Instructions for
Internship Report Drafting

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Important sections in project report

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Internship Report

On

House price prediction using machine learning

Submitted by
K. Bhanu sree
22KN1A4432
NRI Institute of technology

Submitted to
Mallika Srivastava
Head, Training Delivery
EISystems Services

&

Mayur Dev Sewak
Head, Internships & Trainings
EISystems Services

Student's Declaration

I, K. Bhanu sree, a student of B TECH program, Roll No. 22KN1A4432 of the Department of CSE-Data science, NRI Institute of Technology College do hereby declare that I have completed the mandatory internship in Ei systems Technologies.

K . Bhanu sree
06-05-2024

(Signature and Date)

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Table of Content

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

List of Figures

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

List of Tables

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Executive Summary

- People looking to buy a new home tend to be more conservative with their budgets and Market strategies.
- This project aims to analyse various parameters like average income, average area etc. And predict the house price accordingly.
- This application will help customers to invest in an estate without approaching an agent
- To provide a better and fast way of performing operations.
- To provide proper house price to the customers.
- To eliminate need of real estate agent to gain information regarding house prices.
- To provide best price to user without getting cheated.
- To enable user to search home as per the budget.
- The aim is to predict the efficient house pricing for real estate customers with respect To their budgets and priorities. By analyzing previous market trends and price ranges, and alsoup coming developments future prices will be predicted.
- House prices increase every year, so there is a need for a system to predict house prices in the future.
- House price prediction can help the developer determine the selling price of a house and can help the customer to arrange the right time to purchase a house.
- We use linear regression algorithm in machine learning for predicting the house price trends

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Overview of Organization

About EISystems

India's leader in workshops & trainings at IITs, NITs & top engineering colleges

EISystems Services is a leading Indian technology identity with operations across India. EISystems (We call it EISys) offers trainings in Cybersecurity, Machine Learning, Automobiles, Internet of Things, Robotics and Social media for enterprises and student community. Till date we have trained approximately 50000 students and impacted around 2 lakhs students through our various outreach initiatives since our founding.

Our Presence

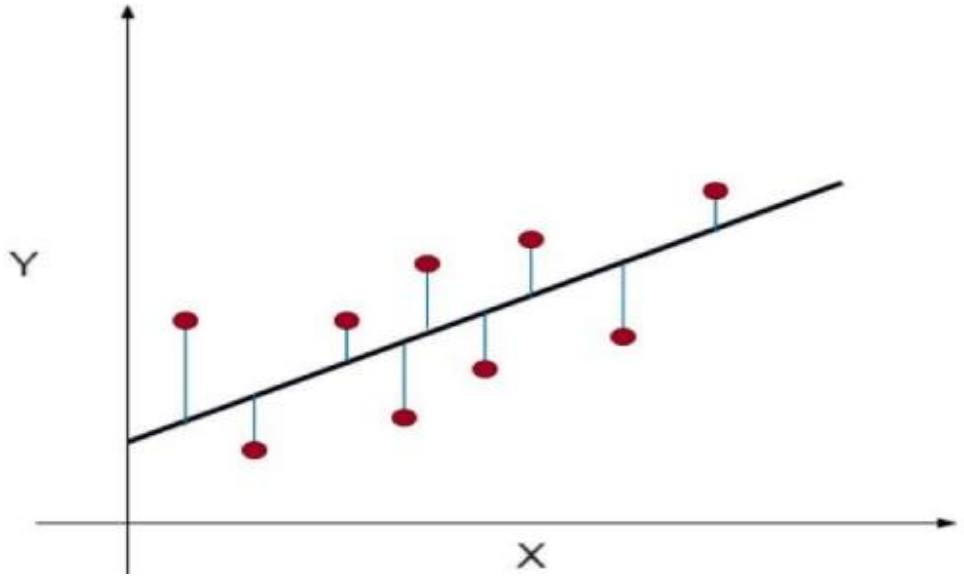
Some of the colleges where we had already felt our presence are given below:-

