PROJECT 4 : COVID 19

Hi everyone I would like to welcome you on this presentation on Machine Learning based predictors on COVI19 patients .

My name is NDEYE NENE CISSE

# First I am going to present a brief introduction (Problem Definition).

# Second I will share with you the Motivation and Project Objectives.

# Then proceed to the Materials along with the Methodology.

# Then move to the Results and Discussions

# And by the end of this presentation we will conclude by summarizing the key point.

**Introduction**

The declaration of Corona virus (COVID 19 ) as a pandemic has emerged a

major challenge for the health sector.

Healthcare generates massive datasets from medical follow-up results of

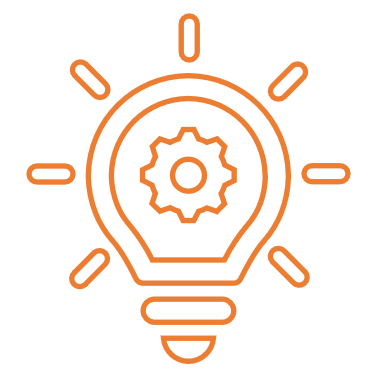
Patients that are updated daily or even hourly.

There are various healthcare aspects that ML can support.

Early prediction of acute cases of covid -19 patients may contribute to

Saving their lives by contributing to the early preparation and management

Of resources in the ICU.

 **Motivation And Project Objective**

**01**

**Explore** patterns in Healthcare records , as well as a

Relationship between vital signs and the ICU admissions.

02

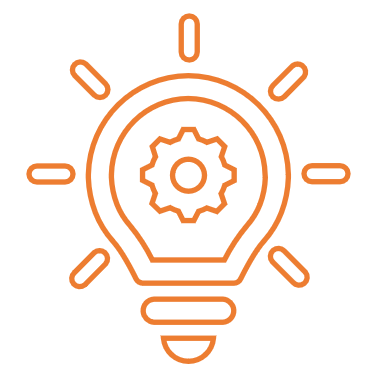
**Identity** suitable preprocessing explore the dataset ,

Understand Features, build and evaluate significant accurate

Model.

03

**Assist** doctors in the ICU

 Materials and Methodology

 **DATA**

\_The data of positive covid-19 patients from the Hospital Sirio-Libanes,

Sao Paulo, and Brasilia were acquired from an open – access repository .

