Analyse User Behaviour & Optimise the User Workflow Using a Machine Learning Algorithm

PROJECT REPORT

1.ABSTRACT

The aim of the project is to Analyse user behaviour & optimise the user workflow using a machine learning algorithm. Firstly I have created an e-commerce website using Html,CSS,Javascript,Bootstrap and then deployed the site in Netlify Platform .

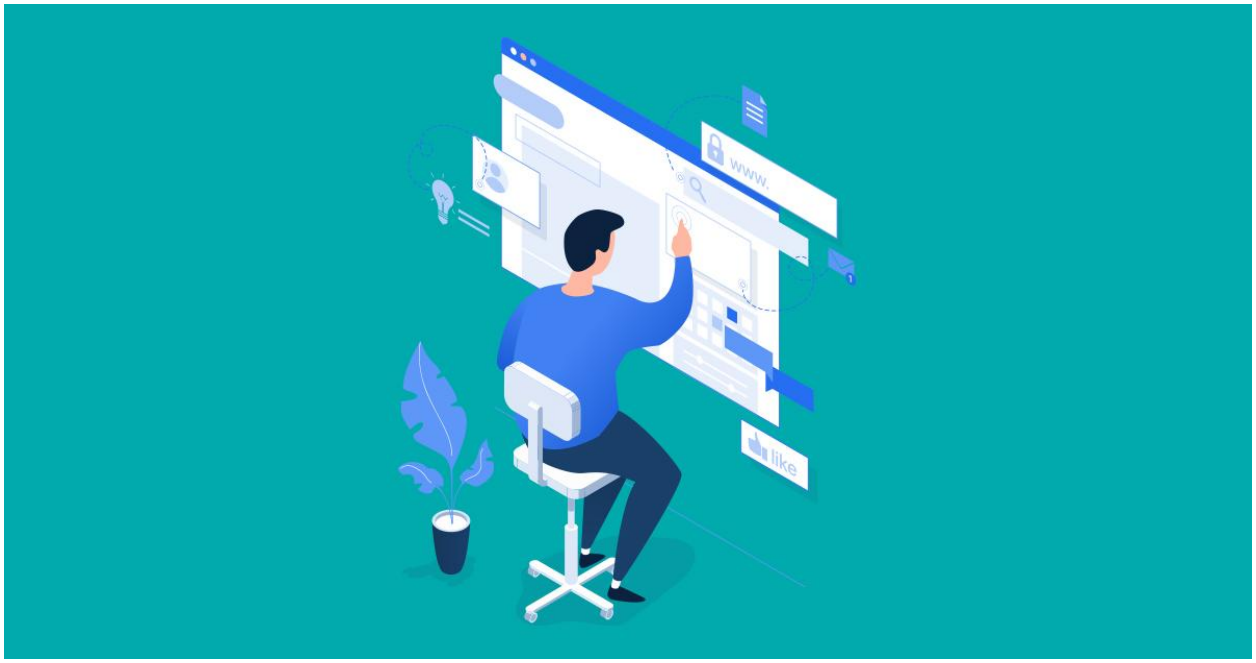
It contains various web pages such as home, store, account, about, blog,contact us. There are various features like product details, add to cart, search products and payment gateway. We have also used filters to change the product list (e.g. category, price range, etc.). I have applied a user guiding tool named Bootstrap Tour that will guide a new user through orderly steps on my website to use it in the most simple way to buy a product. I have integrated my website into an analytics platform named Google Analytics. This tool is able to track the users action and the time spent by the users on the various actions of the website. I have accumulated the analytics data from various users and stored it. I have collected data of around 75 users with various data points such as time spent by users on the different pages on the website, filters being used by the users , time spent by users on particular product pages, etc. Then, Fitting the Markov Chain model gives us the transition probabilities matrices and the lambda parameters of the chain for each one of the three lags along with the Start and End Probabilities.

2.INTRODUCTION

User behavior encompasses all the actions visitors take on a website: where and what they click on, how they scroll down a page, where they stumble, and where they eventually drop off the page and leave. Tracking user activity gives you an inside look at how people interact with your site and what obstacles or hooks they experience in their journey as your customers.

