

E-commerce Customer Attitude predictions with Naive Bayes Machine Learning algorithm

¹Nagarjun Yadav Vanguri, *Research Scholar, CSE, Annamalai University Tamilnadu India*
Nagarjuna.v.yadav@gmail.com

²Dr.S.Pazhanirajan, *Assist.Professor CSE, Annamalai University Tamilnadu India*
pazhanisambandam@gmail.com

³Dr.T Anil Kumar, *Professor HOD(EEE) Anurag University, Telangana, India*
thalluruanil@gmail.com

⁴K. Venkata Rao, *Director, Reachout Analytics Pvt Ltd, Hyderabad, India*
venkat@reachoutanalytics.com

⁵Avvari Sai Saketh, *DATA ANALYST INTERN, Reachout Analytics Pvt Ltd, Hyderabad, India*
sakethprasad111@gmail.com

Abstract:

E-commerce (electronic commerce or EC) is the buying and selling of goods and services, or the transmitting of funds or data, over an electronic network, primarily the Internet. The main purpose of this study is to understand the factors that may influence customer attitude and behaviors towards E-commerce. The motivation is reflected to the customer's desire to meet their own needs. Learning is reflected to the customers' behavior experience arising. Attitudes are reflected to customers' steadily favorable or unfavorable assessments, feelings, and inclinations towards object or idea. Finally, Beliefs is reflected to customers' thoughts about a product or service (Kotler & Armstrong, 1997). The main objective of this study is to investigate the factors influencing the customer attitude towards E-commerce and how demographic variable affect customer attitude towards Ecommerce.

Keywords: E-commerce, Navie Bayes Confusion matrices Customer Perception, Purchase intention, Customer attitude& Customer experience,

1. Introduction

The terms "Electronic Commerce", "Internet marketing" and "on-line shopping" are now commonly used by business executives and consumers throughout the world as businesses are recognizing the potential opportunities for commerce in the on-line business environment (Karakaya and Charlton, 2001). At its core, e-commerce refers to the purchase and sale of goods and/or services via electronic channels such as the Internet. E-commerce was first introduced in the 1960s via an electronic data interchange (EDI) on value-added networks (VANs). The medium grew with the increased availability of Internet access and the advent of popular online sellers in the 1990s and early 2000s. Amazon began operating as a book-shipping business in Jeff Bezos' garage in 1995. EBay, which enables consumers to sell to each other online, introduced online auctions in 1995 and exploded with the 1997 Beanie Babies frenzy. Rising incomes and a greater variety of goods and services that can be bought over the internet is making buying online more attractive and convenient for consumers all over the country.

2. Literature

Review E-commerce is a tool for reducing administrative costs and cycle time, streamlining business processes, and improving relationships with both business partners and customers (Charles, 1998). The Internet and electronic commerce were the two most significant development of information during 1990s. There has been a marked increase in the number of consumers who purchase over the Internet,

as well as an increase in sales worldwide conducted via electronic commerce. Innovation and electronic commerce relationships have resulted in tremendous changes in market competition among various industries (Blosch, 2000; Hamid & Kassim, 2004). Customer purchasing decisions are influenced by perception, motivation, learning, attitudes and beliefs. The perception is reflected to on how the customers select, organize, and interpret information to form knowledge. The motivation is reflected to the customer's desire to meet their own needs. Learning is reflected to the customers' behavior experience arising. Attitudes are reflected to customers' steadily favorable or unfavorable assessments, feelings, and inclinations towards object or idea. Finally, Beliefs is reflected to customers' thoughts about a product or service (Kotler & Armstrong, 1997)

3. Objectives

This study is to identify the attitude of customer towards E Commerce Business. To investigate how demographic variable (age, income and occupation) affect customer buying behavior and E Commerce. To examine the factors those affect the customer buying behavior (perception, motivation, learning, attitudes and beliefs) and E Commerce.

4. Methodology:

For Analyzing the above objectives and hypothesis, Data mining Tools and Techniques i.e *R programming, R Rattle, WEKA and SPSS 20* software's were used for Data validating, summarizing, visualizing, Modeling and Testing. To measure these objectives a cross sectional descriptive study was designed. Questionnaire Design: Based on the research objectives, a structured questionnaire with 30 variables, mainly with a 5-point Likert scale was used, in which 1 = strongly disagree and 5 = strongly agree.