  **Three main classes of information can be extracted :**

2

**Blood Test Results**

3

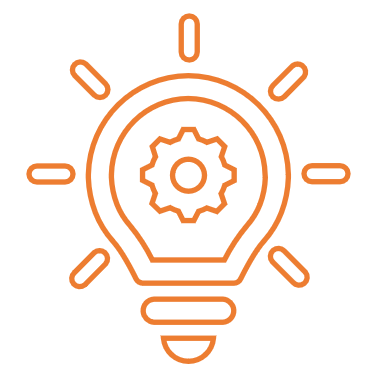
**Vital information**

**1**

**Demographic**

**Information**

**Patient gender age…**

 **Data Preprocessing**

**Handling the Categorical Types**

**01**

**02**

**Data imputation Approach :**

**Mean of all other values of that column**

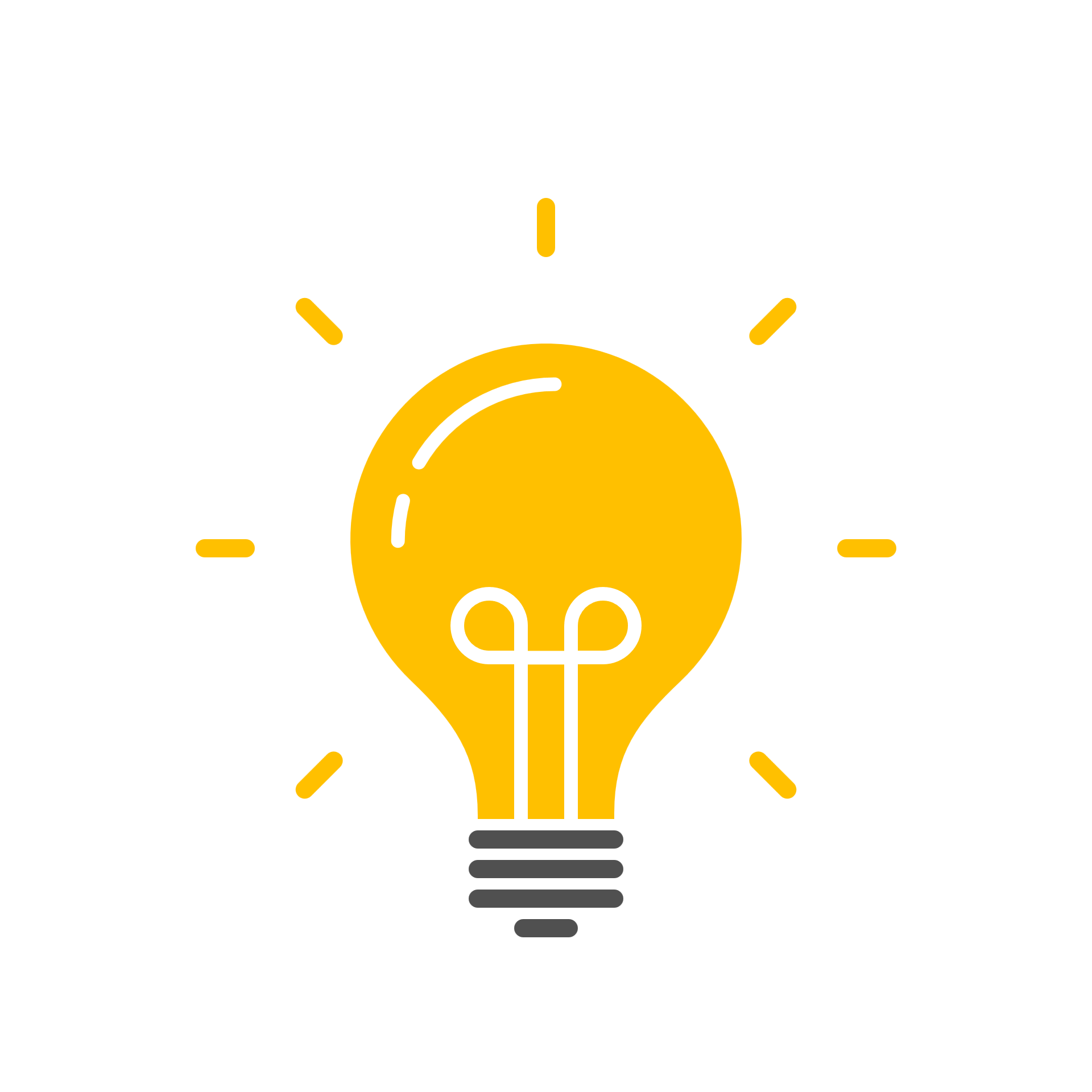
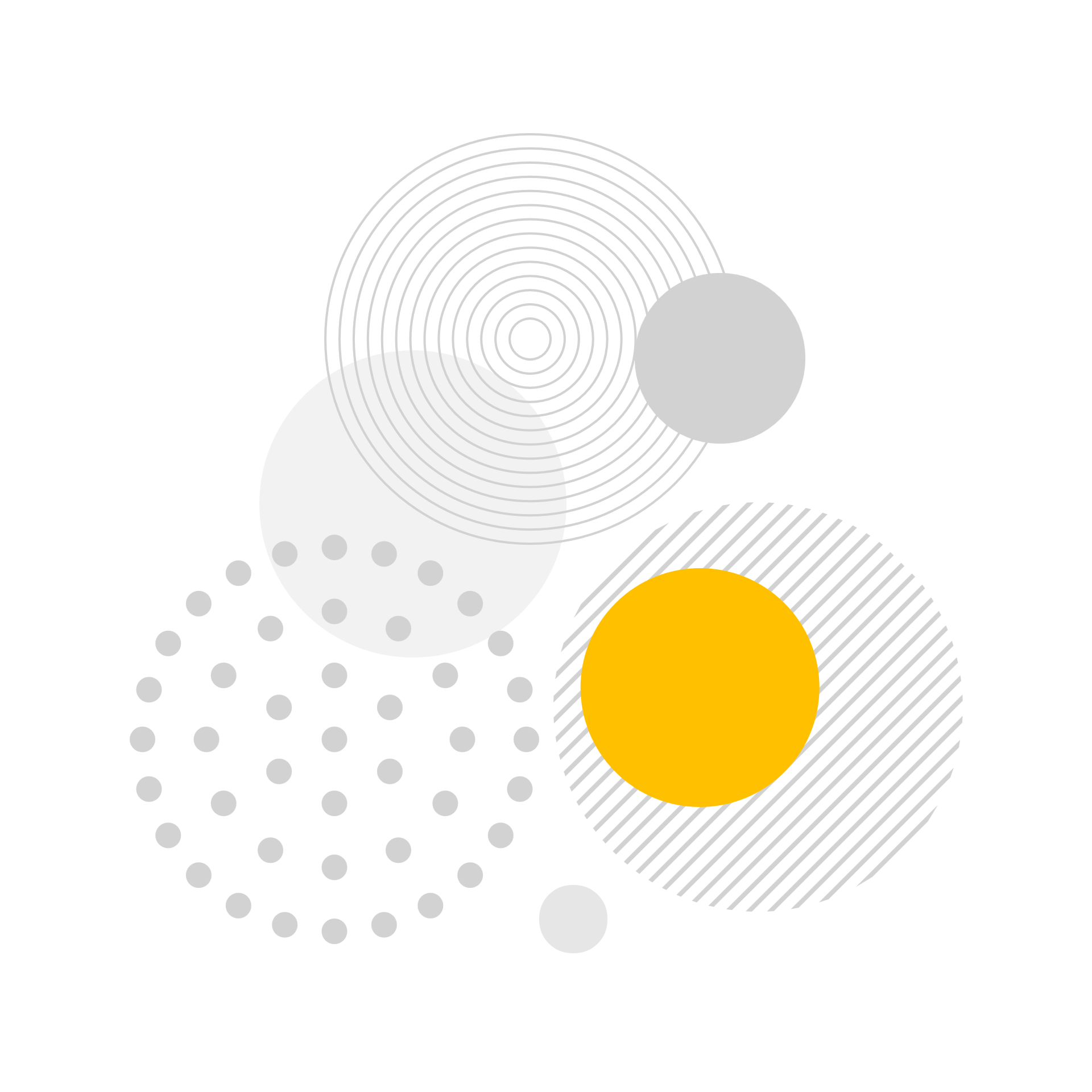
**Filling the missing values of each values with 0.**

