**Fabrication of a Machine Learning Platform**

**Synopsis**

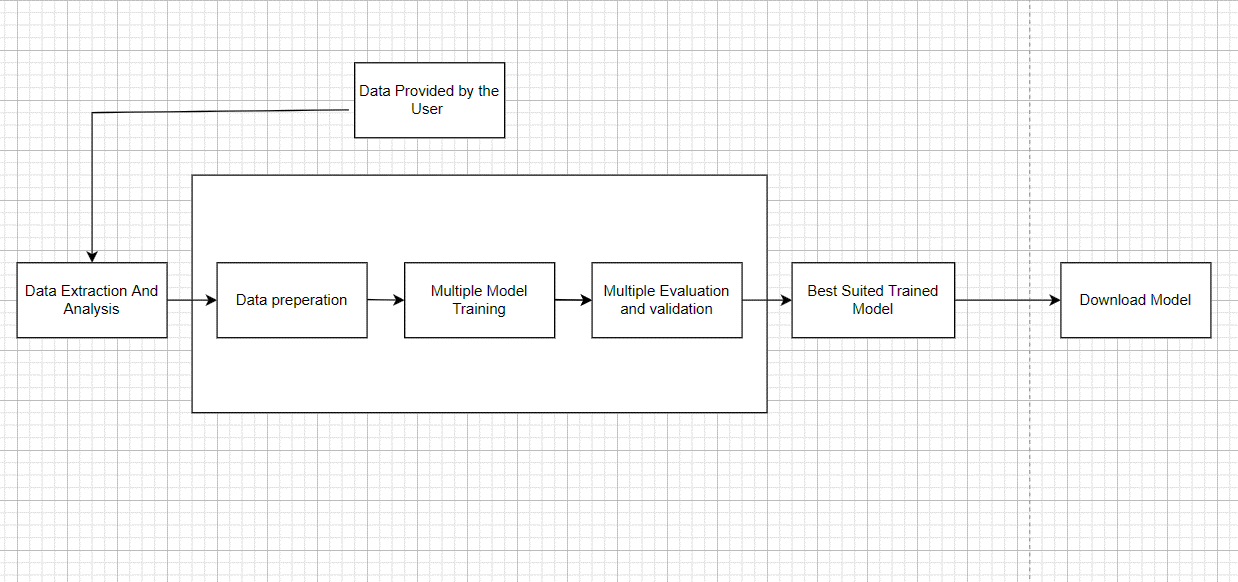
As stated in the GovTech context, Over the last few years, we have seen technology used extensively by the government in various fields such as automation, education, health, energy management, agriculture, etc. Even though we have the right means as well as the quantity in order for the predictive analysis of this data, the lack of time and cost for the creation of best machine model may lead to improper training of data that further results in substandard accuracy. The above problem can be solved by development of a platform where an individual, with limited knowledge of statistics and machine learning, can pre-process and train the data without any code and select the best model by tuning the algorithms and parameters with just a few clicks.

**Motivation and required skills for the project**

The main goal is to make machine learning more accessible to the general public with limited knowledge about machine learning as well as predictive statistics. A user can upload a dataset and the platform will help the customer to preprocess the data and find the best suitable model within a certain period of time. This will not only help in reduction of time but retrenchment too in each and every sector possible.

Skills required for development of the project:

* Understanding of probability and statistics
* Python (Matplotlib, SciKit, NumPy, Pandas, SciPy)
* PyCaret (an open-source, easy-to-learn, low-code machine learning library in Python)
* Front-end Technologies (HTML/CSS, Bootstrap, JavaScript, ReactJS)
* Back-end Technologies (Django)
* Database (MySQL, OracleSQL)

**Project Details**

**Features of the project must include:**

1. Preprocessing of the data – Data cleaning (missing value imputation, removal of noisy data, outlier treatment), Data integration, Data transformation, Data Reduction
2. Data Visualization - This feature can be achieved through Streamlit library incorporated by python.
3. Machine modeling –
   1. For regression: Linear Regression, Logistic regression, Knn, SVM, Decision tree Regressor etc.
   2. For Classification: Logistic classification, Gradient Classifier, Decision Tree Classifier, Naïve Bias etc.
4. Selection of the model – selection of the best suitable model can be accomplished by TOPSIS algorithm.

