NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

CZ4032 Data Analytics & Mining

Group 26 - Project 1 Report

Adrian Goh Jun Wei	U1721134D
Lim Jin Yang	U1921682L
Tan Hu Soon	U1921282A
Vivien Chew Jing Wen	U1922718D

Abstract

This project explores seven different data sets and analyses data with association rule techniques and open source classifiers. Some meaningful patterns and relationships among attributes are discovered, such as the association between different attributes and the prediction of a successful outcome based on the relevant attributes is also learnt through classification.

Association Rule

Apriori algorithm selected is an association rule mining algorithm that optimizes the frequent itemset generation and rule generation based on anti-monotone property.

Classification

Four classifiers, K-Nearest Neighbors, Gaussian Naïve Bayes, Random Forest, Decision Tree classifiers are selected to predict the outcomes of the selected attribute and we will compare the performance of each classifier based on their accuracy. To evaluate these open source classifiers, Accuracy Score, Confusion Matrix, and F-Score are used as the metrics of the performance.

Classification based on Association rules (CBA)

CBA (Classification Based on Associations) consists of 3 parts

1) A process to bin and discretizing continuous attributes
2) A rule generator (called CBA-RG), which is based on algorithm Apriori for finding association rules
3) A classifier builder (called CBA-CB). This is done by ranking all* possible rules in order of their confidence and support, and then selecting the subset with the right rule sequence that gives the least number of errors. We do so by evaluating the ranked rules and eliminating any rule that could not find a single record to tag. After all rules have been evaluated, we tag all remaining records with the default class, which is the majority class in the remaining data.

* In practice, depending on the size of the data set, we might have too many rules generated by CBA-PG that it becomes too computationally and infeasible to rank all of them, given the machines we own. In such cases, we will set a reasonable cut-off for the minimum support required in order to consider a rule.

Dataset Description

Datasets	Goal	Categorical Attributes	Numerical Attributes	No. of Instances
Adult	Predicts whether income exceeds $50K/yr based on census data	*Workclass, education, marital-status, occupation, relationship, race, sex, native-country, income*	*Age, hours-per-week*	48842
Algerian Forest Fires	Predicts if forest fire will occur	*Classes*	*Date, Temp, RH (relative humidity), Ws (wind speed), Rain, Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI), Buildup Index (BUI), Fire Weather Index (FWI)*	244
Heart Failure Clinical Record	Predicts if death occurs due to heart failure	*Death_event, sex, smoking*	*Age, anaemia, high blood pressure, creatinine, diabetes, ejection fraction, platelets, serum creatinine, serum sodium, time*	299
Occupancy Detection	Predicts room occupancy	*Occupancy*	*Date, temperature, relative humidity, light, CO2, humidity ratio*	20560
Bank Marketing	Predict if the client will subscribe a term deposit	*Job, marital status, education, default, housing, loan, contact, month, day_of_week, poutcome (outcome of campaign), y (subscribed)*	*Age, duration, campaign, pdays, previous*	45211
Wine Quality	Predicts wine quality based on physicochemical tests	*Quality*	*Fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, alcohol*	4898
Breast Cancer	Predicts for Breast Cancer	*Diagnosis*	*32*	569

Implementation Data

Preprocessing Data

Preprocessing is implemented using mainly NumPy and Pandas packages, with help from some other Python packages and utilities.

Data Exploration

Data is implemented using Pandas for summary statistics and correlation matrix, and the PyPlot from Matlibplot for visualization, with help from the Seaborn package.

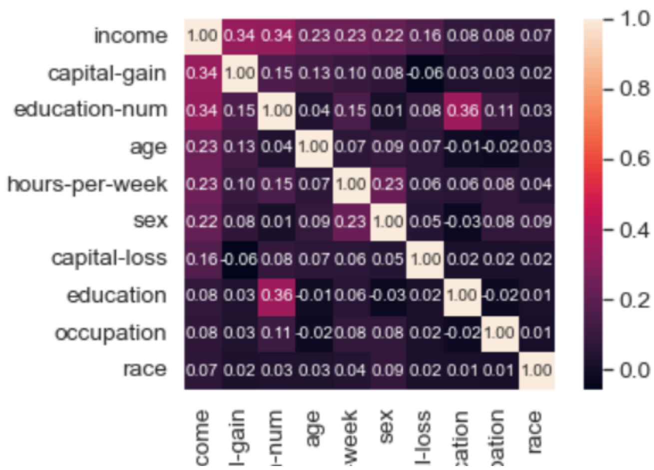


Figure: Example of heatmap using 'adult.csv' dataset

Here, we can remove columns that are not well correlated to the target attribute "income" as they are not going to improve the classification, such as 'race', 'occupation', 'education' and 'capital-loss' attributes.

Experimental Results

*Refer to Appendices for the rules generated by the apriori classifier

Datasets	No. of Rules*	Runtime (seconds)	Dataframe Memory (MB)	Accuracy Score
Adult (ms = 0.20)	12	25.63	433.00	0.7652
Forest Fires (ms = 0.10)	7	0.14	0.23	0.7398
Heart Failure (ms = 0.05)	60	6.36	0.32	0.8395
Occupancy (ms = 0.05)	46	0.81	5.02	0.9606
Bank Marketing (ms = 0.30)	22	4.48	5.56	0.8848
Wine Quality (ms = 0.05)	43	0.45	1.59	0.7530
Breast Cancer (ms = 0.05)	81	36.48	1.56	0.9578

Figure 1: Experiment Results

	KNeigh	Naive Bayes	Random Forest	Decision Tree	Apriori
Adults	0.784892	0.823737	0.830186	0.789959	0.765190
BankMarketing	0.890608	0.879558	0.895028	0.866298	0.884760
BreastCancer	0.912281	0.938596	0.947368	0.938596	0.957840
ForestFires	0.816327	0.836735	0.918367	0.918367	0.739837
HeartFailure	0.650000	0.733333	0.800000	0.683333	0.839532
Occupancy	0.986586	0.966235	0.989362	0.987049	0.960600
Wine	0.646875	0.668750	0.765625	0.715625	0.753055

Figure 2: Classification Results

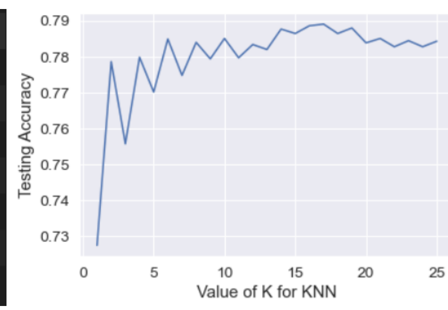


Figure 3: Accuracy vs K-value

The classifier for Apriori algorithm does not work as well in the case where the data is full of discrete data. This may be due to the way binning is being performed on the data.