Indian Institute of Science, Bangalore
Indian Institute of Technology, Bombay
Indian Institute of Technology, Delhi
Indian Institute of Technology, Madras
Indian Institute of Technology, Kanpur
Indian Institute of Technology, Roorkee
Indian Institute of Technology, Guwahati
Indian Institute of Technology (Banaras Hindu University), Varanasi
Indian Institute of Technology, Indore
Indian Institute of Technology, Jodhpur
Indian Institute of Technology, Hyderabad
National Institute of Technology, Tiruchirappalli
National Institute of Technology, Warangal
National Institute of Technology, Calicut

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Project Summary

• Linear Regression is a supervised machine learning model that attempts to model a linear relationship between dependent variables (Y) and independent variables (X). Every evaluated observation with a model, the target (Y)'s actual value is compared to the target (Y)'s predicted value, and the major differences in these values are called residuals. The Linear Regression model aims to minimize the sum of all squared residuals. Here is the mathematical representation of the linear regression:

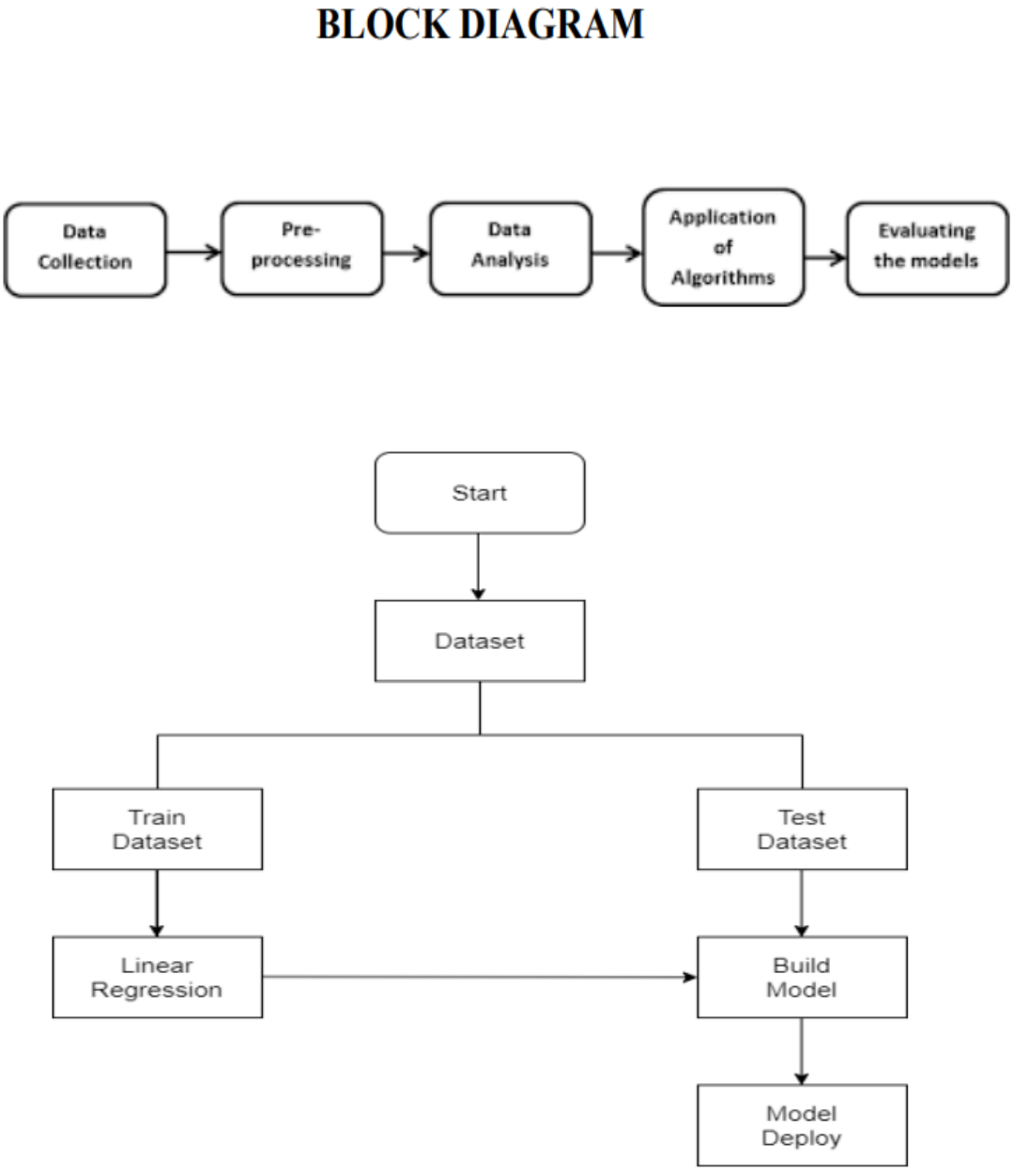


The values of X and Y variables are training datasets for the model representation of Line a regression. When a user implements a linear regression, algorithms start to find the best fit line using a0 and a1. In such a way, it becomes more accurate to actual data points; since we recognize the value of a0 and a1, we can use a model for predicting the response.

- As you can see in the above diagram, the red dots are observed values for both X and Y.
- The black line, which is called a line of best fit, minimizes a sum of a squared error.
- The blue lines represent the errors; it is a distance between the line of best fit and observed values.
- The value of the a1is the slope of the black line.

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Data Flow Diagram / Process Flow



EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Code / Program with Supported Screenshots

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
HouseDF = pd.read_csv('USA_Housing.csv')
HouseDF.head()
HouseDF=HouseDF.reset_index()
HouseDF.head()
HouseDF.info()
HouseDF.describe()