A website user is a person who is accessing, browsing or interacting with a website, and user behavior refers to how people use a website. Behaviors include everything from the journey they take through the site to interactions such as clicks. When it comes to optimising a website, simply monitoring behavior can only get you so far. The real value comes from analysing users' actions to get to the bottom of what makes them behave as they do. Behaviour is complex and varies across different websites depending on the target audience. This means you need to learn specifically about your users. Who are they? What are their needs? Which browsers and devices do they prefer? How often do they purchase? Answering these questions is crucial if

you want to have a competitive edge, meet consumer needs and retain your customers. By researching online behaviours, you can get an idea of what users are trying to achieve, the factors driving certain behaviours, where they experience friction and areas where user experience can be better. Ultimately, learning how visitors behave on your website allows you to provide an enhanced experience that's in line with user needs, which in turn will ensure your business continues to grow. One of the main tools we use to carry out research is Google Analytics. This is free and fairly easy to set up. Once installed, you'll have access to valuable data about how users behave on your site, including where they land and go next, where they drop off and what they interact with. You can also use it to discover overall trends and patterns and source opportunities for growth.



Users may follow multiple different paths and click sequences in accessing a site

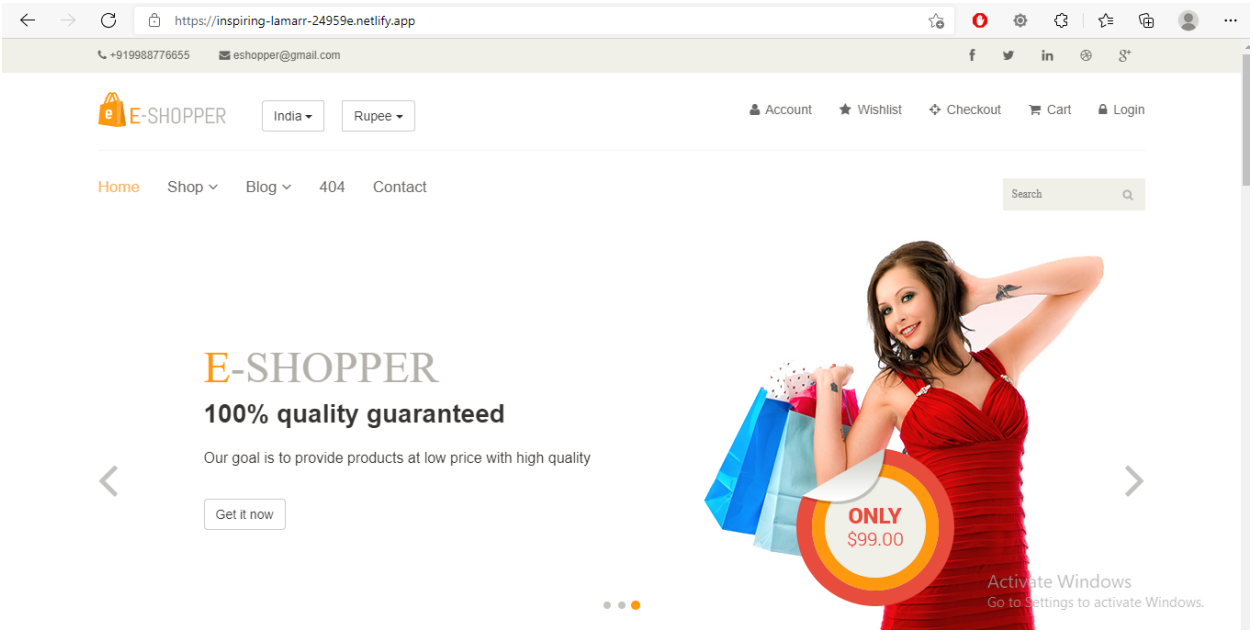
3.WEBSITE

E-commerce websites are online portals that facilitate online transactions of goods and services through means of the transfer of information and funds over the Internet. Now, with a single website, anything and everything that a transaction needs, can be executed online. I have used the following steps to create my E-commerce website: I have used Netlify platform for Web hosting. I deployed my web files in Netlify for my website link(Home | E-Shopper (inspiring-lamarr-24959e.netlify.app)).