The K Nearest Neighbour classifier makes predictions based on local information, therefore the value of K used may result in underfitting or overfitting, hence affecting accuracy. In the code implementation, we first use a ‘For’ loop to return the accuracy for a range of K values displayed using matplotlib (figure 2). We will then select the value of K whilst considering the nature of the attributes in our datasets so that we can prevent overfitting / underfitting.

Random Forest achieves the best results for all the datasets although Apriori is still comparable in terms of performance. Random forest builds multiple decision trees and merges them together to get a more accurate and stable prediction, therefore, it is understandable that random forest consistently beats the decision tree classifier in terms of accuracy.

Naive Bayes classifier is robust to isolated noise points as well as irrelevant attributes, since the probability can be too small to be ignored for certain instances. Due to this, the accuracy is slightly better than K-Nearest Neighbors. However, this algorithm relies on the independent assumption, which may not always hold.

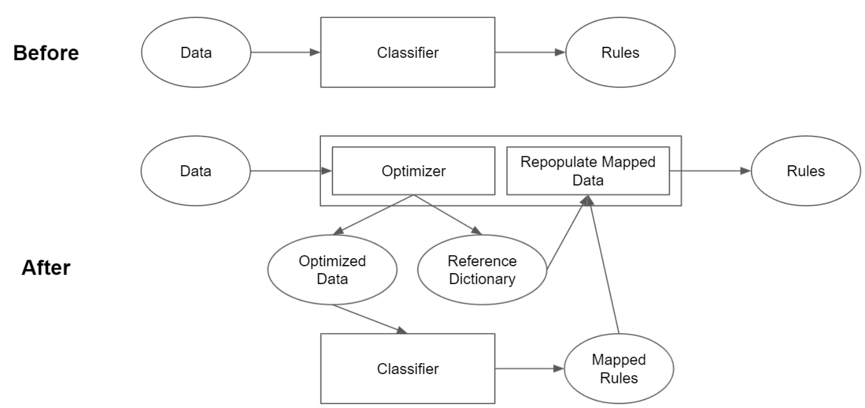
Directions for Improvements

1) Reducing memory needed for classifier

The current implementation of the Association Rule Classifier requires us to pass in all the values in the dataset into the Apriori classifier to determine the association rules.

The issue with this approach is that many duplicate values are being passed in, thus requiring more memory. The impact is negligible if the dataset is fairly small and contains mostly integers. However, the memory needed increases exponentially with the size of the data (number of rows and attributes used) and if the attributes contain long strings.

As such, an improvement that can be made will be to mapped the values into integers, and mapped it back to the original values after computation. This can reduce the size of the dataset being passed into the Apriori classifier.



Theoretical explanation:

Before optimization

Index	Attribute 1	Attribute 2	...	Attribute X
1	“abcdefghij”	“abcdefghij”	....	“abcdefghij”
2	“abcdefghij”	“abcdefghij”	...	“abcdefghij”
...	...	...	...	...
Y	“abcdefghij”	“abcdefghij”	...	“abcdefghij”

Let’s assume a dataset has X number of attributes, each of it containing a string of 10 characters, as well as Y number of rows.

Assuming that each character takes 1 byte, the memory required to store and process can be estimated to be

$$(Y \times X \times 10)\ Bytes$$

After optimization

Index	Attribute 1	Attribute 2	...	Attribute X
1	0	1	....	2
2	0	1	...	2
...	...	...	...	...
Y	0	1	...	2

Assuming that each integer takes 2 byte, the memory required to store and process can be estimated to be the following, where Z is the number of unique attributes (for the mapping of the values)

$$(Y \times X \times 2)\ Bytes + ((10 + 2) \times Z)\ Bytes$$

When the value of Y and X approaches a very large number, the data needed for the mapping will be negligible. As such, the memory saving will be approximately 5x (from 10XY to 2XY).

Implementation:

For the optimization, we decided to do a demonstration based on the Adult dataset as it contains one of the most number of attributes with datatype string, as well as these attributes having values that are very long e.g. *Exec-managerial, Married-civ-spouse*.

Before optimization

Attributes	**Example row #1**	**Example row #2**	**Example row #3**	**...**
Age	39	50	38	...
Workclass	State-gov	Self-emp-not-inc	Private	...
...	...	...	...	...
Native-country	United-States	United-States	United-States	...
Income-group	<=50K	<=50K	<=50K	...

Performance

```
>>>>> Memory used for this dataframe is: 433.531018MB
>>>>> Time to determine assoc rules: 18.21875secs
```

Total memory used: **433MB**
Total time taken: **18.21 seconds**

After optimization

Attributes	**Example row #1**	**Example row #2**	**Example row #3**	**...**
Age	0	3	0	...
Workclass	16	17	18	...
...	...	...	...	...
Native-country	76	76	76	...
Income-group	118	118	118	...

As we can see, instead of being represented by the original strings, the values are now represented by integers. These integers can be mapped back to the original values by referring to the mapping dictionary that we generated. Here's an extract of it:

```
{
   ...
   "18":"workclass___Private",
   "19":"workclass___Federal-gov",
   "20":"workclass___Local-gov",
   "21":"workclass___?",
   "22":"workclass___Self-emp-inc",
   "23":"workclass___Without-pay",
   "24":"workclass___Never-worked",
   "25":"education___Bachelors",
   "26":"education___HS-grad",
   ...
}
```

Performance

```
BEFORE ------------------------------
>>>>> Memory used for this dataframe is: 433.531018MB
>>>>> Time to determine optimize dataframe: 15.515625secs

AFTER ------------------------------
>>>>> Memory used for this dataframe is: 44.257704MB
>>>>> The size of the dictionary is 4704 bytes
>>>>> Time to determine assoc rules: 12.5secs
```

Total memory used: 44MB + 4704B = **48.04MB**
Total time taken: 15.52 + 12.50 = **28.02 seconds**

As seen from the result, we managed to reduce the size of the memory needed from the original 433M to 48MB, reducing it by ~89%.

