**Symbiosis Skills and Professional University**

**Kiwale, Pune**

------------------------------------------------------------------------------------------

**PROJECT REPORT**

**On**

**“Smart HR”**



**Submitted by**

**Prashant Kawle**

**Humaira Nasreen**

**Priyanka Vhadgir**

**(Registration Number)**

**DA-Batch-II**

**Under The Guidance of**

**Kushal Sharma**

**STUDENT DECLARATION AND ATTESTATION BY TRAINER**

This is to declare that this report has been written by me. No part of the report is plagiarized from other sources. All information included from other sources have been duly acknowledged. I aver that if any part of the report is found to be plagiarized, I shall take full responsibility for it.

Signature of student

**Prashant Kawle**

**Humaira Nasreen**

**Priyanka Vhadgir**

Registration Number:

Signature of trainer

**Name of trainer: Kushal Sharma**

**CERTIFICATE**

This is to certify that the report entitled, “**Smart HR**” submitted by **”Prashant Kawle, Humaira Nasreen, Priyanka Vhadgir**” to Symbiosis Skills and Professional University, Pune, Maharashtra, India, is a record of bonafide Project work carried out by him under my supervision and guidance and is worthy of consideration for the completion of certificate course in ‘Machine Learning Engineer”.

Signature of Trainer

Kushal Sharma

Date: / / 2021

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Supervisor Supervisor

Date:

**ACKNOWLEDGEMENTS**

We take this opportunity to express our sincere gratitude to all those who helped us in various capacities in this project. We are privileged to express our gratitude to our respected Trainer Kushal Sharma, whose unparalleled knowledge, moral fiber and judgment along with his Know-how was an immense support in completing project.

Mr.Prashant Kawle Ms.Humaira Nasreen

Ms.Priyanka Vhadgir

**INDEX**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sl. No** | **Index** |  |  |  | **Page No.** |
| 1 | **Acknowledgement** | |  |  |  |
| 2 | **Plan of Capstone Project** | |  |  |  |
|  | 2.1 | Purpose of Project | |  |  |
|  | 2.2 | Period of Project | |  |  |
|  | 2.3 | Problem Statement Detailing | |  |  |
| 3 | **Objective of the Project** | | |  |  |
|  | 3.1 | General Objective | |  |  |
| 4 | **Introduction (About the problem Statement)** | | |  |  |
|  | 4.1 | Problem Statement | |  |  |
|  | 4.2 | Overall Description | |  |  |
|  | 4.3 | Constraints | |  |  |
|  | 4.4 | Proposed Solution | |  |  |
|  | 4.5 | Implementation | | |  |
|  | 4.6 | Benefits | | |  |
| 5 | **Future plans** | | |  |  |
| 6 | Learning from the Project | | |  |  |
| 7 | Suggestions | |  |  |  |
| 8 | Annexure | |  |  |  |
| 9 | Weekly Reports | | | |  |
| 10 | Faculty and Student Feedback | | | |  |

**2. Plan of Capstone Project**

**2.1 Purpose of Project**

Predict the Employee Performance which employee performance is high or low. Our main objective is to Minimizes the HR efforts to hire new employees. To make decision easier. Increase the productivity. First Step we analyze the previous data and Find the result which employee performance is good. See their Schooling, College Details, Certifications, Grade. Based on this details the system predict the performance of new employee.

**2.2 Period of Project**

**2.3 Problem Statement Detailing**

* Design The System For HR Manager To View Employee Performance Easily.
* Which Employee Performance Is Good Or Bad.
* Very Useful To Making The Better Decision.
* This system helps HR managers to hire fresher's and new employees easily.
* By applying Some Condition or constraints i.e. Education details , Grades etc.

**3. Objective of the Project**

**3.1 General Objective**

* Predict the Employee Performance which employee performance is high or low.
* Our main objective is to Minimizes the HR efforts to hire new employees.
* To make decision easier.
* Increase the productivity.

**4. Introduction**

**4.1 Problem Statement**

* To Analyze the Employees Performance.
* To Predict The performance of the New Employee On the basis of Old employee.
* Recommend To the HR which Candidate Should get Hire.