HouseDF.columns
sns.pairplot(HouseDF)
sns.distplot(HouseDF['Price’])
sns.heatmap(HouseDF.corr(), annot=True)
X = HouseDF[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms','Avg. Area
Number of Bedrooms', 'Area Population']]
y = HouseDF['Price’]
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y,test_size=0.4,random_state=101)
from sklearn.linear_model import minmaxscaler
lm = minmaxscaler(feature_range=(0,1))
lm.fit_transform(X_train,y_train)
print(lm.intercept_)
coeff_df = pd.DataFrame(lm.coef_,X.columns,columns=['Coefficient’])
coeff_df
from keras.layersimport Dense,Dropout,LSTM
from keras.models import Sequential
model = Sequential()
model.add(LSTM(units = 50,activation = 'relu',return_sequences = True,input_shape =
(x_train.shape[1], 1)))
model.add(Dropout(0.2))
model.add(LSTM(units = 60,activation = 'relu',return_sequences = True))
model.add(Dropout(0.3))
model.add(LSTM(units = 80,activation = 'relu',return_sequences = True))
model.add(Dropout(0.4))
model.add(LSTM(units = 120,activation = 'relu'))
model.add(Dropout(0.5))
model.add(Dense(units = 1))
model.compile(optimizer='adam', loss = 'mean_squared_error’)
model.fit(x_train, y_train,epochs=50)
```


EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

```
print(lm.intercept_)
coeff_df = pd.DataFrame(lm.coef_,X.columns,columns=['Coefficient’])
coeff_df
predictions = lm.predict(X_test)
scale_factor = 1/0.02099517
y_predicted = y_predicted * scale_factory

y_test = y_test * scale_factor
plt.scatter(y_test,predictions)
sns.distplot((y_test-predictions),bins=50);
plt.figure(figsize=(12,6))
plt.plot(y_test,'b',label = 'Original Price')
plt.plot(y_predicted,'r',label = 'Predicted Price')
plt.xlabel('Time')
plt.ylabel('Price')
plt.legend()
plt.show()
from sklearn import metrics
print('MAE:', metrics.mean_absolute_error(y_test, predictions))
print('MSE:', metrics.mean_squared_error(y_test, predictions))
print('RMSE:', np.sqrt(metrics.mean_squared_error(y_test, predictions)))
```


EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Input / Output with Datasets & Supported Screenshots

```
CRIM       0
ZN         0
INDUS      0
CHAS       0
NOX        0
RM         0
AGE        0
DIS        0
RAD        0
TAX        0
PTRATIO    0
B          0
LSTAT      0
price      0
dtype: int64
```