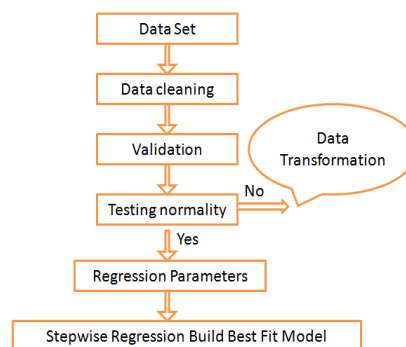
5. Data acquisition

In this research, the survey method was selected to understand the factor affecting customer attitude towards E-commerce applications in India. For data collection, random sampling was adopted. To ensure all questions being answered in a proper way, questionnaires were completed and screened one-by-one. Data Accessing and Cleaning: The data collected through survey is saved in csv format. This data is converted into data table format. Data cleaning and validation process is carried out using R Programming language and Weka Language.

6. Data Validation & Cleaning

A survey is conducted on 1500 customers with a sample error of 5%. After a pilot study we consider 727 to be a good sample. The data is cleaned, validated and reliability process is performed using R Programming language and Weka Programming Language. We have implemented below Data mining machine learning algorithm techniques:

7. Modeling process Plan



. Data mining Tools and Techniques i.e. WEKA, R programming, R Rattle and R Rcmdr have been used to analyze the data. Multinomial Logistics model is applied Machine Learning techniques PCA applied for dimension reduction.

Originally there were 30 factor variables which were reduced to 7 after the technique is applied. Now all the irrelevant variables are removed and the analysis is carried here afterwards with the reduced 7 variables.

8. Model Naive Bayes Algorithm confusion matrices:

Dimension reduction technique like factor analysis is applied to reduce the dimensions of the data without loss of the data. 30 factor variables are reduced to 7. Irrelevant variables are removed and then Logistics Multinomial model is applied. Models are obtained for the 7 demographic variables that are affecting the e commerce. A detailed study is done on each model and conclusions are made about the factors that are influencing consumer buying behavior. All the models are evaluated by the confusion matrix and the respective model diagnostics for each variable which are shown below.

8.1 Customer Marital Status:

	Test Data	Train Data
Sensitivity	0.8023	0.7801
Specificity	0.5	0.45
Pos Pred Value	0.697	0.6683
Neg Pred Value	0.6383	0.5902
Prevalence	0.589	0.5869
Detection Rate	0.4726	0.4578
Detection Prevalence	0.6781	0.685
Balanced Accuracy	0.6512	0.615

	Marital Status Group	
	Test Data	Train Data
Accuracy	0.6781	0.6437
95% CI	(0.5958, 0.7529)	(0.6033, 0.6827)
No Information Rate	0.589	0.5869
P-Value [Acc > NIR]	0.01681	0.002915
Kappa	0.3126	0.2384
Mcnemar's Test P-Value	0.08005	9.93E-05

For the Marital status group, we can see from the above classification table that, the metrics for the test were Sensitivity (80%), Specificity (50%), Positive predictive value (70%), Negative Predictive value (64%), and Balanced Accuracy (65%). The Accuracy of Test data is 68%. The metrics for the training set were Sensitivity (78%), Specificity (45%), Positive predictive value (67%), Negative Predictive value (59%), and Balanced Accuracy (62%). The Overall Accuracy of Training data is 64%.

8.2 Customer Gender:

	Gender Group	
	Test Data	Train Data
Accuracy	0.6849	0.6162
95% CI	(0.6029, 0.7592)	(0.5753, 0.6559)
No Information Rate	0.5137	0.5198
P-Value [Acc > NIR]	2.01E-05	1.80E-06
Kappa	0.376	0.22
Mcnemar's Test P-Value	6.87E-05	0

	Test Data	Train Data
Sensitivity	0.5067	0.4122
Specificity	0.8732	0.8046
Pos Pred Value	0.8085	0.6609
Neg Pred Value	0.6263	0.5971
Prevalence	0.5137	0.4802
Detection Rate	0.2603	0.1979
Detection Prevalence	0.3219	0.2995
Balanced Accuracy	0.69	0.6084

For the Gender group, from the above table it can be drawn that, the metrics for the test set were Sensitivity (51%), Specificity (87%), Positive predictive value (81%), Negative Predictive value (63%), and Balanced Accuracy (69%). The Overall Accuracy of Test data is 68%. The metrics for the training set were Sensitivity (41%), Specificity (80%), Positive predictive value (66%), Negative Predictive value (60%), and Balanced Accuracy (61%). The Overall Accuracy of Training data is 62%.

