ENSA Ecole Nationale des Sciences Appliquées BERRECHID

# Post-operatory complication detection for kidney stones patients


**Presented by :**

- Mouad NID ABDELLAH
- Ali IKKEN

**Supervised by :**

- Hamid HRIMECH


# Outline

01 Problem description & Objective

02 Data source

03 Global workflow

04 Results

01

# Problem descrption & Objective

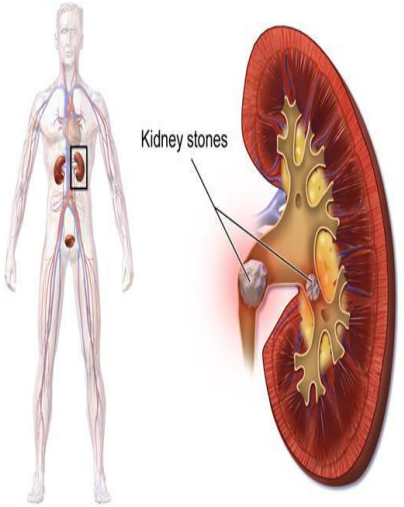
# Problem description

Nearly adults will ultimately need some type of operative procedure.
The urologist should identify and treat a handful of serious complications when performing kidney stone surgery.

So our objective is to detect the complication of the operatory.

# Objective

Our objective is to compare different Machine learning algorithms and different feature selection methods in order to predict the possible post-operatory complication of kidney stones patients.



02

# Data source

# Data source

## Data description

Our dataset contains records of epidemiological characteristics, characteristics of the stones and surgical treatment

## Data size

Our dataset contains :

- 669 row.
- 136 column.

