

## Capstone Project – Synopsis

Project Title	Predicting a patient's LOS in a hospital
Domain of Project	Healthcare
Team members	Debesh Das Saptarshi Sit , Chirag Rai
Mentor Name	P V Subramanian

*P.V. Subramanian*

**Signature of the Mentor**

## Table of Contents

1. OVERVIEW.....	3
2. Business problem statement (Goals).....	3
2.1. Business Problem Understanding .....	3
2.2. Business Objective .....	3
3. Approach .....	4
4. Conclusion .....	4
5. Topic survey in Brief.....	4
5.1. Problem understanding .....	4
5.2. Current solution to the problem.....	5
5.3. Proposed solution to the problem .....	5
5.4. Reference to the problem.....	5
5.5. Critical Assessment of Topic Survey .....	5
5.6. Methodology to be followed .....	6
6. References .....	7

# Project Details

## 1. OVERVIEW

In recent days, Data Science is having wide applications across many industries of the world. It has revolutionized healthcare and the medical industry in large ways. In this project, we are using Data Science techniques to determine Length of Stay (LOS) of a patient in a hospital.

LOS of patients is one of the most closely watched metrics an important factor in the efficiency of hospital management, use of medical services, patient quality of care, and functional evaluation.

Absence of accurate LOS knowledge beforehand, causes problems to both hospital and patient. Wrong or inadequate knowledge of LOS, leads to patient receiving inadequate treatment and more expenses. For a hospital, it results in poor bed management, improper time scheduling & inefficient treatment.

## 2. Business problem statement (GOALS)

### 2.1. Business Problem Understanding

Absence of accurate LOS knowledge beforehand, causes problems to both hospital and patient.

For a patient, whose LOS was wrongly predicted as less, he receives inadequate treatment.

For a patient, whose LOS was wrongly predicted as high, he incurs more expenses

For a hospital, it results in poor bed management, improper time scheduling & inefficient treatment.

Hence, a high degrade in patient satisfaction which leads to decrease in patient footfall and ultimately decrease in hospital's profits.

### 2.2. Business Objective

Our objective is to create a model that predicts the LOS quickly and accurately for each patient at the time of his admission in the hospital.

This will lead to the benefit of both the patient and hospital

For a patient, whose LOS was correctly predicted as less, he does not incur extra expenses.

For a patient, whose LOS was correctly predicted as more, he receives adequate treatment.

For a hospital, it results in better bed management, proper time scheduling & efficient treatment.

Hence, a high upgrade in patient satisfaction which leads to increase in patient footfall and ultimately increase in hospital's profits.

### **3. Approach**

Traditional approach used by hospitals which involved manual consultancy with doctors, resulted in delayed and not-so-accurate predictions of LOS.

Instead, we involve machines to do the tedious and time-taking task at a faster and efficient manner.

We use Predictive modelling technique which involves exploratory data analysis, data wrangling process, statistical tests, different machine learning models to predict LOS range.

### **4. Conclusion**

As, our model gives accurate LOS info at a faster pace without involving manual doctor's consultancy each time for every individual patient. It also leads to increased patient satisfaction which in turn leads to increase in hospital's profits. For these reasons, we would urge the hospital to deploy our model in production environment.

## **5. TOPIC SURVEY IN BRIEF**

### **5.1. Problem understanding**

In India, there are over 23,000 hospitals with around 7 lakh beds. In some states, government hospital beds are being shared by patients.

Absence of accurate LOS knowledge beforehand, causes problems to both hospital and patient.

Wrong or inadequate knowledge of LOS leads to patient receiving inadequate treatment and more expenses. The longer a patient stays in the hospital, the greater the risk they will develop a healthcare-acquired infection (HAI) that they can become vulnerable to.

For a hospital, it results in poor bed management, improper time scheduling & inefficient treatment. Hospitals benefit from a shorter LOS. They do not have to cover the expense of treating an HAI and they free up beds for new patients. LOS of patient impacts hospital's bed availability, which in turn impacts overall patient flow. Inaccurate knowledge of LOS of a patient can disrupt the overall hospital management system.

