# MACHINE LEARNING APPROACHES OF AWARENESS AMONG PATIENTS ON ADR REPORTING SYSTEM IN CHENNAI

Dr.G.Ayyappan,
Associate Professor, Department of Information Technology,
BIHER, BIST, Bharath University, Chennai, India.

K.SivaKumar
SIPS Technologies, Chennai, India

**Abstract - Patient's adverse drug reacting reporting is a brand new idea in Pharmacovigilance which make contributions to the enrichment of current drug protection performs. In this research work focuses on implementation of machine learning techniques for patient awareness on opposing drug reaction reporting system in Chennai. The current study was a cross-sectional study which was showed for a period of one year amongst patients hospitalized at Chennai. Sample size taken was 1000 and the sample size was collected by google forms. Data was collected using a standardized questionnaire. Data entered in MS Excel and analyzed using Weka 3.8.3 and results interpreted. The NaïveBayes classifier has 93.21% accuracy level and it has take time to build the model 0.01 seconds. The SMO(Support Vector Machine) has produced the 96.75% accuracy and it has take time to build the model 0.51 seconds. The IBK machine learning algorithm has 95.28% accuracy and it has take time to build the model 0.00 second. The remaining machine learning algorithms namely ClassificationViaRegression, DecisionTable and J48 classifiers have same accuracy level like 97.34%. But the Classification Via Regression has taken the time to shape the model 1.02 seconds, DecisionTable has taken the time to shape the model 0.23 seconds, J48 classifier has taken the time to build the model 0.09 seconds. The review of consciousness between patients designates low consciousness and it could be upgraded by presenting educational interventional programs.**

**Keywords:** Machine Learning Algorithms, Naïve Bayes, J48, ClassifiactioViaRegression, DecisionTable, Adverse drug reactions reporting, Awareness, and Patients.

## I. Introduction

Pharmacovigilance is the science and activities relating to the detection, monitoring, assessment, understanding and prevention of adverse effects or any other drug-related problem from any pharmaceutical products. Medicines have, beyond any doubt, proved to be a boon for humanity and it fights against disease and suffering. However, like most other useful things, medicines come with inherent risks associated with their use, called Adverse Drug Reactions (ADRs). These reactions, though mild in most cases, have the potential to cause disability and even death. ADRs are often referred to as "any noxious and unintended effects of a drug that occurs at doses normally used in human beings for the prophylaxis, diagnosis or therapy of disease, or for modification of physiological function1. They account for approximately 4.2% to 6.0% of all hospital admissions and they occur in about 10%-20% of all hospitalized patients. The process of identifying and preventing ADRs associated with postmarketed drugs i.e. Pharmacovigilance is becoming increasingly important due to the potential harmful effects of drugs on patient's health, economic burden associated with ADRs and circulation of large numberof over-the-counter and counterfeit drugs in the market**.**

In this research article presents in section 2 materials and methods, in section 3 results and discussions and finally, conclusion of the research.

## II. Materials and Methods

The current study was a cross-sectional study which was showed for a dated of one year between patients hospitalized at Chennai region hospital. Sample size taken was 358 and the sample size was arrived by google forms. Data was collected using a standardized questionnaire. Data entered in MS Excel and analyzed using Weka 3.8.3 and results interpreted.

The below machine learning algorithms have applied in this study for classification approaches of this dataset.

- NaiveBayes
- SMO(Support Vector Machine)
- IBK
- ClassificationViaRegression
- DecisionTable
- J48

## III. Results and Discussions

In this session focuses on the results and discussions of this study. The dataset contains 1000 instances and 11 attributes namely Username, Name, Age, Sex, Educational qualification, marital status, Residential location, Occupation, Working status, Family type, and Accessibility to health care facility. The below table represents the list of attributes involved in current study.

Table 1. List of the Attributes

| S.No | Name of the Attribute | Data type |
|---|---|---|
| 1 | Username | Character |
| 2 | Name | Character |
| 3 | Age | Numerical |
| 4 | Sex | Character |
| 5 | Educational qualification | Character |
| 6 | Marital status | Character |
| 7 | Residential location | Character |
| 8 | Occupation | Character |
| 9 | Working status | Character |
| 10 | Family type | Character |
| 11 | Accessibility to health care facility | Character |

The below table specifies that the demographical distributions of the current study.

