Universidad Politécnica de Madrid

Escuela Técnica Superior de Ingenieros Informáticos

Máster Universitario en Ingeniería Informática

Data Processes

Course project - Project plan

2021-22 | SEPT-FEB

Author/s:

Miguel Alonso, Carlos
Num. matrícula: 170243

Fernández Molleda, Lucía
Num. matrícula: 170312

Leira García-Baamonde, Manuel
Num. matrícula: 170136

January 9, 2022

Table of Contents

1 Executive Summary

COVID19 pandemic has hit society in the last two years, and, one of the main signs of the severity of each wave has been the number of patients in the hospital's ICU[1].

Since the beginning of the pandemic, more than 5 million people have died[1] because of COVID19 (or due to related circumstances caused by COVID19), and the ICU of hospitals has been saturated to levels never before seen and, since there aren't that many beds in the ICU, when they're saturated, many patients do not receive the appropriate treatment, leading to either death or poor treatment, since there's not enough material for all patients[2].

Predicting the curve of potential ICU patients will give ICU's management team a great advantage to prepare more beds, or organize the appropriate response to a sudden spike/surge of critically ill patients, because on ICU, there aren't only COVID19 patients, and all of them need treatment.

2 Scope

This project will be focused on analysing the relationship between the patients that enter the ER[2] and their first health measurements, such as blood pressure, blood O2 saturation, heart rate, etc. Their age will be taken into account as-well, since COVID19 affects severely the elder population.

[1] Intensive Care Units
[2] Emergency Room

3 Goals

3.1 Business goal

The main business goals of this project will be:

BG1 Improving the ICU saturation prediction based on new patients health measurements after arriving at the ER

BG2 Improving the prediction of patients survivability

This goals will help the management of ICU based on new patients, which will enable the preemptive mobilisation of new resources, so the ICU does not collapse.

Table 1: Details of each business goal

Business Goal	Description	Indicator of success
Predict ICU saturation	Predict number of patients that will require ICU beds based on incoming patients health state	???
Predict patients survivability	Predict the probability of survivalof a patient based on it's health state	??

3.2 Data Mining goals

To achieve this goals, we must analyse the data we already have to understand the situation of the patients and their condition when the first arrive at the ER.

Through analysing the data, we must be able to get answers for the next questions/goals in order to give an answer to our business goals:

DM1 Analyse and understand patients conditions through variables on dataset

DM2 Curve of survivability of each patients age group

DM3 Curve of number of days on ICU based on patients condition.

DM4 Likelihood of a patient entering ICU based on patients age.

DM5 Likelihood of a patient entering ICU based on patients condition.

The next table shows this same goals, but with much more detail in regard of their relation with the business goal and the results their achievement:

Table 2: Details of each data mining goal

Goal	Attributes	Dataset	BG it helps to achieve (%)
DM1	*all*	*[3]	BG1 (15 %), BG2 (15 %)
DM2	EXITUS, AGE	*[4]	BG2 (25 %)
DM3	DAYS_ICU TEMP HEART_RATE, SAT_O2, BLOOD_PRES_SYS, BLOOD_PRES_DIAS	*[4]	BG1 (25%)
DM4	DAYS_ICU, EXITUS	*[4]	BG2 (60 %)
DM5	DAYS_ICU TEMP HEART_RATE, SAT_O2, BLOOD_PRES_SYS, BLOOD_PRES_DIAS	*[4]	BG1 (60 %)

Through this smaller goals, we should be able to get a good understanding of the condition of a patient that enters ICU and make predictions ICU occupation.

4 Work Plan

This project must follow a structured and well-though out plan, in which each objective and milestone is specified following a schedule.

4.1 Deliverables

Throughout this project, there will be 2 deliverables:

D1 The final assessment, conclusions and answers to the business goals as a result of the data analysis.

D2 Any code and instructions used to perform said analysis, so anyone with access to this resources could verify this conclusions.

