have to either perform PCA and Factor analysis and perform dimension reduction. After performing all the above test and treating the data accordingly, we have to next proceed with Model building.

## Modeling

There are multiple studies that are required during the progress of this project.

* Item – Item - Collaborative Filtering: Mapping customer preferences based on transaction data and provide recommendation to a given user by matching his/her item of interest with similar items close to the user preferences by using past transactions.
  + Collaborative Filtering based Algorithms provides item recommendations or Prediction based on the opinion obtained from other likeminded users explicitly or implicitly using other measures available in the dataset. Our Project will repurpose the existing data set to derive the opinion score for every user.



Source : <http://homepages.abdn.ac.uk/advaith/pages/teaching/abdn.only/AIS/lectures/abdn.only/CollaborativeFiltering.pdf>

There are 2 types of implantation of Collaborative Filtering Algorithm, Memory Based [User Based] and Model Based [Item based]. This project uses Model Based Recommender Algorithm.

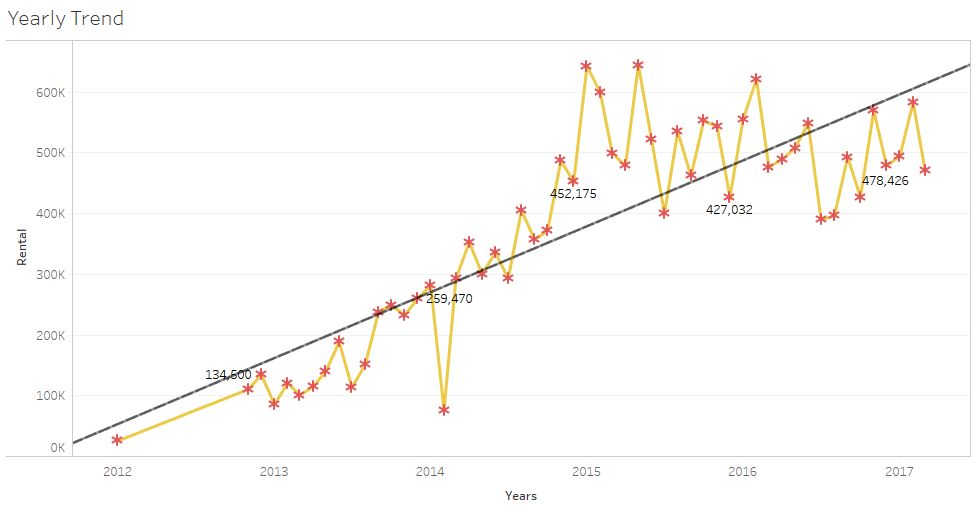
* As this is a transaction based data, we will be using Time Series Forecasting for forecasting sales, rentals. We might use ARIMA (Autoregressive Integrated Moving Average) or Holt-Winters method. This will also help us understanding the pattern and seasonality of specific groups of gears going out for rentals.
* We will be using RFM Model for Customer Segmentation. This will help us target right set of customers with right set of products. This will boost the chances of renting more gears there by improving the revenue.
* Using Web scrapping from <https://www.dpreview.com> and Text mining, we are going to build a “Topic Model” which will then be used to match the results from similar product from K-Means, which in turn will feed the recommendation engine.
* Cohort analysis for customer loyalty.

