CS 771 Presentation

Classifying Heart Sounds Challenge

By: Group 12
Atulya Shivam Shree
Rajesh Shubhankar
Richa Sharma
Rishabh Nigam

Challenge

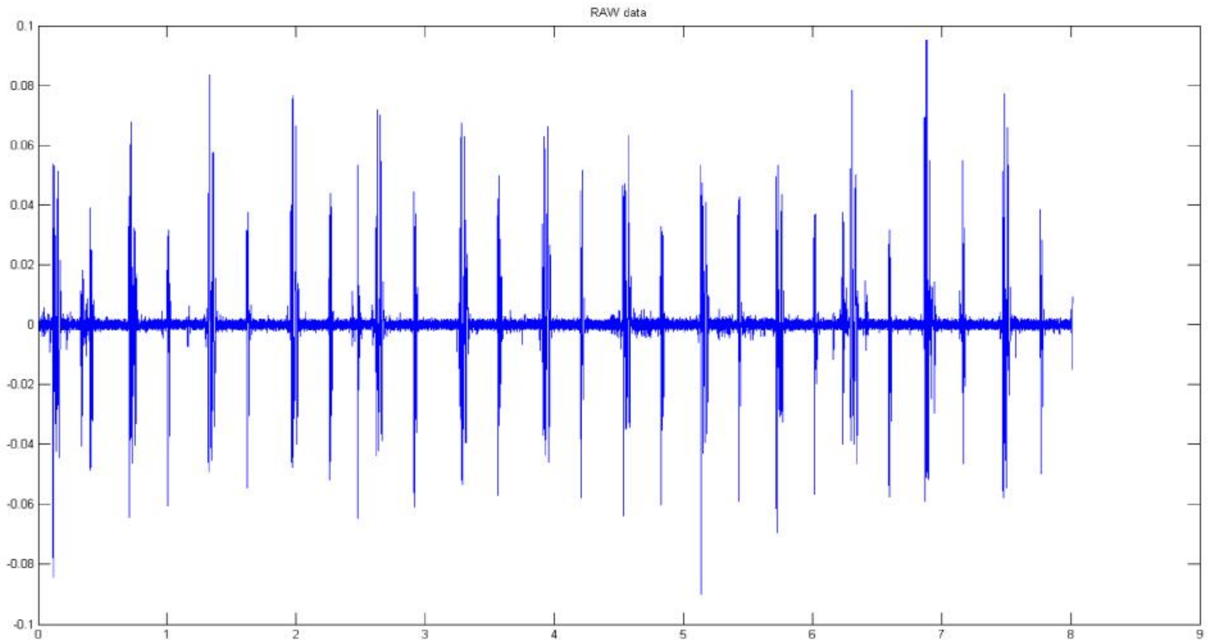
Heart beat data – using

- Clinical Trial – dataset B
- IPhone app – dataset A

Challenge1 – Segmentation of heart data to find S1 and S2 peaks.
Challenge2 – Classification of heart sound into normal, murmur, extra heart sound and Artifact for dataset A and normal, murmur, extrasystole for dataset B.

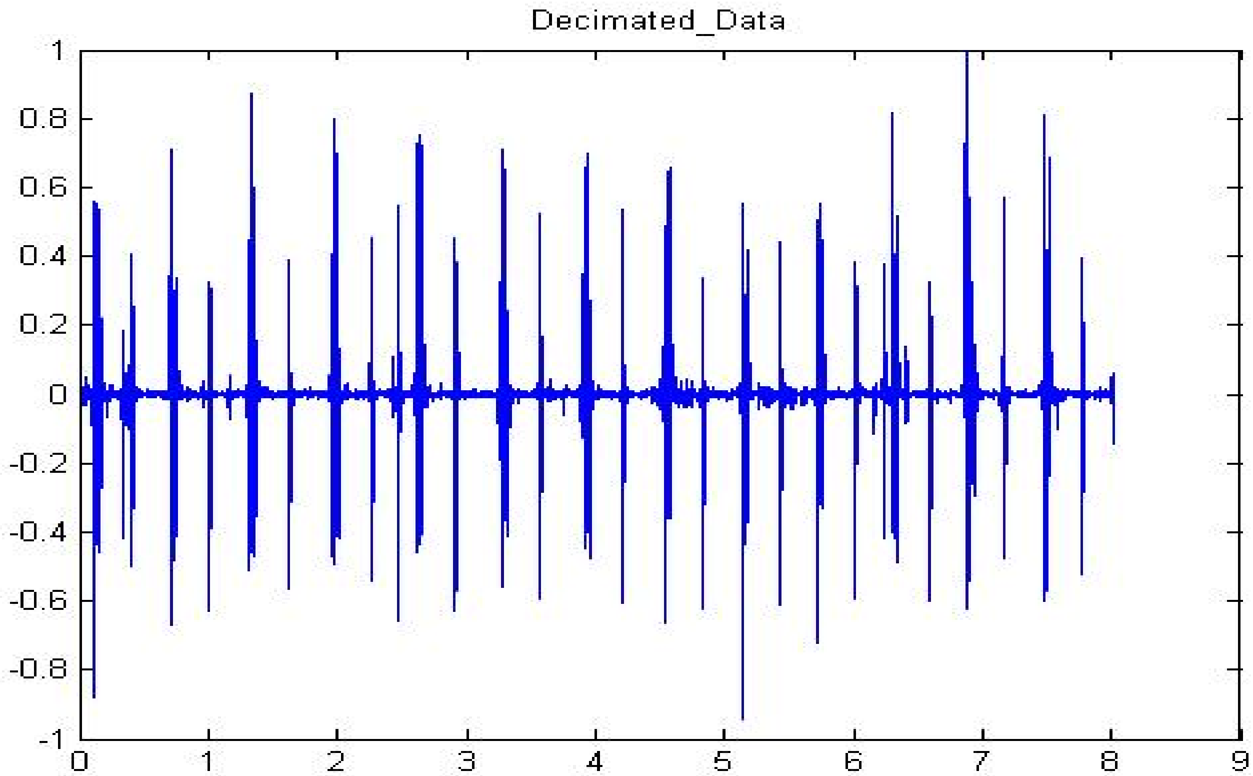
Dataset

- Dataset available in .wav and .aif format. We use data from .wav format.
- Dataset A contains data at 44100 frames per second, Dataset B at 4000 frames per second
- Training File: segmentation data is provided in a csv for 21 and 90 files for dataset A and B respectively



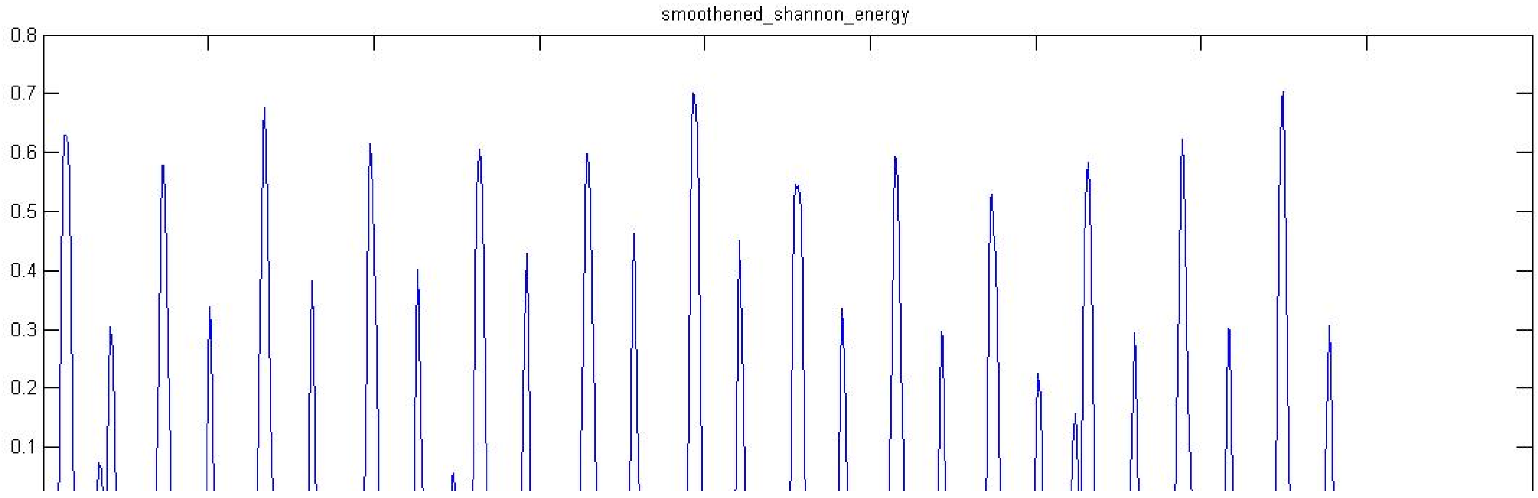
Challenge1 : Step1 decimation

- Decimate the signal by a factor of 20 for dataset A and for 2 for dataset B. This brings down the frequency close to 2000 for both the datasets.
- Normalize the signal to (-1,1)



