Appendix A Alyuda NeuroIntelligence 2.2 (**Tool**)

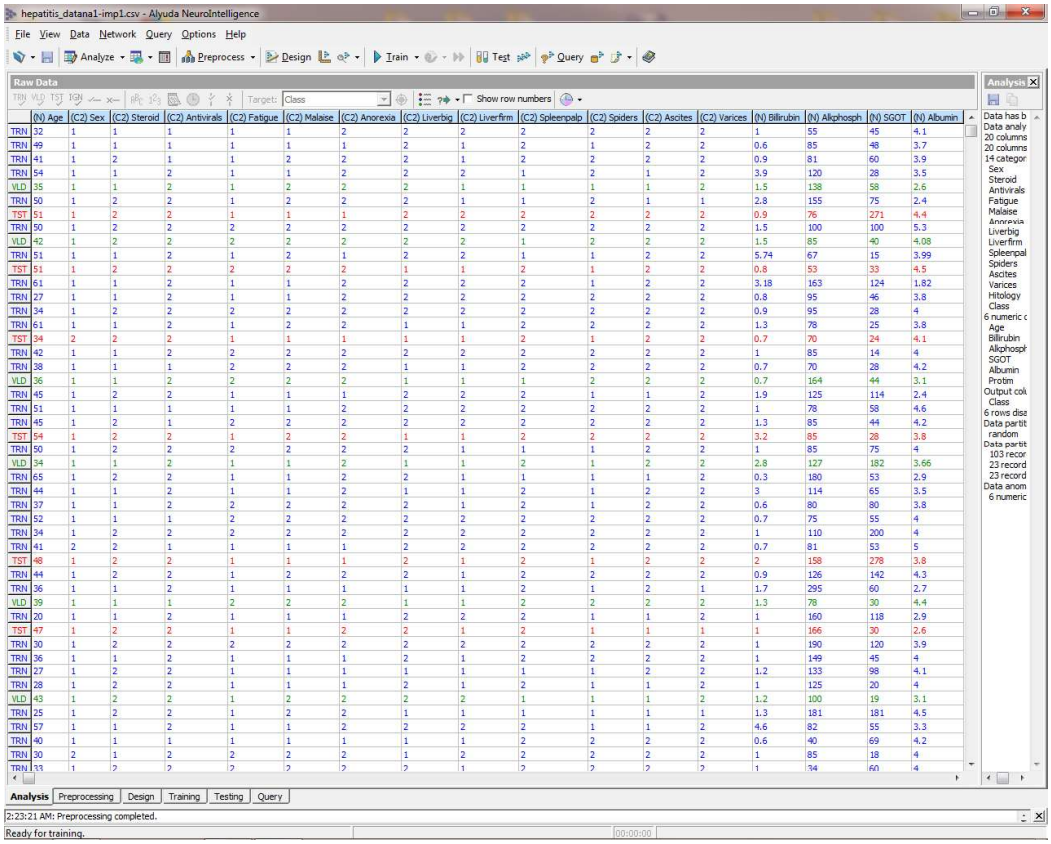


Figure A1. Screen of Alyuda NeuroIntelligence 2.2

Figure A1. is the screen of Alyuda NeuroIntelligence 2.2, which has been used as a tool for implemen ng neural network for solving problems related to classi ca on. It is a neural network so ware. Tasks that can be done using this tool are as given below:

- **Data Analysis**: It is a step before data preprocessing. In this step, one can select target column, specify type of a ributes, iden fy anomalies in data, par on data onto training, valida on, and test set, exclude or reject selected a ribute, and instances.

- **Data Preprocessing**: The internal representa on of data set can be inspected in this step. The sta s cal details about each a ribute can be viewed.

- **Network Design**: This feature helps in designing neural network architecture. The node func ons, classi ca on model, and accept / reject levels can be de ned in this phase.

- **Network training**: This feature helps to train the network based on training data set using di erent training / learning algorithms as chosen by user. Training graph, error distribu on, and real- me training details can also be monitored. One can observe the whole training process visually.