My website contains following webpages:
a) **Home-** It is the homepage of my website.
b) **Store-** It contains various products categorized into different categories.

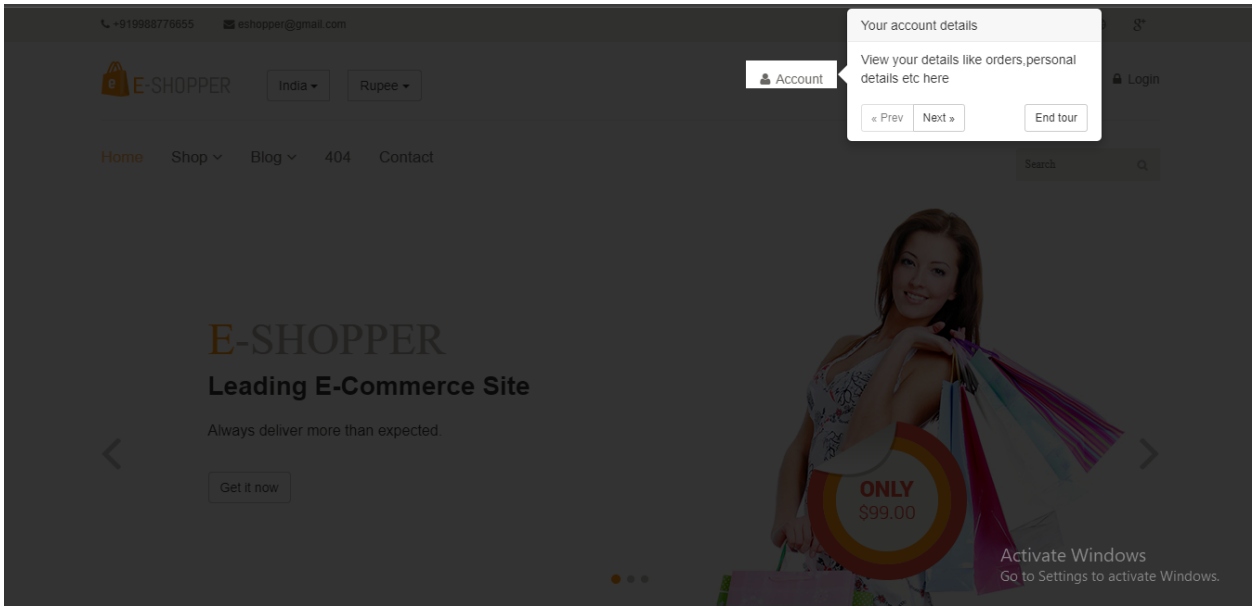
c) **Account-** It contains account details of the user.
d) **Contact Us**- It contains the contact details of the company.
e) **Blog**- It contains customer reviews and ratings for various products.
e) **About-** It contains team details.

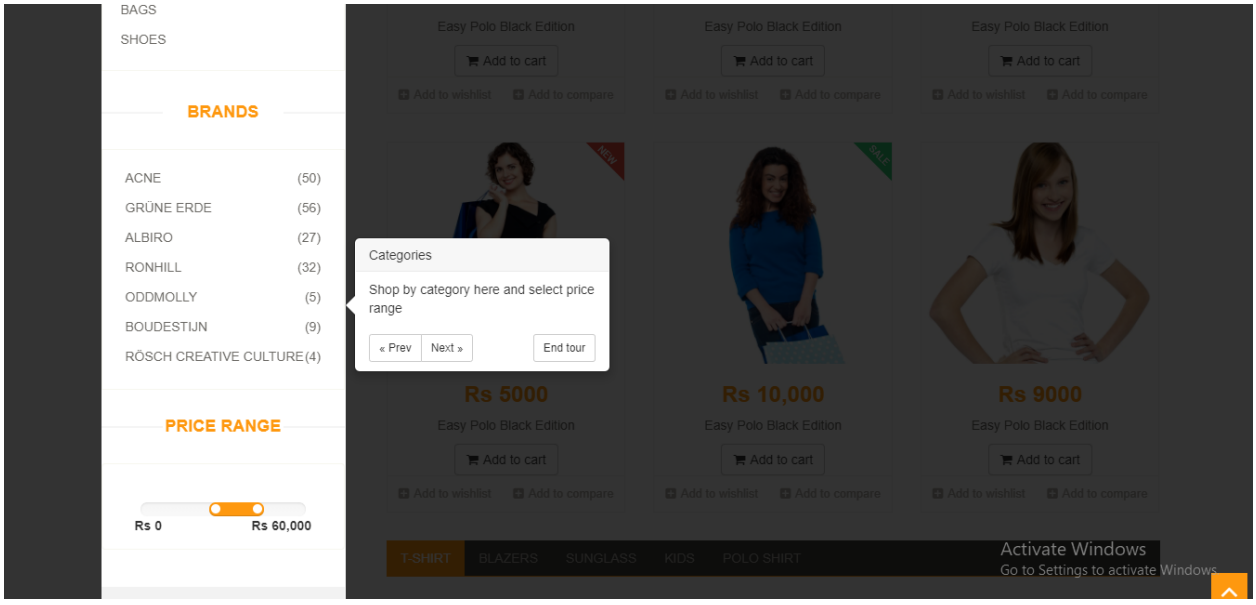
Website contains various features:
a) View a list of products
b) View product details
c) Search products
d) Use filters to change the product list (eg. Category, price range, etc.)
e) Add a product to the cart.