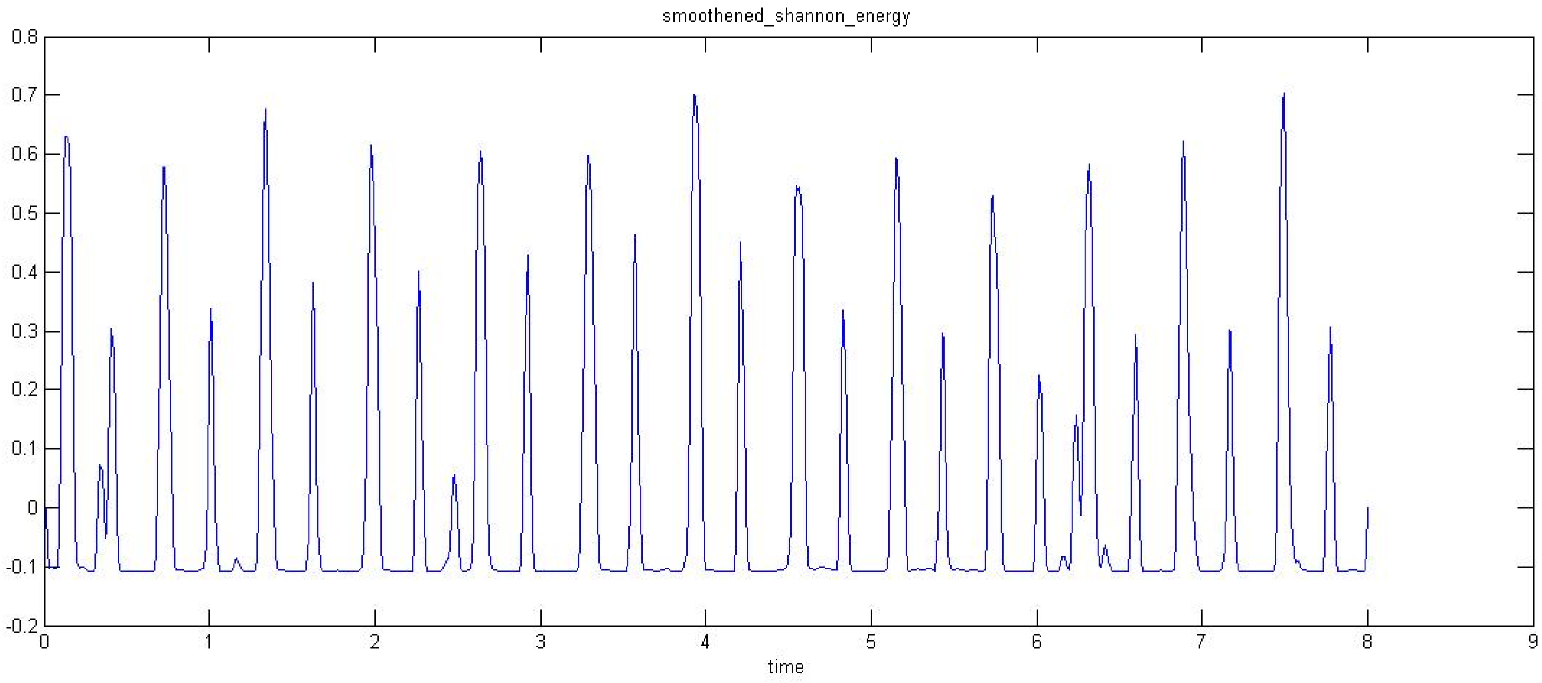
Challenge1: Step2 Shannon Energy

- Shannon energy is calculated for continuous 0.02 seconds with an overlap of 0.01 second.
- Shannon Energy is $-\frac{1}{n}\sum_{i=1}^{N} x^2(i) * \log x^2(i)$

