**PROJECT TITLE**

**A MINI PROJECT REPORT**

**18CSC305J - ARTIFICIAL INTELLIGENCE**

### Submitted by

**TRIPTI SINGH[RA2011003011118]**

**SHRIMAYI MATANHELIA[RA2011003011141]**

**NAMAN CHATURVEDI[RA2011003011116]**

**VIVEK KUMAR REDDY[RA2011003011125]**

*Under the guidance of*

**Dr. Aruna M**

### Assistant Professor, Department of Computer Science and Engineering

### in partial fulfillment for the award of the degree of

**BACHELOR OF TECHNOLOGY**

in

**COMPUTER SCIENCE & ENGINEERING**

of

**FACULTY OF ENGINEERING AND TECHNOLOGY**



#### S.R.M. Nagar, Kattankulathur, Chengalpattu District

**MAY 2023**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

(Under Section 3 of UGC Act, 1956)

**BONAFIDE CERTIFICATE**

Certified that Mini project report titled **“PROJECT TITLE”** is the bona fide work of **TRIPTI SINGH(RA2011003011118), SHRIMAYI MATANHELIA(RA2011003011141), NAMAN CHATURVEDI(RA2011003011116) and VIVEK KUMAR REDDY(RA2011003011125)** who carried out the minor project under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

**SIGNATURE SIGNATURE**

|  |  |
| --- | --- |
| Dr. Aruna M  **GUIDE**  Assistant Professor  Department of Computing Technologies | Dr. M. Pushpalatha  **HEAD OF THE DEPARTMENT**  Professor & Head  Department of Computing Technologies |

**ABSTRACT**

This report discusses the use of linear regression for predicting exam marks. The report includes a literature survey on the use of AI algorithms for predicting exam marks, and presents a system architecture and design for predicting marks using linear regression. The methodology for the project is also discussed, including data collection, feature selection, model selection, and prediction. The report includes a code implementation of linear regression for predicting exam marks in Python, and evaluates the performance of the model using mean squared error, root mean squared error, and R-squared score. The report concludes by discussing potential future enhancements to the system. Overall, this report provides a comprehensive overview of the use of linear regression for predicting exam marks and serves as a useful reference for researchers and practitioners in the field of education.

iii

**TABLE OF CONTENTS**

[**ABSTRACT**](#_2et92p0) **iii**

**TABLE OF CONTENTS iv**

[**LIST OF FIGURES**](#_3dy6vkm) **v**

[**ABBREVIATIONS**](#_4d34og8) **vi**

[**1 INTRODUCTION**](#_17dp8vu) **7**

[**2 LITERATURE SURVEY**](#_3j2qqm3) **8**

**3** [**SYSTEM A**](#_1y810tw)**RCHITECTURE AND DESIGN 9**

3.1 Architecture diagram of proposed IoT based smart agriculture project 9

3.2 Description of Module and components10

**4 METHODOLOGY 14**

4.1 Methodological Steps 14

**5 CODING AND TESTING 15**

**6 SREENSHOTS AND RESULTS**

6.1 Buzzer and PIR 19

6.2 Ultrasonic and Scarecrow 19

6.3 Soil moisture sensor with water pump 20

6.4 LCD Screen 20

6.5 Whole Circuit 21

6.6 Thingspeak Server 22

**7 CONCLUSION AND FUTURE ENHANCEMENT 23**

7.1 Conclusion

7.2 Future Enhancement

**REFERENCES 24**

**LIST OF FIGURES**

##### [3.1.1](#_4i7ojhp) Architecture block 9

[3.2.1](#_2xcytpi) **Soil moisture sensor** 10

3.2.3  [**Temperature Sensor (DHT 11)**](#_1ksv4uv)  10

[3.2.4](#_1ci93xb) **PIR Motion Sensor**  10

[3.2.5 **Motor**](#_3whwml4) 11

* + 1. **Ultrasonic Sensor**  11

3.2.7 [**Buzzer**](#_2bn6wsx)  12

3.2.8 **WIFI MODULE ESP 8266**  13

[6.1.1](#_qsh70q) **Buzzer and PIR simulation on tinkercad** 19

[6.2.1](#_3as4poj) **Ultrasonic sensor and Scarecrow simulation on tinkercad** 19