## **5.2. Current solution to the problem**

Currently there are two solutions to the problem of inadequate / inaccurate LOS prediction.

- Manual consultancy with doctors and hospital staff about LOS for each incoming patient.  
This requires a lot of time of doctors' intervention as the process is executed for each patient.
- Use of spreadsheets in compiling, reporting, and graphically depicting statistical data.  
This requires a lot of time and is not so user-friendly.

## **5.3. Proposed solution to the problem**

Involve machines to do the tedious and time-taking task at a faster and efficient manner

We use Predictive modelling technique which just requires the hospital staff to input the necessary values. Then let our model predict the LOS for each patient.

## **5.4. Reference to the problem**

Recent Covid-19 Pandemic has raised alarms over the efficiency of management of hospitals.

Every Covid-treatment hospital is almost full of patients occupying every available bed.

Hence, LOS is one critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital.

## **5.5. CRITICAL ASSESSMENT OF TOPIC SURVEY**

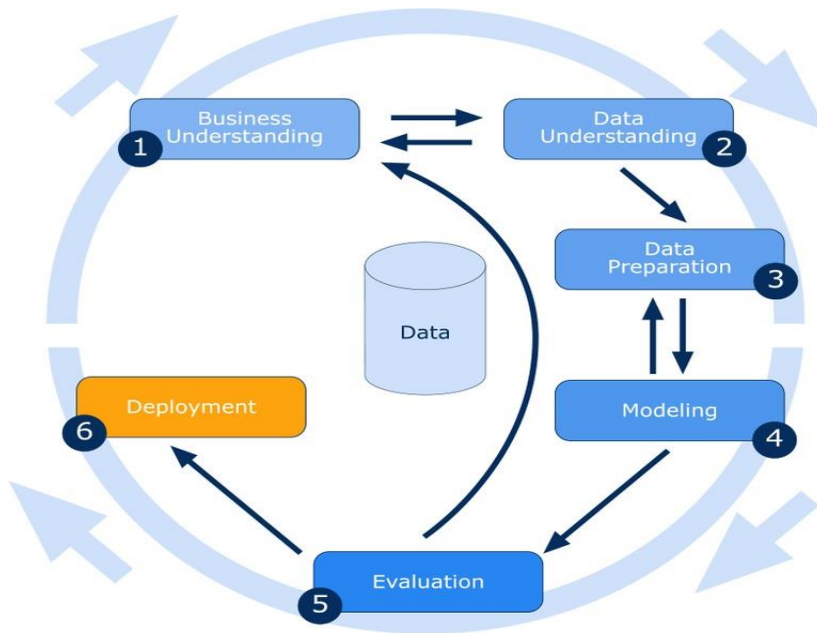
### **1. Key area, gaps identified, where the project can add value to the customers and business**

- Shorter hospital stays reduce the burden of medical fees
- Increase the bed turnover rate
- Increases the profit margin of hospitals, while lowering the overall social costs.

### **2. Key gaps we are trying to solve**

Adequate & accurate knowledge of LOS benefits both the hospital and the patient.

## 5.6. METHODOLOGY TO BE FOLLOWED



### 1. Business Understanding:

Predicting Length of stay (LOS) of patient to increase the efficiency of hospital management and reducing the cost incurred by patient.

### 2. Data Understanding:

Dataset includes numerical and categorical data type, while target variable being categorical. We have to drop many unnecessary columns for target column prediction.

### 3. Data Preparation: Using the following techniques:

- Missing value treatment
- Outliers treatment
- Dropping non-significant features
- Visualizing features using suitable plots

### 4. Modelling

Gathering firsthand information using train test split and gauging the performance and going ahead by performing different classification machine learning models on the data set.

### 5. Evaluation:

Evaluating the model performance on the parameters such as `f1_score` , `classification_report` , `bias-error` , `variance-error`

## 6. REFERENCES

- <https://datahack.analyticsvidhya.com/contest/janatahack-healthcare-analytics-ii/#ProblemStatement>
- <https://www.kaggle.com/nehaprabhavalkar/av-healthcare-analytics-ii>
- <https://towardsdatascience.com/predicting-hospital-length-of-stay-at-time-of-admission-55dfdf69598>