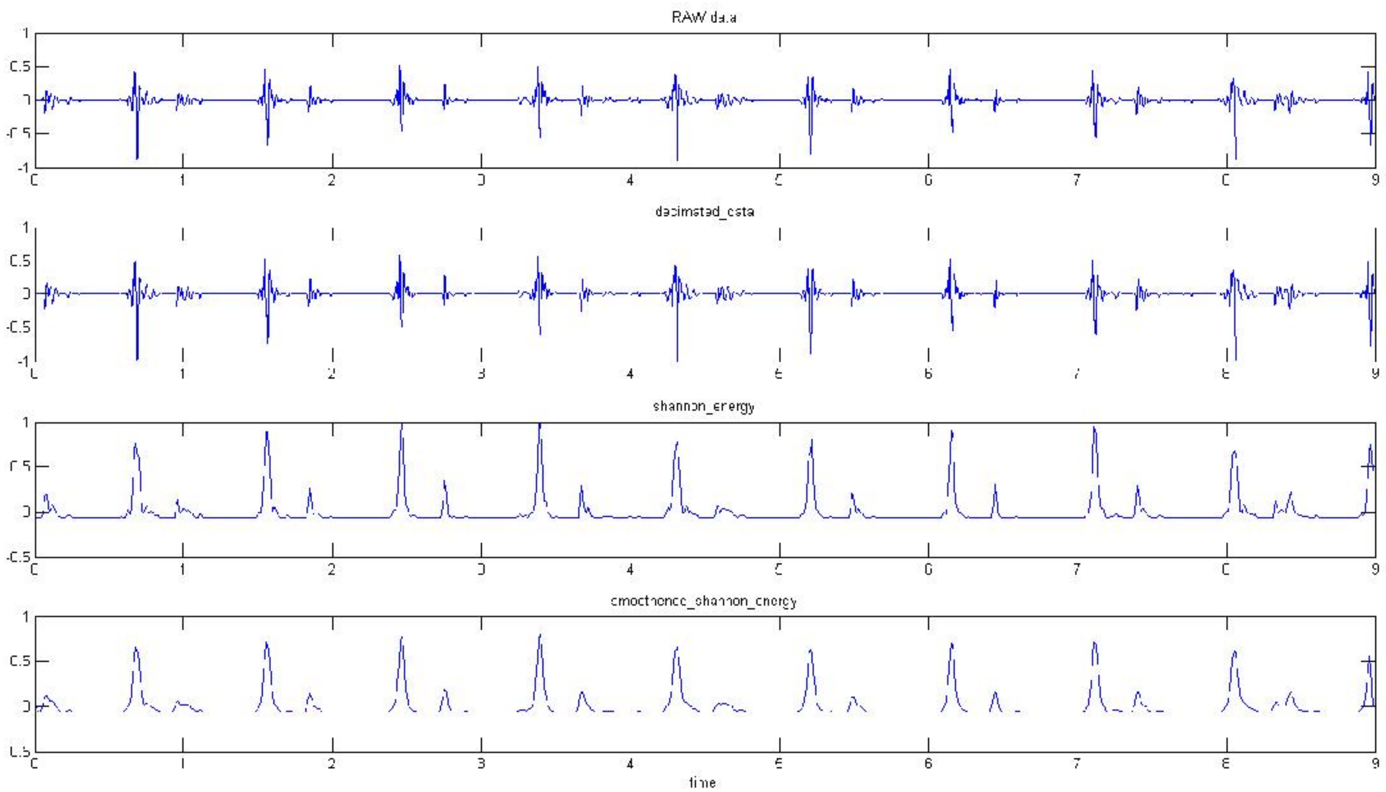


Challenge1: Step3 smoothening

- We use "triangular smooth" to smoothen the signal

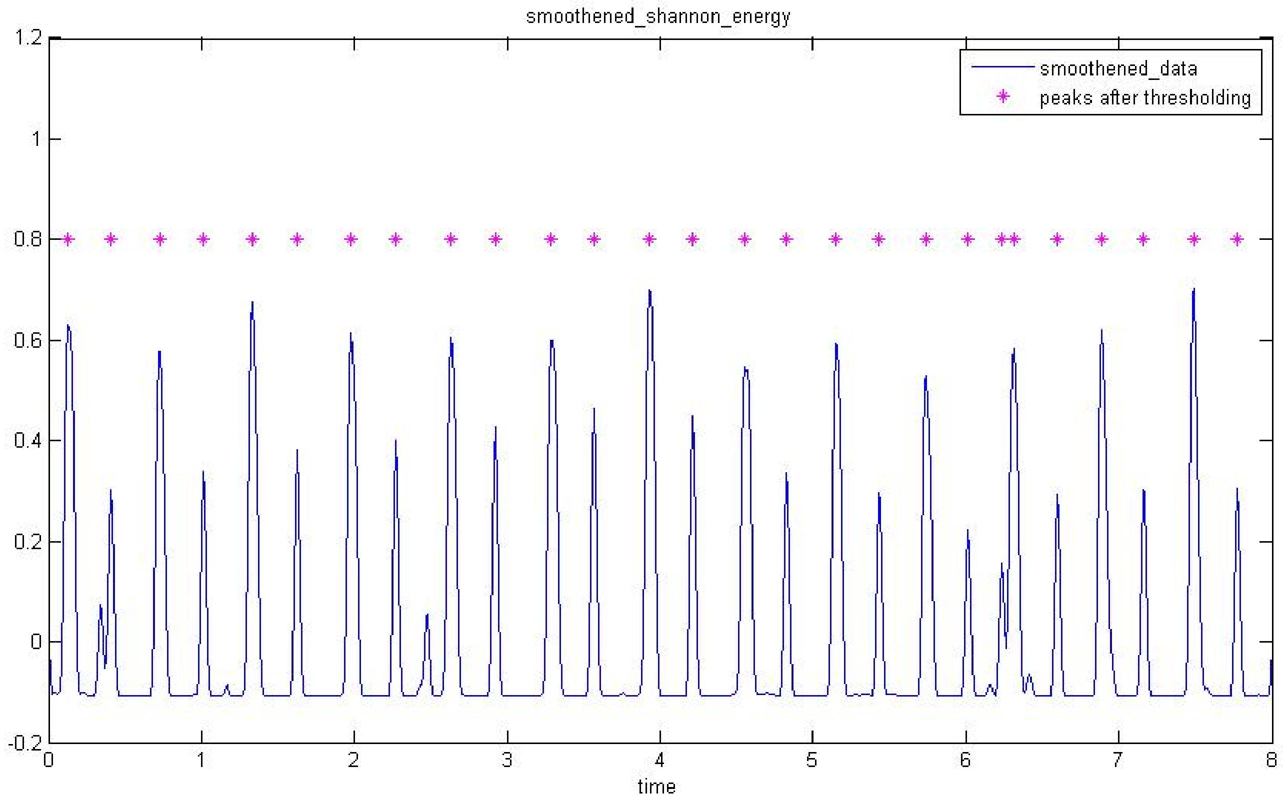


Complete preprocessing

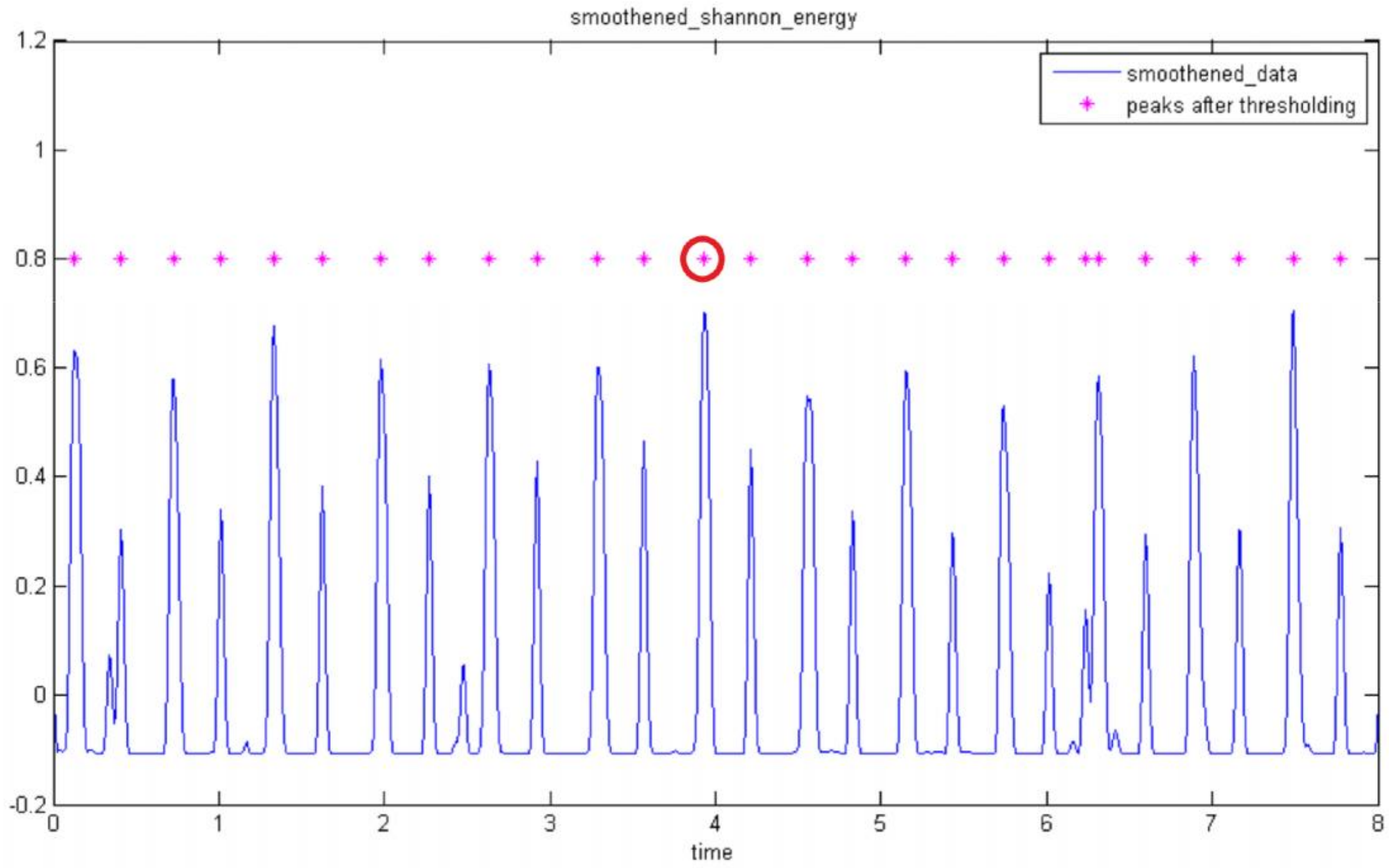


Challenge1: Step4a getPeaks

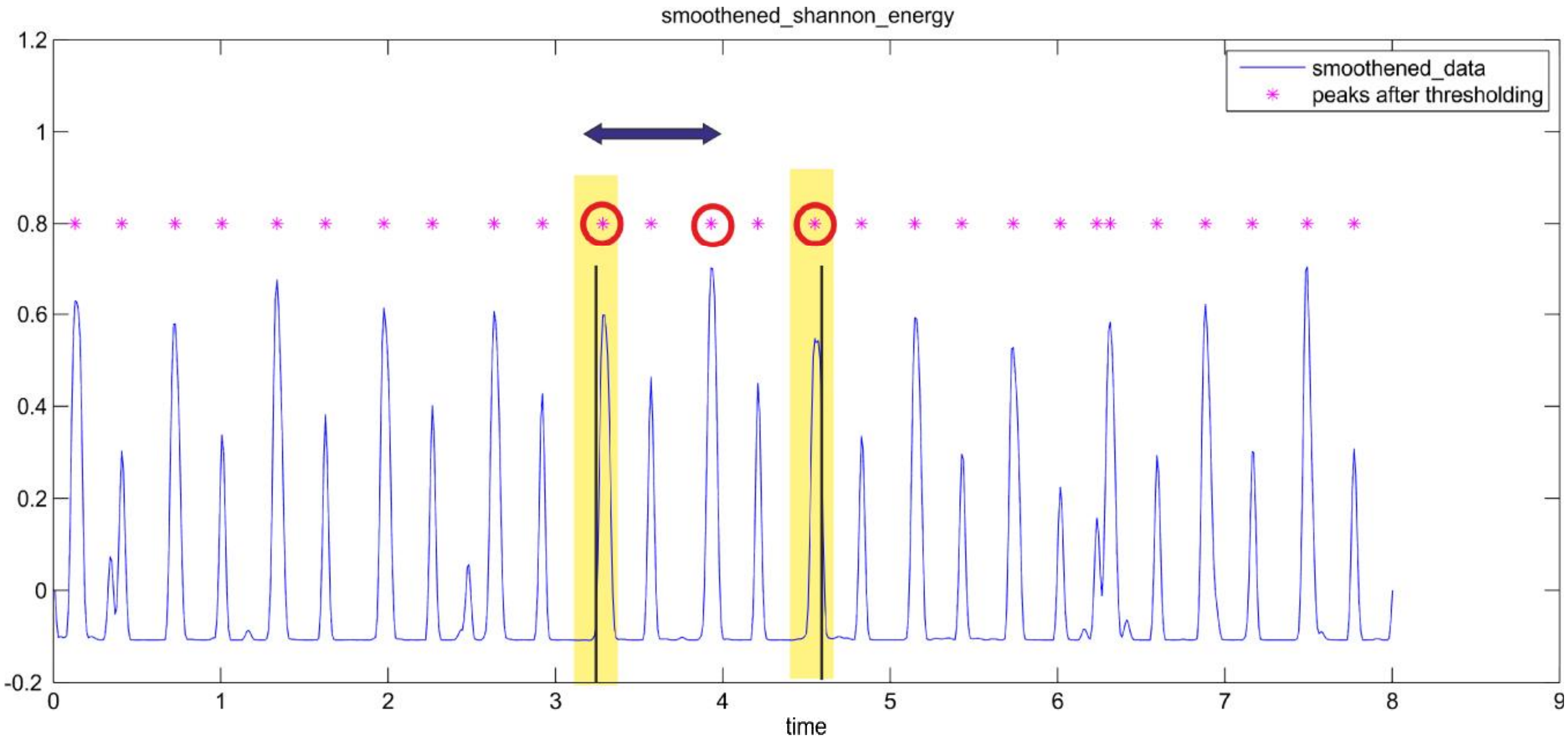
- We use a threshold and extract all values above the threshold. Then we select the best peak among the neighbouring peaks within a range.