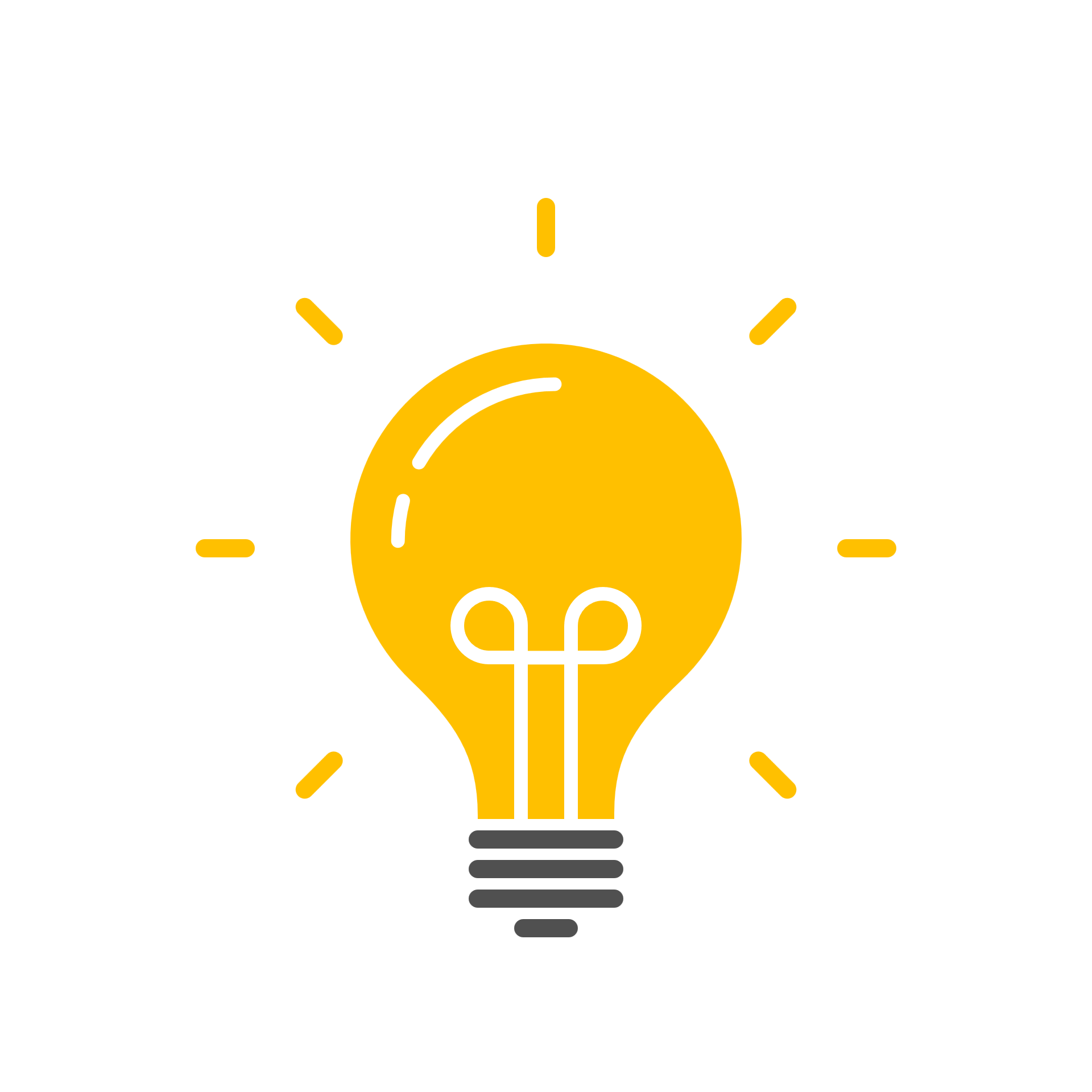
**Based on the Patient’s ICU Experience**