HOME PAGE OF THE ECOMMERCE WEBSITE

User Guide (Bootstrap Tour) - I have applied a user guiding tool named Bootstrap Tour that will guide a new user through orderly steps on my website to use it in the most simple way to buy a product.



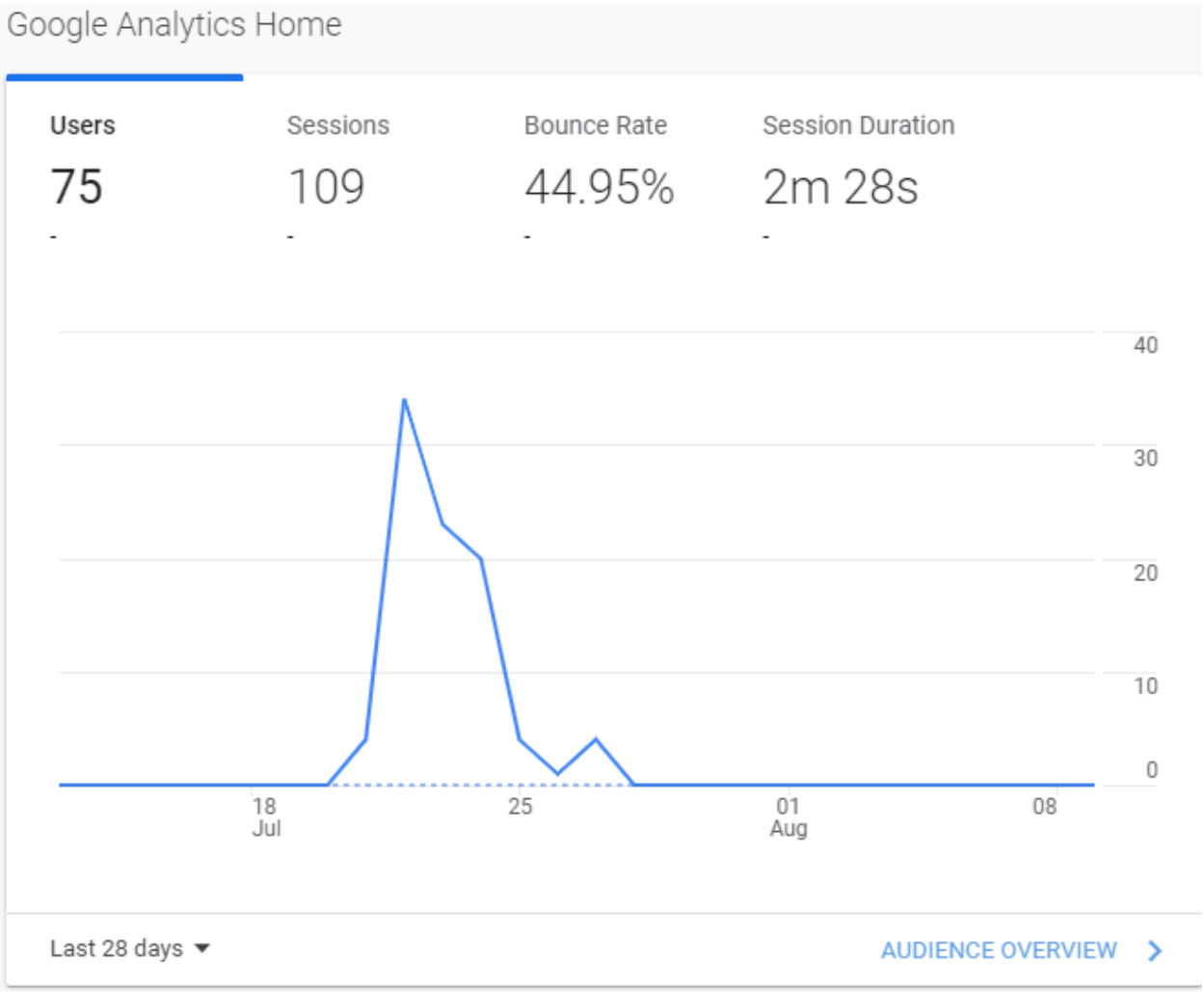


PICTURES DEPICTING GUIDE FOR USERS

4.DATA ANALYTICS

I have integrated my website into an analytics platform named Google Analytics. This tool is able to track the users action and the time spent by the users on the various actions of the website. I have accumulated the analytics data from various users and stored it.
I have collected data of around 75 users with various data points such as:

1) Time spent by users on the different pages on website,
2) Filters being used by the users ,
3)Time spent by users on particular product pages, etc.



SCREENSHOT OF GOOGLE ANALYTICS HOME PAGE(AUDIENCE OVERVIEW)

DATA TRANSFORMATION:

The clickstream data that is collected is often raw. It requires some refinement before it can be used to perform any analysis.

To transform the initially collected raw data into actual clickstream data, we need to identify the actions that are performed by an individual user and group them all together in an appropriate time order.

