Intentions Of Online Shoppers Prediction By Fuzzy Petri Nets Construction

S.Meher Taj, A.Kumaravel

Abstract: In this business field, the growth of online stores makes competition harder. Online stores need to have a website or application ready to calculate and classify consumer spending intentions, so customers will ultimately have an eye on things on the sites and applications to make purchases. Fuzzy Petri networks are dominant features as they cover the efficiency and inaccuracy management of sometimes the software field. Researchers use this tool to encode the processing results achieved. Data mining helps sell professionals to improve their consumer behavior understanding. This higher perceptive, in effect, helps them to concentrate on mercantilism campaigns with specific precision and match campaigns with shoppers and prospects, preferences, desires, and attitudes. Fuzzy Petri networks and classification mining methods are enforced data on intention online shopper victimization throughout this paper.

Index Terms: Data mining, high level fuzzy Petri nets, classifications, selected attributes, search methods, online shoppers intention data.

1. INTRODUCTION:

These days, the Internet is playing a very main role in the business world. The Internet has created an immense volume of transactions throughout the world. Furthermore, it does not just introduce a replacement approach to doing business; it changes the customer's lifestyle together. The transition from expertise in stores to online experience becomes a primary vehicle for e-commerce applications to look at product attributes, compare costs, and then get merchandise from online retailers [1]. How to purchase or sell something fairly simple, by the method use bound applications that are on the market online, shoppers will simply purchase things while not should leave the house, so does the seller can make sales without the need out of the store and a lot too seller who provides goods without have a shop [1]. At present the switch to offline outlets is changing into online stores are terribly huge, several huge firms that don't wish to miss this chance, several too standard people that make the most online search as a way to try to do business sell / purchase. Cannot denied once more, for currently do on-line shopping for and merchandising activities already become a part of the life-style the community [2]. Although there are many online stores that sprung up and plenty of many too people that use services from web site that gives online search facilities for free of charge, not all of them will attract shopper interest. Several retailers online that failing and failed to work for get the specified advantages as a result of shoppers don't feel search online may be a way sell / purchase that appeal to them [2]. Data mining is mainly used by hard-client-focused companies-retail, economic, networking, and organizational advocacy. Because of its tremendous applicability, data mining is very significant. It is widely used for understanding and forecasting valuable data in business applications, such as consumer purchase behavior and buying pattern, customer profiles, market analysis, etc [2].

- *S.Meher Taj, Department of Mathematics,*
- *AMET Deemed to be University,*
- *Chennai, Tamil Nadu, India mehertaj.s@ametuniv.ac.in*
- *Dr.A.Kumaravel, School of Computing, Bharath University, Chennai, Tamil Nadu, India. deancs@bharathuniv.ac.in*

research conducted by Norazah Mohd Suki and Norbayah Mohd Suki concluded that which determines online shopping behavior is how familiar and confident you are consumers when using website and online store applications [4]. Research to evaluate machine learning algorithm to judge online shopper intentions have been made earlier by C. Okan Sakar, Cashew Alpaslan Katircioglu, S. Olcay Polat and Yomi Kastro .That research completed with two modules using several types of algorithms the best machine learning and prediction obtained from a multilayer algorithm perceptron [2].

2. MATERIALS AND METHODS

2.1 Fuzzy Petri Nets

Petri Nets (PN) is a multi-system graphical and mathematical representation method. Promising tools are available to define and discover information science processes that are described as coincidental, asynchronous, distributed, parallel, non-deterministic and/or random [5]. Fuzzy Petri networks containing 2 kinds of nodes: Places and Transitions, where circles represent places and parallelogram represent transitions. Each place stands for an associate degree precedent or resultant and may or may not contain a token associated with a degree of truth between zero and one that speaks the validity of the precedent or sequence for the amount of trust at intervals. Every transition that represents a rule is said to have an issue value of certainty between zero and one. The factor of certainty in rule [6, 7] represents the strength of the idea.

2.2 Data Set:

In this paper, we tend to use UCI Repository's "Online Shoppers Buying Intention Dataset." The database consists of 12,330 sessions function vectors. The database was structured so that in a very1-year amount each session would belong to a different user to prevent any bias toward a particular campaign, big day, user profile, or time. Of the 12,330 dataset sessions, 84.5 million (10,422) were negative category samples that failed to finish searching, and therefore the remainder 1908) were positive category samples that ended with searching [14].