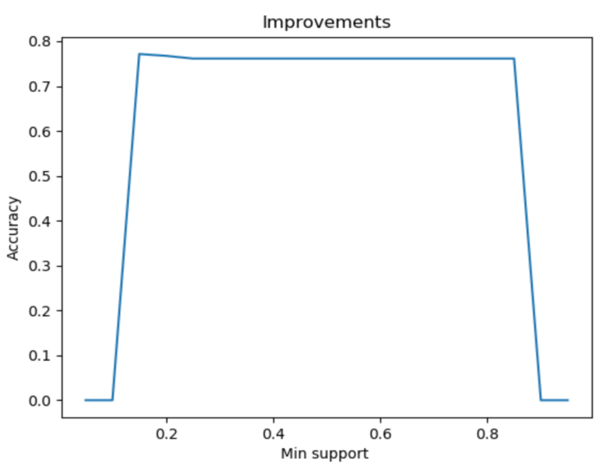
However, while the time needed to generate the association rules decreases from 18.21 seconds to 12.50 seconds, the total time needed for the operation increases to 28.02 seconds. This is because the processing of the data and generating a mapping dictionary itself takes up to 15.52 seconds.

We think that the tradeoff of processing time for memory is worth it as the magnitude of memory saving is far larger than the magnitude of increment in time taken. In addition, should we wish to further improve the processing time, we can simply apply parallelization and multiprocessing.

2) Optimize classifier's accuracy by choosing the best min_support value

With our current implementation of the Association Rule Classifier, we assume a minimum support value which acts as a lower bound to be adhered to across all the different columns in question.

The Minimum Support Level plays an important role in ensuring that our classifier is able to perform with optimal accuracy. Instead of just randomly allowing our user to decide the Minimum Support Level, we decided to add a component to empirically test for the optimal support level, allowing our classifier to perform with the maximum possible accuracy. Attached below are some results we obtained



We can see from the graph that there is a difference in results when Minimum Support Level is being changed. In order to further improve on our new addition, we decided to change it such that each class will be assigned a Minimum support level according to their frequency.This helps in accommodating uneven class frequency distributions to allow for more meaningful rule generation. With this change, we aim to include more rules for infrequent classes and remove meaningless rules generated for frequent classes that cause overfitting and do not add value to our classifier.

Appendices

Association Rule Classifier - Adult


```
[
        {'antecedents': {'marital-status': ['Never-married'],
                'workclass': ['Private']},
         'consequents': '<=50K',
         'index': 0},

        {'antecedents': {'marital-status': ['Never-married']},
         'consequents': '<=50K',
         'index': 3},

        {'antecedents': {'sex': ['Female'], 'workclass': ['Private']},
         'consequents': '<=50K',
         'index': 7},

        {'antecedents': {'relationship': ['Not-in-family']},
         'consequents': '<=50K',
         'index': 8},

        {'antecedents': {'sex': ['Female']}, 'consequents': '<=50K', 'index': 10},

        {'antecedents': {'education': ['HS-grad'], 'workclass': ['Private']},
         'consequents': '<=50K',
         'index': 14},

        {'antecedents': {'education': ['HS-grad']},
         'consequents': '<=50K',
         'index': 15},

        {'antecedents': {'hours-per-week': ['40.0 to 48.0'], 'workclass': ['Private']},
         'consequents': '<=50K',
         'index': 19},

        {'antecedents': {'workclass': ['Private']},
         'consequents': '<=50K',
         'index': 20},

        {'antecedents': {'hours-per-week': ['40.0 to 48.0']},
         'consequents': '<=50K',
         'index': 25},

        {'antecedents': {'marital-status': ['Married-civ-spouse']},
         'consequents': '>50K',
         'index': 26},
```

{'antecedents': {'sex': ['Male']}, 'consequents': '>50K', 'index': 27}
]

Association Rule Classifier - Algerian Forest Fires

[
{'antecedents': {'Rain': ['0'], 'month': ['07']},
'consequents': 'fire   ',
'index': 0},

{'antecedents': {'Rain': ['0'], 'month': ['08']},
'consequents': 'fire   ',
'index': 2},

{'antecedents': {'Rain': ['0']}, 'consequents': 'fire   ', 'index': 4},

{'antecedents': {'month': ['08']}, 'consequents': 'fire   ', 'index': 6},

{'antecedents': {'month': ['07']}, 'consequents': 'fire   ', 'index': 8},

{'antecedents': {'month': ['06']}, 'consequents': 'not fire   ', 'index': 10},

{'antecedents': {'month': ['09']}, 'consequents': 'not fire   ', 'index': 11}
]

Association Rule Classifier - Heart Failure

[
{'antecedents': {'diabetes': ['0.0 to 1.0'],
'sex': ['more than 1.0'],
'time': ['less than 26.8']},
'consequents': 'DEATH_EVENT9',
'index': 0},

{'antecedents': {'time': ['more than 244.0']},
'consequents': 'DEATH_EVENT6',
'index': 1},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
'serum_sodium': ['137.0 to 138.0']},
'consequents': 'DEATH_EVENT6',
'index': 2},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
'high_blood_pressure': ['0.0 to 1.0'],
'serum_sodium': ['140.0 to 141.2']},
'consequents': 'DEATH_EVENT6',

'index': 8},

{'antecedents': {'diabetes': ['more than 1.0'],
 'ejection_fraction': ['40.0 to 47.00000000000003'],
 'smoking': ['0.0 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 11},

{'antecedents': {'serum_creatinine': ['0.8 to 0.9']},
 'consequents': 'DEATH_EVENT6',
 'index': 16},

{'antecedents': {'sex': ['more than 1.0'], 'time': ['187.0 to 210.4']},
 'consequents': 'DEATH_EVENT6',
 'index': 17},

{'antecedents': {'sex': ['more than 1.0'], 'time': ['less than 26.8']},
 'consequents': 'DEATH_EVENT9',
 'index': 18},

{'antecedents': {'age': ['less than 45.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 21},

{'antecedents': {'anaemia': ['more than 1.0'],
 'ejection_fraction': ['40.0 to 47.00000000000003'],
 'high_blood_pressure': ['0.0 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 22},

{'antecedents': {'diabetes': ['0.0 to 1.0'], 'time': ['115.0 to 147.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 23},

{'antecedents': {'ejection_fraction': ['35.0 to 38.0'],
 'high_blood_pressure': ['0.0 to 1.0'],
 'smoking': ['more than 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 24},

