

Summer Internship/ Project (2023)

Group Members:

K Nithya Sree (Reg No: 2021115071)
Shruthi B (Reg No: 2021115106)
Vasumathy V (Reg No: 2021115118)

Project Guide:

Dr. K. Indra Gandhi
(Department Of Information Science
And Technology)

Project Domain: Machine Learning / Deep Learning

Project Abstract:

The following document contains necessary information regarding the project, including the problem statement, tools utilized along with a divisive timeline for the same.

Problem Statement:

To ascertain the scoring methodology used in the NIRF Ranking Parameters and predict the score using the current data value and previous year Data set.

Project Description

The project is based on the National Institutional Ranking Framework (NIRF). This framework outlines a methodology to rank institutions across the country. Each institution is ranked based on certain parameters. These parameters are organized into five broad heads, and have been further elaborated into suitable sub-heads. For the purpose of this project, we focus on a dataset of past and present NIRF Ranking list of the main 5 parameters. The overall score is computed based on the weights allotted to each head. The weightage for each parameter varies.

The overall score can take a maximum value of 100. Based on this overall score, the colleges are allotted a rank.

The project involves various stages. A brief explanation of the stages is given below. The initial step is to understand the NIRF ranking parameters. We then proceed to collect relevant data. Then we apply techniques such as scaling, normalizing to the data to ensure that the indicators are on a comparable scale. We choose an appropriate Machine learning (ML) algorithm based on this specific problem type and train the chosen model.

We then calculate the ranks for the current and previous year. Using the current year and the previous year data, we analyze the changes in the indicators and calculate the predicted scores accordingly. This involves comparing the indicator values, assessing the impact of any changes, and applying appropriate adjustments to the previous year's scores. We then proceed to validate the results by comparing with the actual NIRF rankings to assess the accuracy and reliability of the predictions. Finally, we prepare the model for deployment, integrate the model with the front-end web application and test its functioning.

Tools Utilized:

The machine learning algorithms will be implemented in python, using various libraries such as TensorFlow, SciKitLearn, Pandas and NumPy, which are commonly used modules to implement mathematical functions and machine learning algorithms.

We will build a Rest API to deploy this model onto a Web Application (built using HTML, CSS and ReactJs) using Flask.

The primary database used is SQLite / PostGres.

The “previous year datasets” will be fetched from Kaggle, a common machine learning platform.

Other supplementary libraries used will be constantly added to the list of tools.

Project Timeline:

As a team, we came to the consensus that we will start the project on the 3rd of July, 2023. As the project will be worked on over the course of four weeks, a tentative date of completion is the 30th of July, 2023. This includes the completion of the project, documentation, as well as the corresponding report.

The timeline of this project is as follows. It is to be noted that this is subject to minor changes, due to the convenience of team members, along with the time necessary to make changes that were mentioned by our project guide.

Week 1 - 3rd July to 9th July

During the course of the week, multiple discussions regarding the project topic, discovering various aspects and implementation of the problem statement, framing the required abstract, and building a functional timeline will be performed. Along with this, members of the team will also engage in exploring the various tools (mentioned previously), and learning topics that may seem unfamiliar. Segregation of tasks adhering to the skill set of the members will also be done during this period of time.

Week 2 - 10th July to 16th July

The web application to deploy the model will be built from scratch (utilizing basic client side frameworks and front end tools). In areas

where we feel necessary, bootstrap will also be used. Using ReactJS, client side validation will be done, thus creating a fully functional dynamic web page.

Week 3 - 17th July to 23rd July

Implementing a machine learning model suitable for the given problem statement. This involves the following tasks:

1. Data collection and Data Preprocessing - cleaning, preprocessing and normalization of the data
2. Choosing an appropriate model, and training said model
3. Evaluating the aforementioned model using regression and classification specifics
4. Using the model for prediction

Week 4 - 24th July to 30th July

Deploying the newly created model through the web application created earlier, using necessary back-end frameworks to do so (such as Flask) will be done during this period of time.

We will also ensure that the deployment is continuously monitored.

Feedback will be collected periodically, and the model's performance will be compared against real world data.

Finally, the model will be documented and a project report will be framed based on the successful results of the project.