2.3 Methodology

Feature selection techniques are applied in this study to improve the system's classification performance and/or scalability. Therefore, if higher or comparable classification quality is often achieved with fewer options, we prefer to strive for analysis. An alternative to feature selection is that a feature extraction technique such as the Main Part Analysis is used to reduce spatial properties. In this case, however, the features in the reduced space will be the linear combinations of 18 attributes, which will result in the need to track all features during the visit [12,13].

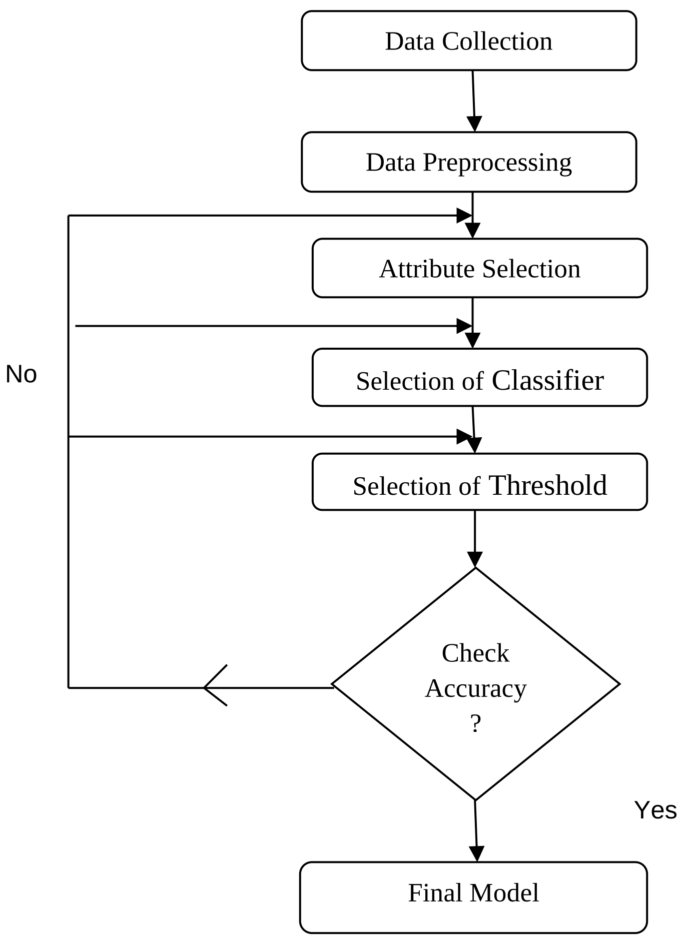


Figure.1: *Main method Proposed*

It has therefore been considered acceptable to use feature choice rather than feature extraction within the scope of this study.
We apply different attribute evaluators, search methods, and classifier techniques in this paper [8].

3. EXPEREIMENTAL RESULT:

The following tables 1 and 2 illustrate that we want to see a reasonable combination of classifier attribute choice with results consists of the level of accuracy, mean absolute error AUC of ROC as maximum, minimum and high

S.No	Search method	Attribute Evaluator	Classifier	Thrust hold value (if any)	Accuracy	No of Rules generated	Mean absolute Error	ROC Area
1.	Best First	Cfs Subset	JRIP	-	88.87	7	0.1708	0.793
2.	Greedy Stepwise	Cfs Subset	JRIP	-	88.87	7	0.1708	0.793
3.	Ranker	Info Gain	JRIP	-	90.24	9	0.1535	0.804
				0.007	90.02	8	0.1554	0.801
				0.008	90.06	8	0.1542	0.806
				0.009	90.26	14	0.1531	0.812
4.	Ranker	Correlation	JRIP	-	90.26	14	0.1538	0.809
				0.07	90.24	8	0.1551	0.807
				0.08	89.42	08	0.1647	0.787
				0.09	89.23	13	0.1640	0.778
5.	Ranker	Gain ratio	JRIP	-	99.06	08	0.1539	0.805
				0.011	89.56	10	0.1626	0.793
				0.012	89.24	14	0.1623	0.789
				0.013	89.34	19	0.1639	0.781
6.	Ranker	Relief F	JRIP	-	90.12	13	0.1533	0.808
				0.004	89.80	12	0.1558	0.799
				0.005	90.05	9	0.1544	0.803
				0.006	90.05	13	0.1542	0.807
7.	Ranker	Symmetrical Uncert	JRIP	-	90.03	11	0.1534	0.809
				0.012	90.16	16	0.1549	0.798
				0.013	90.12	10	0.1543	0.806