{'antecedents': {'diabetes': ['0.0 to 1.0'],
 'ejection_fraction': ['less than 25.0']},
 'consequents': 'DEATH_EVENT9',
 'index': 26},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
 'time': ['less than 26.8']},
 'consequents': 'DEATH_EVENT9',
 'index': 27},

```
{'antecedents': {'high_blood_pressure': ['more than 1.0'],
                'time': ['26.8 to 59.6']},
 'consequents': 'DEATH_EVENT9',
 'index': 28},

{'antecedents': {'ejection_fraction': ['40.0 to 47.00000000000003'],
                'high_blood_pressure': ['0.0 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 29},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
                'serum_sodium': ['140.0 to 141.2']},
 'consequents': 'DEATH_EVENT6',
 'index': 30},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
                'time': ['187.0 to 210.4']},
 'consequents': 'DEATH_EVENT6',
 'index': 31},

{'antecedents': {'diabetes': ['0.0 to 1.0'], 'time': ['less than 26.8']},
 'consequents': 'DEATH_EVENT9',
 'index': 32},

{'antecedents': {'smoking': ['0.0 to 1.0'],
                'time': ['79.39999999999999 to 95.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 33},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
                'serum_sodium': ['140.0 to 141.2'],
                'smoking': ['0.0 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 34},

{'antecedents': {'sex': ['more than 1.0'], 'time': ['115.0 to 147.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 36},

{'antecedents': {'smoking': ['0.0 to 1.0'], 'time': ['210.4 to 244.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 37},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
                'serum_sodium': ['140.0 to 141.2'],
                'sex': ['more than 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 39},
```

{'antecedents': {'time': ['less than 26.8']},
 'consequents': 'DEATH_EVENT9',
 'index': 40},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
                'ejection_fraction': ['35.0 to 38.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 41},

{'antecedents': {'serum_creatinine': ['1.1 to 1.2'],
                'smoking': ['more than 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 45},

{'antecedents': {'anaemia': ['more than 1.0'],
                'ejection_fraction': ['40.0 to 47.00000000000003'],
                'smoking': ['0.0 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 46},

{'antecedents': {'time': ['210.4 to 244.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 48},

{'antecedents': {'ejection_fraction': ['35.0 to 38.0'],
                'high_blood_pressure': ['0.0 to 1.0'],
                'sex': ['more than 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 49},

{'antecedents': {'time': ['187.0 to 210.4']},
 'consequents': 'DEATH_EVENT6',
 'index': 52},

{'antecedents': {'ejection_fraction': ['more than 60.0'],
                'sex': ['more than 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 54},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
                'serum_creatinine': ['0.9 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 60},

{'antecedents': {'creatinine_phosphokinase': ['176.8 to 250.0'],
                'high_blood_pressure': ['0.0 to 1.0']},
 'consequents': 'DEATH_EVENT6',
 'index': 61},

{'antecedents': {'diabetes': ['more than 1.0'],
'serum_sodium': ['140.0 to 141.2']},
'consequents': 'DEATH_EVENT6',
'index': 62},

{'antecedents': {'ejection_fraction': ['35.0 to 38.0'],
'high_blood_pressure': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 63},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
'serum_creatinine': ['0.9 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 64},

{'antecedents': {'serum_creatinine': ['less than 0.8'],
'sex': ['more than 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 65},

{'antecedents': {'serum_creatinine': ['less than 0.8']},
'consequents': 'DEATH_EVENT6',
'index': 68},

{'antecedents': {'ejection_fraction': ['less than 25.0']},
'consequents': 'DEATH_EVENT9',
'index': 70},

{'antecedents': {'age': ['53.0 to 58.0'], 'sex': ['more than 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 71},

{'antecedents': {'serum_creatinine': ['0.9 to 1.0'],
'smoking': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 72},

{'antecedents': {'time': ['79.39999999999999 to 95.0']},
'consequents': 'DEATH_EVENT6',
'index': 74},

{'antecedents': {'platelets': ['237000.0 to 262000.0'],
'smoking': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 78},

{'antecedents': {'smoking': ['0.0 to 1.0'], 'time': ['26.8 to 59.6']},
'consequents': 'DEATH_EVENT9',

'index': 86},

{'antecedents': {'ejection_fraction': ['40.0 to 47.00000000000003'],
'sex': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 89},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
'platelets': ['237000.0 to 262000.0']},
'consequents': 'DEATH_EVENT6',
'index': 91},

{'antecedents': {'diabetes': ['more than 1.0'], 'time': ['26.8 to 59.6']},
'consequents': 'DEATH_EVENT9',
'index': 100},

{'antecedents': {'creatinine_phosphokinase': ['620.4000000000001 to 1203.8'],
'high_blood_pressure': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 103},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
'platelets': ['265000.0 to 285200.0']},
'consequents': 'DEATH_EVENT6',
'index': 109},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
'platelets': ['285200.0 to 319800.0']},
'consequents': 'DEATH_EVENT6',
'index': 110},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
'serum_creatinine': ['1.0 to 1.1']},
'consequents': 'DEATH_EVENT6',
'index': 121},

{'antecedents': {'high_blood_pressure': ['0.0 to 1.0'],
'serum_sodium': ['137.0 to 138.0'],
'sex': ['more than 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 126},

{'antecedents': {'time': ['115.0 to 147.0']},
'consequents': 'DEATH_EVENT6',
'index': 129},

{'antecedents': {'creatinine_phosphokinase': ['less than 67.6'],
'diabetes': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',

'index': 138},

{'antecedents': {'time': ['95.0 to 115.0']},
'consequents': 'DEATH_EVENT6',
'index': 150},

{'antecedents': {'anaemia': ['more than 1.0'],
'diabetes': ['more than 1.0'],
'sex': ['more than 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 159},

{'antecedents': {'anaemia': ['0.0 to 1.0'],
'diabetes': ['more than 1.0'],
'high_blood_pressure': ['0.0 to 1.0'],
'smoking': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 160},

{'antecedents': {'platelets': ['221000.0 to 237000.0']},
'consequents': 'DEATH_EVENT6',
'index': 185},

{'antecedents': {'creatinine_phosphokinase': ['67.6 to 100.2'],
'smoking': ['0.0 to 1.0']},
'consequents': 'DEATH_EVENT6',
'index': 189}
]

Association Rule Classifier - Occupancy Detection

[
{'antecedents': {'CO2': ['less than 434.0']},
'consequents': 'Occupancy7',
'index': 0},

{'antecedents': {'CO2': ['434.0 to 438.666666666667']},
'consequents': 'Occupancy7',
'index': 1},

