International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

ANALYSIS OF DATA MINING TOOL ORANGE

V.Padmavaty[1],C.Geetha[2],.N.Priya[3]

[1].Student, Dep of CSE, Bharath University, Chennai, Tamilnadu, India.
[2.,3]Assistant professor, Dep of CSE, Bharath University, Chennai, Tamilnadu, India.

Abstract

Data mining is the process of sorting through large datasets to identify pattern and establish relationships to solve problems through data analysis. Data mining is the process of discovering pattern in large data sets involving methods at the intersection of machine learning, statistics, and data base systems.There are many tools such as weka, R, orange etc to analyze, visualize and extract data using data mining . All the datamining tools are not compatible to perform all analysis operations. In this paper we have attempted data mining tools for analysis and checking for better tool.

Key words: Data Mining, orange, attribute statistics, Pre-processing.

Introduction

Data analysis is an essential part of data mining and Business Intelligence(BI) and is critical to picking up the understanding that drives business choices. Data analysis is the process of performing three operations which would be major such as cleaning, transforming and data extraction. Comparing with some tools along with features and parameters are decided to use for analysis.

Data Mining Tools

Apache Mahout, Orange, Weka, R, Rapid Miner, Data Melt, Knime

- **Apache Mahout:** Apache mahout focuses mainly on data clustering, classification, and collaborative filtering. Mahout is written in JAVA and includes JAVA libraries to perform mathematical operations like linear algebra and statistics. Mahout is growing continuously as the algorithms implemented inside Apache Mahout are continuously growing. The algorithm of mahout have implemented a level above Hadoop through mapping/reducing templates.

- **Orange:** Orange is a perfect software suite for machine learning and data mining. It is useful for visual programming and explorative data analysis. It is written in python and has multiple components known as widgets. Data mining tool supports macOS, Window and Linux.

- **Weka:** Weka is a machine learning data mining tool in java and best suited for data analysis and predictive modelling. It contains visualization and analysis. Weka supports major data mining tasks including processing, visualization, regression etc. It works on the assumption that data is available in the form of flat file.

- **R:** R is also called open source and a free software environment to perform statistical computing and graphics. It is widely used in academia, research, engineering, and industrial applications.

- **Rapid Miner:** Rapid miner is one of the best predictive analysis system developed by a company with the same name as rapid miner. It is written in java programming language. It provides an integrated environment for deep learning, text mining, machine learning and predictive analysis. It uses client/server model which is used in performing extraction, transformation and data processing operations.

- **Knime:** knime is the best integration platform for data analytics and reporting development by KNIME.com AG .It operates on the concept of the modular data pipeline. KNIME is widely used in

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

pharmaceutical research. In addition ,it performs excellently for customer data analysis, financial data analysis, and business intelligence.

- **Data melt:** Data melt, also known as DMelt is a computation and visualization environment that provides an iterative framework to do data analysis and visualization. Dmelt is written in java and it is a multiplatform utility and a open source data mining software.

Comparision of data mining with parameters:

	Offset
True	1
False	0

(Table1 specific value to set in comparision)

In Table 1 and Table 2 shows all data. mining tool has been specified with their parameters either tool supports or ignore. If it gets positive or supporative value to offset assign true means 1 and if gets negative or not supportive value to offset assign false means 0. Offset is using for to balance one to another.

(Table2 comparison of tools with parameters)

Feautres/Parameters of Data Mining tools	Orange	Weka	R	Rapid miner	Knime	Datamelt	Apache mahout
Open source	1	1	1	1	1	1	1
Data visualization and analysis	1	1	0	1	0	1	1
Interaction and data analysis	1	1	0	1	1	1	0
Large toolbox	1	0	0	1	1	0	1
Scripting interface	1	1	1	0	1	1	0
Platform independence	1	1	1	1	0	0	0
Covering methods	0	1	0	1	0	1	0
Parameters optimized method learning/statistical methods	0	0	0	1	0	1	0
Total	06	05	03	05	04	04	03