The below table represents that the classification approaches of several machine learning algorithms are implemented in our dataset.

Table 2: Various Classification Approaches

| **Category** | **Classifier** | **Accuracy** | **Time taken to build the model(In Seconds)** |
|---|---|---|---|
| Bayes | NaiveBayes | 93.21 | 0.01 |
| Functions | SMO(Support Vector Machine) | 96.75 | 0.51 |
| Lazy | IBK | 95.28 | 0 |
| Meta | ClassificationViaRegression | 97.34 | 1.02 |
| Rules | DecisionTable | 97.34 | 0.23 |
| Trees | J48 | 97.34 | 0.09 |

The above table demonstrates that the Naïve Bayes classification algorithm belongs to under the category of Bayes category. The SMO(Support Vector Machine) algorithm comes under the category of functions. The IBK machine learning algorithm originatesbelow the group of lazy. The ClassificationViaRegression machine learning algorithm comes below the grouping of Collaborative. The DecisionTable classifier comes under the category of Rules.The J48 classifier comes under the category of Trees.

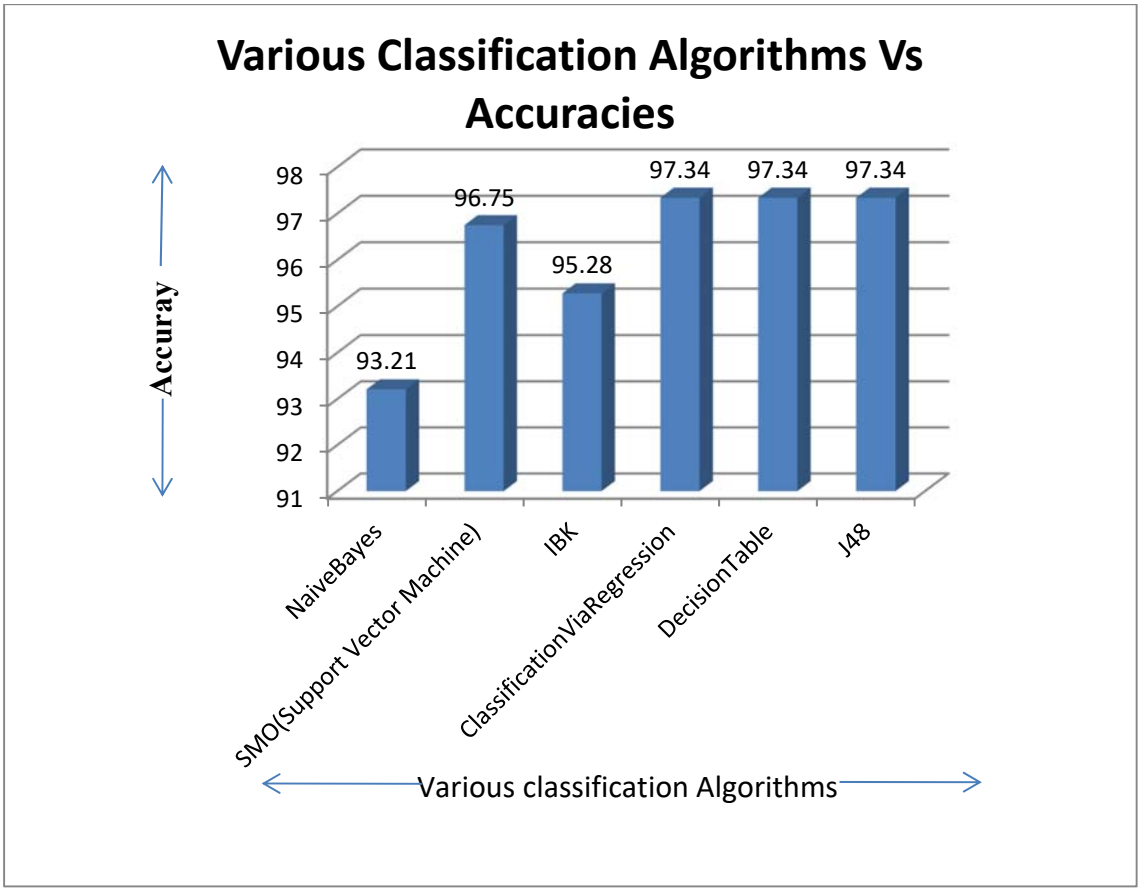


Fig 1: Various Classifiers Vs Accuracies

The above graph clearly demonstrates that there are six leading machine learning classification algorithms applied in this study. The NaïveBayesclassifer has 93.21% accuracy level and it has take time to build the model 0.01 seconds. The SMO(Support Vector Machine) has produced the 96.75% accuracy and it has take time to build the model 0.51 seconds. The IBK machine learning algorithm has 95.28% accuracy and it has take time to build the model 0.00 second. The remaining machine learning algorithms namely ClassificationViaRegression, DecisionTable and J48 classifiers have same accuracy level like 97.34%. But the ClassificationVia Regression has taken the time to build the model 1.02 seconds, DecisionTable has taken the time to build the model 0.23 seconds, J48 classifier has taken the time to build the model 0.09 seconds.

The below table clearly represents that the values of the True Positive(TP) Rate, False Positive(FP) Rate, Precision, Recall, F- Measure, ROC Area and PRC Area.