{'antecedents': {'Temperature': ['less than 19.39']},
'consequents': 'Occupancy7',
'index': 2},

{'antecedents': {'CO2': ['438.666666666667 to 444.0'],
'Light': ['0.0 to 140.44999999999982']},
'consequents': 'Occupancy7',
'index': 5},

```
{'antecedents': {'CO2': ['444.0 to 451.5'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 6},

{'antecedents': {'CO2': ['451.5 to 464.0'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 7},

{'antecedents': {'CO2': ['464.0 to 493.5'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 8},

{'antecedents': {'CO2': ['493.5 to 636.9749999999999'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 9},

{'antecedents': {'HumidityRatio': ['0.0028345206000395303 to '
                                   '0.00312552883074308'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 11},

{'antecedents': {'HumidityRatio': ['0.00312552883074308 to '
                                   '0.00338851894145776'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 12},

{'antecedents': {'HumidityRatio': ['0.00338851894145776 to '
                                   '0.0036987387253513'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 13},

{'antecedents': {'HumidityRatio': ['0.0036987387253513 to 0.0038045317622804'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 14},

{'antecedents': {'HumidityRatio': ['0.0038045317622804 to 0.0040005801555237'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 15},
```

```
{'antecedents': {'HumidityRatio': ['0.0040005801555237 to '
                                   '0.004294721192727219'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 16},

{'antecedents': {'HumidityRatio': ['0.00456507645387648 to '
                                   '0.0049385855553457'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 17},

{'antecedents': {'Humidity': ['25.68 to 26.7'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 23},

{'antecedents': {'Humidity': ['27.525 to 30.55'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 25},

{'antecedents': {'Humidity': ['30.55 to 33.0'],
                 'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 26},

{'antecedents': {'Light': ['0.0 to 140.44999999999982'],
                 'Temperature': ['20.4175 to 20.7']},
 'consequents': 'Occupancy7',
 'index': 31},

{'antecedents': {'Light': ['0.0 to 140.44999999999982'],
                 'Temperature': ['20.7 to 20.89']},
 'consequents': 'Occupancy7',
 'index': 32},

{'antecedents': {'Light': ['0.0 to 140.44999999999982'],
                 'Temperature': ['20.89 to 21.29']},
 'consequents': 'Occupancy7',
 'index': 33},

{'antecedents': {'Light': ['0.0 to 140.44999999999982']},
 'consequents': 'Occupancy7',
 'index': 36},

{'antecedents': {'CO2': ['444.0 to 451.5']},
 'consequents': 'Occupancy7',
 'index': 37},
```

{'antecedents': {'CO2': ['438.666666666667 to 444.0']},
 'consequents': 'Occupancy7',
 'index': 39},

{'antecedents': {'CO2': ['451.5 to 464.0']},
 'consequents': 'Occupancy7',
 'index': 40},

{'antecedents': {'HumidityRatio': ['less than 0.0028345206000395303']},
 'consequents': 'Occupancy7',
 'index': 41},

{'antecedents': {'Temperature': ['19.39 to 19.79']},
 'consequents': 'Occupancy7',
 'index': 42},

{'antecedents': {'Humidity': ['less than 19.1']},
 'consequents': 'Occupancy7',
 'index': 44},

{'antecedents': {'Temperature': ['19.79 to 20.2']},
 'consequents': 'Occupancy7',
 'index': 45},

{'antecedents': {'Light': ['433.0 to 467.333333333333']},
 'consequents': 'Occupancy9',
 'index': 46},

{'antecedents': {'CO2': ['464.0 to 493.5']},
 'consequents': 'Occupancy7',
 'index': 47},

{'antecedents': {'Temperature': ['20.2 to 20.4175']},
 'consequents': 'Occupancy7',
 'index': 48},

{'antecedents': {'Light': ['more than 467.333333333333']},
 'consequents': 'Occupancy9',
 'index': 49},

{'antecedents': {'HumidityRatio': ['0.0036987387253513 to '
 '0.0038045317622804']},
 'consequents': 'Occupancy7',
 'index': 50},

{'antecedents': {'Temperature': ['20.4175 to 20.7']},
 'consequents': 'Occupancy7',
 'index': 51},

```
{'antecedents': {'HumidityRatio': ['0.0028345206000395303 to '
                                  '0.00312552883074308']},
 'consequents': 'Occupancy7',
 'index': 52},

{'antecedents': {'HumidityRatio': ['0.0038045317622804 to '
                                  '0.0040005801555237']},
 'consequents': 'Occupancy7',
 'index': 53},

{'antecedents': {'Humidity': ['22.5 to 24.5'],
                 'HumidityRatio': ['0.00338851894145776 to '
                                  '0.0036987387253513']},
 'consequents': 'Occupancy7',
 'index': 54},

{'antecedents': {'CO2': ['more than 1088.0']},
 'consequents': 'Occupancy9',
 'index': 55},

{'antecedents': {'Humidity': ['22.5 to 24.5']},
 'consequents': 'Occupancy7',
 'index': 56},

{'antecedents': {'HumidityRatio': ['0.0040005801555237 to '
                                  '0.00429472119272721 9']},
 'consequents': 'Occupancy7',
 'index': 58},

{'antecedents': {'CO2': ['853.4333333333333 to 1088.0']},
 'consequents': 'Occupancy9',
 'index': 59},

{'antecedents': {'Temperature': ['20.7 to 20.89']},
 'consequents': 'Occupancy7',
 'index': 60},

{'antecedents': {'Humidity': ['19.1 to 20.5']},
 'consequents': 'Occupancy7',
 'index': 63},

{'antecedents': {'Humidity': ['26.7 to 27.525']},
 'consequents': 'Occupancy7',
 'index': 64},

{'antecedents': {'CO2': ['636.9749999999999 to 853.4333333333333']},
 'consequents': 'Occupancy9',
 'index': 71}
```

]