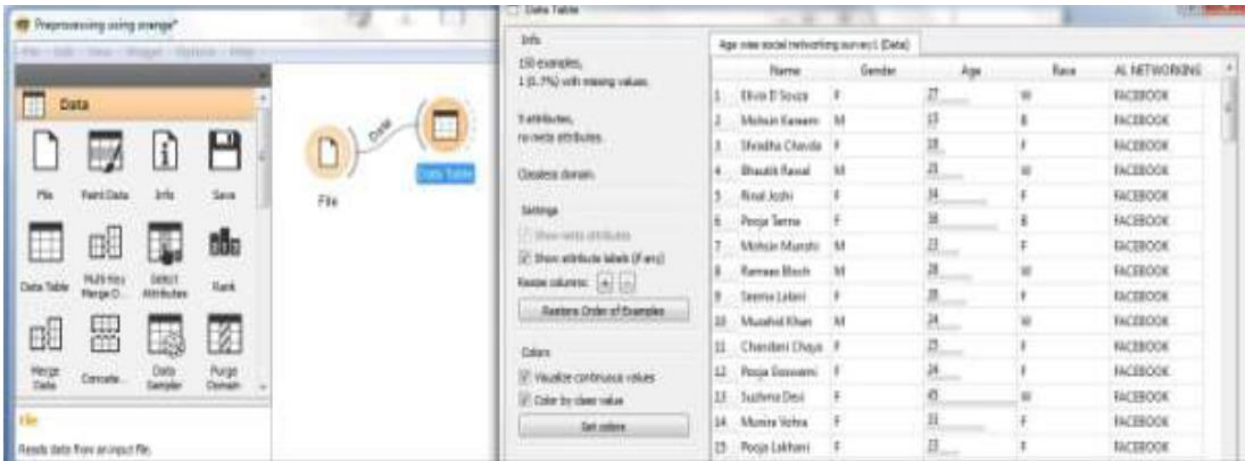
Methodology/Technique: Orange data mining is a component based visual programming software package for data visualization, machine learning , data mining and data analysis. It also helps in supporting programming language like C,C++ and python that also supports data validation , comparision and prediction. Orange is easy to learn

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

and is better and best than other as compare it above

Orange uses for practical Implentation

- **Data Analysis**

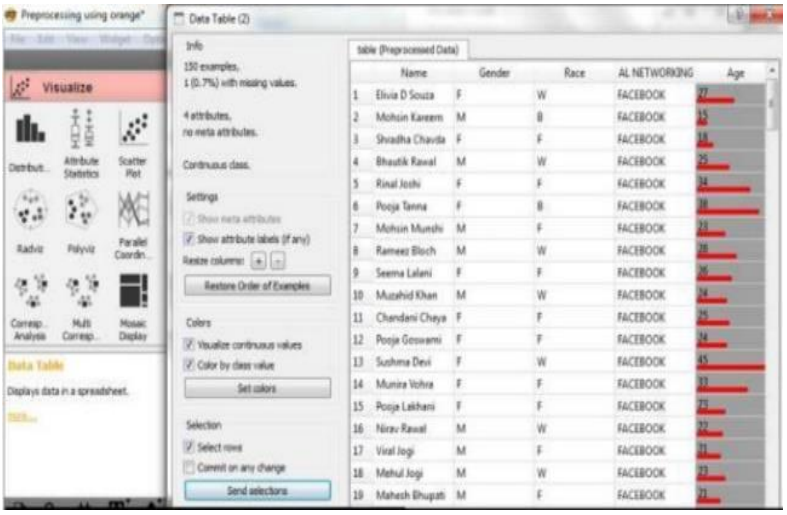


(Table 3)Data analysis using orange

Orange has performed practical of data analysis is a long with some csv (Comma Separated Values) to file data many tool and co- relates with sources which provide to the Data Table.

Data pre-processing:

Data pre-processing is a data mining technique that involves transforming raw data into understandable data.Pre-processing is a crucial for achieving better quality analysis results.The pre-process widget offers several pre-processing methods that can be combined in a single preprocessing pipeline



(Table 5) Data visualization using orange

Result

After analysis practically, orange generates results in numerical or statistical data. It display attribute statistical with mean and median values.

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

Conclusion

In this study of data analysis using data mining tool, comparing their parameters to each other and find out which tool is better to perform best analysis over time. Therefore Orange tool has performed well and easy to use .Moreover after perform practical implementation Orange has done everything as its feature said. This tool makes analysis work easier.

References

1. Kaliyamurthie, K.P., Sivaraman, K., Ramesh, S. Imposing patient data privacy in wireless medical sensor networks through homomorphic cryptosystems 2016, Journal of Chemical and Pharmaceutical Sciences 9 2.

2. Kaliyamurthie, K.P., Balasubramanian, P.C. An approach to multi secure to historical malformed documents using integer ripple transfiguration 2016 Journal of Chemical and Pharmaceutical Sciences 9 2.

3. Kaliyamurthie, K.P., Balasubramanian, P.C. An approach to multi secure to historical malformed documents using integer ripple transfiguration 2016 Journal of Chemical and Pharmaceutical Sciences 9 2.

4. A.Sangeetha,C.Nalini,"Semantic Ranking based on keywords extractions in the web", International Journal of Engineering & Technology, 7 (2.6) (2018) 290-292

5. S.V.GayathiriDevi,C.Nalini,N.Kumar,"An efficient software verification using multi-layered software verification tool "International Journal of Engineering & Technology, 7(2.21)2018 454-457

6. C.Nalini,ShwtambariKharabe,"A Comparative Study On Different Techniques Used For Finger – Vein Authentication", International Journal Of Pure And Applied Mathematics, Volume 116 No. 8 2017, 327-333, Issn: 1314-3395

7. M.S. Vivekanandan and Dr. C. Rajabhushanam, "Enabling Privacy Protection and Content Assurance in Geo-Social Networks", International Journal of Innovative Research in Management, Engineering and Technology, Vol 3, Issue 4, pp. 49-55, April 2018.

