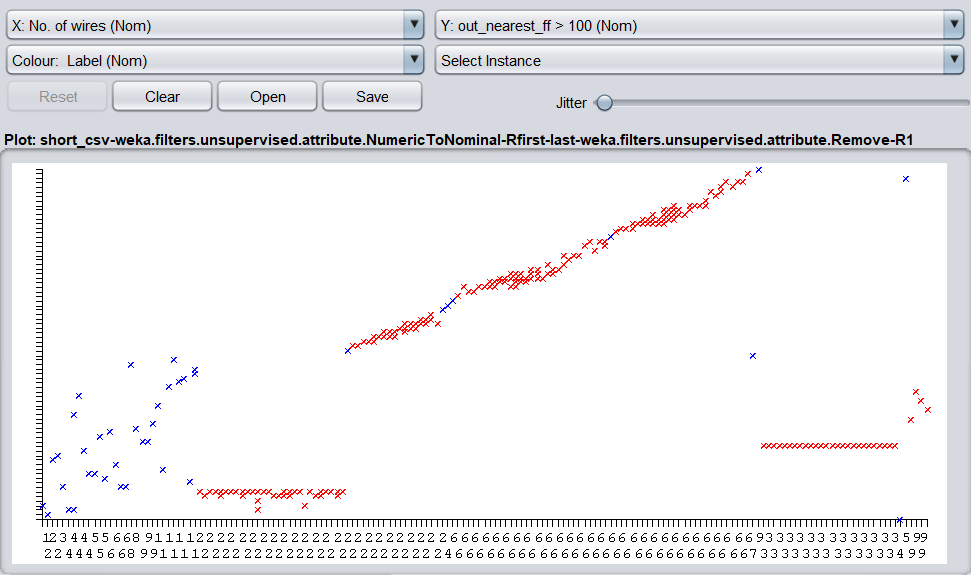
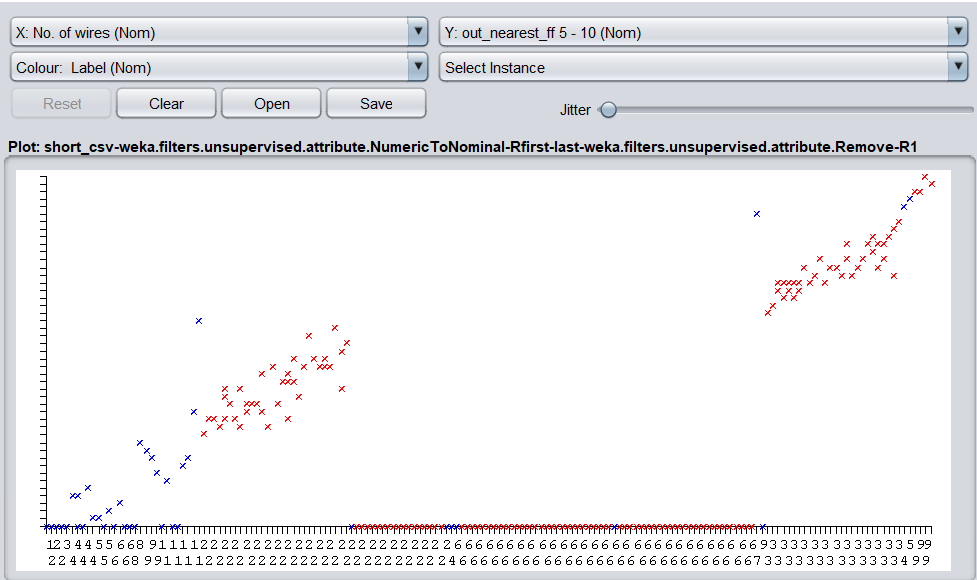
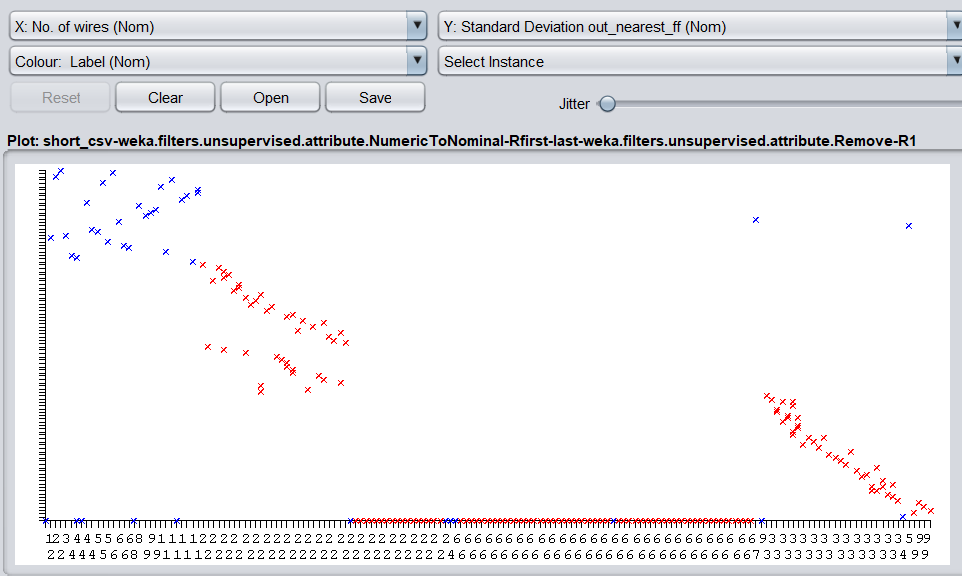
In this phase of the project, features were extracted from circuit bench files. A total of 605 structural, as well as functional features, were extracted from the circuit bench files which included features like rareness value, fan in, fan out, CC, Co etc these features where distilled from a collection of 1571 initial features that we were able to extract from the bench files. The collection of both structural, as well as functional features, were extracted by using multiple circuit simulations, loops and internal functions to extract values as accurate as possible. In our experiments, we used a combination of different simulation techniques and methods to get the most accurate values.