**4.2 Overall Description**

**Analytics:**

* Analytics is not so much about numbers, as it is to do with logic and reasoning .
* Analytics is different from Analysis, which is the equivalent of number crunching.
* Analytics uses analysis but then builds on it to understand the 'why' behind the figures and/or to predict decisions.
* Analytics is the methodology of logical analysis ´ Analytics requires the use of carefully constructed metrics.
* Smart HR is data based; it uses past data to predict the future ´
* It is not about the quantity of data churned; it is about the logic used to link metrics to results

**Introduction**

The field of HR analytics is such a vast subject, you can go down many potential roads. From my earliest ventures in HR analytics, I’ve advocated choosing to concentrate on a clear subject, preferably one you are very interested in! Look for something you can analyse or improve – often the “low hanging fruits” are the best for getting your HR analytics expertise going.

This system helps HR managers to hire freshers and new employees easily .

This project can predict the performance of the new employees using the previous employees recorded data.

We can analyze  on various performance parameters like employment time ,Absenteeism Rate.

We can give opportunity for non-selected candidates those who appear for upcoming positions.

**4.3 Constraints**

These can include:

**Machine learning** is a branch of [artificial intelligence (AI)](https://www.ibm.com/cloud/learn/what-is-artificial-intelligence) focused on building applications that learn from data and improve their accuracy over time without being programmed to do so.

In data science, an algorithm is a sequence of statistical processing steps. In machine learning, algorithms are 'trained' to find patterns and features in massive amounts of data in order to make decisions and predictions based on new data. The better the algorithm, the more accurate the decisions and predictions will become as it processes more data.

Today, examples of machine learning are all around us. Digital assistants search the web and play music in response to our voice commands. Websites recommend products and movies and songs based on what we bought, watched, or listened to before. Robots vacuum our floors while we do . . . something better with our time. Spam detectors stop unwanted emails from reaching our inboxes. Medical image analysis systems help doctors spot tumors they might have missed. And the first self-driving cars are hitting the road.

We can expect more. As big data keeps getting bigger, as computing becomes more powerful and affordable, and as data scientists keep developing more capable algorithms, machine learning will drive greater and greater efficiency in our personal and work lives.

Machine learning methods

i] Supervised machine learning

ii] Unsupervised machine learning

**i] Supervised machine learning**

Supervised machine learning trains itself on a labeled data set. That is, the data is labeled with information that the machine learning model is being built to determine and that may even be classified in ways the model is supposed to classify data. For example, a computer vision model designed to identify purebred German Shepherd dogs might be trained on a data set of various labeled dog images.

Supervised machine learning requires less training data than other machine learning methods and makes training easier because the results of the model can be compared to actual labeled results. But, properly labeled data is expensive to prepare, and there's the danger of overfitting, or creating a model so closely tied and biased to the training data that it doesn't handle variations in new data accurately.

**ii] Unsupervised machine learning**

Unsupervised machine learning ingests unlabeled data—lots and lots of it—and uses algorithms to extract meaningful features needed to label, sort, and classify the data in real-time, without human intervention. Unsupervised learning is less about automating decisions and predictions, and more about identifying patterns and relationships in data that humans would miss. Take spam detection, for example—people generate more email than a team of data scientists could ever hope to label or classify in their lifetimes. An unsupervised learning algorithm can analyze huge volumes of emails and uncover the features and patterns that indicate spam (and keep getting better at flagging spam over time).

**Regression analysis** is a group of statistical processes used in R programming and statistics to determine the relationship between dataset variables. Generally, regression analysis is used to determine the relationship between the dependent and independent variables of the dataset. Regression analysis helps to understand how dependent variables changes when one of the independent variable is changes and other independent variables are kept constant. This helps in building a regression model and further, helps in forecasting the values with respect to a change in one of the independent variables. On the basis of types of dependent variables, number of independent variables and the shape of the regression line, there are 4 types of regression analysis techniques i.e., Linear Regression, Logistic Regression, Multinomial Logistic Regression and Ordinal Logistic Regression.

i] Linear Regression

ii] Logistic Regression

**i] Linear Regression**

Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as **sales, salary, age, product price,** etc.

Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable.

The linear regression model provides a sloped straight line representing the relationship between the variables. Consider the below image:



Mathematically, we can represent a linear regression as:

y= a0+a1x+ ε

**Here,**

Y= Dependent Variable (Target Variable)  
X= Independent Variable (predictor Variable)  
a0= intercept of the line (Gives an additional degree of freedom)  
a1 = Linear regression coefficient (scale factor to each input value).  
ε = random error

The values for x and y variables are training datasets for Linear Regression model representation.

**Types of Linear Regression**

i]Simple Linear Regression

ii]Multiple Linear Regression

**i]Simple Linear Regression:**  
If a single independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Simple Linear Regression.

**ii]Multiple Linear regression:**  
If more than one independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Multiple Linear Regression.