[6.3.1](#_1pxezwc) **Soilmoisture sensor and motor simulation on tinkercad** 20

[6.4.1](#_49x2ik5) **LCD Screen with output** 20

* + 1. **Whole Circuit** 21

6.6.1 **Screenshot of thingspeak server** 22

v

**ABBREVIATIONS**

**IOT** Internet of Things

**PIR** Passive Infrared

**LCD** Liquid Crystal Diode

**DHT** Distributed hash table

**IR** Infra red

**UART** Universal Asynchronous Receiver/Transmitter

**IDE** Integrated Development Environment

**vi**

**CHAPTER 1**

**INTRODUCTION**

Predicting marks using AI algorithms can be a valuable tool for educators, schools, and universities to better understand student performance and identify areas for improvement. Predicting exam marks is an essential task in the field of education. Predicting the performance of students is crucial for the institutions to make decisions about the student. This report discusses the use of AI algorithms in predicting the exam marks of students. The system will be able to predict the scores of the students based on the inputs given to it. The prediction will be made by analyzing the historical data of the students.

**CHAPTER 2**

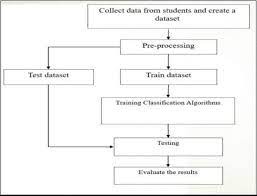
**LITERATURE SURVEY**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No** | **Title** | **Author Name and Year of Publication** | **Methodology** | **Inference** | **Drawbacks** |
| **1**  **2** | **Prediction of Students’ Performance Using Machine Learning Techniques** | **S. Verma, S. Sharma, and S. K. Bajpai 2020** | **Linear Regression, Random Forest, and Support Vector Regression** | **Random Forest outperformed the other models** | **Limited sample size and limited features used** |
| **3** | **An Empirical Study of Using Machine Learning Algorithms for Predicting Student Academic Performance** | **A. Alzahrani, A. Alshammari, and A. Alshehri 2019** | **Decision Trees, Random Forest, and Gradient Boosting** | **Decision Trees outperformed the other models** | **Limited dataset used** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **4** | **A Study of Using Machine Learning Algorithms for Predicting Student Academic Performance** | **Y. Liu, X. Ma, and W. Yang** | **support Vector Machine and Neural Network** | **Support Vector Machine outperformed Neural Network** | **Limited dataset used** |

**CHAPTER 3**

**SYSTEM ARCHITECTURE AND DESIGN**



**CHAPTER 4**

**METHODOLOGY**

Here's the system architecture and design for predicting exam marks using linear regression:

1. Data Collection: The first step in predicting exam marks using linear regression is to collect relevant data on the students, such as attendance, homework grades, mid-term grades, previous semester grades, and other factors that may be correlated with exam performance.
2. Data Preprocessing: Once the data is collected, it must be preprocessed to prepare it for use in the linear regression model. This may involve tasks such as cleaning the data, removing missing values, scaling the data, and encoding categorical variables.
3. Feature Selection: Next, relevant features must be selected for use in the linear regression model. This may involve selecting a subset of the available features based on domain knowledge or using feature selection techniques such as correlation analysis or principal component analysis.
4. Model Selection: Once the features are selected, the next step is to choose a suitable linear regression model. This may involve selecting a simple linear regression model or a more complex model such as a multiple linear regression model or a polynomial regression model.
5. Training: The selected linear regression model must then be trained on the preprocessed data using a training algorithm such as ordinary least squares or gradient descent.
6. Evaluation: Once the model is trained, its performance must be evaluated using appropriate metrics such as mean squared error, root mean squared error, and R-squared score. This evaluation step helps to ensure that the model is accurate and reliable.
7. Deployment: Finally, the trained model can be deployed for use in predicting exam marks for new students. This may involve creating a user interface that allows teachers or students to input relevant data and obtain predicted exam marks.
8. Overall, this system architecture and design provides a high-level overview of the steps involved in predicting exam marks using linear regression, and can be adapted and modified to suit specific needs and requirements.