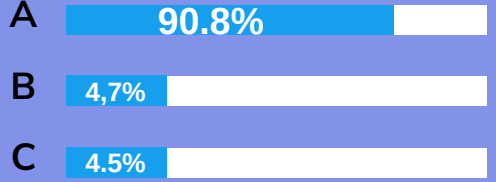
## Target classes

A No-complication

B Class 1 complication

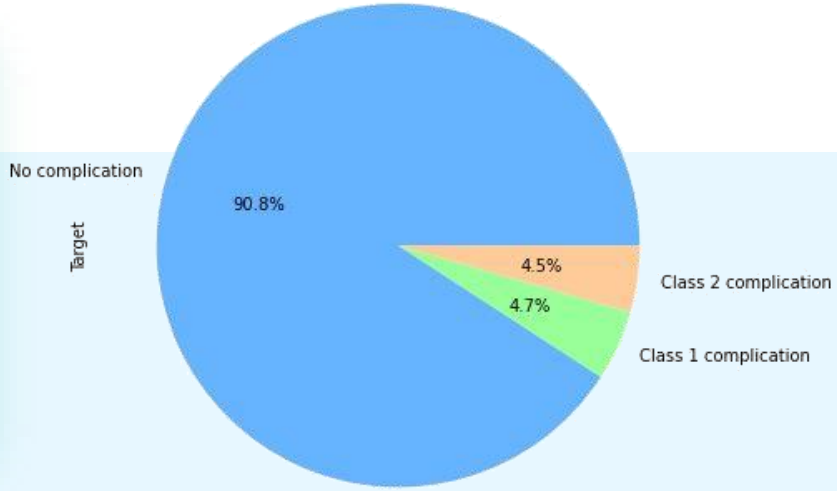
C Class 2 complication

## Class ditribution



# Class distribution

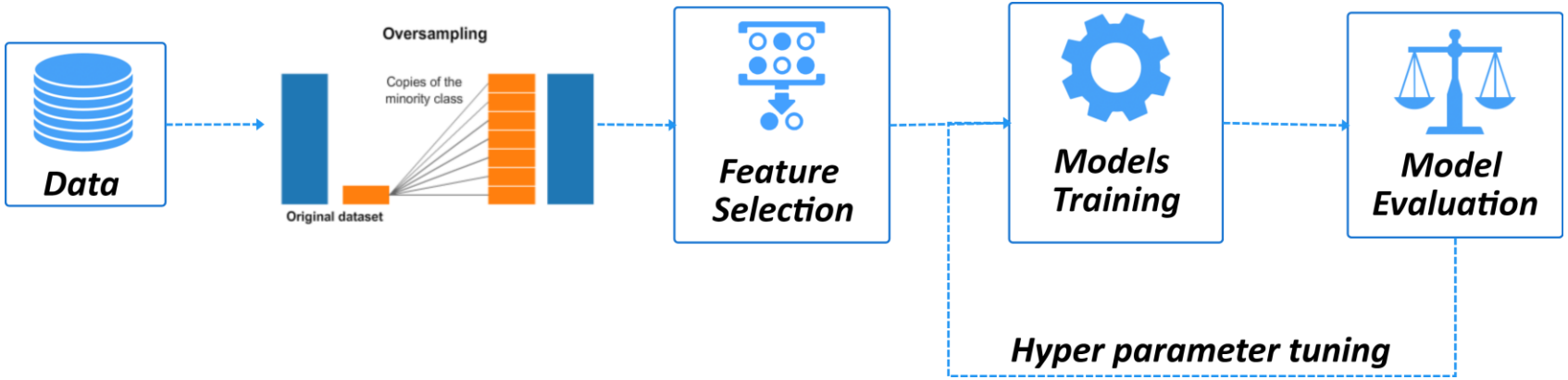
For no-complication we have 387 patient (90,8%) which is the majority of the dataset, for Class 1 complication we have 20 patient (4,7%), and 19 patient for class 2 complication (4,5%).



02

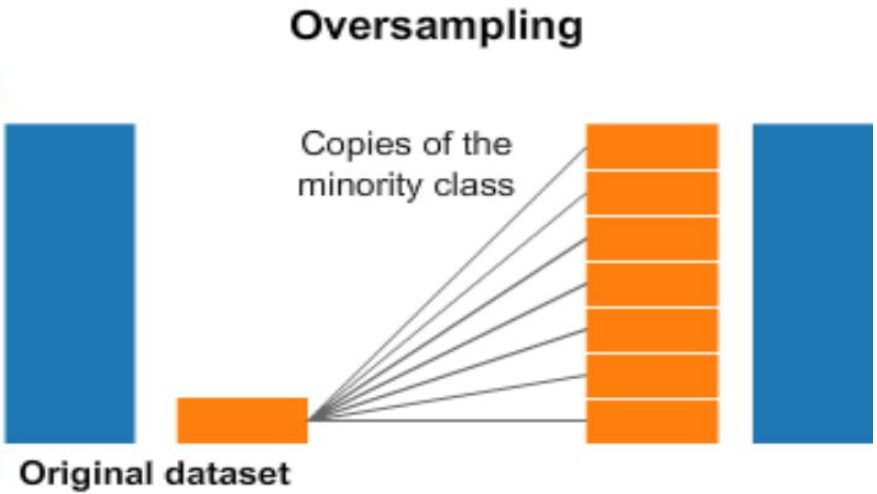
# Global workflow

# Workflow


Data
Oversampling
Copies of the minority class
Original dataset
Feature Selection
Models Training
Model Evaluation
Hyper parameter tuning

# Class balancing

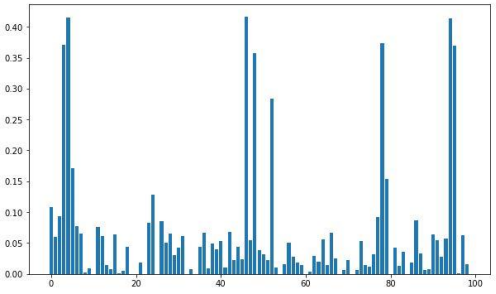
One approach to address imbalanced datasets is to oversample the minority class. The simplest approach involves duplicating examples in the minority class.
We used SMOTE technique to synthesized new records from the existing examples.



# Feature selection

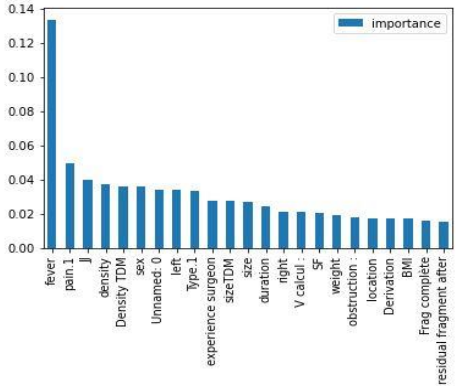
## Mutual information

Between two random variables is a non-negative value, which measures the dependency between the variables