## Tools and Software

* Python
* R
* Tableau
* Microsoft Excel
* Microsoft Word
* Microsoft PowerPoint

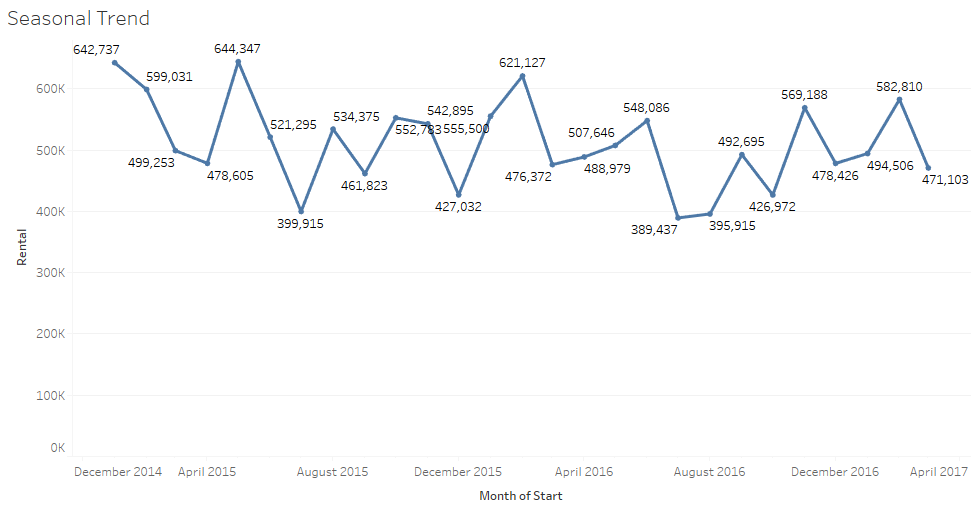
# Exploratory Data Analysis

## Sales Yearly Trend.



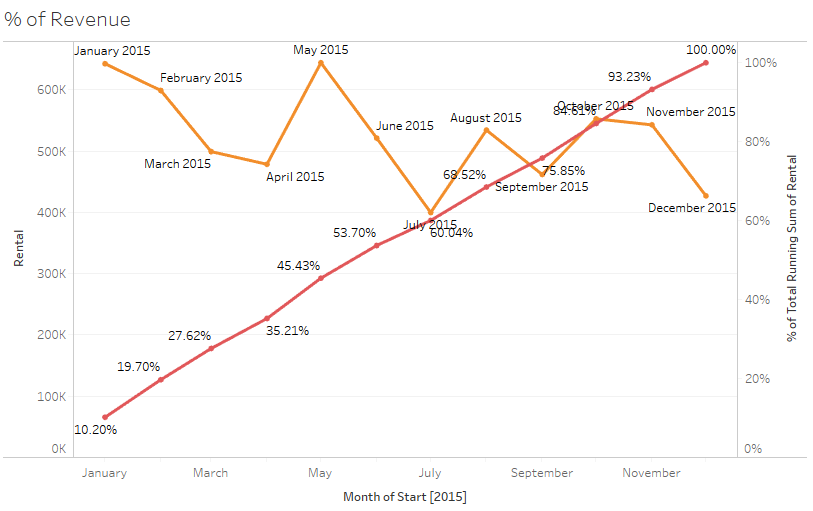
**Inference**: Rentals shows increasing trend whereas we see seasonality only for the last two years. Also, trend seems to become horizontal.

## Seasonal Trend

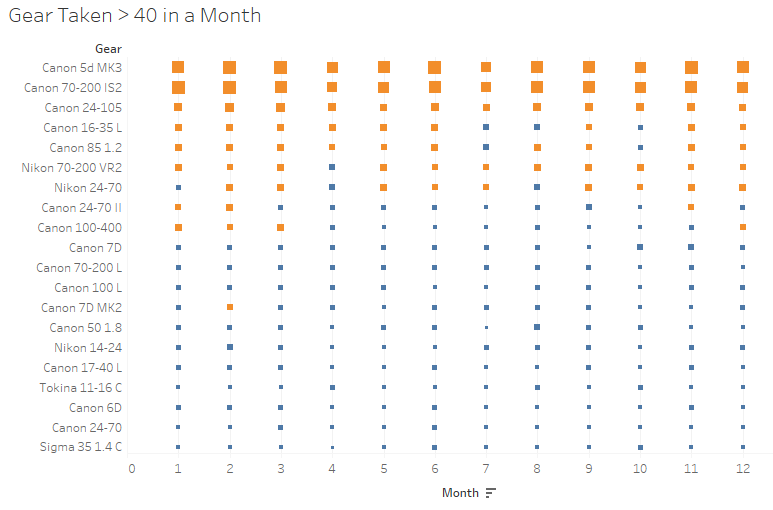


**Inference**: Three seasonal peaks per year in the months as January/February, July/August and October/November.

## Percentage of Revenue



**Inference**: For the Year 2015, 60% of revenue achieved in first six months. This is a pattern that can be observed year over year.

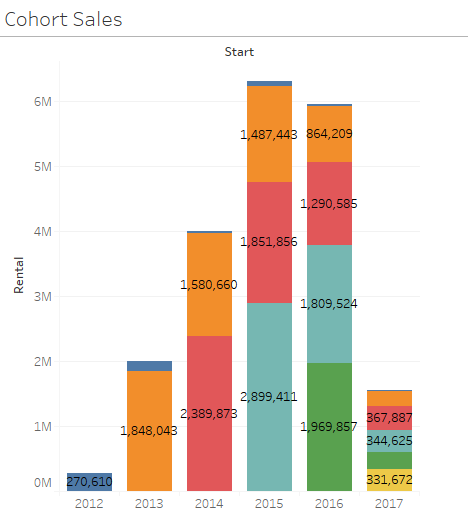


Seasonality and type of cameras rented confirms that most of the transactions were related to marriage events or other festivals. Also, Canon is contributing more for sales. Top5 are Canon Brand.