Table 1- *Estimation of different feature selection methods based on JRIP Classifier*

S.No	Search method	Attribute Evaluator	Classifier	Thrust hold value (if any)	Accuracy	No of Rules generated	Mean absolute Error	ROC Area
1.	Best First	Cfs Subset	PART	-	88.88	5	0.1557	0.854
2.	Greedy Stepwise	Cfs Subset	PART	-	88.88	5	0.1557	0.854
3.	Ranker	Info Gain	PART	-	88.14	212	0.1368	0.862
				0.007	89.09	118	0.1375	0.901
				0.008	89.14	73	0.1385	0.904
				0.009	89.49	62	0.1377	0.911
4.	Ranker	Correlation	PART	-	220	88.17	0.1369	0.856
				0.07	89.29	108	0.1367	0.902
				0.08	88.98	31	0.1450	0.895
				0.09	88.95	35	0.1460	0.895
5.	Ranker	Gain ratio	PART	-	88.07	212	0.1376	0.857
				0.011	89.27	18	0.1457	0.898
				0.012	89.24	12	0.1457	0.897
				0.013	89.25	15	0.1464	0.895
6.	Ranker	Relief F	PART	-	88.02	213	0.1379	0.859
				0.004	89.16	107	0.1379	0.894
				0.005	89.00	103	0.1390	0.899
				0.006	89.24	92	0.1390	0.897
7.	Ranker	Symmetrical Uncert	PART	-	89.12	212	0.1367	0.861
				0.012	89.5	73	0.1384	0.904
				0.013	89.66	72	0.1375	0.914

Table 2*- Estimation of different feature selection methods based on PART Classifier*

It is found from the above table that the JRIP algorithm has the lowest error rate. Consequently, JRIP classification algorithms perform well because they contain the lowest error rate and also the highest accuracy compared to other algorithm[6,10].