**03**

**Data Sampling**

**For our study , sampling is an important consideration whether we will rely on patient’s data as being stateless or will we rely on collecting patient data and grouping them. Here we will try considerations and compare their results:**

**Stateless Data Grouped Data**

**Stateless Data**

**When considering data as stateless means that these records are independent of each other and save no information of the previous records. Thus, the identifier for each patient is neglected. Each sample is considered as belonging to a different patient. However, only 1925 samples were taken into account based on our Data Imputation Approach.**

**Reason that encourage us to handle it as a stateless data:**

Main goal for this study is to use the features

**a**

Without adopting the patient’s demographic

Information such as gender , age , or any medical

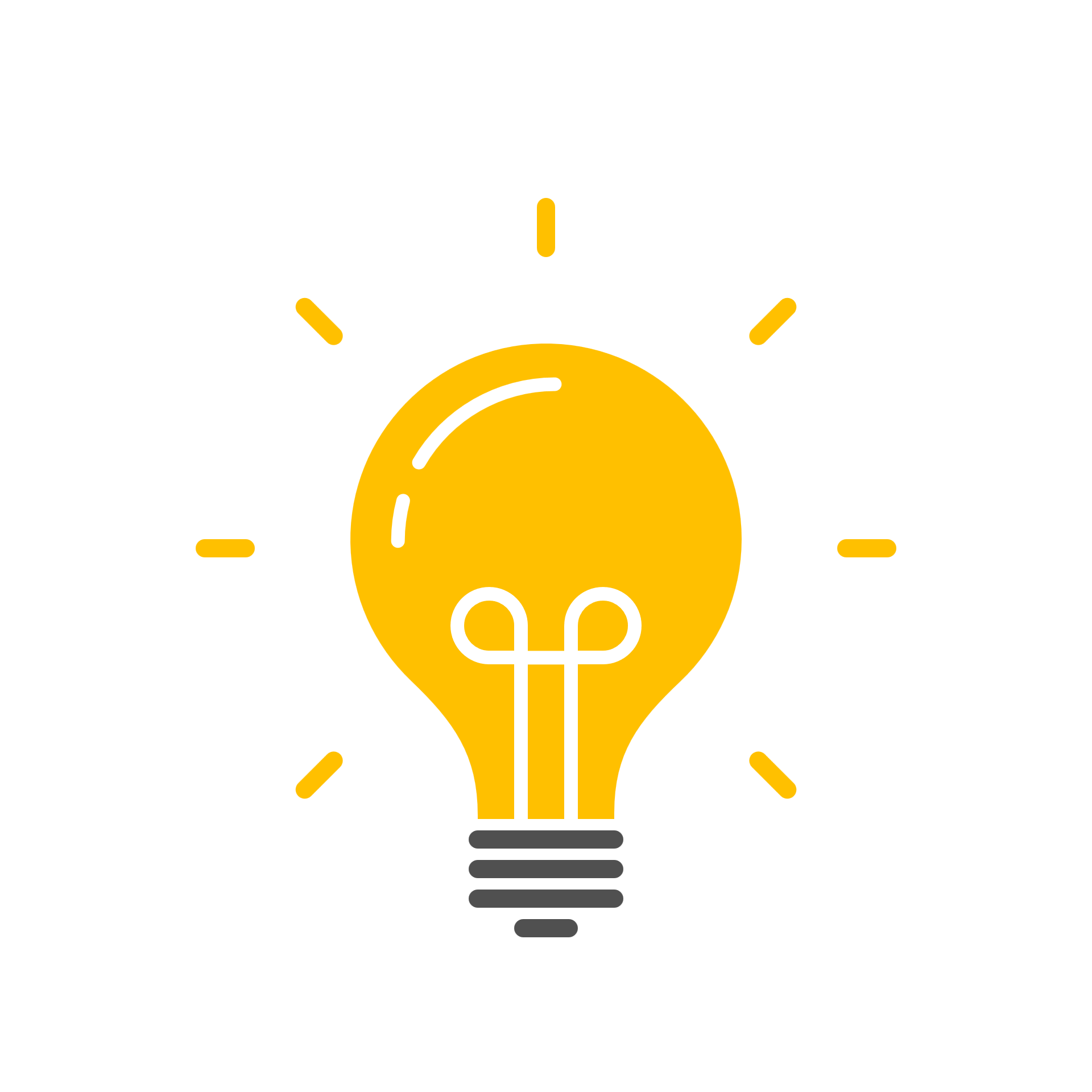
History records

Based on the literature .

**b**

We drop from possibility that the previous record d affects the current record state.