**4.4 Proposed Solution**

* In Very First Step we analyze the Previous data and Find the result which employee performance is Good. See their Schooling, College Details, Certifications, Grade.
* Based on this details the system predict the performance of new employee.

To Analyze the Employees Performance.

Previous Data

Apply Constraints

Visualize the data

To Predict The Performance of Candidate

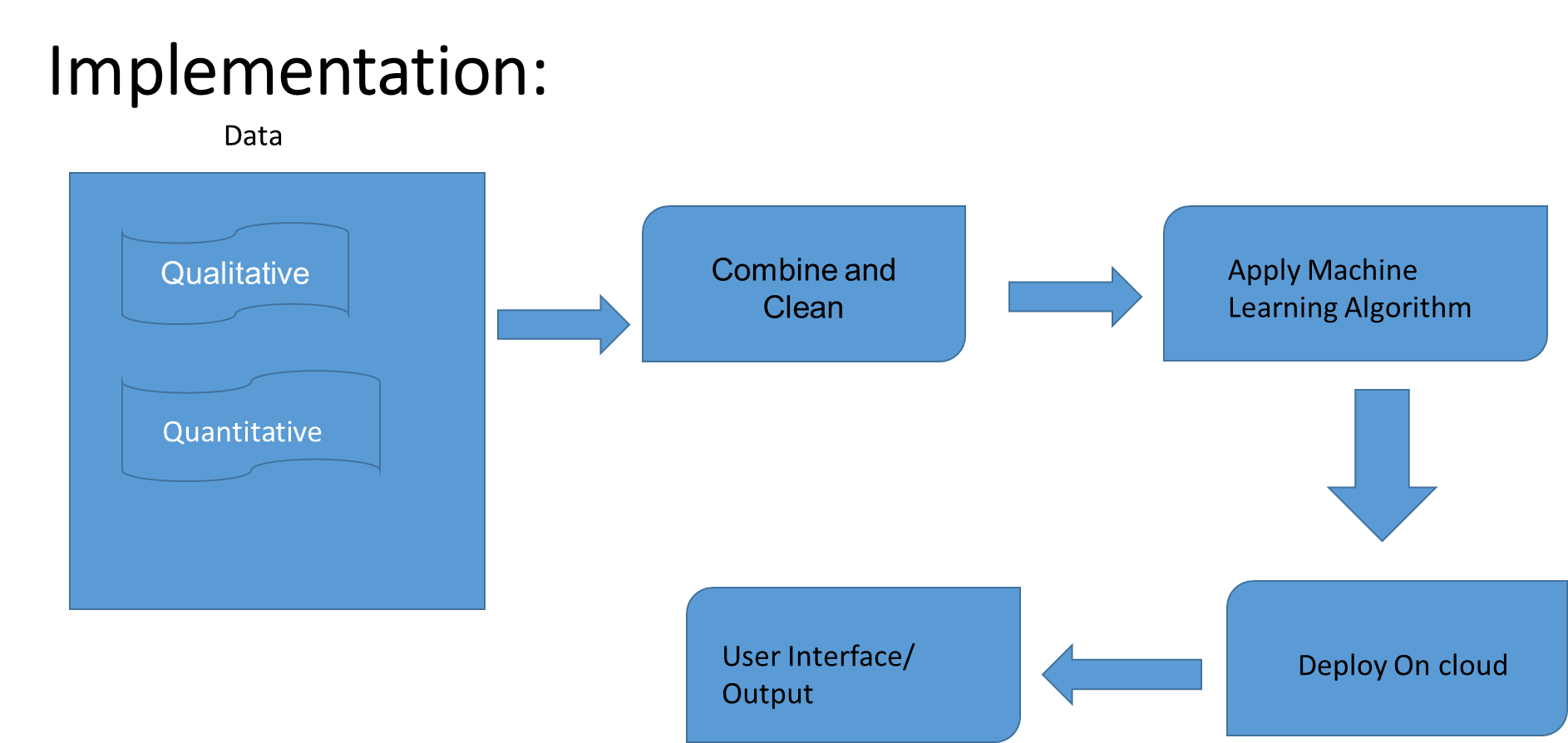
Enter New Employee Details

Previous Employee Detail

Machine Learning

Model

Display the Prediction



**4.6 Benefits**

* Minimize the Efforts of HR manager.
* Less time required.
* Easy to Handle.

**5. Future Plans**

* Using qualitative analysis performance of the model can be improved.
* Also using other machine learning algorithm we can increase the efficiency of model.
* It can also implemented in an organization to real time training and predication of model.