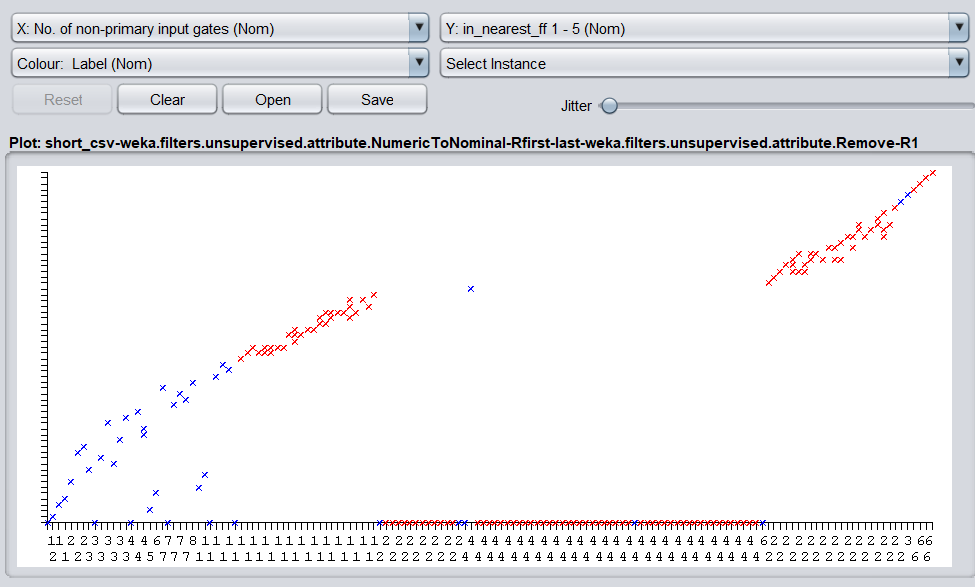
The accuracy and effectiveness of extracted features were verified by comparing the extracted features against our simulation results. In our experimentation, the accuracy of our results was high enough to be able to perform further analysis.

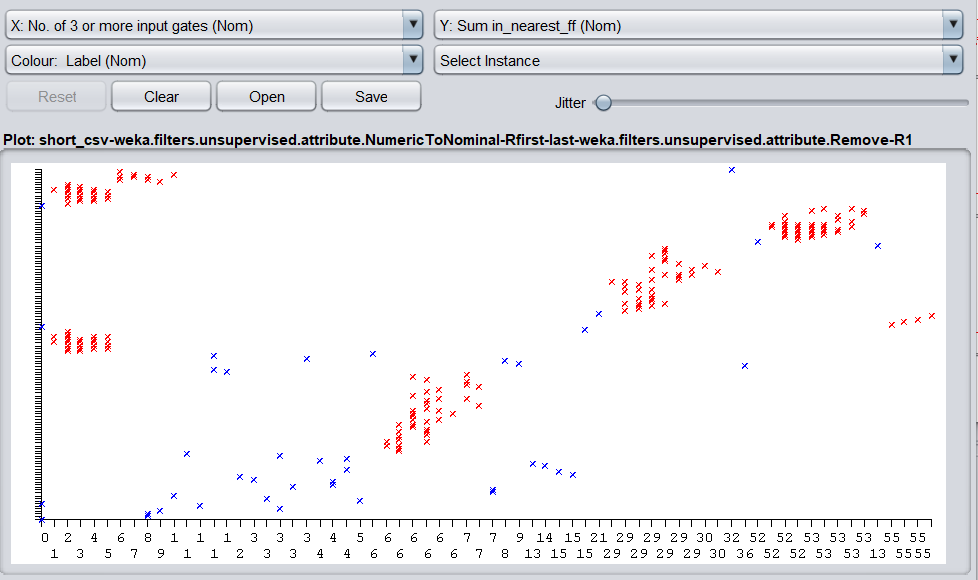
A visualization of the dataset was also done in order to get further insight

