

## study case

This is a study case which develop the methodology outlined by John Rollins, to answer this question.

How can we predict the number of patients in a hospital to have an efficient staff allocation?



### 1. Business understanding

A big problem in hospitals is the staff capacity you need that change each day. In fact, every staff manager in a hospital faces this problem in a given period. This is a big issue because if they are unable to predict workforces needed it can lead to staff overwork, worker burnout, and to poor patient care. On the other hand, it can result in losing money for the hospital if they have too many workers for the number of patients.

So, we are going to try anticipating patient volumes in a hospital to help the staffing process.

### 2. Analytic approach

We can get now to the second step which is the analytics approach. We need to define the method we want to use to solve our problem. We want to make a prediction on the future to know when we will have patients in the hospital. To do so, we are going to use a machine learning algorithm. The algorithm will use the data of the previous patients but also the data on the external dataset such as weather, national epidemics, holidays and many more. It will then help us determine at what time of the year the hospital need more capacity.

### 3. Data requirements

We need to define which type of data we want to use in our machine learning algorithm. If we focus on one hospital, we need the day to day history of all the patients for at least the last five years. Once we have collected all the data, we need to determine the contents and the formats of the data to put in our algorithm. We require on record per day with all the information such as the number of patients, the hours when they go to the hospital, the information's about holidays, weather...

### 4. Data collection

For the data collection, we had no problem to get what we want, the hospital keeps medical records of his patients. It's also easy to get weather and holiday records.

## 5. Data understanding

In this step we try to understand the content the quality and the insights of our dataset. To do so, we run descriptive statistics on each variable of our dataset. For example, we can look when we have the maximum patients in the hospital, or what is the mean. We also try to observe correlation between variable for example between number of patients and epidemics.

## 6. Data preparation

The goal of this step is to process the data in a state where it is easier to work with. In our case we first combine all the data. We organize it in a way where we have for each day different attributes that give us information on the patients, the weather etc. Then, we deal with the missing values, for example when we don't have the number of patients in a certain day, we choose to skip it, otherwise it would change the result.

## 7. Modeling

Now that the data is prepared, we can finally develop the model. In our case we need to develop a predictive model to know in advance when there will be lots of patients in the hospital. To do that, we first work on a training set which is 70% of all our dataset. With this dataset, we are going to train a neural network to find out for the next month the number of patients. In our neural network we need to adapt the weight of each neuron to make it more accurate.

## 8. Evaluation

Once the model is developed, we test it with the test dataset which is 30% of the full dataset. We then plot the accuracy of our model and adapt it according to the results. The modelling and evaluation steps are highly iterative, and we need to go from one to another to get the best accuracy.

## 9. Deployment

Now that we are satisfy with the model, we can finally deploy it. By working with developers and database administrators we can make an application where all the staff can see the previsions for the next month of the numbers of patients, they can also directly put in the application the real number of patients they get each day.

## 10. Feedback

Now that the model is deployed, we need to incorporate the news information's we get. By analysing the feedback enables the we can refine the model, increase the accuracy. For example, if we were wrong on the prediction on certain day we need to understand why and review our model. For example, we might have forgotten some external factors like big event that can change the number of patients.