Algorithm: getS1Peaks

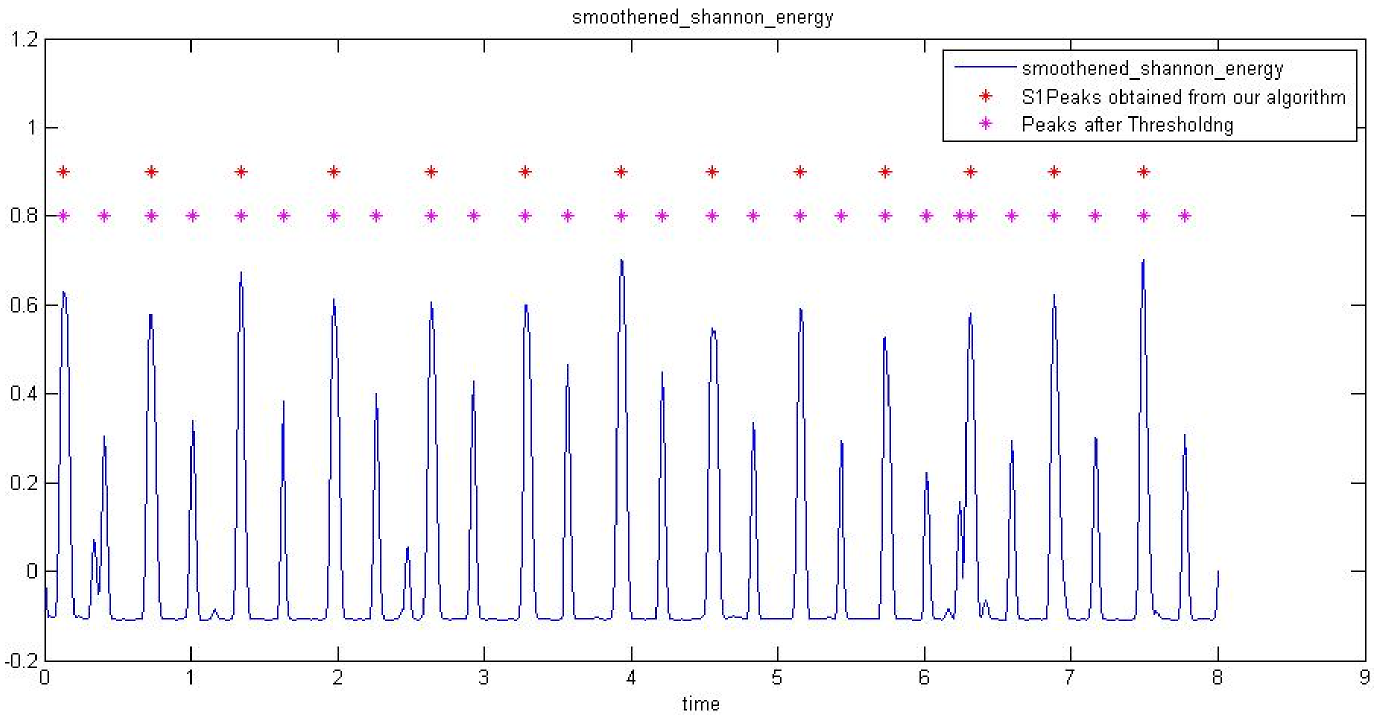


Algorithm: getS1Peaks



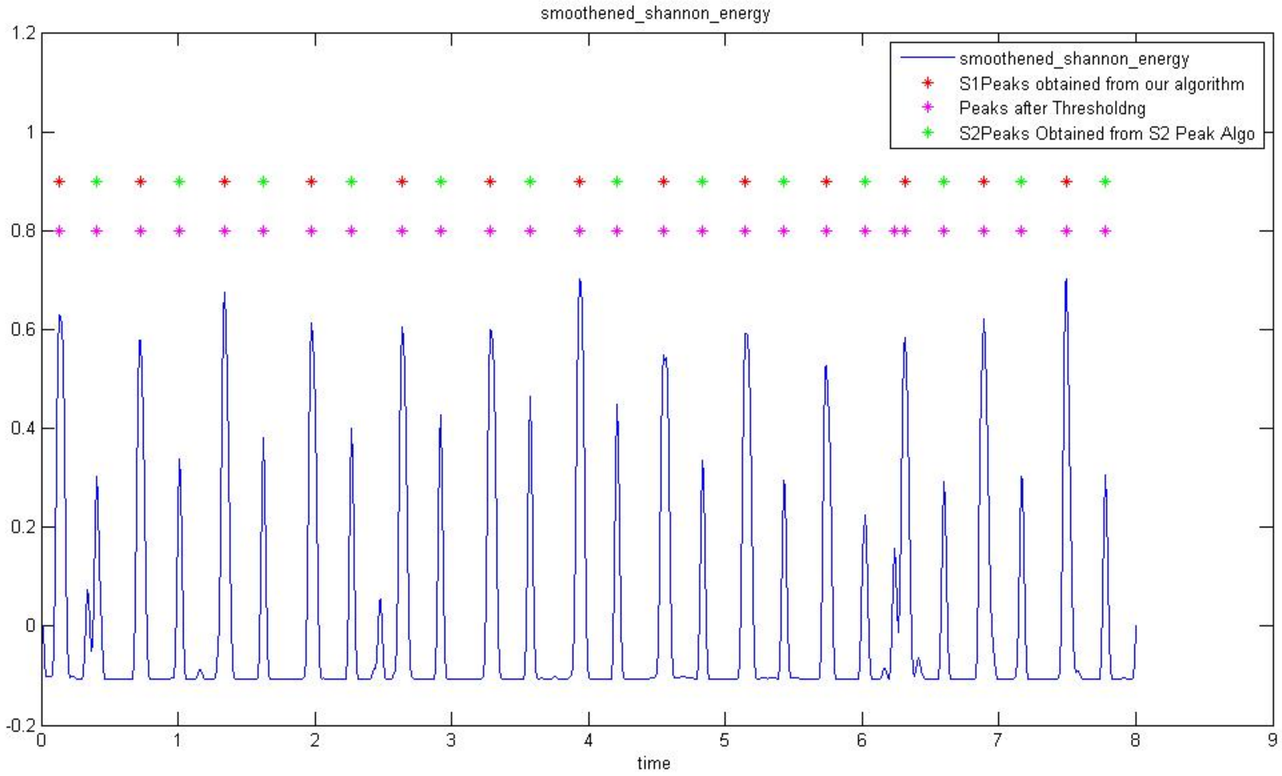
Challenge1: Step4b getS1Peaks

- In this we start with the assumption the S1peaks have higher energy. We start by picking the peak with the maximum energy. Then we look at peaks within distance of one time period of it, and so on. We vary the time period and pick the best set of S1Peaks.



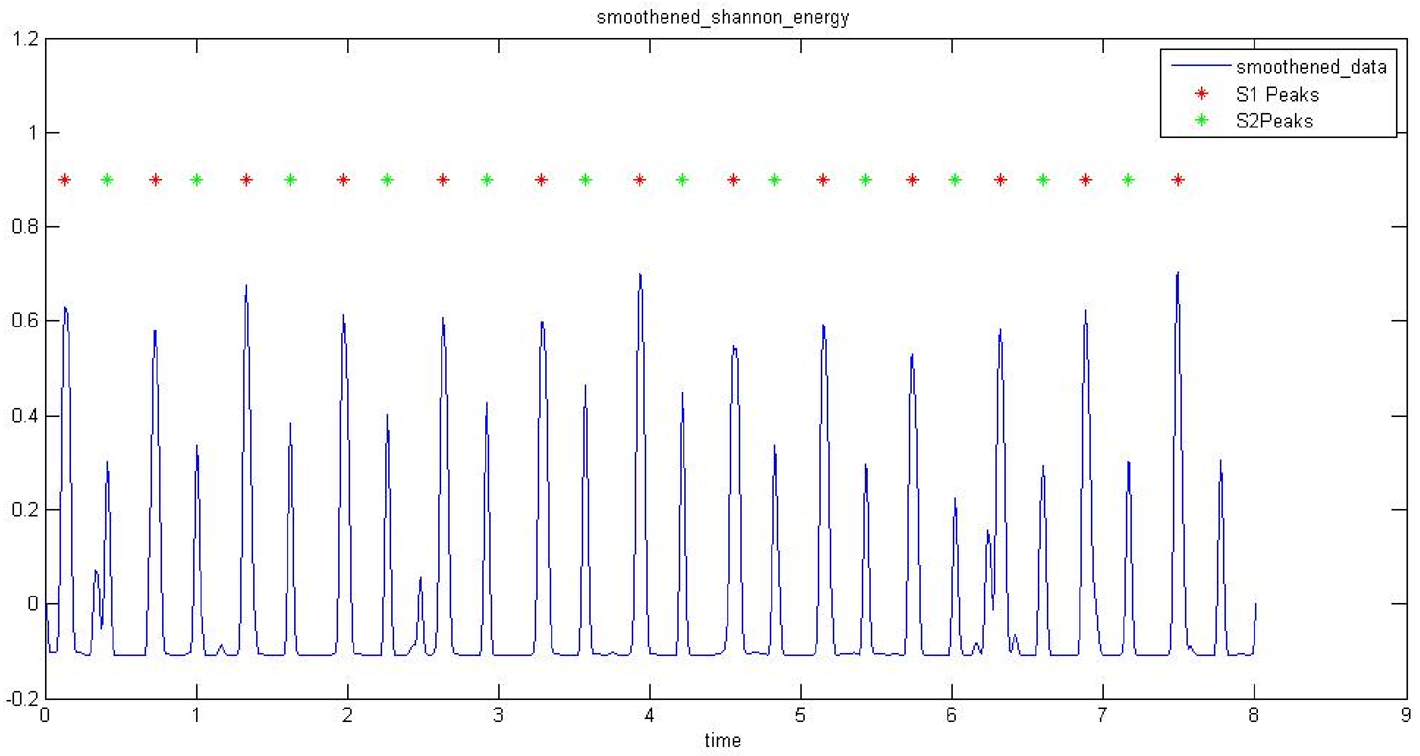
Challenge1: Step4c getS2Peaks

- Look between 2 S1 peaks, and pick the best peak. If no peak is found we lower the threshold and look again.