Association Rule Classifier - Bank Marketing


```
[
        {'antecedents': {'housing': ['yes'], 'poutcome': ['unknown']},
         'consequents': 'no',
         'index': 0},

        {'antecedents': {'housing': ['yes'],
                         'pdays': ['-1.0 to 183.0'],
                         'previous': ['0.0 to 2.0']},
         'consequents': 'no',
         'index': 8},

        {'antecedents': {'housing': ['yes'], 'previous': ['0.0 to 2.0']},
         'consequents': 'no',
         'index': 10},

        {'antecedents': {'marital': ['married'], 'poutcome': ['unknown']},
         'consequents': 'no',
         'index': 12},

        {'antecedents': {'housing': ['yes'], 'marital': ['married']},
         'consequents': 'no',
         'index': 20},

        {'antecedents': {'marital': ['married'],
                         'pdays': ['-1.0 to 183.0'],
                         'previous': ['0.0 to 2.0']},
         'consequents': 'no',
         'index': 32},

        {'antecedents': {'default': ['no'],
                         'marital': ['married'],
                         'previous': ['0.0 to 2.0']},
         'consequents': 'no',
         'index': 36},

        {'antecedents': {'housing': ['yes'], 'pdays': ['-1.0 to 183.0']},
         'consequents': 'no',
         'index': 37},

        {'antecedents': {'default': ['no'],
                         'education': ['secondary'],
                         'poutcome': ['unknown']},
         'consequents': 'no',
         'index': 48},
```

```
{'antecedents': {'education': ['secondary'], 'poutcome': ['unknown']},
 'consequents': 'no',
 'index': 52},

{'antecedents': {'default': ['no'],
                'education': ['secondary'],
                'previous': ['0.0 to 2.0']},
 'consequents': 'no',
 'index': 57},

{'antecedents': {'housing': ['yes']}, 'consequents': 'no', 'index': 61},

{'antecedents': {'default': ['no'], 'poutcome': ['unknown']},
 'consequents': 'no',
 'index': 76},

{'antecedents': {'poutcome': ['unknown']}, 'consequents': 'no', 'index': 82},

{'antecedents': {'marital': ['married'], 'pdays': ['-1.0 to 183.0']},
 'consequents': 'no',
 'index': 88},

{'antecedents': {'default': ['no'],
                'pdays': ['-1.0 to 183.0'],
                'previous': ['0.0 to 2.0']},
 'consequents': 'no',
 'index': 92},

{'antecedents': {'default': ['no'], 'previous': ['0.0 to 2.0']},
 'consequents': 'no',
 'index': 94},

{'antecedents': {'marital': ['married']}, 'consequents': 'no', 'index': 104},

{'antecedents': {'default': ['no'],
                'education': ['secondary'],
                'pdays': ['-1.0 to 183.0']},
 'consequents': 'no',
 'index': 106},

{'antecedents': {'default': ['no'], 'education': ['secondary']},
 'consequents': 'no',
 'index': 114},

{'antecedents': {'default': ['no'], 'pdays': ['-1.0 to 183.0']},
 'consequents': 'no',
 'index': 118},

{'antecedents': {'default': ['no']}, 'consequents': 'no', 'index': 122}
```

]

Association Rule Classifier - Wine Quality

[
{'antecedents': {'alcohol': ['more than 12.0']},
'consequents': 'good',
'index': 0},

{'antecedents': {'volatile_acidity': ['less than 0.31']},
'consequents': 'good',
'index': 1},

{'antecedents': {'alcohol': ['11.3 to 12.0']},
'consequents': 'good',
'index': 2},

{'antecedents': {'total_sulfur_dioxide': ['more than 93.20000000000005']},
'consequents': 'bad',
'index': 3},

{'antecedents': {'sulphates': ['0.76 to 0.85']},
'consequents': 'good',
'index': 4},

{'antecedents': {'alcohol': ['9.3 to 9.5']}, 'consequents': 'bad', 'index': 5},

{'antecedents': {'sulphates': ['less than 0.5']},
'consequents': 'bad',
'index': 6},

{'antecedents': {'density': ['less than 0.994556']},
'consequents': 'good',
'index': 7},

{'antecedents': {'volatile_acidity': ['0.31 to 0.37']},
'consequents': 'good',
'index': 8},

{'antecedents': {'chlorides': ['0.06 to 0.067']},
'consequents': 'good',
'index': 9},

{'antecedents': {'density': ['0.994556 to 0.99534']},
'consequents': 'good',
'index': 10},

{'antecedents': {'sulphates': ['0.5 to 0.54']},
'consequents': 'bad',

'index': 11},

{'antecedents': {'alcohol': ['9.6 to 9.9']},
'consequents': 'bad',
'index': 12},

{'antecedents': {'sulphates': ['0.7 to 0.76']},
'consequents': 'good',
'index': 13},

{'antecedents': {'citric_acid': ['0.39 to 0.46']},
'consequents': 'good',
'index': 14},

{'antecedents': {'total_sulfur_dioxide': ['24.0 to 30.0']},
'consequents': 'good',
'index': 15},

{'antecedents': {'alcohol': ['9.5 to 9.6']},
'consequents': 'bad',
'index': 16},

{'antecedents': {'chlorides': ['less than 0.06']},
'consequents': 'good',
'index': 17},

{'antecedents': {'citric_acid': ['more than 0.5220000000000005']},
'consequents': 'good',
'index': 18},

{'antecedents': {'volatile_acidity': ['more than 0.745']},
'consequents': 'bad',
'index': 19},

{'antecedents': {'fixed_acidity': ['more than 10.7']},
'consequents': 'good',
'index': 20},

{'antecedents': {'sulphates': ['more than 0.85']},
'consequents': 'good',
'index': 21},

{'antecedents': {'alcohol': ['10.5 to 10.9']},
'consequents': 'good',
'index': 22},

{'antecedents': {'volatile_acidity': ['0.66 to 0.745']},
'consequents': 'bad',
'index': 23},

{'antecedents': {'sulphates': ['0.65 to 0.7']},
 'consequents': 'good',
 'index': 24},

{'antecedents': {'alcohol': ['less than 9.3']},
 'consequents': 'bad',
 'index': 25},

{'antecedents': {'alcohol': ['10.9 to 11.3']},
 'consequents': 'good',
 'index': 26},

{'antecedents': {'total_sulfur_dioxide': ['14.0 to 19.0']},
 'consequents': 'good',
 'index': 27},

{'antecedents': {'volatile_acidity': ['0.37 to 0.415']},
 'consequents': 'good',
 'index': 28},

{'antecedents': {'fixed_acidity': ['9.7 to 10.7']},
 'consequents': 'good',
 'index': 29},

{'antecedents': {'volatile_acidity': ['0.61 to 0.66']},
 'consequents': 'bad',
 'index': 30},

{'antecedents': {'density': ['0.9976 to 0.998174']},
 'consequents': 'bad',
 'index': 31},

