Dataset (or) Data Frame.

→ What is a Dataset?

--> A dataset is a collection of data in which data is arranged in some order.

Data Types:

--> Structured data

--> Unstructured data

Structured Data:

--> Structured data refers t highly organized and formatted data.

How Does ML work?

1. Data collection
2. Data preprocessing
3. Model fit.

→ML → Machine learning, often Abbreviated As ML is A Branch of Artificial Intelligence (AI).

→ Machine learning enables a machine to automatically learn from data, improve performance from experiences, and.

<u>Subsets of AI</u>

→ AI (Artificial Intelligence):
A broad field that enables
nachines to simulate human
telligence.

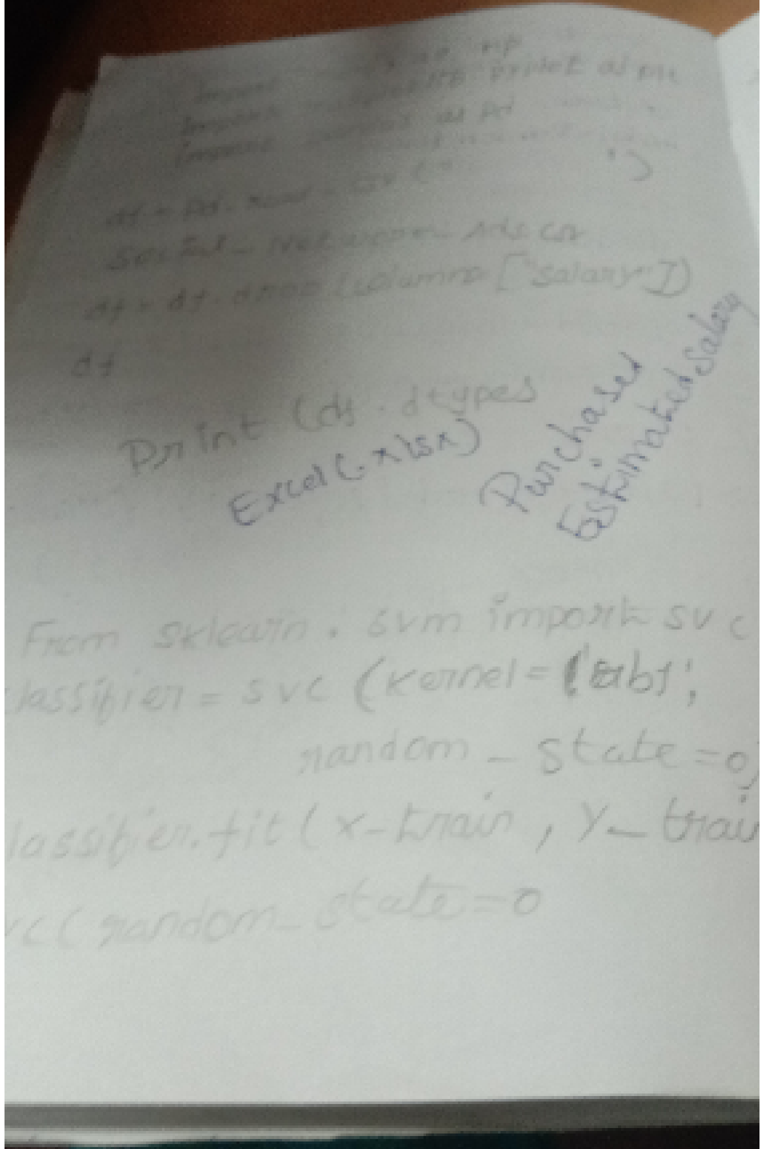
→ ML (Machine Learning):
A Subset of AI that allows
them to learn from data and
ove over time without explicit
amming.

```
y_Pred = Classifier. Predict (xtest)
import sclearn. metrics im

age_input = float (input
( enter the prediction inputvalue)
bmi_input =
children_input =

future_prediction = regresson.
  Predict ([[age_input, bmi_input
children_input]])
Print ("future prediction = {
    format(future_predic

X = data  drop
```



```
import [illegible] as np
import [illegible] pyplot as plt
import pandas as pd
[illegible]
df = pd.read_csv("[illegible]")
Social_Network_Ads.csv
df = df.drop(columns=["salary"])
df
```

Print (df.dtypes

Excel(.xlsx)

Purchased

EstimatedSalary

```
From sklearn.svm import svc
classifier = svc (kernel='[illegible]',
                  random_state=0)
classifier.fit(x_train, y_train
svc(random_state=0
x = df . drop (columns = ["status"])
y= df [" status"]
~~x-train,~~
--> model = RandomForestclassifier
    (n_ estimators =100 , random_
    state = 42)
--> model . fit (x- train , y- train)

--> y_pred = model. predict (x_test)
--> accurancy = accuracy_ score (y_test,
y_ pred)
--> print (f" Model Accuracy: {accuracy
                        .2f}")

-->     user_input = [ ]
```

```
[illegible] data data
    df=df.drop(columns=[illegible])
    data

from sklearn.model_selection
    import train_test_split
from sklearn.preprocessing
    import LabelEncoder
from sklearn.ensemble import
    RandomForestClassifier
from sk. metrics im
    accuracy_score
```

Insurance

nent_Data_Full_Classify)

[illegible]

8.4.22

AI Based Building Survey

→ from sklearn.preprocessing import standardscaler # import standard scaler function from sklearn library

→ sc = standardscaler() # create object for standard scaler function

→ X_train = sc.fit_transform(x_train

→ x_test = sc.transform(x_test)

→ from sklearn.ensemble import RandomForestRegresson

→ from sklearn.datasets import make_regression

→ regresson = RandomForestRegressor(criterion="squared_error", max_features="sqrt", random_state=0)

→ regresson.fit(x_train, y_train)