8.3 Customer Region:

	Region Group	
	Test Data	Train Data
Accuracy	0.8219	0.8485
95% CI	(0.7501, 0.8802)	(0.8168, 0.8767)
No Information Rate	0.3699	0.3029
P-Value [Acc > NIR]	< 2.2e-16	< 2.2e-16
Kappa	0.758	0.7963
McNemar's Test P-Value	NA	NA

	Test Data					Training Data				
	Metro	Rural	Town	Urban	Village	Metro	Rural	Town	Urban	Village
Sensitivity	0.9268	NA	0.51852	0.8333	0.9583	0.9277	0	0.623	0.892	0.9138
Specificity	0.9524	1	0.98319	1	0.8443	0.9687	1	0.9847	0.9654	0.8839
+Pred Value	0.8837	NA	0.875	1	0.5476	0.9222	NaN	0.9157	0.9181	0.6625
- Pred Value	0.9709	NA	0.9	0.9109	0.9904	0.971	0.998279	0.9076	0.9537	0.9762
Prevalence	0.2808	0	0.18493	0.3699	0.1644	0.2857	0.001721	0.21	0.3029	0.1997
Detection	0.2603	0	0.09589	0.3082	0.1575	0.2651	0	0.1308	0.2702	0.1824
Prevalence	0.2945	0	0.10959	0.3082	0.2877	0.2874	0	0.1429	0.2943	0.2754
Balanced	0.9396	NA	0.75086	0.9167	0.9482	0.9482	0.5	0.8039	0.9287	0.8988

For the Region group, it is evident from the above table that, under Metro, the metrics for the test set were Sensitivity (93%), Specificity (95%), Positive predictive value (88%), Negative Predictive value (97%), and Balanced Accuracy (94%). Under Town, the metrics for the test set were Sensitivity (52%), Specificity (98%), Positive predictive value (88%), Negative Predictive value (90%), and Balanced Accuracy (75%). Under Urban, the metrics for the test set were Sensitivity (83%), Specificity (100%), Positive predictive value (100%), Negative Predictive value (91%), and Balanced Accuracy (92%). Under Village, the metrics for the test set were Sensitivity (96%), Specificity (84%), Positive predictive value (55%), Negative Predictive value (99%), and Balanced Accuracy (90%). The Overall Accuracy of Test data is 82%.

For the Region group, it is evident from the above table that, under the Metro, the metrics for the training set were Sensitivity (93%), Specificity (97%), Positive predictive value (92%), Negative Predictive value (97%), and Balanced Accuracy (95%). Under Town, the metrics for the training set were Sensitivity (62%), Specificity (98%), Positive

predictive value (92%), Negative Predictive value (91%), and Balanced Accuracy (80%). Under Urban, the metrics for the training set were Sensitivity (89%), Specificity (97%), Positive predictive value (92%), Negative Predictive value (95%), and Balanced Accuracy (93%). Under Village, the metrics for the training set were Sensitivity (91%), Specificity (88%), Positive predictive value (66%), Negative Predictive value (98%), and Balanced Accuracy (90%). The Overall Accuracy of Training data is 85%.

8.4 Customer Income Group:

	Income Group	
	Test Data	Train Data
Accuracy	0.8699	0.8675
95% CI	(0.8043, 0.9198)	(0.8372, 0.894)
No Information Rate	0.3699	0.3391
P-Value [Acc > NIR]	<2e-16	< 2e-16
Kappa	0.8186	0.8191
McNemar's Test P-Value	0.2147	0.01274