```
        CRIM    ZN  INDUS  CHAS    NOX     RM   AGE     DIS  RAD    TAX
0    0.00632  18.0   2.31   0.0  0.538  6.575  65.2  4.0900  1.0  296.0
1    0.02731   0.0   7.07   0.0  0.469  6.421  78.9  4.9671  2.0  242.0
2    0.02729   0.0   7.07   0.0  0.469  7.185  61.1  4.9671  2.0  242.0
3    0.03237   0.0   2.18   0.0  0.458  6.998  45.8  6.0622  3.0  222.0
4    0.06905   0.0   2.18   0.0  0.458  7.147  54.2  6.0622  3.0  222.0
..       ...   ...    ...   ...    ...    ...   ...     ...  ...    ...
501  0.06263   0.0  11.93   0.0  0.573  6.593  69.1  2.4786  1.0  273.0
502  0.04527   0.0  11.93   0.0  0.573  6.120  76.7  2.2875  1.0  273.0
503  0.06076   0.0  11.93   0.0  0.573  6.976  91.0  2.1675  1.0  273.0
504  0.10959   0.0  11.93   0.0  0.573  6.794  89.3  2.3889  1.0  273.0
505  0.04741   0.0  11.93   0.0  0.573  6.030  80.8  2.5050  1.0  273.0

     PTRATIO       B  LSTAT
0       15.3  396.90   4.98
1       17.8  396.90   9.14
2       17.8  392.83   4.03
3       18.7  394.63   2.94
4       18.7  396.90   5.33
..       ...     ...    ...
501     21.0  391.99   9.67
502     21.0  396.90   9.08
503     21.0  396.90   5.64
504     21.0  393.45   6.48
505     21.0  396.90   7.88