The NaiveBayes Classifier has the TP Rate was found 0.93, FP Rate was found 0.76, Precision Value was found 0.96, Recall value was found 0.93, F Measure Values was found 0.99, ROC Area was found 0.69 and PRC Area was found 0.96.

The SMO(Support Vector Machine) Classifier has the TP Rate was found 0.97, FP Rate was found 0.97, Precision Value was found 0, Recall value was found 0.97, F Measure Values was found 0, ROC Area was found 0.22 and PRC Area was found 0.93.

In the IBK Classifier, the TP Rate was found 0.95, FP Rate was found 0.97, Precision Value was found 0.95, Recall value was found 0.95, F Measure Values was found 0.95, ROC Area was found 0.54 and PRC Area was found 0.95.

In the ClassificationViaRegression classifier, the TP Rate was found 0.97, FP Rate was found 0.97, Precision Value was found 0, Recall value was found 0.97, F Measure Values was found 0, ROC Area was found 0.22 and PRC Area was found 0.93.

In the DecisionTable classifier, the TP Rate was found 0.97, FP Rate was found 0.97, Precision Value was found 0, Recall value was found 0.97, F Measure Values was found 0, ROC Area was found 0.22 and PRC Area was found 0.93.

In the J48 classifier, the TP Rate was found 0.97, FP Rate was found 0.97, Precision Value was found 0, Recall value was found 0.97, F Measure Values was found 0, ROC Area was found 0.22 and PRC Area was found 0.93.

Table 3: Various Matrices of the Algorithms

| Category | Classifier | TP RATE | FP RATE | Precision | Recall | F Measure | ROC Area | PRC Area |
|---|---|---|---|---|---|---|---|---|
| Bayes | NaiveBayes | 0.93 | 0.76 | 0.96 | 0.93 | 0.99 | 0.69 | 0.96 |
| Functions | SMO(Support Vector Machine) | 0.97 | 0.97 | 0 | 0.97 | 0 | 0.22 | 0.93 |
| Lazy | IBK | 0.95 | 0.97 | 0.95 | 0.95 | 0.95 | 0.54 | 0.95 |
| Meta | ClassificationViaRegression | 0.97 | 0.97 | 0 | 0.97 | 0 | 0.22 | 0.93 |
| Rules | DecisionTable | 0.97 | 0.97 | 0 | 0.97 | 0 | 0.22 | 0.93 |
| Trees | J48 | 0.97 | 0.97 | 0 | 0.97 | 0 | 0.22 | 0.93 |