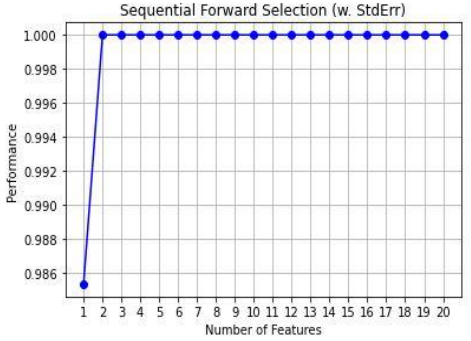

## Embedded RF

Embedded methods are highly accurate. Generalize better, and they are interpretable

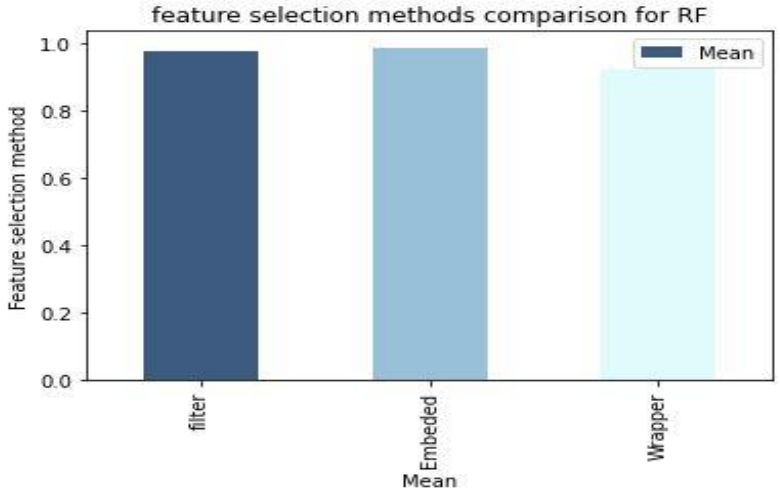


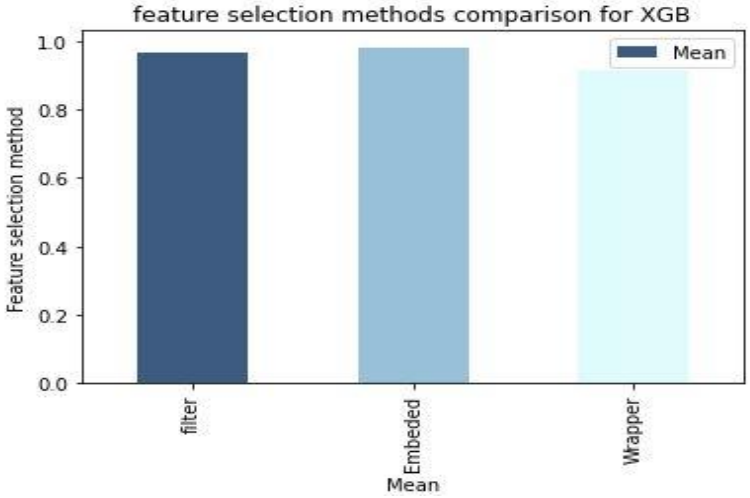

## Wrapper SFS

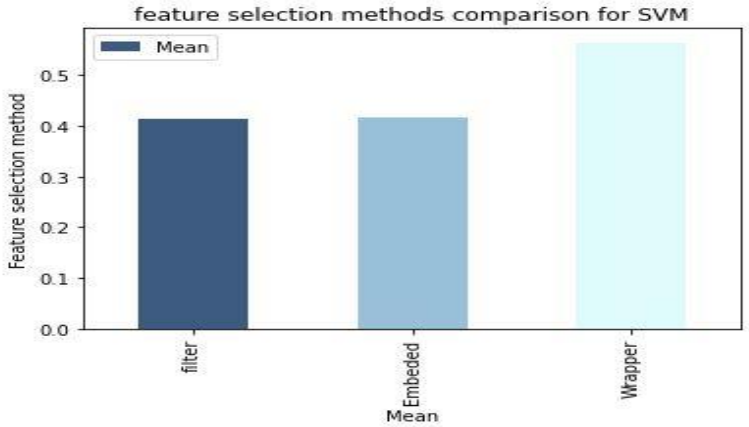
SFS are a family of greedy search algorithms that are used to reduce the initial dimensional feature space.