	Test Data				Train Data			
	Above 60,000	30,000 to 40,000	40,000 to 60,000	60,000 to 80,000	60,000	30,000 to 40,000	40,000 to 60,000	60,000 to 80,000
Sensitivity	0.9111	0.7826	0.875	0.8704	0.8477	0.8879	0.7941	0.92
Specificity	0.9307	0.9675	0.9672	0.9565	0.9714	0.9536	0.9582	0.9409
+ Pred Value	0.8542	0.8182	0.84	0.9216	0.9382	0.812	0.802	0.8703
-Pred alue	0.9592	0.9597	0.9752	0.9263	0.9256	0.9741	0.9563	0.9646
Prevalence	0.3082	0.1575	0.1644	0.3699	0.3391	0.1842	0.1756	0.3012
Detection	0.2808	0.1233	0.1438	0.3219	0.2874	0.1635	0.1394	0.2771
Prevalence	0.3288	0.1507	0.1712	0.3493	0.3064	0.2014	0.1738	0.3184
Balanced	0.9209	0.875	0.9211	0.9134	0.9095	0.9207	0.8762	0.9304

Under Income group, it can be concluded from the above table that, for the income group (Above 80,000), the metrics for the test set were Sensitivity (91%), Specificity (93%), Positive predictive value (85%), Negative Predictive value (96%) and Balanced Accuracy (92%). For Income group (Bet 30,000 to 40,000) the metrics for the test set were Sensitivity (78%), Specificity (97%), Positive predictive value (82%), Negative Predictive value (96%), and Balanced Accuracy (88%). For Income group (Bet 40,000 to 60,000) the metrics for the test set were Sensitivity (88%), Specificity (97%), Positive predictive value (84%), Negative Predictive value (98%), and Balanced Accuracy (92%). For Income group (Bet 60,000 to 80,000) the metrics for the test set were Sensitivity (87%), Specificity (96%), Positive predictive value (92%), Negative Predictive value (93%), and Balanced Accuracy (91%). The Overall Accuracy of Test data is 87%.

Under Income group, can be concluded from the above table that, For Income group (Above 80,000), the metrics for the training set were Sensitivity (85%), Specificity (97%), Positive predictive value (94%), Negative Predictive value (93%), and Balanced Accuracy (91%). For Income group (Bet 30,000 to 40,000), the metrics for the training set were Sensitivity (89%), Specificity (95%), Positive predictive value (81%), Negative Predictive value (97%), and Balanced Accuracy (92%). For Income group (Bet 40,000 to 60,000), the metrics for the training set were Sensitivity (79%), Specificity (96%), Positive predictive value (80%), Negative Predictive value (96%), and Balanced Accuracy (88%). For Income group (Bet 60,000 to 80,000), the metrics for the training set were Sensitivity (92%), Specificity (94%), Positive predictive value (87%), Negative Predictive value (96%), and Balanced Accuracy (93%). The Overall Accuracy of Training data is 87%.

8.5 Customer Income Group:

	Income Group	
	Test Data	Train Data
Accuracy	0.8699	0.8675
95% CI	(0.8043, 0.9198)	(0.8372, 0.894)
No Information Rate	0.3699	0.3391
P-Value [Acc > NIR]	<2e-16	< 2e-16
Kappa	0.8186	0.8191
McNemar's Test P-Value	0.2147	0.01274

	Test Data				Train Data			
	Employed	Home Maker	Professional	Self-Employee	Employed	Home Maker	Professional	Self-Employee
Sensitivity	0.45	0.81	0.71	0.45	0.52	0.9	0.74	0.59
Specificity	0.85	0.79	0.84	0.94	0.91	0.86	0.83	0.93
+ Pred	0.62	0.47	0.59	0.71	0.77	0.54	0.6	0.71
- Pred alue	0.74	0.95	0.90	0.8	0.78	0.97	0.888	0.90
Prevalence	0.34	0.18	0.23	0.2	0.34	0.15	0.29	0.20
Detection	0.15	0.15	0.17	0.10	0.18	0.13	0.21	0.12
Prevalence	0.25	0.31	0.28	0.14	0.23	0.25	0.33	0.17
Balanced	0.65	0.80	0.78	0.70	0.72	0.88	0.78	0.76

Under Income group, it can be concluded from the above table that, for the income group (Above 80,000), the metrics for the test set were Sensitivity (91%), Specificity (93%), Positive predictive value (85%), Negative Predictive value (96%) and Balanced Accuracy (92%). For Income group (Bet 30,000 to 40,000) the metrics for the test set were Sensitivity (78%), Specificity (97%), Positive predictive value (82%), Negative Predictive value (96%), and Balanced Accuracy (88%). For Income group (Bet 40,000 to 60,000) the metrics for the test set were Sensitivity (88%), Specificity (97%), Positive predictive value (84%), Negative Predictive value (98%), and Balanced Accuracy (92%). For Income group (Bet 60,000 to 80,000) the metrics for the test set were Sensitivity (87%), Specificity (96%), Positive predictive value (92%), Negative Predictive value (93%), and Balanced Accuracy (91%). The Overall Accuracy of Test data is 87%.