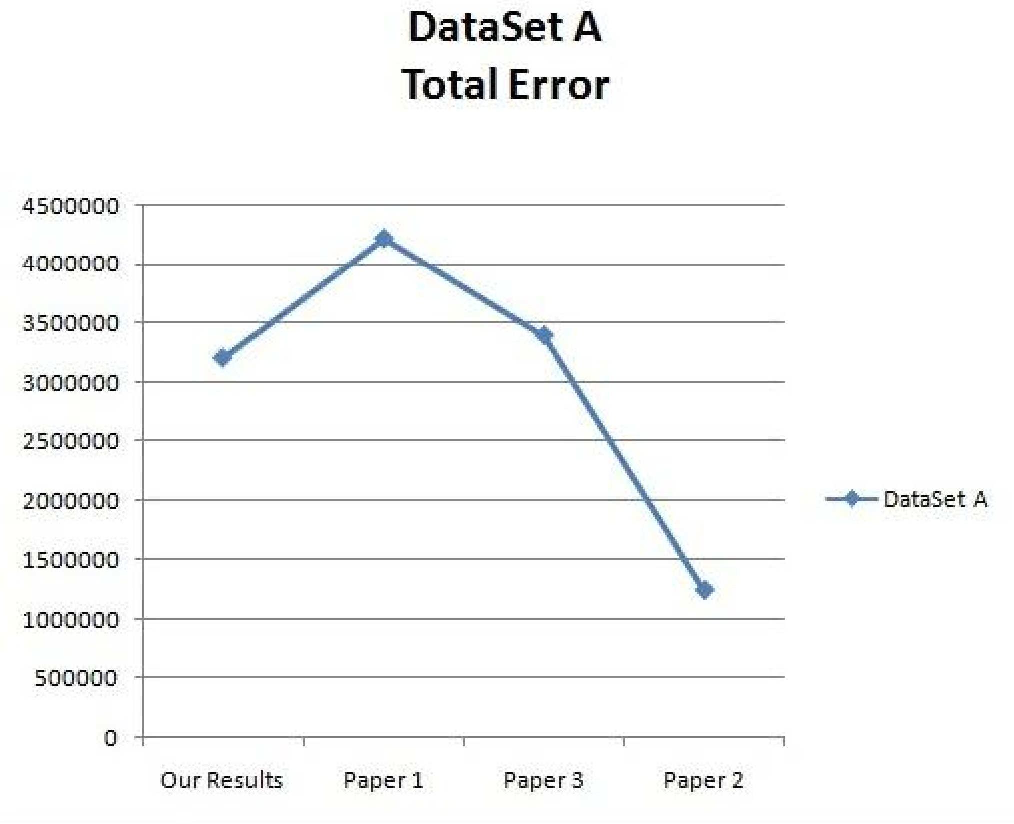


Challenge1: Step4d FixPeaks

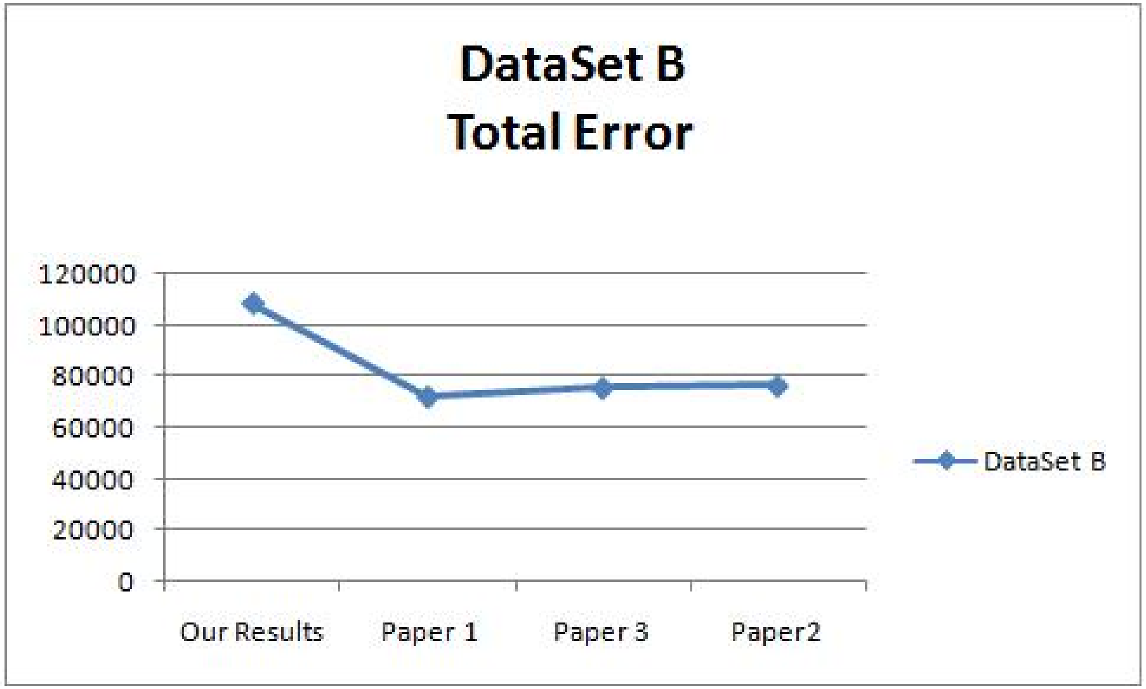
- Fix peaks uses the length of diastolic and systolic period to fix which is S1 and S2.



Result comparison



Result comparison



Challenge2: AttributeSet

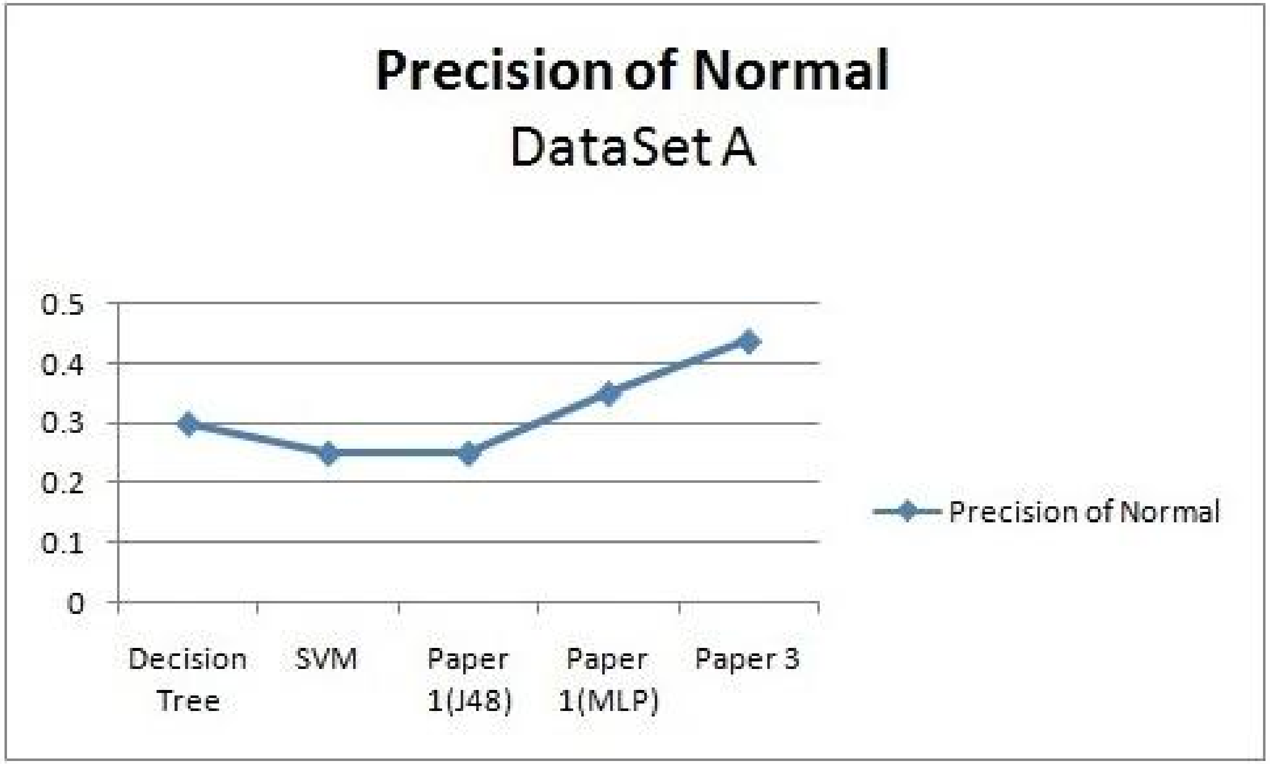
We use 14 attributes to classify heart data.