{'antecedents': {'free_sulfur_dioxide': ['6.0 to 9.0']},
 'consequents': 'good',
 'index': 32},

{'antecedents': {'total_sulfur_dioxide': ['38.0 to 45.799999999999955']},
 'consequents': 'good',
 'index': 34},

{'antecedents': {'citric_acid': ['0.32 to 0.39']},
 'consequents': 'good',
 'index': 36},

{'antecedents': {'pH': ['3.31 to 3.35']}, 'consequents': 'good', 'index': 37},

{'antecedents': {'volatile_acidity': ['0.47 to 0.52']},
 'consequents': 'good',

'index': 38},

{'antecedents': {'chlorides': ['0.072 to 0.076']},
'consequents': 'good',

'index': 39},

{'antecedents': {'total_sulfur_dioxide': ['30.0 to 38.0']},
'consequents': 'good',
'index': 41},

{'antecedents': {'residual_sugar': ['2.3 to 2.5']},
'consequents': 'good',
'index': 42},

{'antecedents': {'pH': ['3.18 to 3.23']}, 'consequents': 'good', 'index': 43},

{'antecedents': {'total_sulfur_dioxide': ['45.799999999999955 to 55.0']},
'consequents': 'good',
'index': 44},

{'antecedents': {'density': ['0.99534 to 0.99586']},
'consequents': 'good',
'index': 46}
]

Association Rule Classifier - Breast Cancer

[
{'antecedents': {'area_mean': ['more than 1177.3999999999999']},
'consequents': 'M',
'index': 0},

{'antecedents': {'area_se': ['more than 91.31399999999998']},
'consequents': 'M',
'index': 1},

{'antecedents': {'area_worst': ['1269.0 to 1673.0']},
'consequents': 'M',
'index': 2},

{'antecedents': {'area_worst': ['more than 1673.0']},
'consequents': 'M',
'index': 3},

{'antecedents': {'concave points_mean': ['more than 0.10042']},
'consequents': 'M',
'index': 4},

{'antecedents': {'concave points_worst': ['0.17754 to 0.20894000000000001']},
 'consequents': 'M',
 'index': 5},

{'antecedents': {'concave points_worst': ['more than 0.20894000000000001']},
 'consequents': 'M',
 'index': 6},

{'antecedents': {'perimeter_worst': ['133.5 to 157.73999999999998']},
 'consequents': 'M',
 'index': 9},

{'antecedents': {'radius_worst': ['20.294 to 23.682']},
 'consequents': 'M',
 'index': 12},

{'antecedents': {'area_se': ['54.088 to 91.31399999999998'],
 'radius_se': ['0.5461400000000001 to 0.7488799999999999']},
 'consequents': 'M',
 'index': 27},

{'antecedents': {'compactness_mean': ['more than 0.17545999999999998'],
 'compactness_worst': ['more than 0.44783999999999996']},
 'consequents': 'M',
 'index': 48},

{'antecedents': {'compactness_worst': ['more than 0.44783999999999996'],
 'symmetry_worst': ['more than 0.36008']},
 'consequents': 'M',
 'index': 51},

{'antecedents': {'area_mean': ['less than 321.6']},
 'consequents': 'B',
 'index': 253},

{'antecedents': {'area_se': ['less than 13.16']},
 'consequents': 'B',
 'index': 254},

{'antecedents': {'area_worst': ['less than 384.71999999999997']},
 'consequents': 'B',
 'index': 255},

{'antecedents': {'area_worst': ['384.71999999999997 to 475.98']},
 'consequents': 'B',
 'index': 256},

{'antecedents': {'concave points_mean': ['less than 0.011158']},
 'consequents': 'B',

'index': 257},

{'antecedents': {'concave points_mean': ['0.011158 to 0.017866']},
 'consequents': 'B',
 'index': 258},

{'antecedents': {'concave points_worst': ['0.03846 to 0.058086000000000006']},
 'consequents': 'B',
 'index': 259},

{'antecedents': {'concavity_mean': ['less than 0.013686']},
 'consequents': 'B',
 'index': 260},

{'antecedents': {'concavity_se': ['less than 0.0077258']},
 'consequents': 'B',
 'index': 261},

{'antecedents': {'perimeter_se': ['less than 1.2802']},
 'consequents': 'B',
 'index': 263},

{'antecedents': {'perimeter_worst': ['72.178 to 81.402']},
 'consequents': 'B',
 'index': 265},

{'antecedents': {'radius_se': ['less than 0.18308']},
 'consequents': 'B',
 'index': 267},

{'antecedents': {'radius_worst': ['11.234 to 12.498000000000001']},
 'consequents': 'B',
 'index': 269},

{'antecedents': {'area_mean': ['396.56 to 444.06'],
 'area_worst': ['475.98 to 544.14']},
 'consequents': 'B',
 'index': 278},

{'antecedents': {'area_mean': ['396.56 to 444.06'],
 'radius_worst': ['12.498000000000001 to 13.314']},
 'consequents': 'B',
 'index': 279},

{'antecedents': {'area_worst': ['475.98 to 544.14'],
 'perimeter_worst': ['81.402 to 86.328']},
 'consequents': 'B',
 'index': 290},

{'antecedents': {'area_worst': ['544.14 to 599.7'],
                'perimeter_worst': ['86.328 to 91.304']},
 'consequents': 'B',
 'index': 292},

{'antecedents': {'area_worst': ['599.7 to 686.5'],
                'perimeter_mean': ['81.938 to 86.24']},
 'consequents': 'B',
 'index': 293},

{'antecedents': {'compactness_se': ['less than 0.009169'],
                'compactness_worst': ['less than 0.09367600000000001']},
 'consequents': 'B',
 'index': 302},

{'antecedents': {'concavity_mean': ['0.013686 to 0.02493'],
                'concavity_worst': ['0.045652000000000005 to 0.091974']},
 'consequents': 'B',
 'index': 323},

{'antecedents': {'perimeter_mean': ['81.938 to 86.24'],
                'radius_worst': ['14.008000000000001 to 14.97']},
 'consequents': 'B',
 'index': 330},

{'antecedents': {'perimeter_worst': ['86.328 to 91.304'],
                'radius_worst': ['13.314 to 14.008000000000001']},
 'consequents': 'B',
 'index': 337},

{'antecedents': {'radius_mean': ['12.012 to 12.726'],
                'radius_worst': ['13.314 to 14.008000000000001']},
 'consequents': 'B',
 'index': 341},