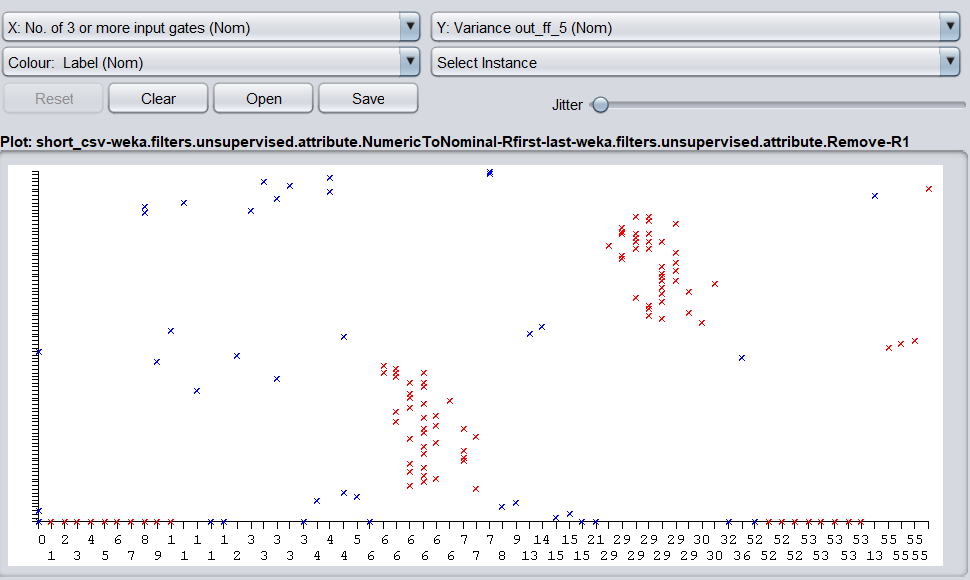












Apart from the ones shown above, many more features were analyzed

**Machine Learning:**

After extracting the features, we performed some data preprocessing by applying an unsupervised attribute filter to convert numeric values in the dataset to nominal ones. Because of the lack of a non-trojan dataset when compared to trojan datasets, we also applied a class balancer filter, but that change was reverted because it negatively impacted the model. After that, we subjected this dataset to multiple classifiers namely SMO, J48 and Zero R. We also used a ranking algorithm to identify the top features of the dataset and then created multiple feature sets and tested them out on these classifiers. These sets included top 50 features, top 605 features, Top 45 features etc.

The train-test ratio was kept to a standard of 0.2, where 20% of the dataset was used for testing while 80% was used for training. We also performed a 10 fold cross-validation on these classifiers as well.

**Deep Learning:**

The neural network trained for the model is made up of the following structure:

Over 600 parameters/features of the dataset are passed into the model, these inputs are then trained on a deep learning structure 5 layers deep and containing 20 neurons each. Each of these layers utilizes the ‘Relu’ activation function and the output is a single layer with the ‘Sigmoid’ activation function that returns the probabilities of the circuit being infected by Trojan or not.

The model utilises is Adam optimizer along with a binary cross-entropy loss structure the model is strange on a batch size of 32 and an overall epoch count of 24.

The model produces an overall accuracy of 98.95% over the data set.

The model can be found on the following collab link

**>>>NEW LINK**

**https://colab.research.google.com/drive/1Fp-7f4gUp7XZg1r55U-Ggp\_BEgcJpw8W**

**>>>OLD LINK**

[**https://colab.research.google.com/drive/1\_cQuIfYGFKbx\_3KSPbP0nvRoBbh-Ud4p?usp=sharing**](https://colab.research.google.com/drive/1_cQuIfYGFKbx_3KSPbP0nvRoBbh-Ud4p?usp=sharing)