

A **PROJECT PHASE I** REPORT **(17ISP78)** ON

“No Code ML”

submitted in partial fulfillment of the Requirements for 7th Semester of

### BACHELOR OF ENGINEERING

**in**

**INFORMATION SCIENCE & ENGINEERING**

**by**

**Vishal Kundar <1CR17IS110>**

**Viraj Patel <1CR17IS107>**

**Vedant Barbhaya <1CR17IS106>**

**Under the guidance of**

**Prof. Vidya U**

**Assistant Professor,**

**Dept. of ISE, CMRIT, Bengaluru**



**CMR INSTITUTE OF TECHNOLOGY**

**DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING**



**DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING**

Certificate

This is to certify that the project work entitled, **“No Code ML”**, is a bonafide work carried out by **Vishal Kundar <1CR17IS110>, Viraj Patel <1CR17IS107> and Vedant Barbhaya <1CR17IS106>** in partial fulfillment for the award of the degree of **Bachelor of Engineering in Information Science & Engineering** of Visvesvaraya Technological University, Belgaum, during the year 2020-21. It is certified that all corrections/suggestions indicated during reviews have been incorporated in the report. The project report has been approved as it satisfies the academic requirements in respect of the project work prescribed for the said Degree.

|  |  |
| --- | --- |
| **Name & Signature of Guide** | **Name & Signature of HOD** |
| **(Prof. Vidya U)** | **(Dr. Shreekanth Mooroor Prabhu)** |
| **Dept. of ISE, CMRIT** | **Dept. of ISE, CMRIT** |
|  |  |



### DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING

DECLARATION

We, the bonafide students of 7th semester, Information Science & Engineering, CMR Institute of Technology, Bangalore, hereby declare that the report entitled **“No Code ML”** has been carried out by us under the guidance of **Prof. Vidya U, Assistant Professor**, Information Science & Engineering Department, CMRIT Bangalore. This dissertation work is submitted in partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering in **Information Science Engineering**, of the Visvesvaraya Technological University, Belgaum during the academic year 2020-2021. Further the matter embodied in the project report has not been submitted previously by anybody for the award of any degree or diploma to any university.

Place:

Date:

##### Team members:

1. Viraj Patel (1CR17IS107) **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
2. Vishal Kundar (1CR17IS110) **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
3. Vedant Barbhaya (1CR17IS106) **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

# ABSTRACT

* No Code ML creates a platform where a user can upload a data set and get personalized predictions with just one click. This eliminates the need to clean or process the data or create a machine learning model.
* The entire machine learning pipeline from data preprocessing is automated, giving the users an easy way to build machine learning models.
* Everything will be handled by our platform and the user will get a personalized dashboard to analyze the outcomes.

# Acknowledgment

We take this opportunity to express our sincere gratitude and respect to **CMR Institute of Technology, Bengaluru** for providing us a platform to pursue our studies and carry out our final year project

We have great pleasure in expressing our deep sense of gratitude to **Dr. Sanjay Jain,**

Principal, CMRIT, Bangalore, for his constant encouragement.

.

We would like to thank **Dr. Shreekanth Mooroor Prabhu**, Professor & HOD, Department of Information Science and Engineering, CMR IT, Bangalore, who has been a constant support and encouragement throughout the course of this project.

.

We consider it a privilege and honor to express our sincere gratitude to our guide **Prof. Vidya U,** Assistant Professor, Department of Information Science & Engineering, for her valuable guidance throughout the tenure of this review.

Finally, we would like to thank all our family members and friends for all their moral support they have given us during the completion of this work.

#### Viraj Patel (1CR17IS107)

#### Vishal Kundar (1CR17IS110)

#### Vedant Barbhaya (1CR17IS106)

|  |  |  |
| --- | --- | --- |
| **TITLE** | | **Page No.** |
|  | |  |
|  | |  |
| **Abstract** | | **4** |
| **Acknowledgement** | | **5** |
| **Table of contents** | | **6** |
| **1** | **INTRODUCTION** | **7** |
| **2** | **LITERATURE SURVEY** | **8** |
|  | **Paper 1** |  |
|  | **Paper 2** |  |
|  | **Paper 3** |  |
| **3** | **PROBLEM DEFINITION**  **3.1 Problem Statement**  **3.2 Data Flow diagram** | **11** |
| **4 OBJECTIVES & METHODOLOGY** | | **12** |
| **4.1 Work break down**  **4.2 Timeline** | |  |

## CHAPTER 1

**INTRODUCTION**

**Summary:**

A GUI platform for handling ML needs for small and medium businesses, by automating the entire machine learning pipeline from data preprocessing to visualization.