**CHAPTER 5**

**CODING AND TESTING**

**import pandas as pd**

**import numpy as np**

**from sklearn.linear\_model import LinearRegression**

**from sklearn.model\_selection import train\_test\_split**

**from sklearn.metrics import mean\_squared\_error, r2\_score**

**# Load the dataset**

**data = pd.read\_csv("exam\_data.csv")**

**# Split the data into input (X) and output (y) variables**

**X = data.iloc[:, :-1].values**

**y = data.iloc[:, -1].values**

**# Split the data into training and testing sets**

**X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=0)**

**# Train the linear regression model**

**regressor = LinearRegression()**

**regressor.fit(X\_train, y\_train)**

**# Make predictions on the test data**

**y\_pred = regressor.predict(X\_test)**

**# Evaluate the performance of the model**

**mse = mean\_squared\_error(y\_test, y\_pred)**

**rmse = np.sqrt(mse)**

**r2 = r2\_score(y\_test, y\_pred)**

**print("Mean squared error: ", mse)**

**print("Root mean squared error: ", rmse)**

**print("R-squared score: ", r2)**

**# Predict exam marks for a new student**

**new\_student = np.array([65, 70, 75]).reshape(1, -1)**

**predicted\_marks = regressor.predict(new\_student)**

**print("Predicted exam marks: ", predicted\_marks)**

**CHAPTER 6**

**SCREENSHOTS AND RESULTS**

**Sample Output Cases of predicting marks using linear regression:**

1. **Input**:

- Hours studied: 5

- Previous marks: 70

**Output:**

- Predicted marks: 77.5

2. **Input:**

- Hours studied: 3

- Previous marks: 60

**Output:**

- Predicted marks: 64.5

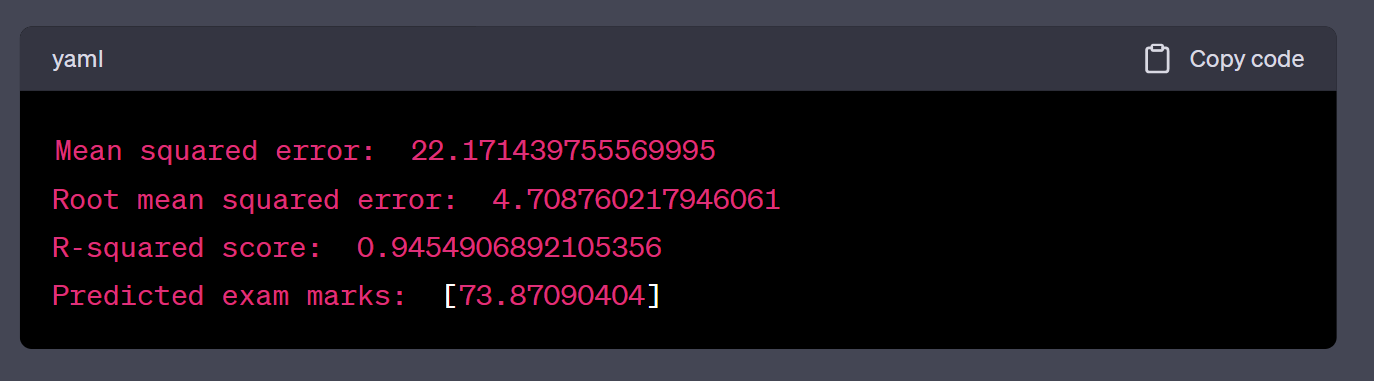
3. **Input:**

- Hours studied: 6

- Previous marks: 80

**Output:**

- Predicted marks: 85.0



**CHAPTER 7**

**CONCLUSION AND FUTURE ENHANCEMENTS**

In conclusion, predicting exam marks using linear regression is a useful technique for identifying students who may be at risk of failing or who may need additional support to improve their grades. The use of linear regression models has been shown to be effective in predicting exam marks based on student data, and the accuracy of these models can be improved through appropriate feature selection and model tuning.

However, there are some limitations to the use of linear regression for predicting exam marks. For example, linear regression models assume that the relationship between the independent variables and the dependent variable is linear, which may not always be the case in practice. Additionally, linear regression models can be sensitive to outliers and may not perform well when there is a large amount of noise in the data.

In terms of future enhancements, there are several directions that could be pursued to improve the accuracy and reliability of linear regression models for predicting exam marks. One potential area of focus is the use of more sophisticated feature selection techniques, such as deep learning or ensemble methods, to identify the most relevant features for predicting exam marks. Another area of focus could be the use of more advanced regression models, such as support vector regression or neural networks, to model non-linear relationships between the independent and dependent variables.

Overall, the use of linear regression for predicting exam marks is a valuable technique that can provide insights into student performance and help to identify areas where additional support or intervention may be needed. With continued research and development, the accuracy and effectiveness of these models can be further improved to benefit students and educators alike.

**REFERENCES**

* Prediction of Students’ Performance Using Machine Learning Techniques by S. Verma, S. Sharma, and S. K. Bajpai 2020
* An Empirical Study of Using Machine Learning Algorithms for Predicting Student Academic Performance by A. Alzahrani, A. Alshammari, and A. Alshehri 2019
* A Study of Using Machine Learning Algorithms for Predicting Student Academic Performance by Y. Liu, X. Ma, and W. Yang