{'antecedents': {'area_mean': ['915.0600000000003 to 1177.3999999999999']},
 'consequents': 'M',
 'index': 512},

{'antecedents': {'area_se': ['54.088 to 91.31399999999998']},
 'consequents': 'M',
 'index': 513},

{'antecedents': {'concave points_mean': ['0.08425400000000002 to 0.10042']},
 'consequents': 'M',
 'index': 514},

{'antecedents': {'compactness_mean': ['less than 0.0497']},
 'consequents': 'B',

'index': 515},

{'antecedents': {'concave points_mean': ['0.017866 to 0.022788000000000003']},
 'consequents': 'B',
 'index': 516},

{'antecedents': {'concave points_worst': ['0.071656 to 0.083914']},
 'consequents': 'B',
 'index': 517},

{'antecedents': {'concavity_worst': ['0.045652000000000005 to 0.091974']},
 'consequents': 'B',
 'index': 520},

{'antecedents': {'concavity_worst': ['0.091974 to 0.13688']},
 'consequents': 'B',
 'index': 521},

{'antecedents': {'perimeter_mean': ['65.83 to 73.292']},
 'consequents': 'B',
 'index': 522},

{'antecedents': {'perimeter_worst': ['86.328 to 91.304']},
 'consequents': 'B',
 'index': 524},

{'antecedents': {'area_worst': ['599.7 to 686.5']},
 'consequents': 'B',
 'index': 527},

{'antecedents': {'perimeter_mean': ['77.36 to 81.938'],
 'radius_mean': ['12.012 to 12.726']},
 'consequents': 'B',
 'index': 540},

{'antecedents': {'perimeter_worst': ['91.304 to 97.66'],
 'radius_worst': ['14.008000000000001 to 14.97']},
 'consequents': 'B',
 'index': 544},

{'antecedents': {'compactness_mean': ['more than 0.17545999999999998'],
 'concavity_mean': ['more than 0.20304']},
 'consequents': 'M',
 'index': 545},

{'antecedents': {'concave points_se': ['less than 0.0054928']},
 'consequents': 'B',
 'index': 566},

{'antecedents': {'concave points_worst': ['0.058086000000000006 to 0.071656']},
'consequents': 'B',
'index': 567},

{'antecedents': {'concavity_mean': ['0.03440000000000001 to 0.04507']},
'consequents': 'B',
'index': 568},

{'antecedents': {'concavity_worst': ['0.13688 to 0.17718000000000003']},
'consequents': 'B',
'index': 569},

{'antecedents': {'radius_worst': ['13.314 to 14.008000000000001']},
'consequents': 'B',
'index': 572},

{'antecedents': {'texture_mean': ['less than 14.078']},
'consequents': 'B',
'index': 573},

{'antecedents': {'area_se': ['13.16 to 16.64']},
'consequents': 'B',
'index': 576},

{'antecedents': {'smoothness_mean': ['less than 0.079654'],
'smoothness_worst': ['less than 0.10296']},
'consequents': 'B',
'index': 580},

{'antecedents': {'symmetry_mean': ['less than 0.14958000000000002'],
'symmetry_worst': ['less than 0.22612000000000002']},
'consequents': 'B',
'index': 594},

{'antecedents': {'area_se': ['16.64 to 19.038']},
'consequents': 'B',
'index': 595},

{'antecedents': {'compactness_worst': ['more than 0.44783999999999996']},
'consequents': 'M',
'index': 597},

{'antecedents': {'perimeter_worst': ['115.9 to 133.5']},
'consequents': 'M',
'index': 599},

{'antecedents': {'perimeter_se': ['1.2802 to 1.5166']},
'consequents': 'B',
'index': 602},

{'antecedents': {'radius_mean': ['12.012 to 12.726']},
 'consequents': 'B',
 'index': 604},

{'antecedents': {'smoothness_worst': ['less than 0.10296']},
 'consequents': 'B',
 'index': 605},

{'antecedents': {'area_worst': ['926.9599999999999 to 1269.0'],
 'radius_worst': ['17.386 to 20.294']},
 'consequents': 'M',
 'index': 607},

{'antecedents': {'symmetry_worst': ['more than 0.36008']},
 'consequents': 'M',
 'index': 614},

{'antecedents': {'compactness_worst': ['0.16140000000000002 to 0.18462']},
 'consequents': 'B',
 'index': 616},

{'antecedents': {'concave points_se': ['0.0054928 to 0.0068934']},
 'consequents': 'B',
 'index': 617},

{'antecedents': {'concave points_se': ['0.0068934 to 0.008465']},
 'consequents': 'B',
 'index': 618},

{'antecedents': {'symmetry_mean': ['less than 0.14958000000000002']},
 'consequents': 'B',
 'index': 621},

{'antecedents': {'texture_worst': ['17.8 to 20.176000000000002']},
 'consequents': 'B',
 'index': 622},

{'antecedents': {'concave points_worst': ['0.1508 to 0.17754']},
 'consequents': 'M',
 'index': 645},

{'antecedents': {'concavity_se': ['0.013212000000000002 to 0.01698']},
 'consequents': 'B',
 'index': 652},

{'antecedents': {'area_se': ['19.038 to 21.078000000000003']},
 'consequents': 'B',
 'index': 661},

{'antecedents': {'compactness_mean': ['0.059638000000000004 to 0.07039']},
 'consequents': 'B',
 'index': 662},

{'antecedents': {'concave points_mean': ['0.0335 to 0.048318']},
 'consequents': 'B',
 'index': 663},

{'antecedents': {'texture_worst': ['20.176000000000002 to 21.988']},
 'consequents': 'B',
 'index': 667},

{'antecedents': {'concave points_mean': ['0.027982 to 0.0335']},
 'consequents': 'B',
 'index': 668},

{'antecedents': {'area_mean': ['700.6599999999999 to 915.0600000000003'],
 'perimeter_mean': ['98.2 to 111.68']},
 'consequents': 'M',
 'index': 683},

{'antecedents': {'concavity_mean': ['0.11191999999999998 to '
 '0.14978000000000005']},
 'consequents': 'M',
 'index': 687},

{'antecedents': {'texture_worst': ['21.988 to 23.584']},
 'consequents': 'B',
 'index': 708}
]

Reference

1) Data Exploration: https://www.kaggle.com/pmarcelino/comprehensive-data-exploration-with-python
2) KNN: https://www.kaggle.com/edwincaleb/eda-breast-cancer-classification