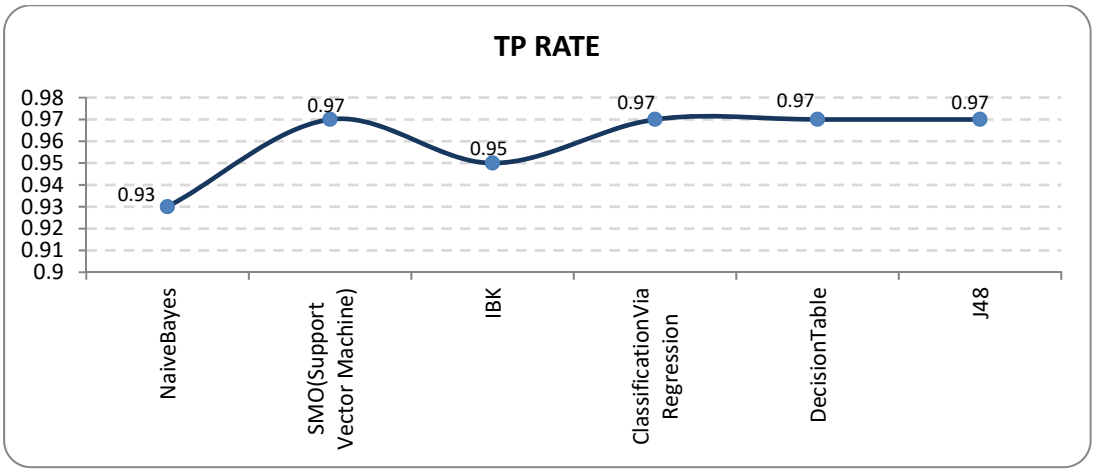


Fig 2: Various Classifiers Vs TP Rate

The above graph shows the true positive rates for different machine learning algorithms and it was found that for Naivebayes the TP rate is minimum with 93 % and for IBK it's about 95%. For all other classifiers taken into consideration the TP rate is similar.

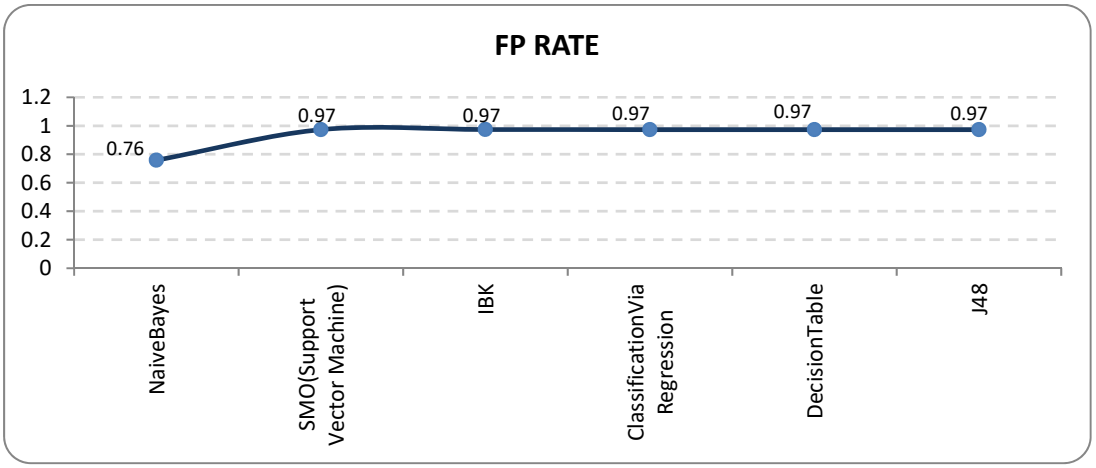


Fig 3: Various Classifiers Vs FP Rate

The above graph shows the false positive rates for different machine learning algorithms and its found to similar results for all types of classifiers except Naïve bayes and is about 76%.