**C**

****Feature Selection

**1**

Field Knowledge

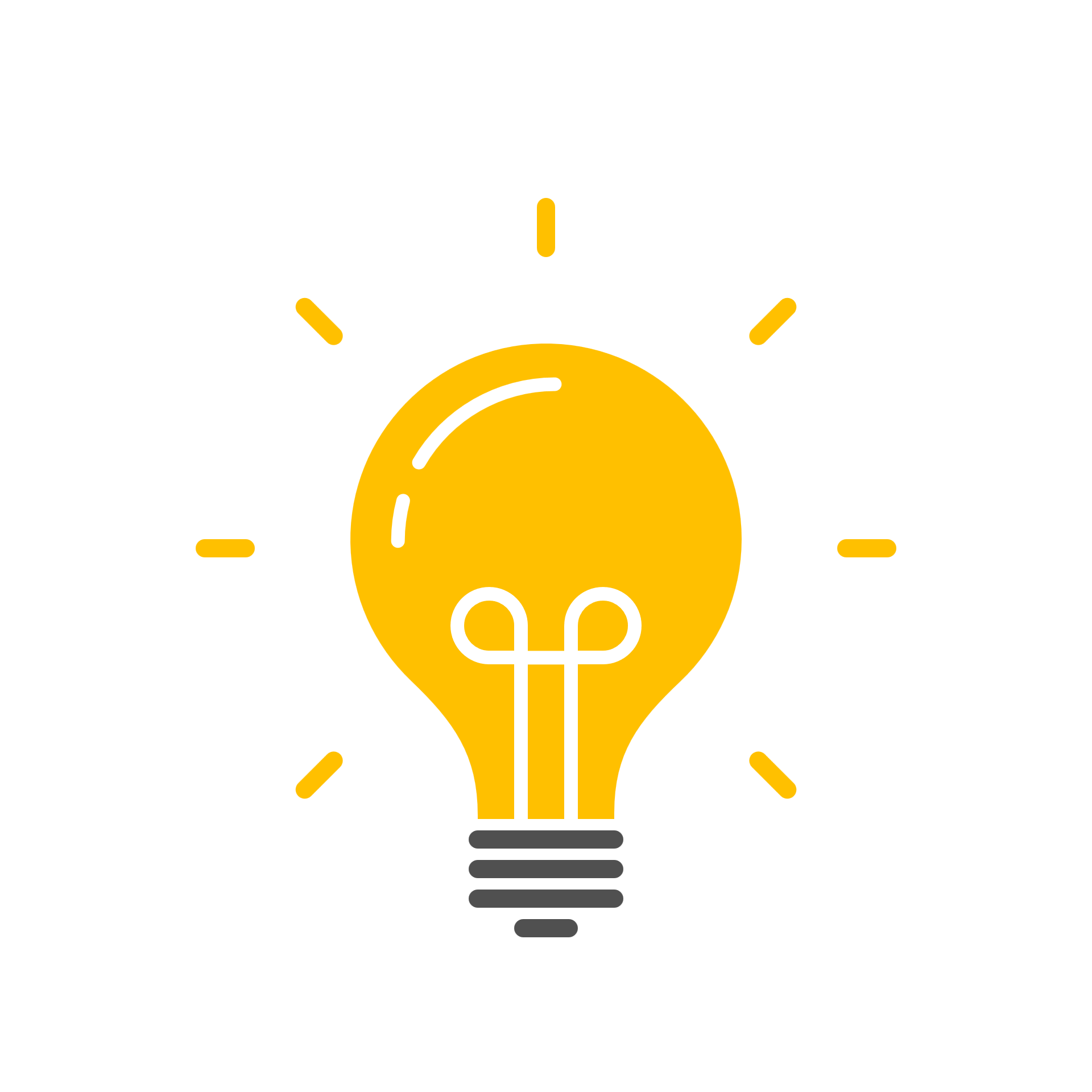
Vital signs that are available in the dataset that we used are the Body temperature, Respiration and Oxygen Saturation.

**2**

Permutation importance

1\_ Large numbers of features in the dataset did not contribute to any impact on the prediction.