8.6 Customer Occupation

	Education	
	Test Data	Train Data
Accuracy	0.7466	0.7748
95% CI	(0.668, 0.8149)	(0.7383, 0.8079)
No Information Rate	0.3151	0.2651
P-Value [Acc > NIR]	<2.2e-16	<2.2e-16
Kappa	0.66	0.6999
McNemar's Test P-Value	NA	2.82E-10

	Test Data				Train Data			
	Employed	Home Maker	Professional	Self-Employee	Employed	Home Maker	Professional	Self-Employee
Sensitivity	0.45	0.81	0.71	0.454	0.52	0.9	0.7	0.59
Specificity	0.85	0.79	0.84	0.94	0.91	0.86	0.83	0.93
+ Pred	0.62	0.47	0.59	0.71	0.77	0.54	0.642	0.71
- Pred alue	0.74	0.95	0.90	0.8	0.7	0.97	0.88	0.90
Prevalence	0.34	0.184	0.23	0.2	0.349	0.15	0.29	0.20
Detection	0.15	0.15	0.17	0.1	0.18	0.13	0.21	0.12
Prevalence	0.25	0.3	0.28	0.14	0.23	0.25	0.33	0.17
Accuracy	0.65	0.80	0.78	0.70	0.72	0.88	0.78	0.7

Under the Occupation group, the above table delivers that, For Employment group (Employed), the metrics for the test set were Sensitivity (45%), Specificity (85%), Positive predictive value (62%), Negative Predictive value (74%), and Balanced Accuracy (65%). For Employment group (Home Maker), the metrics for the test set were Sensitivity (81%), Specificity (80%), Positive predictive value (48%), Negative Predictive value (95%), Balanced Accuracy (81%). For Employment group (Professional), the metrics for the test set were Sensitivity (71%), Specificity (85%), Positive predictive value (60%), Negative Predictive value (90%), and Balanced Accuracy (78%). For Employment group (Self Employed), the metrics for the test set were Sensitivity (45%), Specificity (95%), Positive predictive value (71%), Negative Predictive value (86%), Balanced Accuracy (70%).

Under the Occupation group, the above table delivers that, For Employment group (Employed), the metrics for the training set were Sensitivity (52%), Specificity (92%), Positive predictive value (77%), Negative Predictive value (78%), and Balanced Accuracy (72%). For Employment group (Home Maker), the metrics for the training set were Sensitivity (90%), Specificity (86%), Positive predictive value (55%), Negative Predictive value (98%) and Balanced Accuracy (88%). For Employment group (Professional), the metrics for the training set were Sensitivity (75%), Specificity (83%), Positive predictive value (64%), Negative Predictive value (89%), and Balanced Accuracy (79%). For Employment group (Self Employed), the metrics for the training set were Sensitivity (60%), Specificity (94%), Positive predictive value (71%), Negative Predictive value (90%), Balanced Accuracy (77%). The Overall Accuracy of Training data is 66%.

8.7 Customer Education

	Education	
	Test Data	Train Data
Accuracy	0.7466	0.7748
95% CI	(0.668, 0.8149)	(0.7383, 0.8079)
No Information Rate	0.3151	0.2651
P-Value [Acc > NIR]	<2.2e-16	<2.2e-16
Kappa	0.66	0.6999
Mcneamar's Test P-Value	NA	2.82E-10

	Test Data				Train Data			
	Graduation	Inter	PG	other	UG	Inter	PG	other
Sensitivity	0.34375	0.6875	0.8696	1	0.39189	0.8254	0.8954	0.9805
Specificity	0.90351	0.8246	0.97	0.9727	0.94226	0.8462	0.9533	0.9625
+ Pred lue	0.5	0.5238	0.9302	0.9231	0.6988	0.5977	0.8726	0.9042
Pred alue	0.83065	0.9038	0.9417	1	0.81928	0.9459	0.9623	0.9928
Prevalence	0.21918	0.2192	0.3151	0.2466	0.25473	0.2169	0.2633	0.2651
Detection	0.07534	0.1507	0.274	0.2466	0.09983	0.179	0.2358	0.2599
Prevalence	0.15068	0.2877	0.2945	0.2671	0.14286	0.2995	0.2702	0.2874
Accuracy	0.62363	0.756	0.9198	0.9864	0.66708	0.8358	0.9243	0.9715