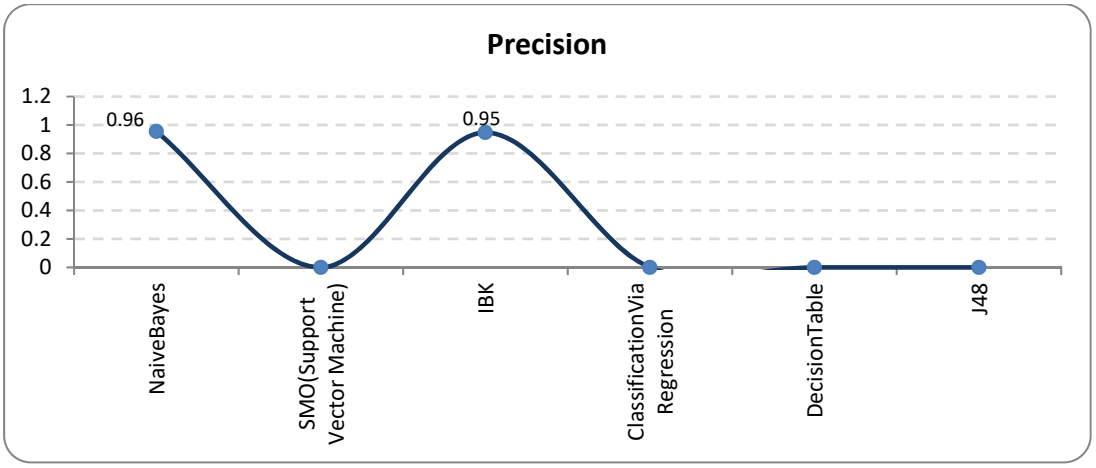


Fig 4: Various Classifiers Vs Precision

The above graph shows the precision values for different machine learning algorithms. Precision is aamount of positive class forecasts to the total positive classes. It was found that for SVM, Regression, decision table and J48 was undefined since the total positive classes were found to be zero. For Naïve Bayes and IBK it was above 95%.

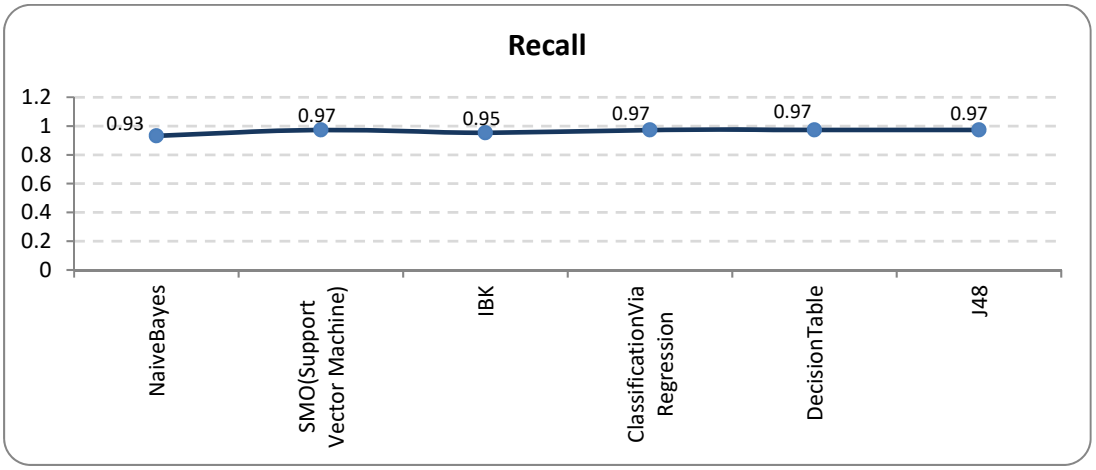


Fig 5: Various Classifiers Vs Recall

The above graph shows the Recall values for different machine learning algorithms. Recall is aamount of the number of positive class forecasts made out of all positive examples in the dataset. It was found to be more or less same for types of classifiers.