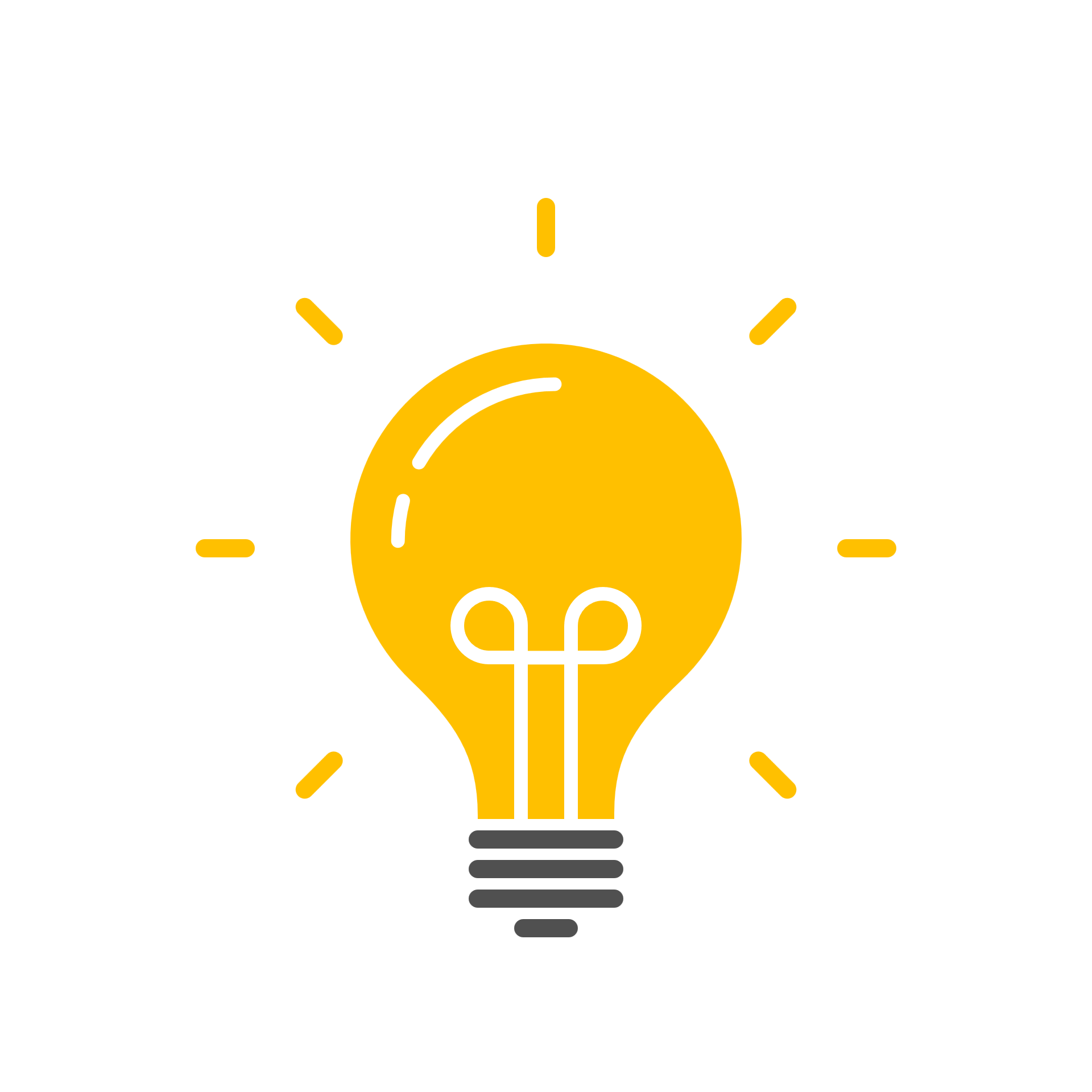
2\_Vital signs are the main influence

****Model Selection

The dataset prosperities are crucial aspect and has a huge impact on the selected models.

After taking into consideration type of problem, size of the dataset, the number of features.