- **Tes ng the Network**: Di erent benchmarking parameters to judge the performance of the network are represented in this phase. Confusion matrix, correct classi ca on accuracy, and area under ROC curve are some of the parameters to be recorded to judge the predic ve performance of the model.

- **Query**: In this phase the outcomes of a trained model can be viewed by entering query either manually or by loading in speci ed format. Results are provided in tabular form including response curve.

Appendix B ROSETTA (**Tool**)

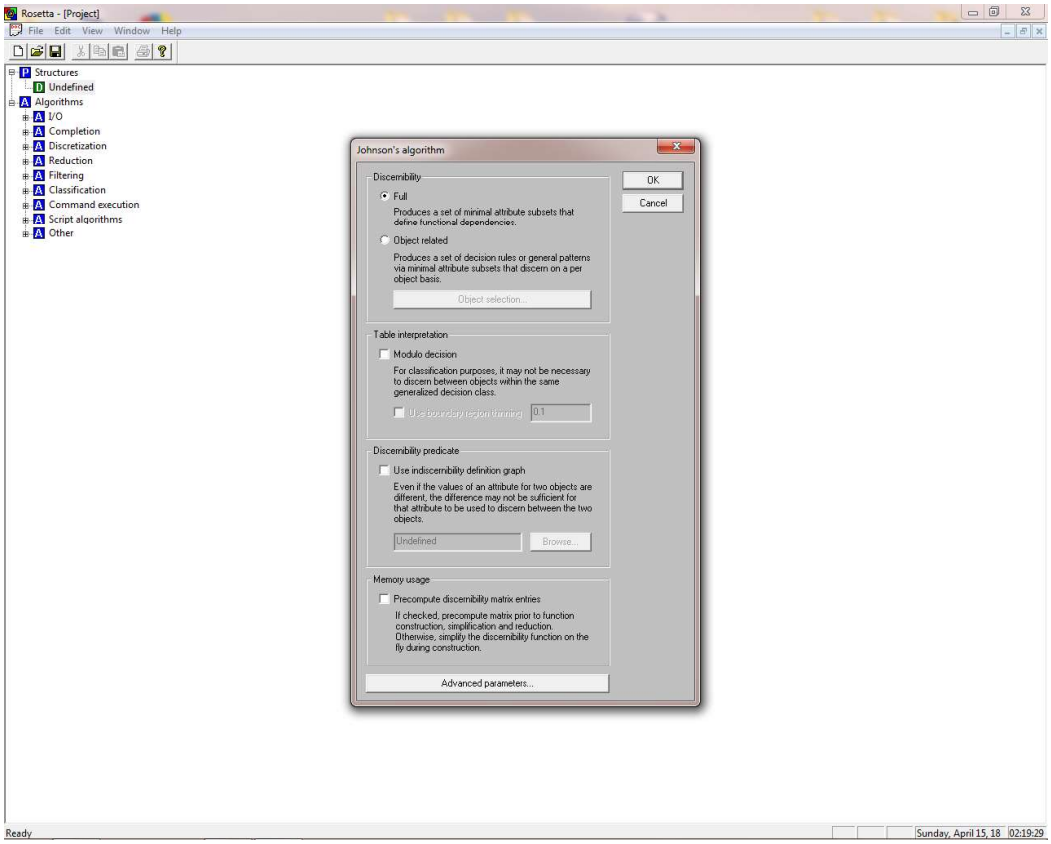


Figure B1. Screen of ROSETTA

Figure B1. represents screen of ROSETTA, a so ware system, which was used as a tool to implement rough set to extract reducts from data set. Some of the features of ROSETTA are as follows:

- Discre sa on of data set using di erent algorithms can be done.
- Data set can be reduced by using genetic algorithm, Johnson's algorithm, and Johnson's algorithm (RSES). Some other algorithms are also available to reduce data set.
- This so ware can also be used to classify data set using di erent algorithms, such as standard vo ng, object tracking voter, Naive Bayes etc. Outcomes are provided by construc ng confusion matrix and by returning ROC curve.