Under the Education group, For Education group (Graduation), the metrics for the test set were Sensitivity (34%), Specificity (90%), Positive predictive value (50%), Negative Predictive value (83%) and Balanced Accuracy (62%). For Education group (Intermediate/10+2), the metrics for the test set were Sensitivity (69%), Specificity (82%), Positive predictive value (52%), Negative Predictive value (90%), Balanced Accuracy (76%). For Education group (Post Graduation), the metrics for the test set were Sensitivity (87%), Specificity (97%), Positive predictive value (93%), Negative Predictive value (94%) and Balanced Accuracy (92%). For Education group (Professional Degree), the metrics for the test set were Sensitivity (100%), Specificity (97%), Positive predictive value (92%), Negative Predictive value (100%) and Balanced Accuracy (99%). The Overall Accuracy of Test data is 75%.

Under the Education group, For Education group (Graduation), the metrics for the training set were Sensitivity (39%), Specificity (94%), Positive predictive value (70%), Negative Predictive value (82%) and Balanced Accuracy (67%). For Education group (Intermediate/10+2), the metrics for the training set were Sensitivity (83%), Specificity (85%), Positive predictive value (60%), Negative Predictive value (95%) and Balanced Accuracy (84%). For Education group (Post Graduation), the), the metrics for the training set were Sensitivity (90%), Specificity (95%), Positive predictive value (87%), Negative Predictive value (96%) and Balanced Accuracy (92%). For Education group (Professional Degree), the metrics for the training set were Sensitivity (98%), Specificity (96%), Positive predictive value (90%), Negative Predictive value (99%) and Balanced Accuracy (97%).

8.8 Customer Age Group:

	Age Group	
	Test Data	Train Data
Accuracy	0.8699	0.8675
95% CI	(0.8043, 0.9198)	(0.8372, 0.894)
No Information Rate	0.3699	0.3391
P-Value [Acc > NIR]	<2e-16	< 2e-16
Kappa	0.8186	0.8191
McNemar's Test P-Value	0.2147	0.01274

	Test Data					Train Data				
	26-35	36-45	46-55	>56	<25	26-35	36-45	46-55	Above 56	Below 25
Sensitivity	0.46	0.18	0.24	0.63	1	0.48	0.29688	0.40426	0.688	0.88889
Specificity	0.88	0.85	0.99	0.91	0.62	0.92	0.87404	0.96818	0.9364	0.66783
+ Pred alue	0.5	0.34	0.90	0.67	0.01	0.60	0.53774	0.80282	0.7478	0.0404
- Pred alue	0.875	0.71	0.77	0.89	1	0.87	0.71579	0.83529	0.9163	0.99739
Prevalence	0.19	0.29	0.28	0.22	0.07	0.19	0.33046	0.24269	0.2151	0.01549
Detection	0.08	0.05	0.06	0.14	0.07	0.09	0.09811	0.09811	0.148	0.01377
Prevalence	0.17	0.15	0.07	0.21	0.37	0.15	0.18244	0.1222	0.1979	0.34079
Accuracy	0.67	0.52	0.61	0.76	0.81	0.70	0.58546	0.68622	0.8122	0.77836

Under the Age group, For Age group (26-35), the metrics for the test set were Sensitivity (46%), Specificity (89%), Positive predictive value (50%), Negative Predictive value (88%) and Balanced Accuracy (68%). For Age group (36-45), the metrics for the test set were Sensitivity (19%), Specificity (85%), Positive predictive value (35%), Negative Predictive value (72%) and Balanced Accuracy (52%). For Age group (46-55), the metrics for the test set were Sensitivity (24%), Specificity (99%), Positive predictive value (91%), Negative Predictive value (77%), Balanced Accuracy (62%). For Age group (Above 56), the metrics for the test set were Sensitivity (64%), Specificity (91%), Positive predictive value (68%), Negative Predictive value (90%) and Balanced Accuracy (77%). For Age group (Below 25), the metrics for the test set were Sensitivity (100%), Specificity (63%), Positive predictive value (2%), Negative Predictive value (100%) and Balanced Accuracy (81%). The Overall Accuracy of Test data is 36%.