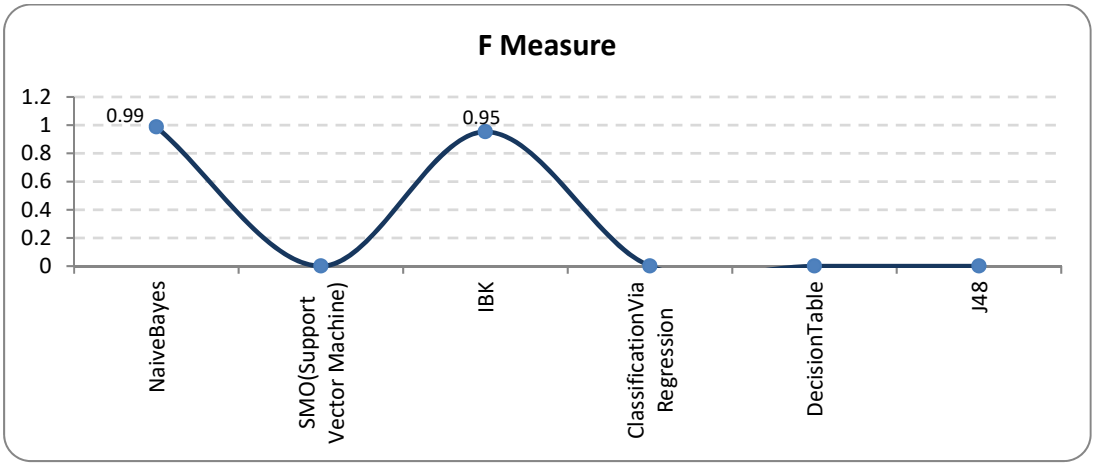


Fig 6: Various Classifiers Vs F Measure

The above graph shows the F Measure values for different machine learning algorithms. F-Measure delivers a single score those equilibrium both the anxieties of exactness and recall in one number. Since F-Measure depends on the value of precision we get similar results for precision and F Measure for all type of classifiers.

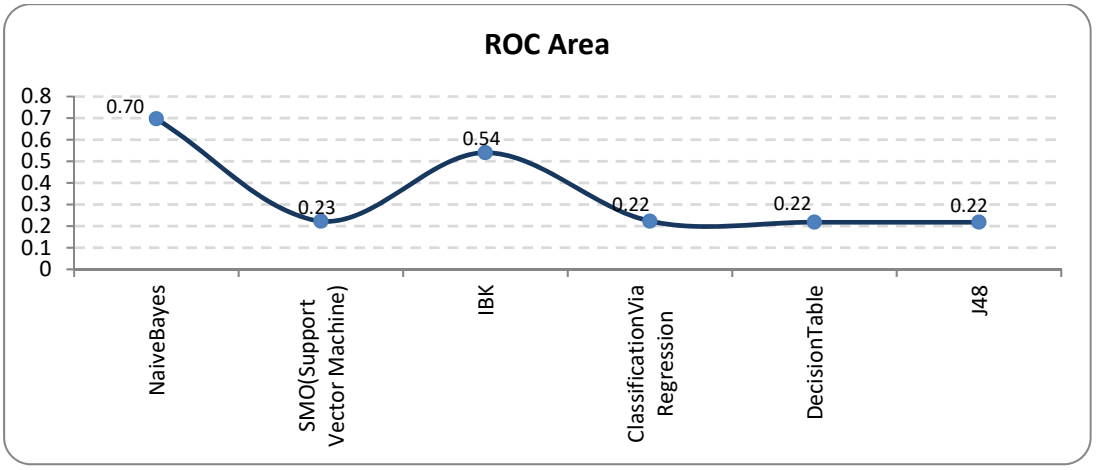


Fig 7: Various Classifiers Vs ROC Area

The above graph shows the ROC Area values for different machine learning algorithms and found that similar results for SVM, Regression, Decision Table and J48 .For Naïve Bayes its about 70% and for IBK it is about 54%.

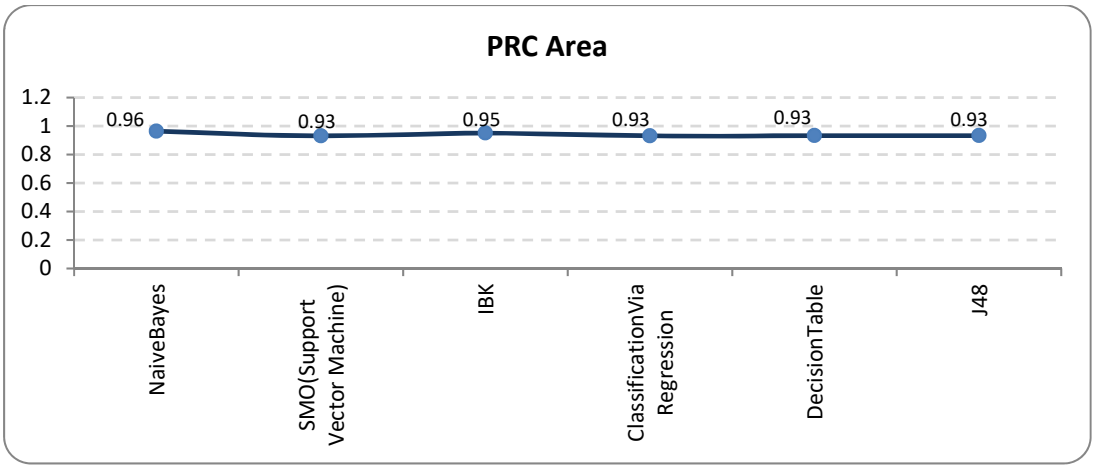


Fig 8: Various Classifiers Vs PRC Area

The above graph shows the PRC Area values for different machine learning algorithms. It was found to be more or less same for types of classifiers.