- Sta s cal informa on about each a ribute in the data set such as mean, median, standard devia on, and correla on can be obtained. Based on internal representa on the correla ons of string a ributes are also computed.
- The so ware provides the facility to write commands in a script le for execu on.
- The lter op on in the so ware is used to remove elements from reduct sets based on di erent evalua on criteria.
- This so ware can also be used to generate rules from a set of reducts.

Appendix C WEKA (Tool)

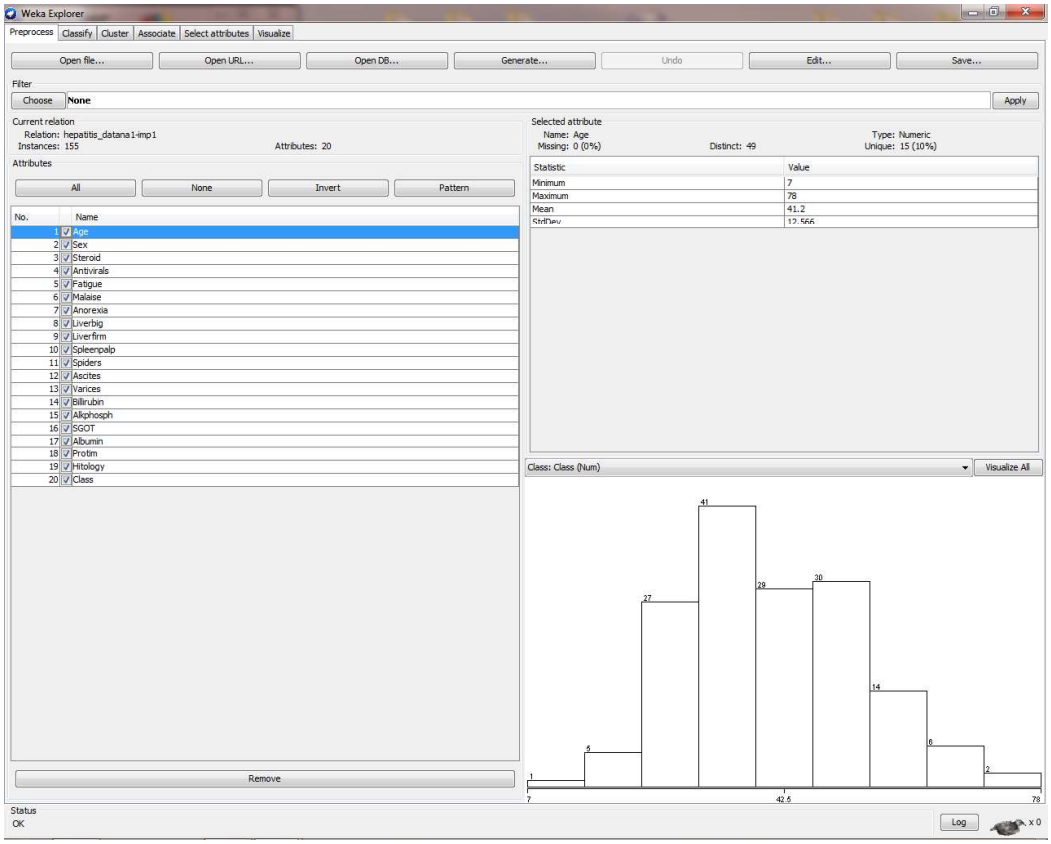


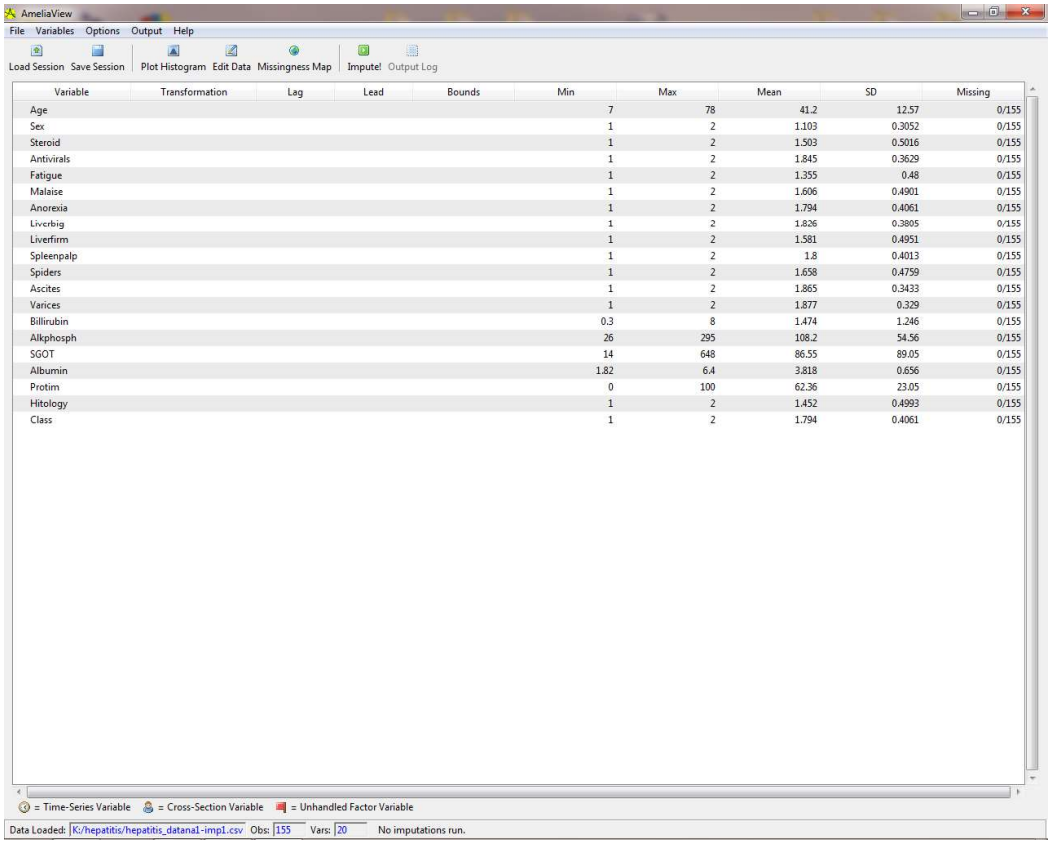
Figure C1. Screen of WEKA

Figure C1. shows screen of WEKA, which has been used to transform input data, e.g. data preprocessing, data transforma on, feature selec on and so on. Classi ers are also available in WEKA. Some features of WEKA are as follows:

- It supports data transforma on, which is an important step in machine learning, e.g. It can be used to discre ze numeric data, transform nominal a ributes to binary.
- It can be used to select important features from the given data set using di erent feature selec on methods, such as correla on based feature subset selec on, principal component analysis, and many more.