1. best_frequency;
2. Systolic_period;
3. Diastolic_period;
4. Diastolic_period_variance;
5. Systolic_period_variance;
6. number_of_peaks_after_findmaxpeak;
7. number_of_peaks_finally;

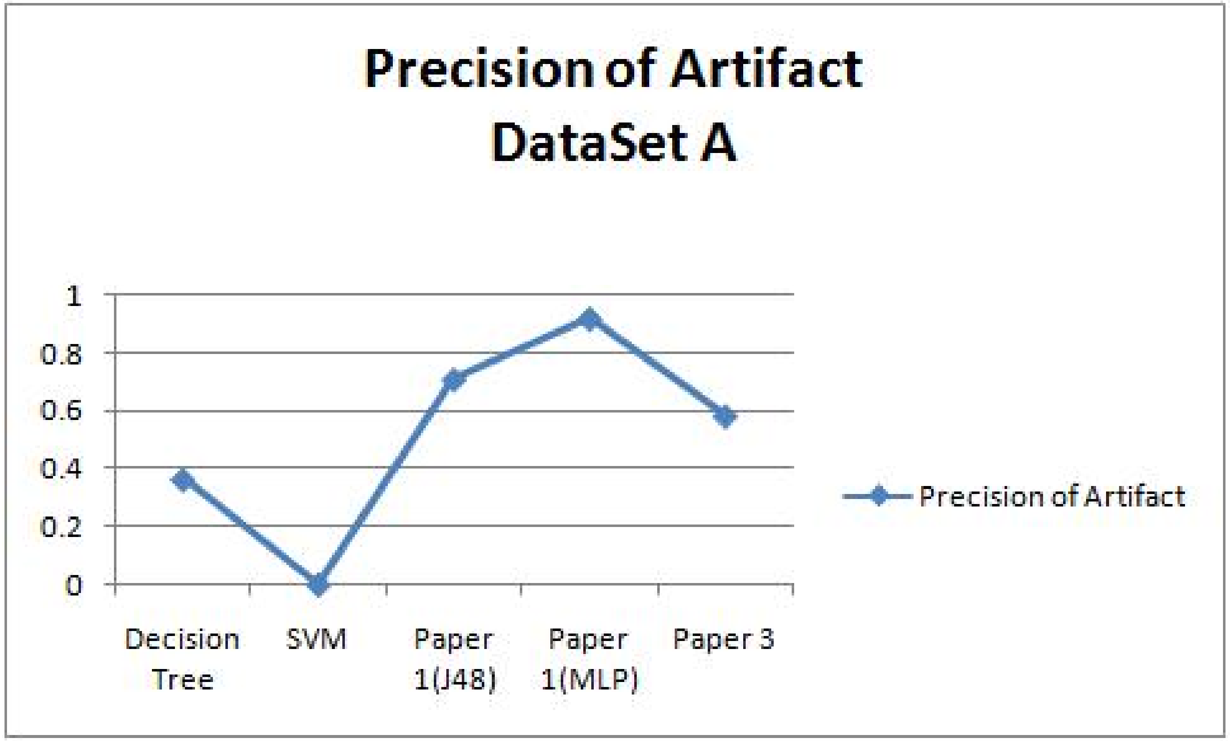
Challenge2: AttributeSet

8. S1Peaks_Energy;
9. S2Peaks_Energy;
10. Extra_Peaks_Energy;
11. systolicPeriodEnergy
12. diastolicPeriodEnergy
13. time_period_variance
14. ratio_threshhold_peaks_to_final_peaks