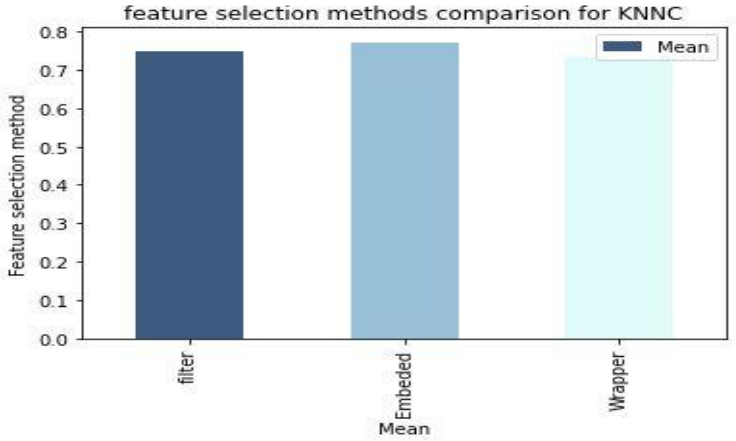


# Feature selection method comparison



# Model training

- XGBoost
- SVM
- Random Forest
- kNN

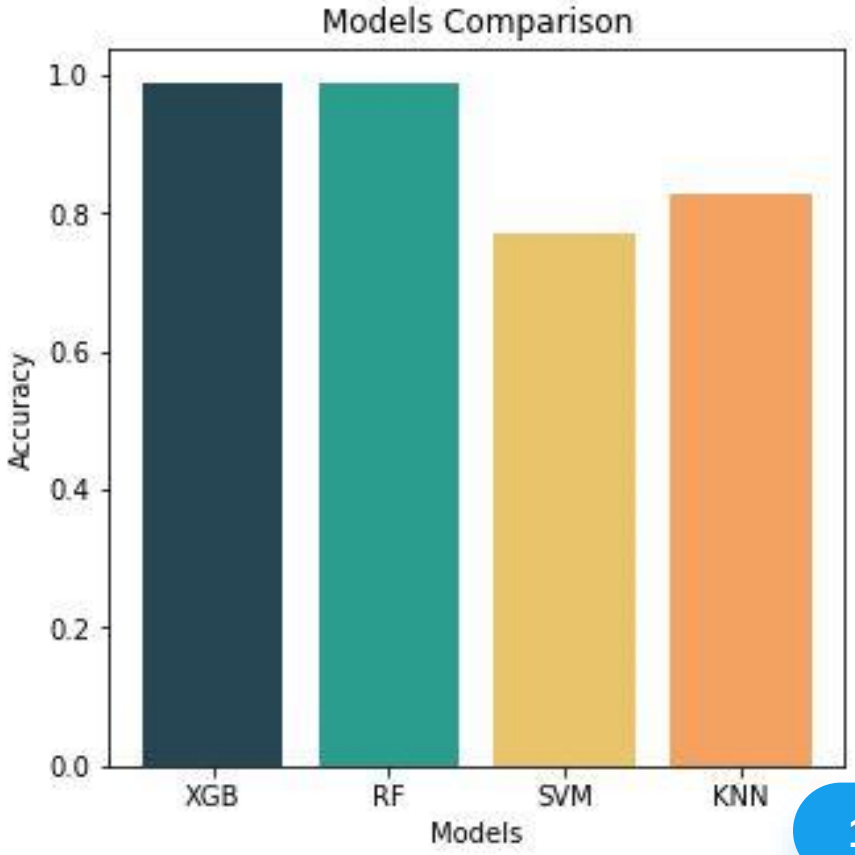
02

# Results

# Results

After choosing the optimal parameter for each model we trained it using the variables given by the feature selection method that is most suited for the model.

The result is summarized in this figure :



# Results

| | SVM | kNN | RF | XGB |
|---|---|---|---|---|
| **Accuracy** | 0.8114 | 0.8457 | 0.96 | 0.9657 |
| **Precision** | 0.8166 | 0.8463 | 0.9627 | 0.9677 |
| **Recall** | 0.8113 | 0.8463 | 0.9598 | 0.9655 |

# Conclusion

To improve the accuracy of the classification of post-operatory complications for kidney stones patients, we could use a voting mechanism that uses two or more algorithms in order to give us an accurate result.

# Thanks!

Do you have any questions?