***Table .3**: Classifier Output of the JRIP Model-I*

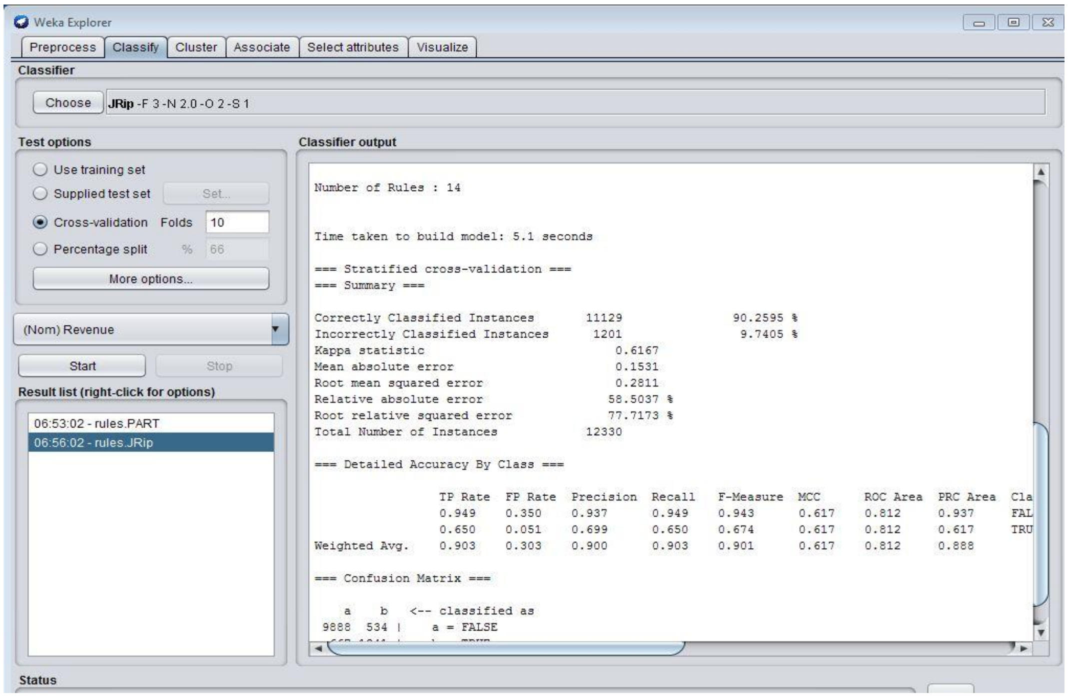
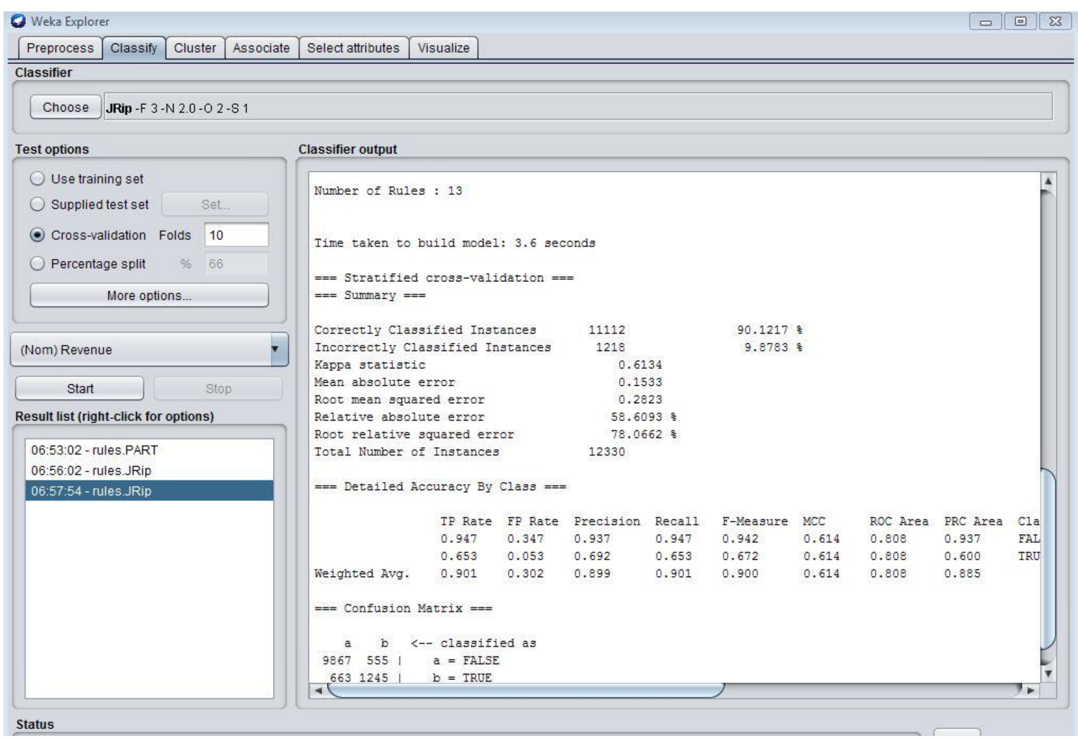


***Table 4:** Classifier Output of the JRIP Model-II*



We take 2 choices of a classifier attribute combination. First, high precision 90.2%, low error 0.1531 and high ROC 0.812. Second, high precision 90.12%, low error 0.1533 and high ROC 0.808. The two configurations above produce the laws of j rip.

Using CPN tool to indicate the below 2 combinations rules.

The first combination produces the subsequent rule

R1: If (Page Values >= 19.448) and (Bounce Rates < =0.001075) then Revenue=TRUE (758.0/124.0).

R2: If (Page Values >= 0.951934) and (Bounce Rates <= 0.000388) and (Product Related <= 22) then Revenue=TRUE (136.0/34.0)

R3: If (Page Values >= 15.357603) and (Exit Rates <= 0.014672) and (Administrative Duration<=100.916667) then Revenue=TRUE (95.0/25.0)

R4: If (Month = Nov) and (Exit Rates <= 0.019118) and (Page Values >= 3.651726) then Revenue=TRUE (217.0/64.0)

R5: If (Page Values >= 0.06705) and (Product Related Duration <= 1285.583333) and (Exit Rates <= 0.010556)
then Revenue = TRUE (42.0/11.0)

R6 : If (Page Values >= 0.06705) and (Product Related Duration <= 1263.864951) and (Administrative <= 4) and (Bounce Rates <= 0.00084) then Revenue=TRUE (41.0/16.0)