## IV. Conclusions

In this research work concludes that the ClassificationViaRegression, DecisionTable and J48 classifiers have same accuracy level 97.34%. But the ClassificationViaRegression has taken the time to build the model 1.02 seconds, DecisionTable has taken the time to build the model 0.23 seconds, J48 classifier has taken the time to build the model 0.09 seconds. Our proposed system recommends that the J48 machine learning classifier has produced better accuracy with less time consumption for building a model for classifying by the novel approaches of the Knowledge on side-impact or damaging impact of drug treatments, the share of defendantspracticedopposing drug responses, whether individuals suggested damaging drug reactions, their notion toward reporting opposing drug responses, consciousness of the present machine of Pharmacovigilance in Chennai region hospitals, and their most efficient mode of reporting unfavorable drug responses in destiny.

## References

[1] Ritu Pahuja, Birendra Shrivastava,Pankaj Kumar Sharma,Kamal Kishore,Sandeep Mahajan and Radhika Sood, Awareness on Adverse Drug Reaction Reporting System in India: A consumer Survey. American Journal of Phytomedicine and Clinical Therapeutics. (AJPCT), Vol 2,Issue 12,2014, Page 1361-1369.

[2] Muehlberger N, Schneeweiss S, Hasford J. Adverse drug reaction monitoring-cost and benefit considerations. Part I: frequency of adverse drug reactions causing hospital admissions. Pharmacoepidemiology & drug safety 1997; 6 (S3): 71-77.

[3] Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients-excess length of stay, extra costs and attributable mortality. Journal of the American Medical Association 1997; 277: 301-6.

[4] Moore TJ, Psaty BM, Furberg CD. Time to act on drug safety. Journal of the American Medical Association, 1998; 279:1571-1573.

[5] Blenkinsopp A, Wilkie P, Wang M, Routledge P.A. Patient reporting of suspected adverse drug reactions: a review of published literature and international experience. British Journal of Clinical Pharmacology. Feb 2007; 63 (2): 148-156.

[6] Avery AJ, Anderson C, Bond CM, Fortnum H, Gifford A, Hannaford PC et al. Evaluation of patient reporting of adverse drug reactions to the UK 'Yellow Card Scheme': literature review, descriptive and qualitative analyses, and questionnaire surveys. Health Technology Assess 2011 May; 15(20):1-234.

[7] Ahmed AM, Izham IM, Subish P. The Importance of the Consumer Pharmacovigilance System in Developing Countries: A Case of Malaysia. Journal of Clinical and Diagnostic Research 2010 Aug; 4:2929-2935.
[8] Margraff F, Bertram D. Adverse drug reaction reporting by patients: an overview of fifty countries. Drug Safety 2014. Jun; 37 (6): 409-19.
[9] Herxheimer A, Crombag MR, Alves C TL. Direct patient reporting of adverse drug reactions. A twelve-country survey and literature review. Briefing Paper Health Action International. January 2010.
[10] Langen J, van Hunsel, Passier, van den Berg L, van Grootheest K. Adverse Drug ReactionsReporting by Patients in the Netherlands. Drug Safety 2008; 31:515-24.
[11] Hazell L, Shakir SA. Under-reporting of adverse drug reactions: a systemic review. Drug Safety, 2006; 29:385-396.
[12] National Statistics Opinions (Omnibus) Survey. User guide. URL: www.statistics. gov.uk/about/services/omnibus/downloads/User_Guide.pdf (accessed 8 October 2014).
[13] Amrita P, Kharbanda B. Knowledge, attitude and skills of nurses of Delhi towards adverse drug reaction reporting. Indian J Pharm Pract 2012; 5: 45-51.
[14] Jha N, Rathore DS, Shankar PR, Gyawali S. Pharmacovigilance Knowledge among Patients at a Teaching Hospital in Lalitpur District, Nepal. Journal of Clinical Diagnostic Research. 2014 March: 8(3): 32-4.