The dataset upon transformation will look like this which can be used for the analysis:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	User	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	1	A1	A2	A1	A2	A3	A3	A3	A3	A1	A3	A4	A3	A1	A2	A3	A3	A3	A4	A1	A3	A5	
3	2	A1	A4	A4	A2	A2	A2	A2	A2	A2	A2	A6	A2	A7	A2	A3							
4	3	A1	A8	A9	A1	A8	A3	A5	A1	A8	A9	A2	A6										
5	4	A1	A1	A1	A1																		
6	5	A1	A4	A1	A1																		
7	6	A1	A1																				
8	7	A1	A1	A1	A1	A1	A1																
9	8	A1	A2	A3	A5	A1	A3	A7	A6	A10	A10	A1											
10	9	A1	A1	A1	A4	A2	A9																
11	10	A1	A8	A9	A4	A2	A1	A4	A10	A8	A11												
12	11	A1	A8	A8	A11	A3	A2																
13	12	A1																					
14	13	A1																					
15	14	A1	A4																				
16	15	A1	A1																				
17	16	A1	A8	A9	A2	A3																	
18	17	A1	A8	A9	A10	A11	A12																
19	18	A1	A1	A2	A4																		
20	19	A1																					
21	20	A1																					
22	21	A1	A4	A10	A8	A8	A9	A9	A7	A1	A2	A8	A9	A7									
23	22	A1	A4	A1	A4	A4																	
24	23	A1																					
25	24	A1	A4	A11	A11	A11																	
26	25	A1	A4	A2	A3	A7																	

ACTIONS PERFORMED BY USERS

In the above dataset, each row corresponds to a user.The first column contains the user number, while the rest of the columns denote the actions (represented by A1, A2, A3, and so on) performed by the user.