R7: If (Page Values >= 0.06705) and (Product Related <= 32) and (Exit Rates <= 0.023222) then Revenue=TRUE (149.0/73.0)

R8: If (Page Values >= 0.06705) and (Month = Nov) and (Product Related Duration >= 4409.3) then Revenue=TRUE (84.0/29.0)

R9: If (Page Values >= 12.837024) and (ProductRelated_Duration <= 504.083333) then Revenue=TRUE (54.0/18.0)

R10: If (Page Values >= 6.324477) and (Traffic Type >= 3) and (Bounce Rates <= 0.013158) and (Product Related <= 51) and (Bounce Rates >= 0.007143) then Revenue=TRUE (52.0/18.0)

R11:If(Page Values >= 0.093547) and (Exit Rates <=0.007692)and(Product Related<=52) then Revenue=TRUE (12.0/3.0)

R12: If (Month = Nov) and (Exit Rates <= 0.027766) and (Product Related < = 56) and (Page Values > =14.028639) then Revenue = TRUE (22.0/6.0)

R13:If (Page Values >= 0.093547) and (Administrative <= 6)
and (Traffic Type <= 1) and (Product Related Duration <= 1038.916667) and (Administrative Duration <= 27.466667) = then Revenue=TRUE (21.0/5.0)

R14: then Revenue=FALSE (10647.0/651.0).

The second combination produces the subsequent rule

R1: If (Bounce Rates <= 0.000375) and (Page Values >= 23.575371) then Revenue=TRUE (659.0/93.0)

R2: If (Page Values >= 14.943793) and (Bounce Rates <= 0) and (Exit Rates <= 0.007456) then Revenue=TRUE (60.0/10.0)

R3: If (Page Values >= 0.06705) and (Product Related <= 35)
and (Bounce Rates <= 0.001111) and (Exit Rates <= 0.012644) then Revenue=TRUE (94.0/20.0)

R4: If (Page Values >= 0.06705) and (Month = Nov) and (Exit Rates <= 0.020721) then Revenue=TRUE (330.0/106.0)

R5: If (Page Values >= 0.602233) and (Month = Nov) and (Product Related Duration >= 4409.3) then Revenue=TRUE (51.0/15.0)

R6: If (Page Values >= 0.093547) and (Administrative<=1) and (Product Related Duration <= 512.716667) => Revenue=TRUE (83.0/19.0)

R7:If (Page Values>=0.093547)and (Product Related Duration <= 1289.25) and(Administrative Duration <= 30.25) then Revenue=TRUE (171.0/73.0)

R8: If(Page Values >= 0.093547) and (Product Related Duration <= 1321.059195) and (Month = Mar) then Revenue=TRUE (51.0/15.0)

R9: If (Page Values >= 10.3572) and (Administrative_Duration<= 78.3) and (Administrative_Duration > =66) and (Informational Duration <= 58) => Revenue=TRUE (32.0/6.0)

R10: If (Page Values >= 0.06705) and (Product Related <= 35)
and (Bounce Rates <= 0.011828) and (Administrative <= 5) and (Administrative_Duration >= 127) => Revenue=TRUE (43.0/12.0)

R11: If (Page Values >= 0.1207) and (Product Related <= 23)
and (Exit Rates <= 0.042708) and (Administrative_Duration <= 60.366667) then Revenue=TRUE (37.0/16.0

R12:If (Page Values >= 10.600892) and (ProductRelated_Duration <= 724.066667) and (Product Related >= 15) then Revenue=TRUE (38.0/16.0)

R13: then Revenue=FALSE (10681.0/660.0).