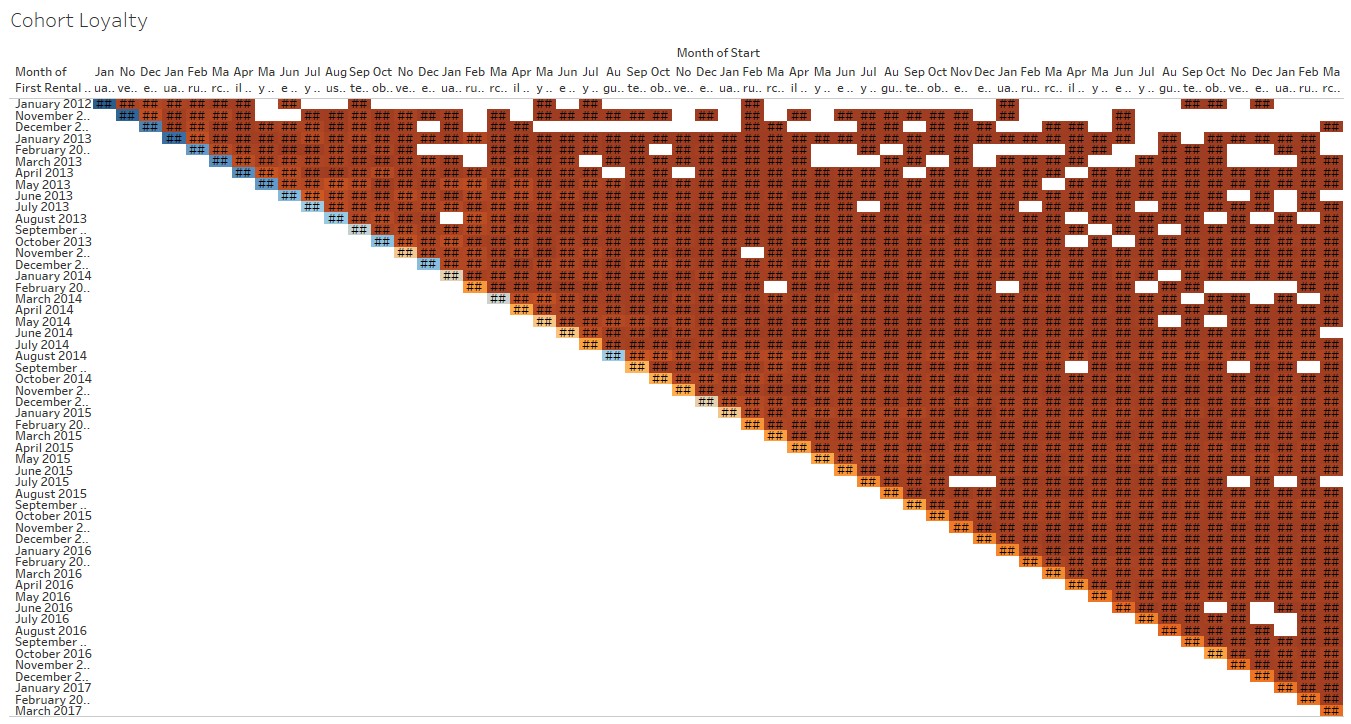
Recommendation:

1. **Replace cameras rented very few times like Nikon 200-400 VR2, Nikon 60 F4, Canon 1D MK4, etc. (extensive list provided separately) with frequently rented cameras like Cannon 5d MK3, Canon 70-200 IS2, Canon 24-105 (preferred mostly for marriage events) to increase revenue. This purely based on the Recommender that is to be built which will also consider the purpose of the rent.**
2. **Promotional exercise in the later part of the year between August to December will increase revenue during off-season**

## Cohort Analysis



Inference: Customer who rented in 2013 or 2014 rented frequently and they are most loyal ones



Inference: Very few new customers are renting in the last two years and high share of revenue generated by customers from initial years. Does that mean, people who become customers of Klachaks always stays as customer because of the value they see in it? It can be true, what can also be true is, Kalachaks is a near to monopoly in Chennai when it comes to Camera Rentals.

Recommendation

Recommending any advertisement/promotional exercises to bring new customers which is not done till now by Klachak. Also, Klachacks needs to be proactive and focus on retaining its customer base and also aggressively look for adding new customers to it Loyalty list.

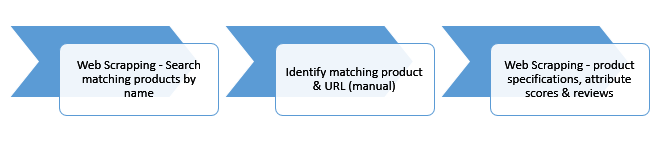
# Data Gathering Techniques

## Web Scrapping – Using Scrapy

Klachak inventory sources very limited products metadata information. Hence, we scrapped dprewview.com for full product specifications, attribute scorings, products reviews by dpreview.com and users.