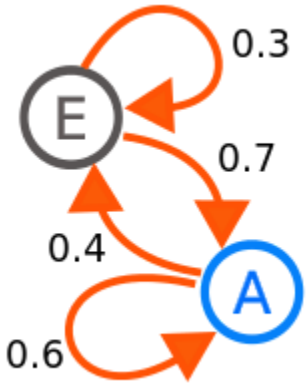
5.ML MODEL AND DATA ANALYSIS

Markov Chains

A Markov chain is a stochastic process that experiences transitions from one state to another according to certain probabilistic rules.Markov processes are stochastic processes because they have the Markov property, which means the next value of the Markov process depends on the current value, but it is conditionally independent of the previous values of the stochastic process.
Markov models are often used to model the probabilities of different states and the rates of transitions among them.
Markov models can also be used to recognize patterns, make predictions and to learn the statistics of sequential data.
The process takes the state from a finite set at each time .The order of a Markov Chain is derived from the number of recent states on which the current state, we assume, depends. Based on this, zero-order chains imply that the probability of being in a state in the next step is

independent of all previous states. Higher order Markov Chain introduced by the Raftery (1985) will lead to more realistic models. At the same time, the parameters needed for the representation increase exponentially and so it is important to find a right balance between these two.

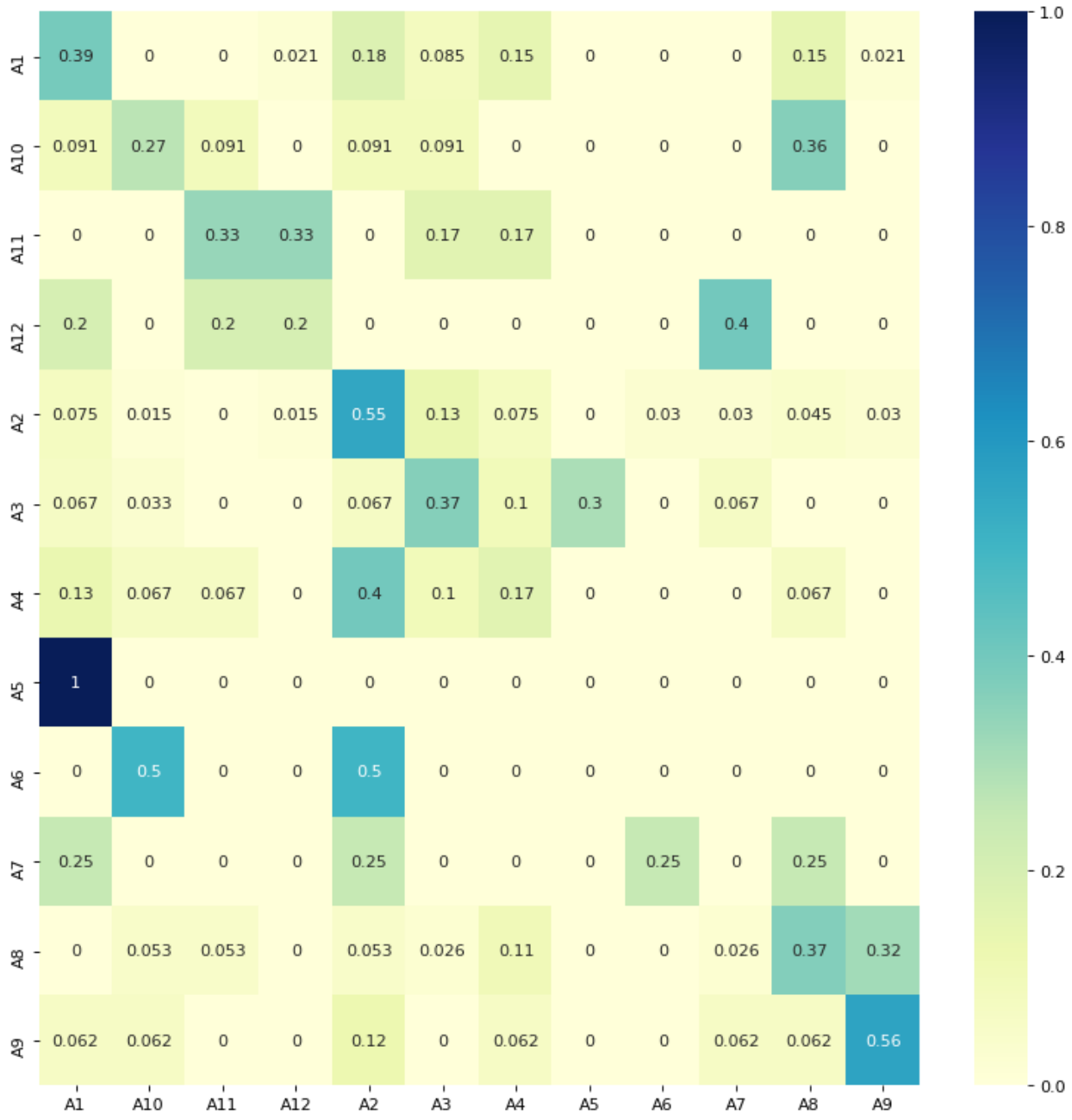
For example,



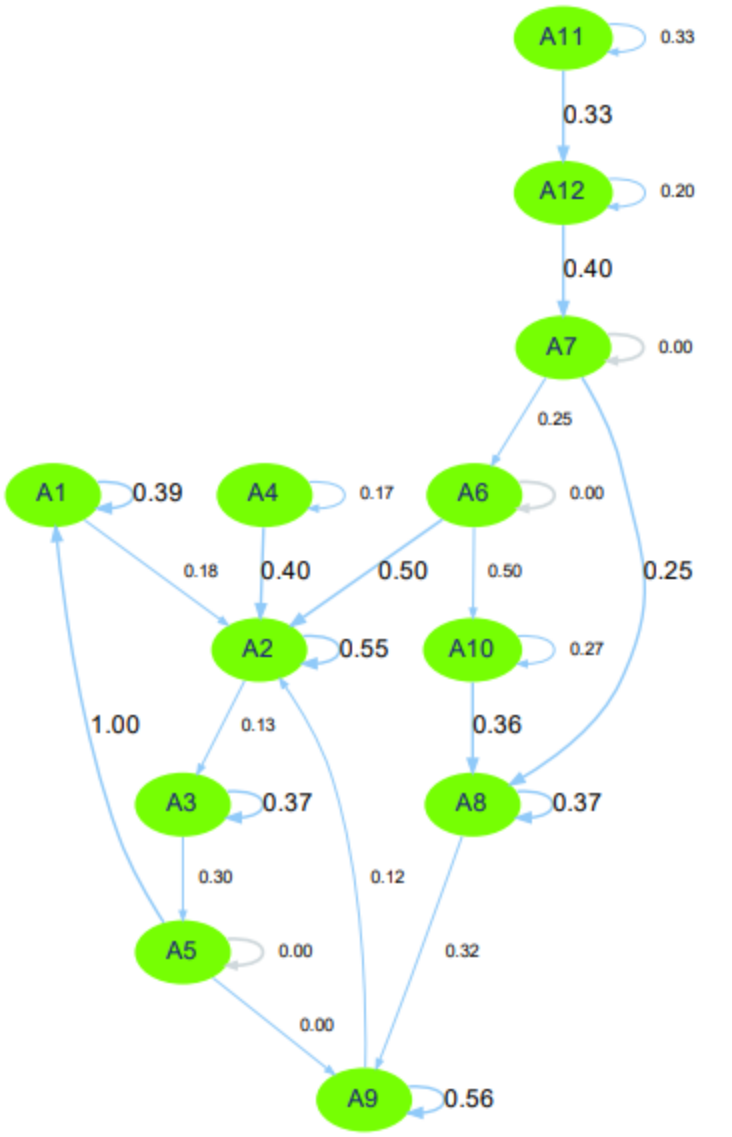
Consider the above diagram representing a two-state Markov process, with the states labelled E and A. Each number represents the probability of the Markov process changing from one state to another state, with the direction indicated by the arrow. For example, if the Markov process is in state A, then the probability it changes to state E is 0.4, while the probability it remains in state A is 0.6.

FITTING MODEL THE ABOVE DATASET

Fitting the Markov Chain model gives us the transition probabilities matrices and the lambda parameters of the chain for each one of the three lags along with the Start and End Probabilities. Start and End probabilities correspond to the probability that a clickstream will start or end with this specific event. The transition probability matrix can be represented as a heat map with the y-axis representing the current state and x-axis the next one. The more blu-ish the color, the more probable the indicated transition will occur.