4.2 Milestones

Throughout this project, there are a series of objectives (or milestones) we have to achieve to continue with the development of said project. This milestones are related to both the Business Objective and Data Mining Goals, and are:[o]

M1 Study the dataset and its attributes

M2 Clean dataset of inconsistencies and errors

M3 Get results for **DM1**

M4 Get results for **DM2** and **DM3**

M5 Get results for **DM4**

M6 Get results for **DM5**

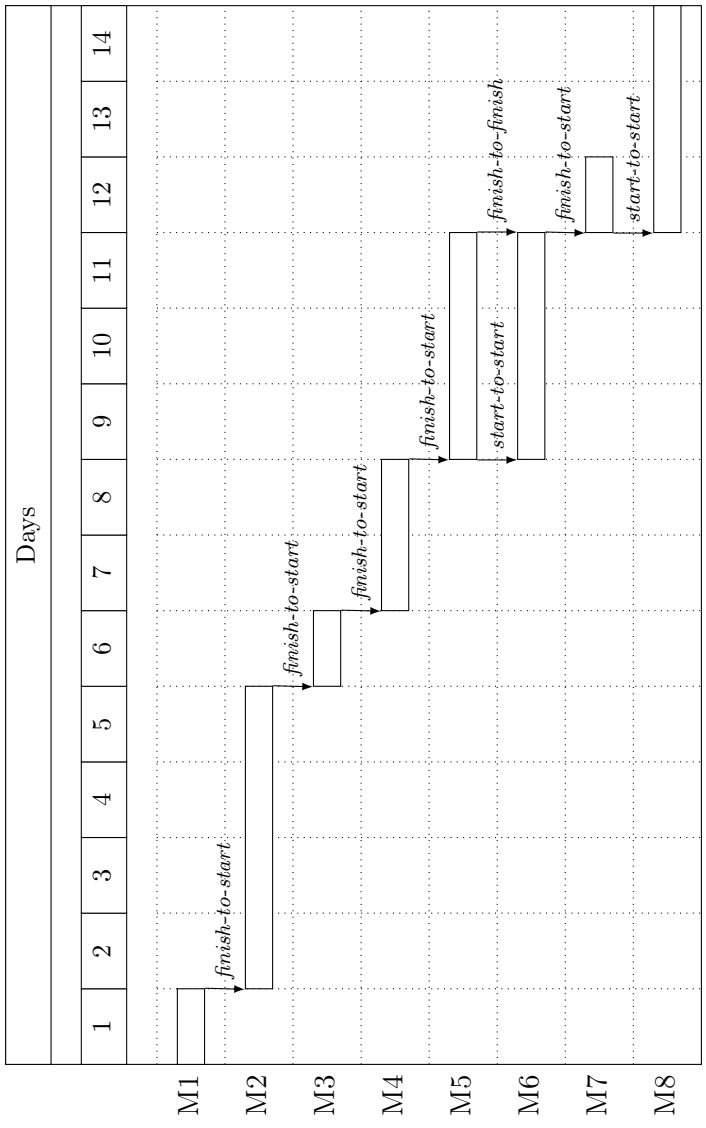
M7 Give an answer to each **BG**

[4] COVID19_data.csv

M8 Write down results on the *Technical Report*

This milestones will be used to plot *Gannt* diagram on the subsection 4.3, in which the work towards a certain milestone throughout time will be shown. This time is not strict and, if able, milestones could start ahead of schedule.

4.3 Gannt Diagram



5 Budget

This project will have enough budget to count with 4 software engineers, which will carry-on this analysis and develop all the code and documentation required.

If a 4th cannot be found, then the 3 remaining will develop this project in the established window of time.

References

[1] WorldOMeter. *Coronavirus Statistics.* URL: https://www.worldometers.info/coronavirus/.

[2] Reuters - New York Post. *Paris ICUs 'at near saturation' as severe COVID-19 cases hit 3-month high.* URL: https://nypost.com/2021/03/09/paris-region-icus-at-near-saturation-as-severe-covid-19-cases-hit-3-month-high/.