[506 rows x 13 columns]
0      24.0
1      21.6
2      34.7
3      33.4
4      36.2
       ...
501    22.4
502    20.6
503    23.9
504    22.0
505    11.9
Name: price, Length: 506, dtype: float64
```

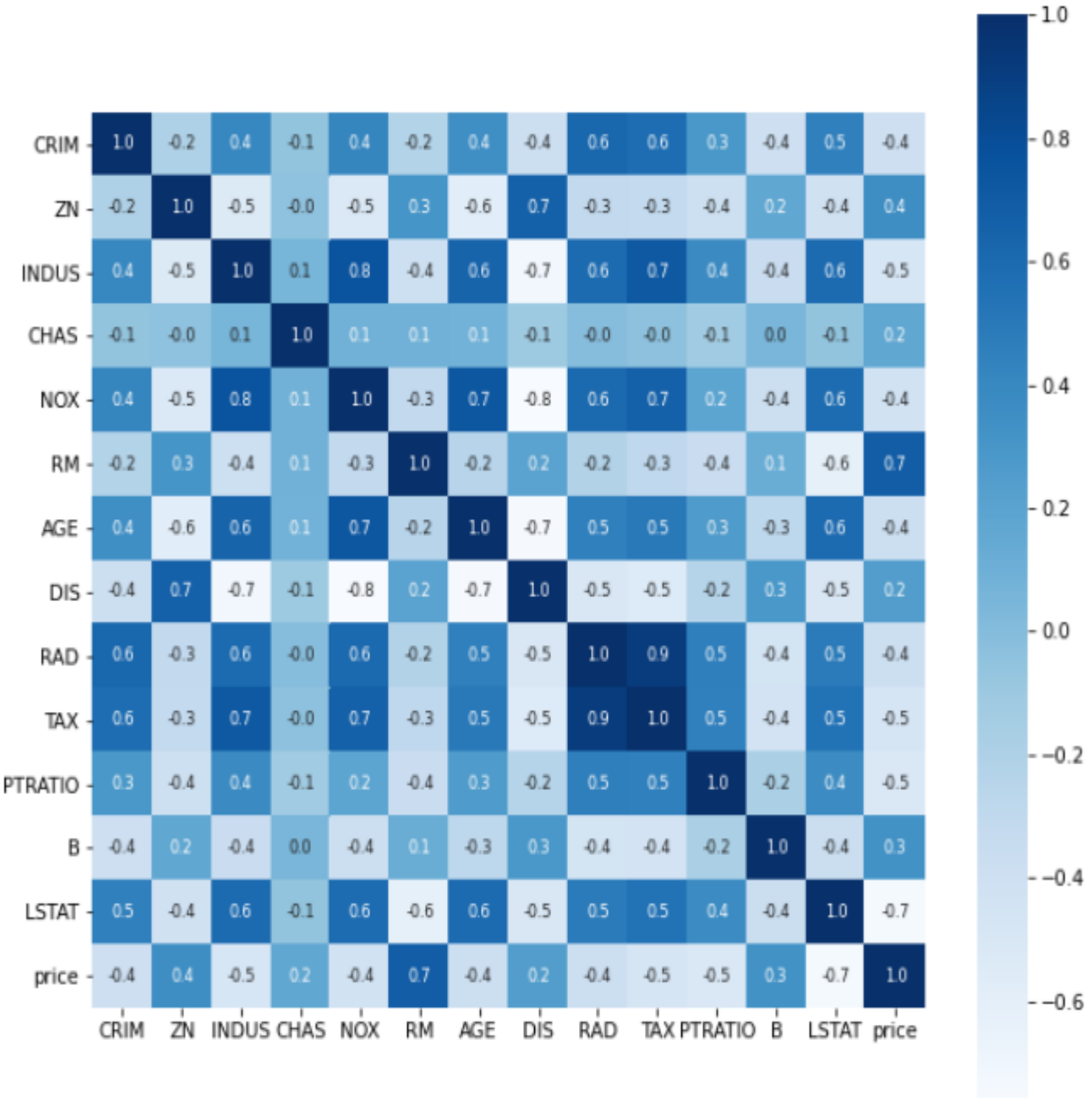

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

1. CRIM	per capita crime rate by town
2. ZN	proportion of residential land zoned for lots over 25,000 sq.ft.
3. INDUS	proportion of non-retail business acres per town
4. CHAS	Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
5. NOX	nitric oxides concentration (parts per 10 million)
6. RM	average number of rooms per dwelling
7. AGE	proportion of owner-occupied units built prior to 1940
8. DIS	weighted distances to five Boston employment centres
9. RAD	index of accessibility to radial highways
10. TAX	full-value property-tax rate per $10,000
11. PTRATIO	pupil-teacher ratio by town
12. B	1000(Bk - 0.63)^2 where Bk is the proportion of blacks by town
13. LSTAT	% lower status of the population
14. MEDV	Median value of owner-occupied homes in $1000's

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	price
count	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000
mean	3.613524	11.363636	11.136779	0.069170	0.554695	6.284634	68.574901	3.795043	9.549407	408.237154	18.455534	356.674032	12.653063	22.532806
std	8.601545	23.322453	6.860353	0.253994	0.115878	0.702617	28.148861	2.105710	8.707259	168.537116	2.164946	91.294864	7.141062	9.197104
min	0.006320	0.000000	0.460000	0.000000	0.385000	3.561000	2.900000	1.129600	1.000000	187.000000	12.600000	0.320000	1.730000	5.000000
25%	0.082045	0.000000	5.190000	0.000000	0.449000	5.885500	45.025000	2.100175	4.000000	279.000000	17.400000	375.377500	6.950000	17.025000
50%	0.256510	0.000000	9.690000	0.000000	0.538000	6.208500	77.500000	3.207450	5.000000	330.000000	19.050000	391.440000	11.360000	21.200000
75%	3.677083	12.500000	18.100000	0.000000	0.624000	6.623500	94.075000	5.188425	24.000000	666.000000	20.200000	396.225000	16.955000	25.000000
max	88.976200	100.000000	27.740000	1.000000	0.871000	8.780000	100.000000	12.126500	24.000000	711.000000	22.000000	396.900000	37.970000	50.000000

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Images / Video Links



EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

References

• Real Estate Price Prediction with Regression and Classification, CS 229 Autumn2016 Project Final Report

• Gongzhu Hu, Jinping Wang, and Wenying Feng Multivariate Regression Modelling for Home Value Estimates with Evaluation using Maximum Information Coefficient

• Byeonghwa Park , Jae Kwon Bae (2015). Using machine learning algorithms for housing price prediction , Volume 42, Pages 2928-2934 [4] Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, 2015. Introduction to Linear Regression Analysis.

• Iain Pardoe, 2008, Modelling Home Prices Using Realtor Data

• Aaron Ng, 2015, Machine Learning for a London Housing Price Prediction Mobile Application

• Wang, X., Wen, J., Zhang, Y .Wang, Y. (2014). Real estate price forecasting based on SVM optimized by PSO. Optik-International Journal for Light and Electron Optics, 125(3), 14391443

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Student *Self Evaluation of the Short-Term Internship*

Please rate your performance in the following areas:

1)	Oral communication	1	2	3	4	5
2)	Written communication	1	2	3	4	5
3)	Initiative	1	2	3	4	5
4)	Interaction with staff	1	2	3	4	5
5)	Attitude	1	2	3	4	5
6)	Dependability	1	2	3	4	5
7)	Ability to learn	1	2	3	4	5
8)	Planning and organization	1	2	3	4	5
9)	Professionalism	1	2	3	4	5
10)	Creativity	1	2	3	4	5