A1 0.39 0 0 0.021 0.18 0.085 0.15 0 0 0 0.15 0.021
A10 0.091 0.27 0.091 0 0.091 0.091 0 0 0 0 0.36 0
A11 0 0 0.33 0.33 0 0.17 0.17 0 0 0 0 0
A12 0.2 0 0.2 0.2 0 0 0 0 0 0.4 0 0
A2 0.075 0.015 0 0.015 0.55 0.13 0.075 0 0.03 0.03 0.045 0.03
A3 0.067 0.033 0 0 0.067 0.37 0.1 0.3 0 0.067 0 0
A4 0.13 0.067 0.067 0 0.4 0.1 0.17 0 0 0 0.067 0
A5 1 0 0 0 0 0 0 0 0 0 0 0
A6 0 0.5 0 0 0.5 0 0 0 0 0 0 0
A7 0.25 0 0 0 0.25 0 0 0 0.25 0 0.25 0
A8 0 0.053 0.053 0 0.053 0.026 0.11 0 0 0.026 0.37 0.32
A9 0.062 0.062 0 0 0.12 0 0.062 0 0 0.062 0.062 0.56
A1 A10 A11 A12 A2 A3 A4 A5 A6 A7 A8 A9
1.0 0.8 0.6 0.4 0.2 0.0

HEAT MAP



TRANSITION DIAGRAM

ACTION DETAILS:

/	A1
/cart.html	A2
/checkout.html	A3
/login.html	A4
/payment.html	A5
/404.html	A6
/contact-us.html	A7
/shop.html	A8

/product-details.html	A9
/index.html	A10
/blog.html	A11
/blog-single.html	A12

Suggestions to UX Designer

The following are the insights from the Heat Map generated by the Markov chain model:

1. The transitions from Action 5 to Action 1,Action 9 to Action 9,Action 2 to Action 2,Action 6 to Action 10,Action 6 to Action 2 are more correlated to each other.

2.The transitions from Action 1 to Action 10,Action 3 to Action 6 are not correlated to each other.

Suggestion taken from transition diagram along with the corresponding probability are:

3.All users must start from A1(homepage)

4.Only 32% of the users are going from A8(shop.html) to A9(product-details.html),rest of the users are staying in the same page or quitting the application.So, the UX designer should guide the users to go to product-details.html or cart.html

5.Only 12% of the users are going from A9(product-details.html) to A2(cart.html) . Hence, the UX designer should guide them to go to cart.html.

6.The UX designer should guide the user to go from A2(cart.html) to A3(checkout.html) because only 13% of the users are going there.

7.From A3(checkout.html) only 30% of the users are going for payment i.e.,A5(payment.html),others are moving to other pages or quitting the application.Hence the UX designer must guide the users to place order and make payment.

8. 40% of the users are going from A12(blog.html) to A7(contact-us.html).The reason might be the users are not finding required information. Hence the UX designer should improve the quality of posts in the blog.

9.The users going to A6(404.html) from A7(contact-us.html) are 25%.The UX should find what's wrong in the contact-us.html page.

10.56% of the users are visiting the same page A9(product-details.html) again. The UX designer should guide the users to take further step in their shopping process. Similar is the case with A2(cart.html).

6)CONCLUSION

APPLICATION OF MARKOV CHAINS:

By fitting the Markov Chain model to the clickstream data i was able to get:

a) a transition probability matrix which was in the form of a Heat Map.
A heat map is a useful tool for representing the transition probability matrix.

b) a Markov Chain which can graphically represented as a transition diagram along with the corresponding probabilities,

IMPACT OF MARKOV CHAINS:

In clickstream analysis, it is very useful to predict where the customers are more likely to click next, given their previously followed patterns(through Clickstream data collected).

With Markov Chains we can analyse clickstream data and predict what the customer's next click could be.

Markov Chain Model is the machine learning model used here with which i was able to:

a)Know the transition probabilities(probability of transitioning from one state to another state) through heat map and transition diagram.

b)Know through the transition probabilities, which actions performed by the users are in the right way and which are causing them to deviate.

b)Finally, Suggest the UX designer the ways in which they can improve the website so that customers can ultimately buy a product without any deviation.