8. Dr. C. Rajabhushanam, V. Karthik, and G. Vivek, "Elasticity in Cloud Computing", International Journal of Innovative Research in Management, Engineering and Technology, Vol 3, Issue 4, pp. 104-111, April 2018.

9. K. Rangaswamy and Dr. C. Rajabhushanamc, "CCN-Based Congestion Control Mechanism In Dynamic Networks", International Journal of Innovative Research in Management, Engineering and Technology, Vol 3, Issue 4, pp. 117-119, April 2018.

10. Kavitha, R., Nedunchelian, R., "Domain-specific Search engine optimization using healthcare ontology and a neural network backpropagation approach", 2017, Research Journal of Biotechnology, Special Issue 2:157-166

11. Kavitha, G., Kavitha, R., "An analysis to improve throughput of high-power hubs in mobile ad hoc network" , 2016, Journal of Chemical and Pharmaceutical Sciences, Vol-9, Issue-2: 361-363

12. Kavitha, G., Kavitha, R., "Dipping interference to supplement throughput in MANET" , 2016, Journal of Chemical and Pharmaceutical Sciences, Vol-9, Issue-2: 357-360

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

13. Michael, G., Chandrasekar, A.,"Leader election based malicious detection and response system in MANET using mechanism design approach", Journal of Chemical and Pharmaceutical Sciences(JCPS) Volume 9 Issue 2, April - June 2016 .

14. Michael, G., Chandrasekar, A.,"Modeling of detection of camouflaging worm using epidemic dynamic model and power spectral density", Journal of Chemical and Pharmaceutical Sciences(JCPS) Volume 9 Issue 2, April - June 2016 .

15. Pothumani, S., Sriram, M., Sridhar, J., Arul Selvan, G., Secure mobile agents communication on intranet,Journal of Chemical and Pharmaceutical Sciences, volume 9, Issue 3, Pg No S32-S35, 2016

16. Pothumani, S., Sriram, M., Sridhar , Various schemes for database encryption-a survey, Journal of Chemical and Pharmaceutical Sciences, volume 9, Issue 3, Pg NoS103-S106, 2016

17. Pothumani, S., Sriram, M., Sridhar, A novel economic framework for cloud and grid computing, Journal of Chemical and Pharmaceutical Sciences, volume 9, Issue 3, Pg No S29-S31, 2016

18. Priya, N., Sridhar, J., Sriram, M. "Ecommerce Transaction Security Challenges and Prevention Methods-New Approach" 2016 ,Journal of Chemical and Pharmaceutical Sciences, JCPS Volume 9 Issue 3.page no:S66-S68 .

19. Priya, N.,Sridhar,J.,Sriram, M."Vehicular cloud computing security issues and solutions" Journal of Chemical and Pharmaceutical Sciences(JCPS) Volume 9 Issue 2, April - June 2016 .

20. Priya, N., Sridhar, J., Sriram, M. "Mobile large data storage security in cloud computing environment-a new approach" JCPS Volume 9 Issue 2. April - June 2016

21. Anuradha.C, Khanna.V, "Improving network performance and security in WSN using decentralized hypothesis testing "Journal of Chemical and Pharmaceutical Sciences(JCPS) Volume 9 Issue 2, April - June 2016 .

22. Anuradha.C, Khanna.V, "A novel gsm based control for e-devices" Journal of Chemical and Pharmaceutical Sciences(JCPS) Volume 9 Issue 2, April - June 2016 .

23. Anuradha.C, Khanna.V, "Secured privacy preserving sharing and data integration in mobile web environments " Journal of Chemical and Pharmaceutical Sciences(JCPS) Volume 9 Issue 2, April - June 2016 .

24. Sundarraj, B., Kaliyamurthie, K.P.Social network analysis for decisive the ultimate classification from the ensemble to boost accuracy rates 2016 International Journal of Pharmacy and Technology 8

25. Sundarraj, B., Kaliyamurthie, K.P.A content-based spam filtering approach victimisation artificial neural networks2016 International Journal of Pharmacy and Technology 8 3.

26. Sundarraj, B., Kaliyamurthie, K.P. Remote sensing imaging for satellite image segmentation 2016 International Journal of Pharmacy and Technology 8 3.

27. Sivaraman, K., Senthil, M. Intuitive driver proxy control using artificial intelligence 2016 International Journal of Pharmacy and Technology 8 4.

28. Sivaraman, K., Kaliyamurthie, K.P. Cloud computing in mobile technology 2016 Journal of Chemical and Pharmaceutical Sciences9 2.

29. 29.Sivaraman, K., Khanna, V. Implementation of an extension for browser to detect vulnerable elements on web pages and avoid clickjacking2016 Journal of Chemical and Pharmaceutical Sciences