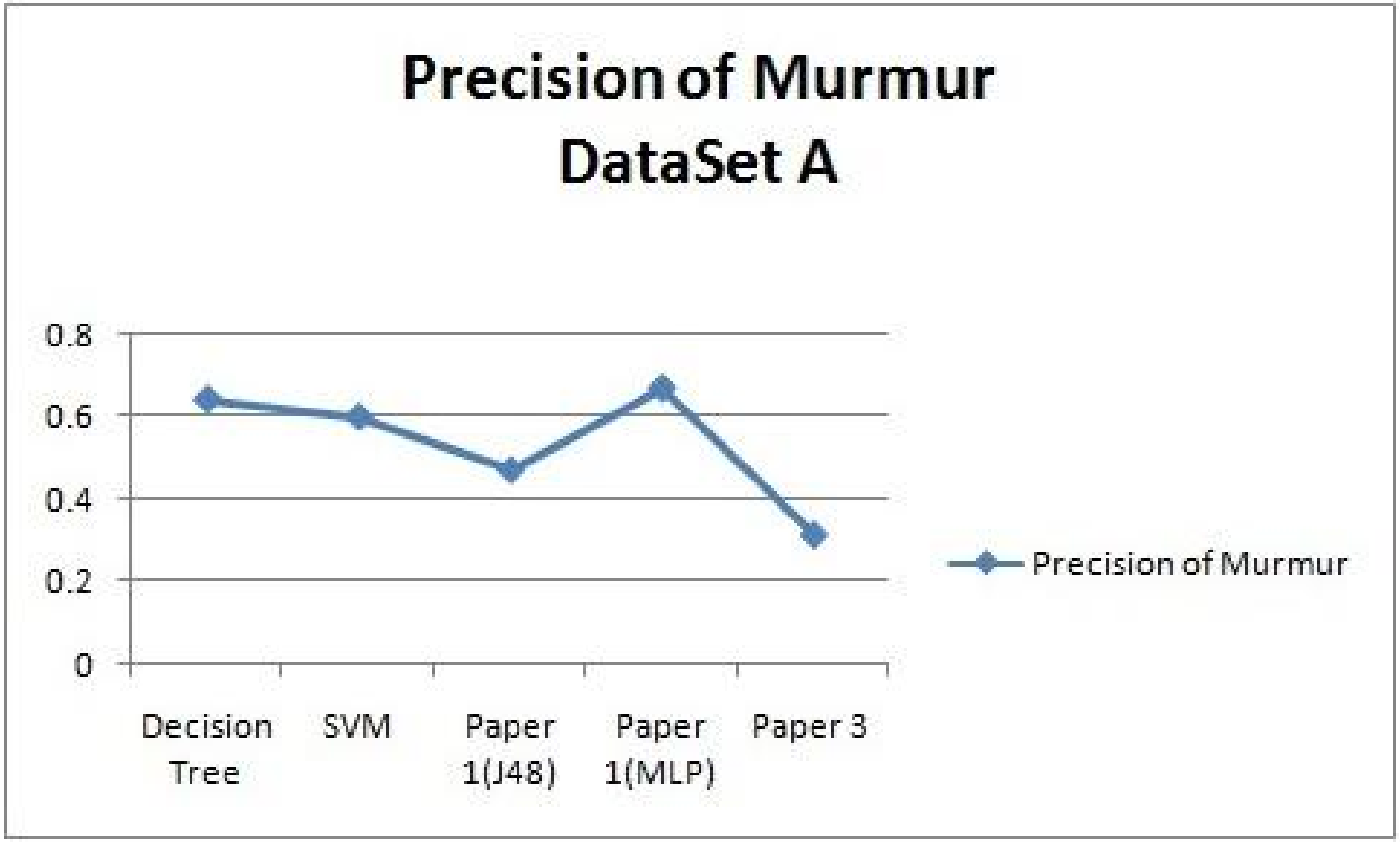
Challenge2: Results dataset1



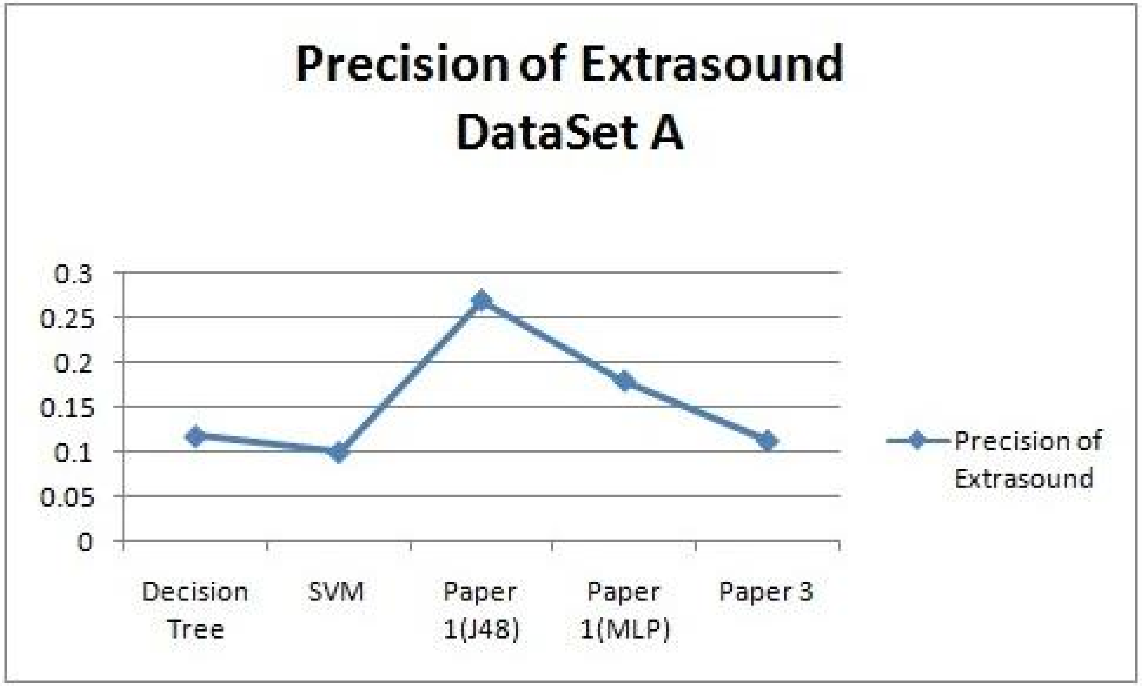
Challenge2: Results dataset1



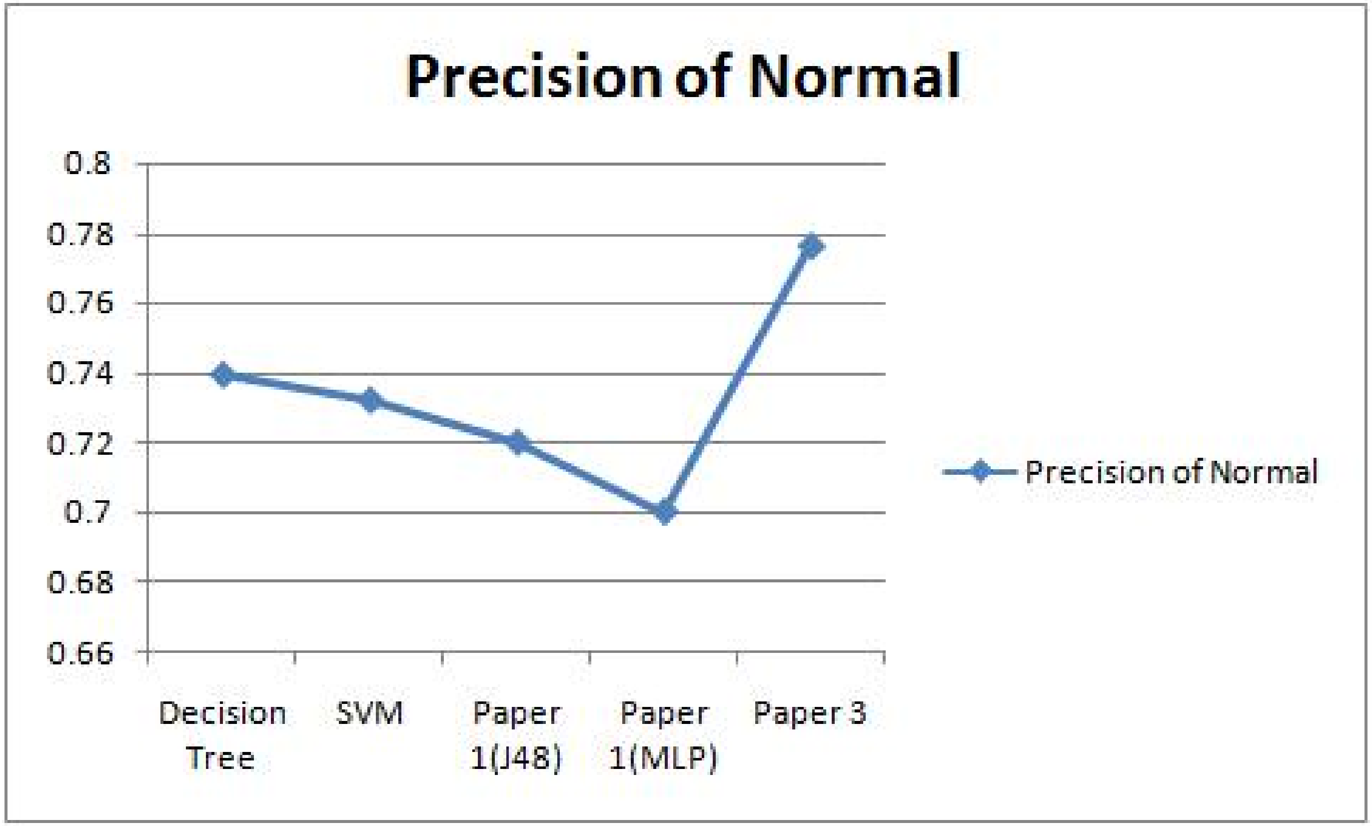
Challenge2: Results dataset1



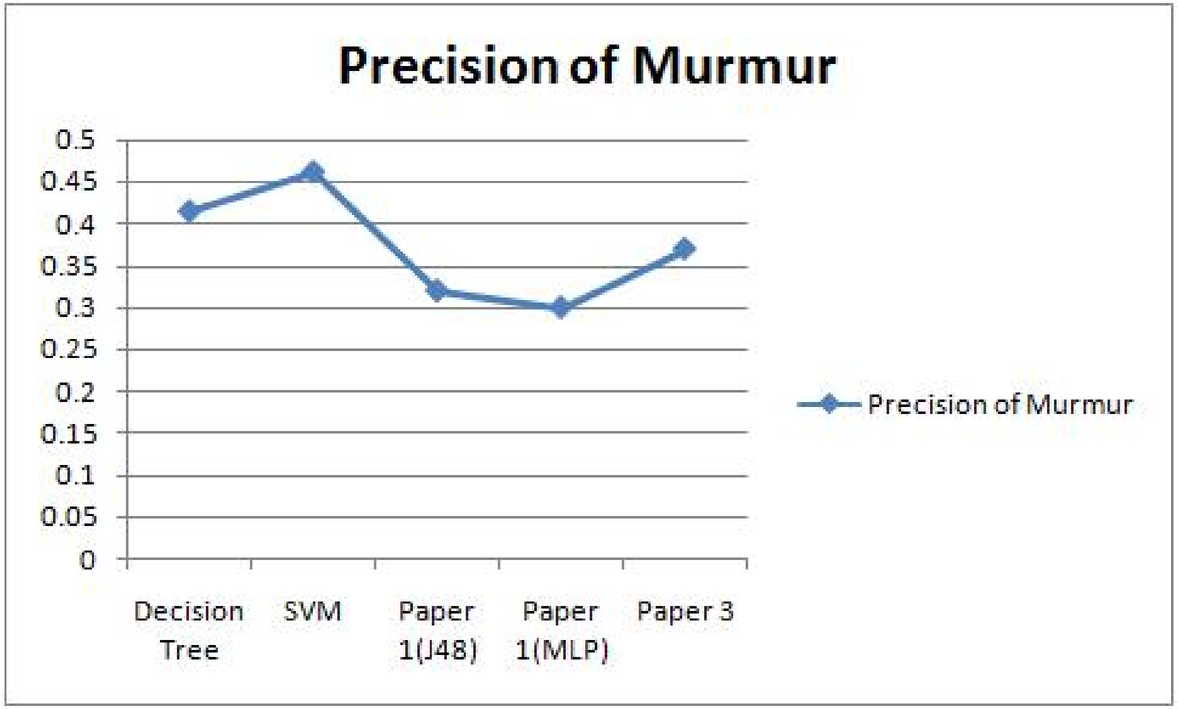
Challenge2: Results dataset1



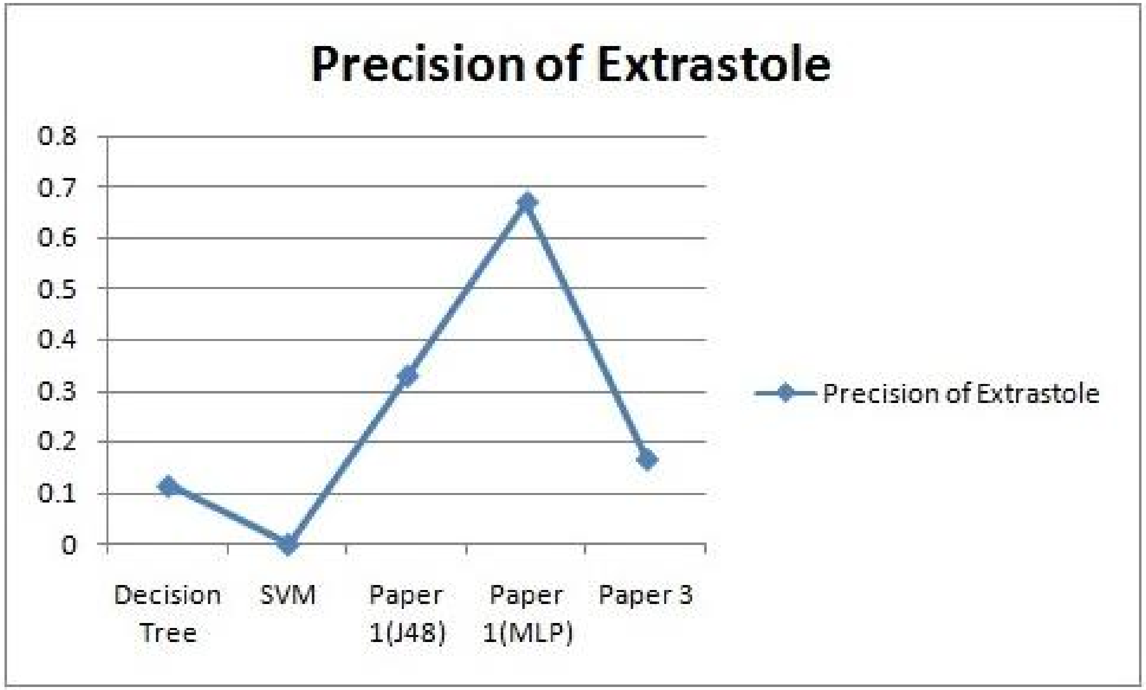
Challenge2: Results dataset2



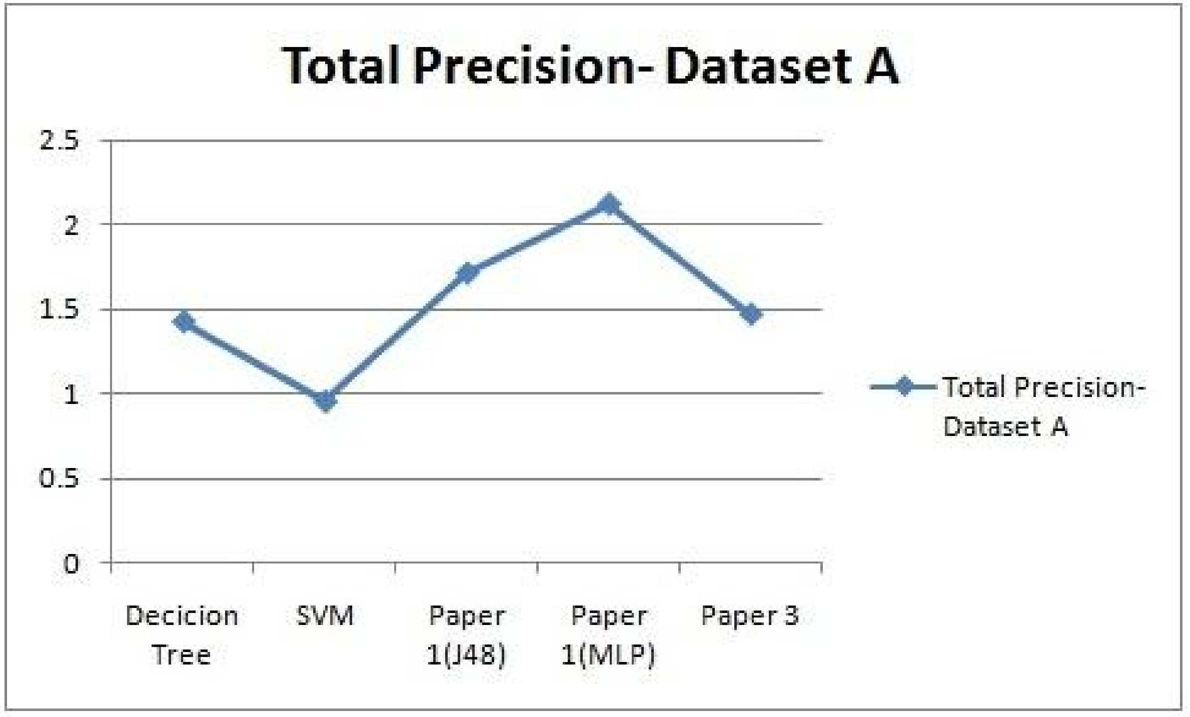
Challenge2: Results dataset2



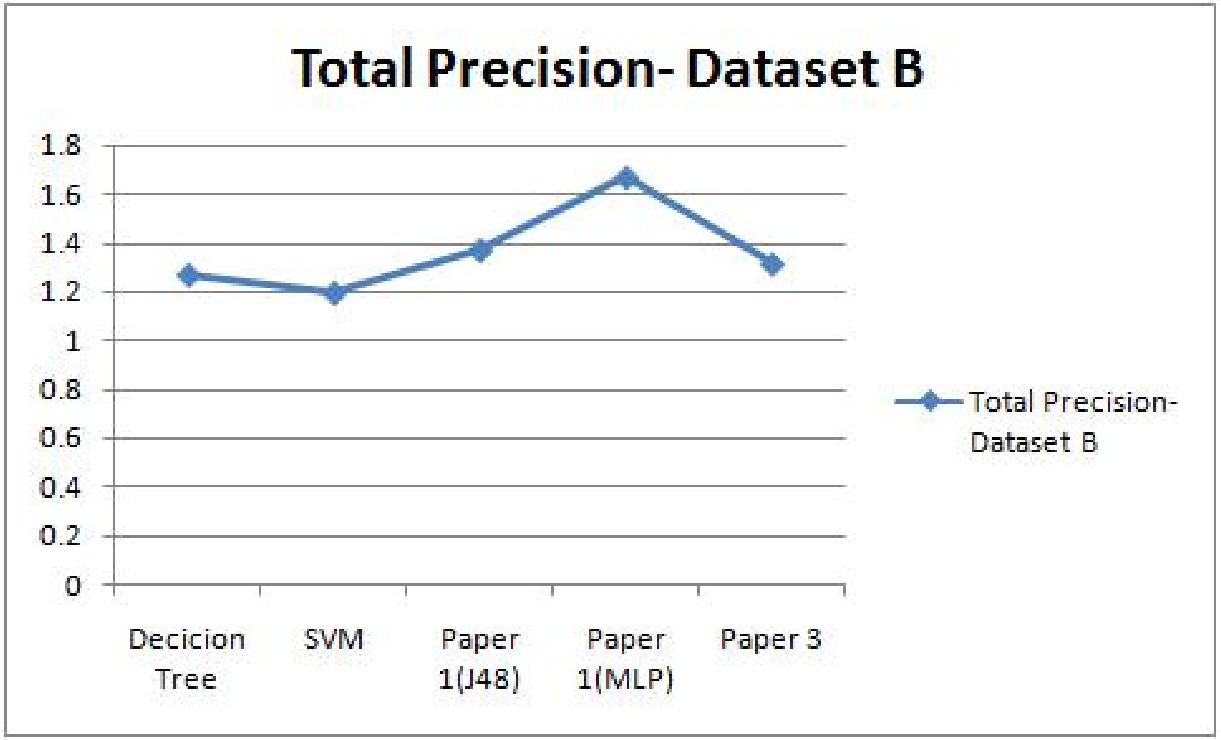
Challenge2: Results dataset2



Challenge2: SummaryResults



Challenge2: SummaryResults



Results using 5 fold validations: Support Vector Machine

Dataset B

- Correctly Classified Instances 69.2308 %
- Incorrectly Classified Instances 30.7692 %
- Mean absolute error 0.3177
- Root mean squared error 0.4117
- Relative absolute error 90.8775 %
- Root relative squared error 98.6397 %
- Total Number of Instances 312

Dataset A

- Correctly Classified Instances 54.8387 %
- Incorrectly Classified Instances 45.1613 %
- Mean absolute error 0.3051
- Root mean squared error 0.3894
- Relative absolute error 82.9609 %
- Root relative squared error 90.8359 %
- Total Number of Instances 124

Results using 5 fold validations: Neural Networks

Dataset B

- Correctly Classified Instances 64.7436 %
