

**Project Design Phase-I**  
**Proposed Solution Template**

Date	19 September 2022
Team ID	PNT2022TMID23263
Project Name	Project – 10738-1659200545 Analytics For Hospitals' Health-Care Data
Maximum Marks	2 Marks

**Proposed Solution Template:**

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	To predict the Length of Stay(LOS) of the patient to get information for optimal resource allocation and better functioning.
2.	Idea / Solution description	The Length of Stay(LOS) of the patients depends on the major factors such as type of disease, age and severity . The data is pre-processed first according to the most important details from the dataset. The dataset is explored and visualized and then using the techniques of ensemble algorithms consisting of many decision trees prediction model is developed.
3.	Novelty / Uniqueness	The problem in real time is to find the availabilities. The uniqueness of our proposal is to convey the availabilities to the consumer with maximal accuracy.
4.	Social Impact / Customer Satisfaction	It helps to identify patients of high LOS risk (patients who will stay longer) at the time of admission. Once identified, patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics such as room and bed allocation planning. The problem is to manage the functioning of Hospitals in a professional and optimal manner.

5.	Business Model (Revenue Model)	<p><b>Training Process</b></p> <p>Outcome variable      Prediction Type      Evaluation Metric</p> <p><b>Step 1: Develop a classifier with best AUPRC (Select based on performance as a classifier)</b></p> <pre> graph LR     A[Short Stay(&lt;=7 days) or prolonged Stay(&gt;7 days)] --&gt; B[Classifier]     B --&gt; C{Predicted Value &gt;= 7 days}     C -- Yes --&gt; D{True Value &gt;= 7 days}     C -- No --&gt; E{True Value &gt;= 7 days}     D -- Yes --&gt; F[No error]     D -- No --&gt; G[Error between 7 days and True Value]     E -- Yes --&gt; H[Error between 7 days and Predicted Value]     E -- No --&gt; I[Error between predicted values and true values]   </pre> <p><b>Step 2: Develop one regressor (select data based on true outcome)</b></p> <pre> graph LR     J[Numerical Outcome (only &lt;=35 days)] --&gt; K[Regressor]     L[Log Transformation] --&gt; J   </pre> <p><b>Testing Process</b></p> <pre> graph TD     M[A testing sample] --&gt; N[Classifier]     N -- "Result: Short LOS" --&gt; O[Regressor (Short LOS)]     N -- "Result: Long LOS" --&gt; P[Set it to the 7 days]     O --&gt; Q[Exponentiate]   </pre>
6.	Scalability of the Solution	<p>The primary model of our solution is to target only less consumers. So it is sufficient to implement with local servers.</p> <p>In future, it can be extended to the large scale on needs .At that time the usage servers is considerably high. So it can be extended further to the Cloud services.</p>