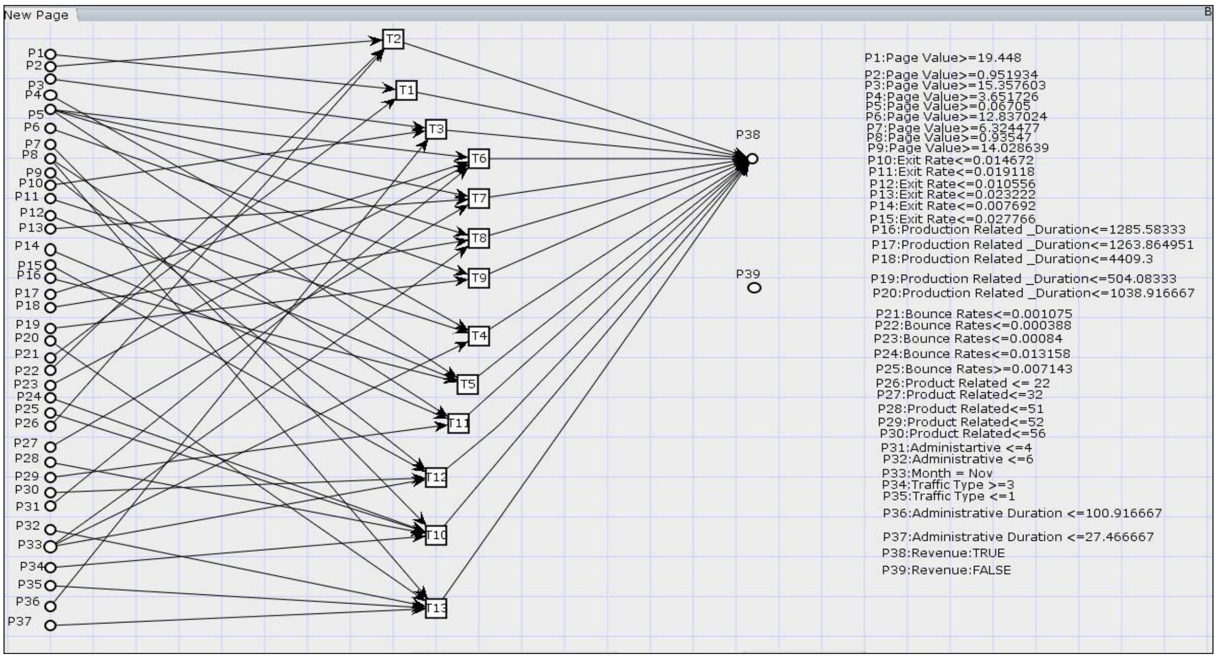


Figure: 2 *CPN Tool Snapshot-I for*

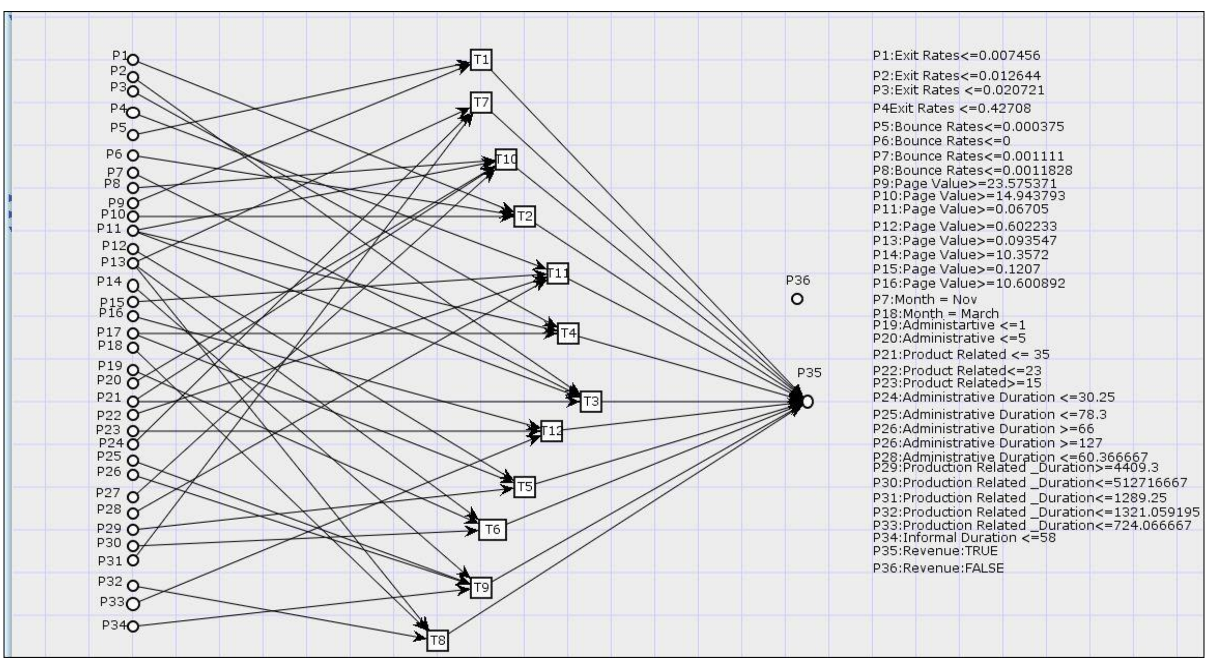


Figure: 3 *CPN Tool Snapshot-II for online shopper intention*

The Petri net functional model is shown in Figure. 2. Among the Petri net model [7, 11] in accordance with the proportions committed to each location, transitions 1 to 13 represent Rules 1 to 14 separately. In Figure.3, the Petri net model, in accordance with the proportions committed to each place, transitions 1 to 12 represent separately rules 1 to 13 between the established rule on high and firing each transition implies completion of the corresponding rule.

4. CONCLUSION

We found the association of Petri net's behavior for executing the decision rules. The context of on-line shopper intention is chosen for this purpose and attained the results for predicting client getting intention. The requirement of produced rules supported the information set collected by colored Petri networks, thus enhancing the clarity. In terms of support and confidence factors, data mining tool such as WEKA is used to fit the correct set of rules with metrics. Besides, the option of attributes achieves maximum precision of around 90 percent by tuning parameters. For additional useful attributes, one could identify entirely different data sizes for the future.

ACKNOWLEDGMENT

The first author would like to thank AMET Deemed University's management for supporting and encouraging this research study.

REFERENCES:

[1] Christian, Y,. Comparison of Machine Learning Algorithms Using WEKA and Sci-Kit Learn in Classifying Online Shopper Intention. JITE (Journal of Informatics and Telecommunication Engineering). Vol.3,pp 58-66, 2019.

[2] Sakar, C. O., Polat, S. O., Katircloglu, M. Real-Time Prediction of Online Shopper' Purchasing Intention Using Multilayer Perceptron and LSTM Recurrent Neural Networks. In: Neural Computing and Applications. Springer, 2018.