Under the Age group, For Age group (26-35), the metrics for the training set were Sensitivity (48%), Specificity (92%), Positive predictive value (60%), Negative Predictive value (88%) and Balanced Accuracy (70%). For Age group (36-45), the metrics for the training set were Sensitivity (30%), Specificity (87%), Positive predictive value (54%), Negative Predictive value (72%), and Balanced Accuracy (59%). For Age group (46-55), the metrics for the training set were Sensitivity (40%), Specificity (97%), Positive predictive value (80%), Negative Predictive value (84%) and Balanced Accuracy (69%). For Age group (Above 56), the metrics for the training set were Sensitivity (69%), Specificity (94%), Positive predictive value (75%), Negative Predictive value (92%), and Balanced Accuracy (81%). For Age group (Below 25), the metrics for the training set were Sensitivity (89%), Specificity (67%), Positive predictive value (4%), Negative Predictive value (100%), and Balanced Accuracy (78%). The Overall Accuracy of Training data is 45%.

9. References

1. Akinci, S., Aksoy, S., & Atilgan, E. (2004). Adoption of internet banking among sophisticated consumer segments in an advanced developing country. *International Journal of Bank Marketing*, 22 (3), 212-232.
2. Bakos, J.Y. (1997), "Reducing buyer search costs: implications for electronic marketplaces", *Management Science*, 43(12), pp. 1676-1692.
3. Chellappa, K.R and Pavlou A.P. (2002), "Perceived information security, financial liability and consumer trust in electronic commerce transactions", *Logistics Information Management*, 15(5), pp. 358-368. 21.
4. Chiles, T.H. and McMackin, J.F. (1996), "Integrating variable risk preferences, trust, and transaction cost economics". *Academy of Management Review*, 21, 73-99. 22.
5. Clarke, K. (2001), "What price on loyalty when a brand switch is just a click away?" *Qualitative Market Research*, 4(3), pp. 160-168.
6. Conner M. and Armitage J.C. (1998), "Extending the theory of planned behavior: A review and avenues for further research", *Journal of Applied Social Psychology*, Vol. 28, pp. 1430- 1464.
7. Constantinides, E. (2004)., "Influencing the online consumer's behavior: the Web experience". *Internet Research*. Vol. 14, No. 2, pp. 111-126. 27.
8. Currall, S. and Judge, T. (1995), "Measuring Trust Between Organizational Boundary Role Persons," *Organizational Behavior and Human Decision Processes* (64), p. 151-170.
9. Cuneyt, K. Gautam, B. (2004). The impacts of quickness, price, payment risk, and delivery issues on on-line shopping, *Journal of Socio-Economics*, Vol.33, PP.241–251.
10. Dawson, J. (2000). Retailing at Century End: Some Challenges for Management and Research. *International Review of Retail Distribution and Consumer Research* vol (2):119-148.
11. Dekimpe, M.G., Steenkamp, J.-B.E.M. and Mellens, M. (1997), "Decline and variability in brand loyalty", *International Journal of Research in Marketing*, Vol. 14, pp. 405-20.
12. Kesh, S., Ramanujan, S. and Nerur, S. (2002) "A framework for analyzing ecommerce security", *Information Management & Computer Security*, 10 (4)pp. 149-458.
13. Liao, Z. and Cheung, M.T. (2002), "Internet-based E-Banking and Consumer Attitudes: An Empirical Study". *Information and Management*, Vol. 39, pp. 283–295.
14. Rotchanakitumnuai, S., & Speece, M. (2003). Barriers to internet banking adoption: a qualitative study among corporate consumers in Thailand. *International Journal of Bank Marketing*, 21 (6/ 7), 312-323.
15. Roehl, W.S. and Fesenmaier, D.R. 1992. Risk Perceptions and pleasure travel: An exploratory analysis. *Journal of Travel Research*, 30(4): 17–26.
16. Schlosser, A.E. (2003a). Experiencing products in the virtual world: the role of goal and imagery in influencing attitudes. *Journal of Consumer Research*, Vol. 30, pp. 184-198.
17. Tan, M. & Teo, T. S. H. 2000. Factors influencing the adoption of Internet banking. *Journal of the Association for Information Systems*. Vol. 1 (5), pp. 1-42