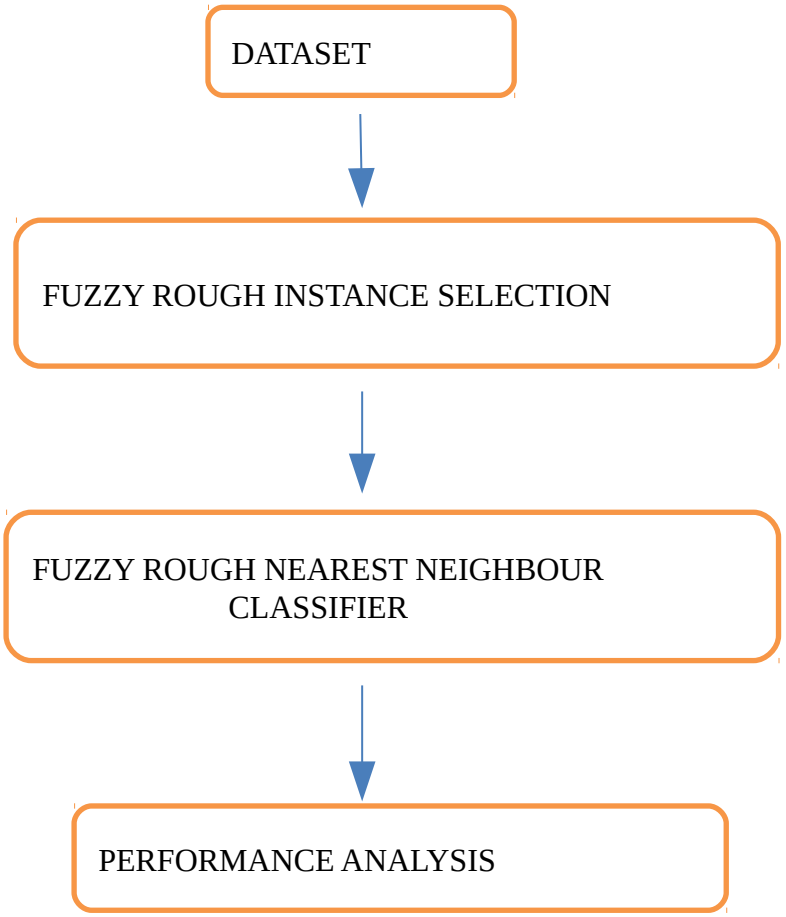
FUZZY ROUGH APPROACH FOR BREAST CANCER CLASSIFICATION

This approach aims to build an automatic diagnostic system for breast cancer based on fuzzy rough nearesr neighbour classifier.

The proposed classification model consists of 2 phases:
1.Instance Selection
2.Classification using fuzzy rough nearest neighbour classifier



INSTANCE SELECTION

Instance selection is a viable tool in the machine learning which aims to reduce the number of instances in the learning set by either extracting bad instances or extracting as much instanes as possible so that the noise in the original data set can be reduced and reduction in the training time of learning algorithm can be achieved.

This algorithm evaluates the degree of membership of each instance to the fuzzy positive region. If there is a instance less than the threshold, then the instance can be removed. Additionally, it uses a fuzzy indiscernibility relation Ra to express the approximate equality between objects x and y on attribute a in the training set.

Its parameters include:
decision.table: a "DecisionTable" class representing the decision table.
control : a list of other parameters which are

- threshold.tau: a value determining whether an object can be removed or not. The object can be removed if it is less than the threshold.
- alpha: a parameter determining the granularity of the fuzzy similarity measure, which has positive values (>= 0). The default value is 1.
- type.aggregation: a list representing the type of aggregation and its value. The default value is type.aggregation = c("t.tnorm", "lukasiewicz").
- t.implicator: a string representing the value of implicator function. The default value is "lukasiewicz".

```
> res.1 <- IS.FRIS.FRST(decision.table = decision.table, control =  list(threshold.tau = 0.75, alpha = 0.8))
> decision.table <- SF.applyDecTable(decision.table, res.1)
```

FUZZY ROUGH NEAREST NEIGHBOR CLASSIFIER

Fuzzy Rough Nearest Neighbour Classifier extends K- Nearest Neighbor algorithm by exploiting the fuzzy rough uncertainly. In the fuzzy-rough nearest neighbor algorithm, the initial memberships of training patterns can be crisp, constrained fuzzy or possibilistic values. The algorithm differs from fuzzy counterpart in not requiring fixing K parameter. The aforementioned uncertainties in the conventional K-nearest neighbor algorithm are captured with the use of a fuzzy-rough ownership function.

The method used uses the fuzzy lower and upper approximations to improve the fuzzy nearest neighbor (FNN) algorithm.
This algorithm assigns a class to a target instance t as follows.

- Determine k nearest neighbors considering their similarity to new patterns.
- Assign new patterns to the class based on maximal value of fuzzy lower and upper approximations. If a value of fuzzy lower approximation is high, it shows that neighbors of newdata belong to a particular class, e.g. C. On the other hand, a high value of fuzzy upper approximation means that at least one neighbor belongs to that class.

Its parameter includes:
decision.table: a "DecisionTable" class representing the decision table.
control: a list of other parameters such as type.LU ,value of k, type.aggregation, type.relation, type.implicator etc

```
> control <- list(type.LU = "implicator.tnorm", k = 3,type.aggregation = c("t.tnorm", "lukasiewicz"),
+              type.relation = c("tolerance", "eq.1"), t.implicator = "lukasiewicz")
>
> res.1 <- C.FRNN.FRST(decision.table = decision.table, newdata = cancer.testing, control = control)
```

PERFORMANCE ANALYSIS

Confusion Matrix:

	test.	class
res.1	0	1
0	107	2
1	1	65

Accuracy: 0.9828571