- It can be used to learn clusters for the data set.

- Associa on rules for the data set can be learnt by using this so ware.
- Classify tab in WEKA so ware allows to implement various classi ers such as Naive Bayes, support vector machine etc. for classi ca on.
- Di erent 2D plots of the input data set can be viewed using the tool.
- The func onali es of the package can be accessed by both via the graphical user interface (GUI) and command line interface (CLI).

Appendix D Amelia View (Tool)



Variable	Transformation	Lag	Lead	Bounds	Min	Max	Mean	SD	Missing
Age					7	78	41.2	12.57	0/155
Sex					1	2	1.103	0.3052	0/155
Steroid					1	2	1.503	0.5016	0/155
Antivirals					1	2	1.845	0.3629	0/155
Fatigue					1	2	1.355	0.48	0/155
Malaise					1	2	1.606	0.4901	0/155
Anorexia					1	2	1.794	0.4061	0/155
Liverbig					1	2	1.826	0.3805	0/155
Liverfirm					1	2	1.581	0.4951	0/155
Spleenpalp					1	2	1.8	0.4013	0/155
Spiders					1	2	1.658	0.4759	0/155
Ascites					1	2	1.865	0.3433	0/155
Varices					1	2	1.877	0.329	0/155
Billirubin					0.3	8	1.474	1.246	0/155
Alkphosph					26	295	108.2	54.56	0/155
SGOT					14	648	86.55	89.05	0/155
Albumin					1.82	6.4	3.818	0.656	0/155
Protim					0	100	62.36	23.05	0/155
Hitology					1	2	1.452	0.4993	0/155
Class					1	2	1.794	0.4061	0/155