[3] Khan, M. A., & Khan,. Service Convenience and Post-Purchase Behavior of Online Buyers: An Empirical Study. In: Innovation in the High-Tech Economy. Springer. 2018

[4] Suki, N. M., & Suki,. Examining Factors Correlated with Consumer Online Shopping Behaviour. In: Innovation in the High-Tech Economy. Springer. 2013.

[5] Murata, T.,Petri nets: Properties, analysis and applications. Proceeding of the IEEE, Vol. I7, pp541-580. 1989.

[6] Fei Liu, Ming Yang and Peng Shi, Verification and Validation of Fuzzy Rules-Based Human Behavior Models, 7th International Conference on System Simulation and Scientific Computing, pp 813 – 819, 2018.

[7] Kouzehgar, M., M.A.Badamchizadeh and S. Khanmohammadi, 2011. Fuzzy Petri nets for human behavior verification and validation. Int. J. Adv. Comput. Sci. Appl., Vol.2, pp106-114, 2011.

[8] Lakshmi Devasena.C, Sumathi.T, Gomathi.V.V and. Hemalatha.M, .Effectiveness Evaluation of Rule-Based Classifiers for the Classification of Iris Data Set, Bonfring International Journal of Man-Machine Interface.Vo1,pp 5-9,2011.

[9] Murata, T,Petri nets: Properties, analysis and applications. Proceeding of the IEEE, Vol.I7 pp 541-580, 1989.

[10] Meher Taj.S and Kumaravel.A, 2015.Measuring Employee performance key indicators by Fuzzy Petri Nets , Research journal of applied science and Engineering Technology,Vol.9,pp 485-490,2015.

[11] Chen S.M, Ke J.S and Chang J.F. 1990 .Knowledge Representation using Fuzzy Petri Nets, IEEE Trans. Knowledge and Data Eng.Vol.3,pp 11-319,1990.

[12] Sudha, Kumaravel.A, 2017 .Comparative Analysis between Rough set theory and Data mining algorithms on their prediction, Global J. Pure Appl. Math. Vol 13: ,pp73249–3260,2017.

[13] M. Sudha and A. Kumaravel,Performance Comparison based on Attribute Selection Tools for Data Mining, Indian Journal of Science and Technology, Vol 7,pp 61–65,2014.

[14] C. Okan SakarS. Olcay Polat,Mete Katircioglu,Yomi Kastro ,Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks, Neural Comput & Applic,Vol.31,pp6893–6908,2018.