**MACHINE LEARNING**

**-SIDDHIKA PAGARE**

**TYCS 8323**

Mahatma Education Society’s

**Pillai College of Arts, Commerce & Science**



**(Autonomous)**

**Affiliated to University of Mumbai**

NAAC Accredited 'A' grade (3 cycles)

Best College Award by University of Mumbai ISO 9001:2015 Certified

**CERTIFICATE**

*This is to certify that Mr. /Miss. Siddhika Pagare* *of*

***T.Y B.Sc. C.S. Semester VI*** *has completed the project work in the* *Subject of* ***Machine learning*** *during the academic year 2022-23 under the guidance of Prof****. Sanjana bhangale*** *being* *the partial requirement for the fulfilment of the curriculum of* ***Degree of Bachelor of Science in Computer Science****,* ***University of Mumbai****.*

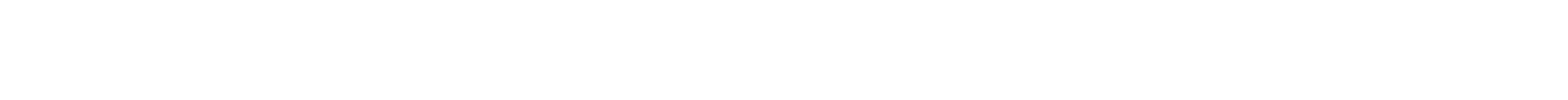
***Place:***

***Date:***

*Name & Signature of faculty Name & Signature of external*

*Name & Signature of Co-ordinator*

Dr. K.M. Vasudevan Pillai Campus, Sector 16, New Panvel - 410206.



Tel: 27456100 / 1700 | Fax: 27483208 | Website: [www.pcacs.ac.in](http://www.pcacs.ac.in/)

ACKNOWLEDGEMENT

I, Siddhika Pagare student of PILLAI COLLEGE OF ARTS, COMMERCE & SCIENCE would like to express my sincere gratitude towards Computer Science Department. I would like to thank lecturer Mrs. Sanjana Bhangale, for her constant support during the project.

Last but not least I thank all my classmates for being with me throughout the project, which leads to a successful completion of my project. The project would not have been completed without creativity and energy, which our friends provided.

Yours faithfully,

Siddhika Pagare

(Third Year Computer Science)

INTRODUCTION

Housing in India varies from palaces of erstwhile maharajas to modern apartment buildings in big cities to tiny huts in far-flung villages. There has been tremendous growth in India's housing sector as incomes have risen. The Human Rights Measurement Initiative finds that India is doing 60.9% of what should be possible at its level of income for the right to housing.

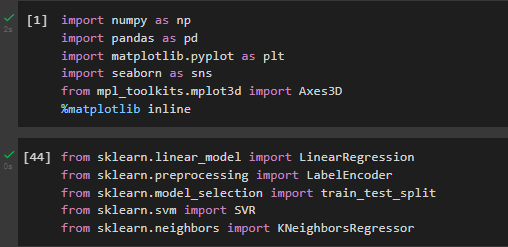
Renting, also known as hiring or letting, is an agreement where a payment is made for the temporary use of a good, service, or property owned by another. A gross lease is when the tenant pays a flat rental amount and the landlord pays for all property charges regularly incurred by the ownership. Renting can be an example of the sharing economy.

In this Dataset, we have information on almost 4700+ Houses/Apartments/Flats Available for Rent with different parameters like BHK, Rent, Size, No. of Floors, Area Type, Area Locality, City, Furnishing Status, Type of Tenant Preferred, No. of Bathrooms, Point of Contact.

**Dataset Glossary (Column-Wise)**

* **BHK**: Number of Bedrooms, Hall, Kitchen.
* **Rent**: Rent of the Houses/Apartments/Flats.
* **Size**: Size of the Houses/Apartments/Flats in Square Feet.
* **Floor**: Houses/Apartments/Flats situated in which Floor and Total Number of Floors (Example: Ground out of 2, 3 out of 5, etc.)
* **Area Type**: Size of the Houses/Apartments/Flats calculated on either Super Area or Carpet Area or Build Area.
* **Area Locality**: Locality of the Houses/Apartments/Flats.
* **City**: City where the Houses/Apartments/Flats are Located.
* **Furnishing Status**: Furnishing Status of the Houses/Apartments/Flats, either it is Furnished or Semi-Furnished or Unfurnished.
* **Tenant Preferred**: Type of Tenant Preferred by the Owner or Agent.
* **Bathroom**: Number of Bathrooms.
* **Point of Contact**: Whom should you contact for more information regarding the Houses/Apartments/Flats.

## **Importing Libraries**



## **Reading Dataset**

Graphical user interface, website

Description automatically generated

Graphical user interface

Description automatically generated

**HEATMAP**

**Chart, bar chart

Description automatically generated**

**PAIRPLOT**

**A screen shot of a computer

Description automatically generated with low confidence**

**NULL DATA**

Text

Description automatically generated

**MEAN, MEDIAN, LOWEST AND HIGHEST HOUSE RENT**

**Text

Description automatically generated**

#### **Bar Plot for Number of House in Each City which is Available for Rent**

Graphical user interface

Description automatically generated with medium confidence

#### **Scatter Plot on House Rents vs House Sizes**

**A picture containing chart

Description automatically generated**

#### **Bar Plot for City vs House Rent**

Chart, box and whisker chart

Description automatically generated

#### **Histogram on House Sizes**

Graphical user interface

Description automatically generated

### Label Encoding:

### In machine learning, we usually deal with datasets that contain multiple labels in one or more than one columns. These labels can be in the form of words or numbers. To make the data understandable or in human-readable form, the training data is often labelled in words.

### Label Encoding refers to converting the labels into a numeric form so as to convert them into the machine-readable form. Machine learning algorithms can then decide in a better way how those labels must be operated. It is an important pre-processing step for the structured dataset in supervised learning.

### A picture containing table Description automatically generated

### Train-Test Split

### Train test split technique is used to estimate the performance of machine learning algorithms which are used to make predictions on data not used to train the model. Train\_test\_split splits the data into 75% training data and 25% test data.

### 

### Model

### LINEAR REGRESSION

**Text

Description automatically generated**

### SVR MODEL

### Graphical user interface Description automatically generated with medium confidence

### KNN REGRESSOR

### Text Description automatically generated