- Incorrectly Classified Instances 35.2564 %
- Mean absolute error 0.2705
- Root mean squared error 0.4267
- Relative absolute error 77.3767 %
- Root relative squared error 102.232 %
- Total Number of Instances 312

Dataset A

- Correctly Classified Instances 54.8387 %
- Incorrectly Classified Instances 45.1613 %
- Mean absolute error 0.2406
- Root mean squared error 0.4187
- Relative absolute error 65.4115 %
- Root relative squared error 97.6673 %
- Total Number of Instances 124

References

- Heart Sound Segmentation Algorithm Based on Heart Sound Envelolgram .H Liang, S Lukkarinen, I Hartimo .Helsinki University of Technology, Espoo, Finland.
- Sapire DW. Understanding and diagnosing paediatric heart disease: Heart sounds and murmurs. Norwalk, Connecticut, Applcton & Langc 1992: 27-43.
- A Robust Heart Sound Segmentation and Classification Algorithm using Wavelet Decomposition and Spectrogram. Yiqi Deng , Peter J Bentley. Dept. of Computer Science, UCL Malet Place, London.
- Classifying heart sounds using peak location for segmentation and feature construction. Emanuel Pereira,Elsa Ferreira Gomes. Institute of Engineering (ISEP/IPP) Porto,Portugal

Code

- Will be available at
  http://github.com/rishabhnigam31/heartSegmentation.git

- Thank You