**Objectives:**

Interactive platform to get personalized predictions with just one click with no need to clean or process the data or create a machine learning model and get a personalized dashboard to analyze the outcomes.

Advanced settings to allow the user to choose a model or compare various model results, and feature relationship.

**Outcome:**

A simple to use website for creating machine learning models and analyzing data.

### CHAPTER 2

**LITERATURE SURVEY**

##### 

##### **Paper 1**

<<https://sci-hub.do/10.1109/ijcnn.2004.1380929r>>

Peng-Wei Chen, Jung-Ying Wang and Hahn-Ming Lee, “Model Selection of SVMs Using GAApproach ” In *IEEE International Joint Conference on Neural Networks*, 2004.

A new automatic search methodology for model selection of support vector machines, based on the CA-based tuning algorithm, is proposed to search for the adequate hyper parameters of SVMs. In this method, each chromosome indicates a group of hyperparameters, and the population is a collection of chromosomes. Experimental results show that this method performs superiorly on time cost, performance and stability. This algorithm only requires the evaluation of an objective function to guide its search with no additional derivative or auxiliary knowledge required. In addition, the encoding of chromosomes makes the implementation of multiple hyperparameters tuning simpler. This paper marked the beginning the first efficient strategies for hyperparameter optimization.

##### **Paper 2**

<<https://www.isi.edu/~gil/papers/gil-etal-automl18.pdf>>

Yolanda Gil, Ke-Thia Yao, Varun Ratnakar, Daniel Garijo, Greg Ver Steeg, Pedro Szekely, Rob Brekelmans, Mayank Kejriwal, Fanghao Luo and I-Hui Huang, “A Phased Performance-Based Pipeline Planner for Automated Machine Learning”

While many problems could benefit from recent advances in machine learning, significant time and expertise are required to design customized solutions to each problem. Prior attempts to automate machine learning have focused on generating multi-step solutions composed of primitive steps for feature engineering and modeling, but using already clean and featurized data and carefully curated primitives. However, cleaning and featurization are often the most time-consuming steps in a data science pipeline. This paper puts forward a novel approach that works with naturally occurring data of any size and type, and with diverse third-party data processing and modeling primitives that can lead to better quality solutions. The key idea is to generate multi-step pipelines (or workflows) by factoring the search for solutions into phases that apply a diiferent expert-like strategy designed to improve performance. This approach is implemented in the P4ML system, and demonstrates superior performance over other systems on a variety of raw datasets.

##### **Paper 3**

<<https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/46180.pdf>>

Daniel Golovin, Benjamin Solnik, Subhodeep Moitra, Greg Kochanski, John Karro, D. Sculley, “Google Vizier: A Service for Black-Box Optimization”

Any sufficiently complex system acts as a black box when it becomes easier to experiment with than to understand. Hence, black-box optimization has become increasingly important as systems have become more complex. In this paper researchers describe Google Vizier, a Google-internal service for performing black-box optimization that has become the de facto parameter tuning engine at Google. Google Vizier is used to optimize many of Google’s machine learning models and other systems, and also provides core capabilities to Google’s Cloud Machine Learning HyperTune subsystem. This paper discusses the requirements, infrastructure design, underlying algorithms, and advanced features such as transfer learning and automated early stopping that the service provides.

## CHAPTER 3

# PROBLEM DEFINITION

**3.1 Problem Statement:**

Data is the oil of the 21st century. Essentially if you are running a service/business, every interaction with a customer generates some kind of valuable data. With the advancements of intelligent systems, it has become possible to extract very valuable information from any given data.

But not every business has the ability or means to develop an in house team to analyze vast amounts of data generated by them. No code ML helps businesses to analyse their data using ML without actually needing the knowledge of how to code it. Our system will give the user a personalised dashboard to analyse their outcomes.

**3.2 Data Flow Diagram:**

## Chapter 4

**PROJECT ORGANISATION & TIMELINE**

**4.1 Work Breakdown**

The project development team have each been assigned responsibilities. At certain phases of development, the clear demarcation of responsibilities may disappear, if necessary, to keep the project development up to speed with the decided project schedule.

1. Viraj Patel

* GUI development
* GUI testing

1. Vedant Barbhaya

* ML Developer
* Backend Developer
* Automation of classification pipeline

1. Vishal Kundar

* ML Developer
* Backend Developer
* Automation of regression pipeline

**4.2 Timeline**