11)	Quality of work	1	2	3	4	5
12)	Productivity	1	2	3	4	5
13)	Progress of learning	1	2	3	4	5
14)	Adaptability to organization's culture/policies	1	2	3	4	5
15)	OVERALL PERFORMANCE	1	2	3	4	5

Rating Scale: 5 will be Best while 1 will be Worst

K . Bhanu sree

Signature of the Student

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 1
Daily Activity Report

Week No: 1

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Foundation of python	Foundation of python
Day 2	Variable and constant naming	Variable and constant naming
Day 3	Print functions and comments	Print functions and comments
Day 4	Starting with datatypes-number	Starting with datatypes-number
Day 5	Starting with datatypes-string	Starting with datatypes-string

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 2

Daily Activity Report

Week No: 2

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	list	list
Day 2	list	list
Day 3	Method	Method
Day 4	Method	Method
Day 5	Method	Method

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 3
Daily Activity Report

Week No: 3

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Tuple	Tuple
Day 2	Tuple	Tuple
Day 3	Dictionary	Dictionary
Day 4	Dictionary	Dictionary
Day 5	Set	Set

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 4
Daily Activity Report

Week No: 4

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Set	Set
Day 2	Boolean	Boolean
Day 3	Boolean	Boolean
Day 4	User input type casting	User input type casting
Day 5	Control statements	Control statements

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 5
Daily Activity Report

Week No: 5

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Control statements	Control statements
Day 2	Control statements	Control statements
Day 3	Loops in python	Loops in python
Day 4	Loops in python	Loops in python
Day 5	Loops in python	Loops in python

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 6
Daily Activity Report

Week No: 6

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Loops in python	Loops in python
Day 2	Loops in python	Loops in python
Day 3	File handling	File handling
Day 4	File handling	File handling
Day 5	File handling	File handling

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 7
Daily Activity Report

Week No:7

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Functions	Functions
Day 2	Functions	Functions
Day 3	Packages and Modules	Packages and Modules
Day 4	Packages and Modules	Packages and Modules
Day 5	Packages and Modules	Packages and Modules

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 8
Daily Activity Report

Week No: 8

Day & Date	Brief Description of Daily Activity	Learning Outcome
Day 1	Exception handling	Exception handling
Day 2	Exception handling	Exception handling
Day 3	Oops concept	Oops concept
Day 4	Oops concept	Oops concept
Day 5	Oops concept	Oops concept

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000

Annexure 9
Weekly Progress Report

Week No:
(1/2/3/4/5/6/7/8)

Week(s)	Summary of Weekly Activity
Week 1	Foundation of python and Variable and constant naming and print fucnctions and comments and starting with datatypes-number and string
Week 2	List,Method
Week 3	Tuple and Dictionary and Set
Week 4	Set and Boolean, User input casting, Control statements
Week 5	Control statements, Loops in python
Week 6	Loops in python and File handling
Week 7	Functions, packages and modules, Exception handling
Week 8	Exception handling, oops concept

EISYSTEMS SERVICES
FF-110, Express Greens Plaza, Sector 1
Vaishali – Delhi NCR – India 201010
W: www.eisystems.in | E: info@eisystems.in | T: +91 92122-51000