Figure D1. Screen of Amelia View

Figure D1. gives the screen of Amelia View, which has been used to handle missing values in the data set. It is a vital step to avoid distorted analysis and to avoid biased results. The package was used to implement mul ple imputa ons using EMB approach to impute unobserved or missing values in the data set. It generates more than one imputed data set from a data set containing missing values. Missing map and two dimensional plot of data can also be viewed. The so ware also provides a report on the imputa on task.

Appendix E Publica ons of Malay Mitra

1. R. K. Samanta, **Malay Mitra**, "*An Intelligent Hepa s Disease Classi ca on System Using So Compu ng*", ECSS-12, Na onal Conference on Electronics, Communica on, Sensing & Signal Processing Technologies, SIEM, Siliguri, November' 2012.

2. R. K. Samanta, **Malay Mitra**, "*A Neural Network Based Intelligent System for Breast Cancer Diagnosis*", Proc. of International Conference on Intelligent Infrastructure, CSI-2012, Kolkata, Published by Tata McGraw Hill Educa on pp. Pvt. Ltd., pp. 20 – 25, December' 2012.

3. **Malay Mitra**, R. K. Samanta, "*Cardiac Arrhythmia Classi ca on using Neural Network with Selected Features*", International Conference on Computational Intelligence: Modeling, Techniques and Applica ons, CIMTA-2013, Procedia Technology (ELSEVIER), Vol. 10, pp. 76 – 84, September' 2013.

4. **Malay Mitra**, R. K. Samanta, "*Hepa s Disease Classi ca on using So Compuing and Mul ple Imputa on*", Seminar Proc. National Seminar on AI & It's Impact on Modern IT World, Gour Mahavidyalaya, Malda & Ananda Chandra College, Jalpaiguri, pp. 117 – 135, December' 2013.

5. **Malay Mitra**, R. K. Samanta, "*An Intelligent Hybrid System for Breast Cancer Classi ca on*", Proc. National Conference on Research Trends in Computer Science & Applica on, Published by BONFIRING, pp. 21 – 27, February' 2014.

6. **Malay Mitra**, R. K. Samanta, "*Hepa s Disease Diagnosis using Mul ple Imputa on and Neural Network with Rough Set Feature Reduc on*", Proc. Interna onal Conference on Fron ers in Intelligent Compu ng Theory and Applica on, FICTA – 2014, Bhubaneswar Engineering College, Springer, Vol. 327, pp. 285 – 293, November' 2014.

7. Rajesh Misir, **Malay Mitra**, R. K. Samanta, "*A Study on Benchmarking Parameters for Intelligent Systems*", International Journal of Computer Sciences and Engineering, Vol. 3, Issue 1, pp. 10 – 17, February' 2015.

8. **Malay Mitra**, R. K. Samanta, "*A Survey on UCI Hepa s Disease Dataset using So Compu ng*", Proc. International Conference on Computing and Systems, ICCS 2016, University of Burdwan, pp. 45 – 50, January' 2016.

9. **Malay Mitra**, R. K. Samanta, "*A Study on Missing Data Management*", Interna onal Journal of Computer Sciences and Engineering, Vol. 5, Issue 2, pp. 30 – 33, February' 2017.

10. Rajesh Misir, **Malay Mitra**, R. K. Samanta, "*A Reduced Set of Features for Chronic Kidney Disease Predic on*", International Journal of Pathology Informa cs USA, Vol. 8, pp. 1 – 5, June' 2017.

11. **Malay Mitra**, R. K. Samanta, *"A Study on UCI Hepatitis Disease Dataset using Soft Compu ng*", A.M.S.E. France Journals – AMSE IIETA Publica on 2017 Series, Modeling C, Vol. 78, Issue 4, pp. 467 – 477, July' 2017.