Scrapped Product specifications and attribute scorings were used for building recommendation model and reviews are converted into scored using sentiment analysis.

## **Information Gathering Flow**



## **Scrapy Data Flow**



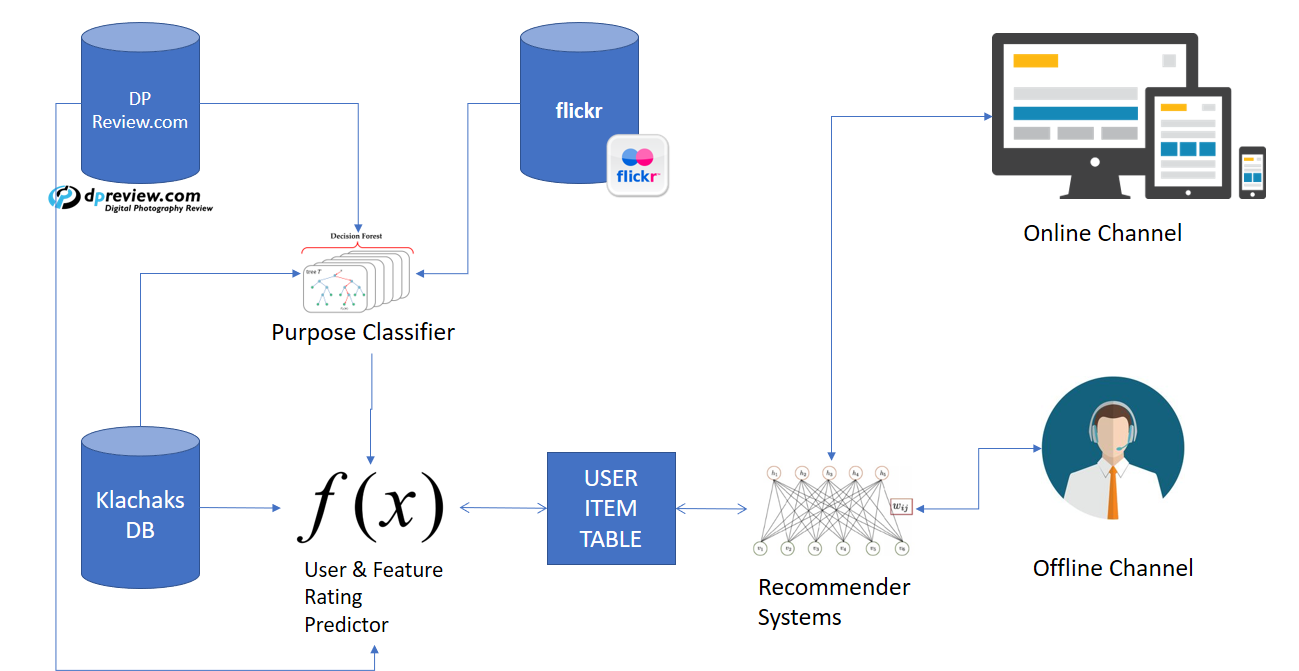
Data flow in Scrapy is controlled by the execution engine, and the steps are

1. The Engine gets the initial Requests to crawl from the Spider.
2. The Engine schedules the Requests in the Scheduler and asks for the next Requests to crawl.
3. The Scheduler returns the next Requests to the Engine.
4. The Engine sends the Requests to the Downloader, passing through the Downloader Middlewares.
5. Once the page finishes downloading, the Downloader generates a Response (with that page) and sends it to the Engine, passing through the Downloader Middlewares.
6. The Engine receives the Response from the Downloader and sends it to the Spider for processing, passing through the Spider Middleware.
7. The Spider processes the Response and returns scraped items and new Requests (to follow) to the Engine, passing through the Spider Middleware.
8. The Engine sends processed items to Item Pipelines, then send processed Requests to the Scheduler and asks for possible next Requests to crawl.
9. The process repeats (from step 1) until there are no more requests from the Scheduler.

# Recommendations & Applications

Based on the output of the studies, Klachak will implement our recommendation engine algorithm in there ordering channels and also revise the pricing structure.

# Overall Architecture



# Status

|  |  |  |
| --- | --- | --- |
| **Not Started** | **In-Progress** | **Completed** |
| Project Document | Purpose Classifier Model | Synopsis |
| Project Presentation | Data Qualification Test | Exploratory Data Analysis |
|  | User & Feature Rating Predictor Model | Model Identification |
|  | User Item Table | In-Term Report |
|  | Recommender Systems Algorithm |  |
|  | APIs |  |