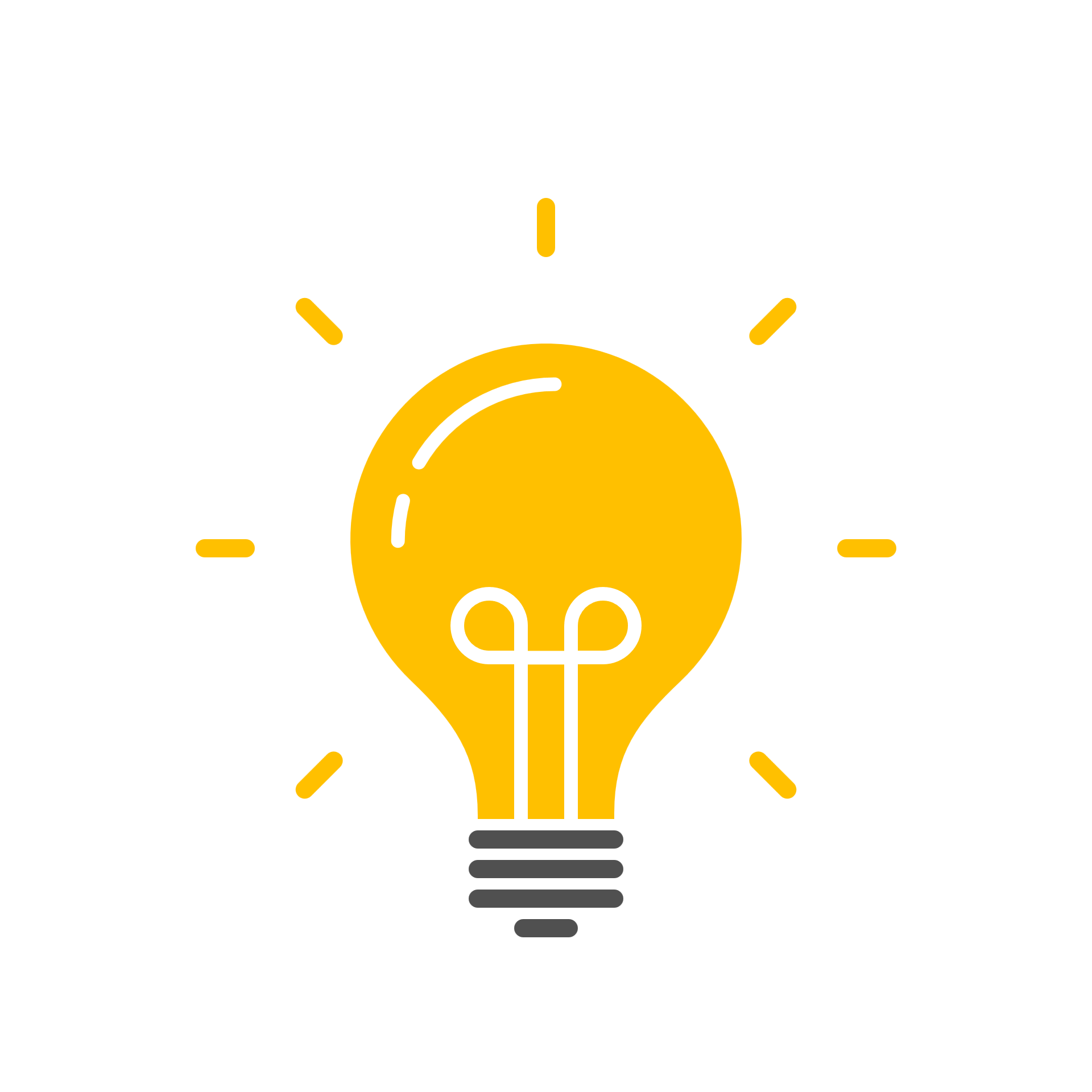
We chose different machine learning algorithms to build classifier which are Logistic Regressions, Random Forest, and Decision Tree.

****Training phase

We have effectively able to build a model that predicts the need for intensive care based on vital information.

Overall, the trained models results range from moderate to accurate prediction results.

One important observation about the models is that we noticed that the Random Forest give us the best result. if a patient of a confirmed COVID-19 case will require admission to the ICU.

**Conclusion**

Vital information provides great opportunities to

Enhance decision-making in hospital and clinics,

Which may relieve pressure, especially at this stage.

For this study , we make a critical study of the

Process of selecting the features and even verifying

The correctness of our hypothesis to use only the

Vital signs. Then , we trained ML algorithms using

Information’s after processing the data.

We have achieved good results for the for the Data

With accuracy up to 88% and this indicates that we

Can rely on the vital signs for the covid-19 patients As a strong indicator that they need ICU admission.