**Problem faced**

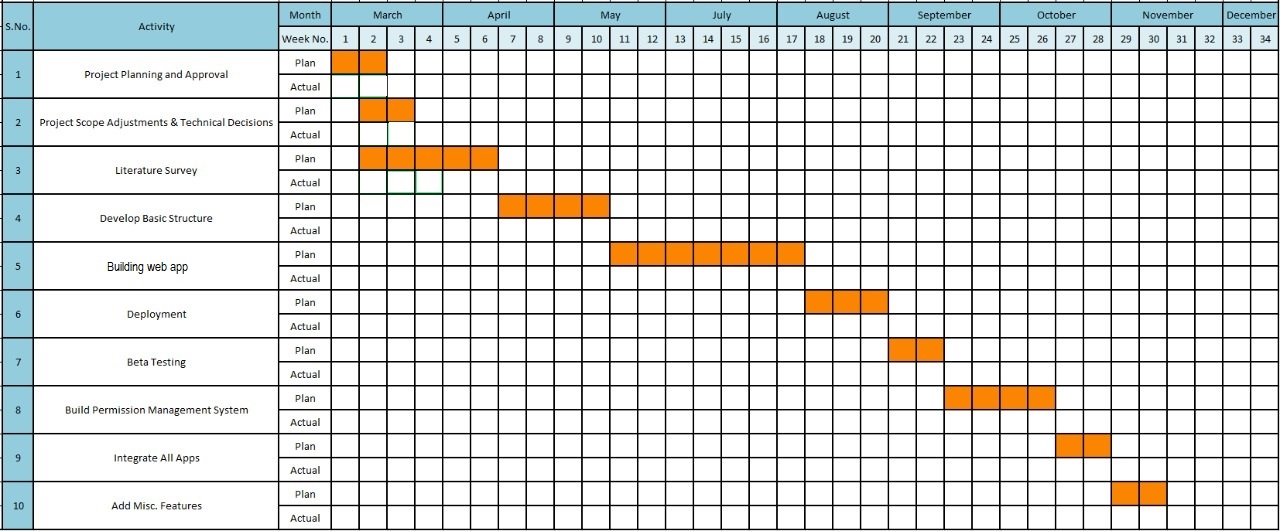
Since the project will be on a small scale at first, traffic and uploading of large quantity data may lead to crashing of the server

**Solution**

Upgrading web hosting through installing caching plugin, switching to WordPress, inclusion of load balancer (AWS).

**Timeline:**

The project will require around 40-60 days depending upon the size and knowledge of team.



\*\*The units may depict 1.5-2 days.

**Future Development**

For future development, Maintenance of the website and customer service should be integrated for increasing users and helping them.

**Availability**

I am available for June-July, 2022 and I can devote 30-40 hours per week to the project

**Personal Background**

My name is Rahul Kapoor and I am a pre-final year student, pursuing Bachelor of Engineering (C.O.E) from Thapar Institute of Engineering and Technology. My short-term goal is to enhance my academic skillsets to prepare me for the professional challenges. My specific areas of interest are software development, front-end development and machine learning.

**Further details:**

Resume: <https://drive.google.com/file/d/149k-HciqMAZRcTIGr31YUowAodyvqY3G/view?usp=sharing>

Portfolio: <https://rahul21200.github.io/My-Page/>

GitHub: <https://github.com/rahul21200>

LinkedIn: <https://www.linkedin.com/in/rahul-kapoor-6714821b3/>

Email: [rahulkapoor902@gmail.com](mailto:rahulkapoor902@gmail.com)

Contact Information: +91 8860030195, +91 8178241817

Given the skills and experience outlined in my enclosed resume, I believe I will be a worthy asset to your team. I assure that I’ll give my full potential to this project in order to impact the society and create awareness about the importance of machine learning in today’s time period.

Thank you for reading my proposal, I sincerely hope that my idea will be appreciated by the team.

Thanks and regards,

Rahul Kapoor

[rahulkapoor902@gmail.com](mailto:rahulkapoor902@gmail.com)

+91 8860030195, +91 8178241817