Vodafone Idea Foundation

_VOIS

Connecting Dreams

Introduction to Machine Learning & its types

Vodafone Idea Foundation

_VOIS

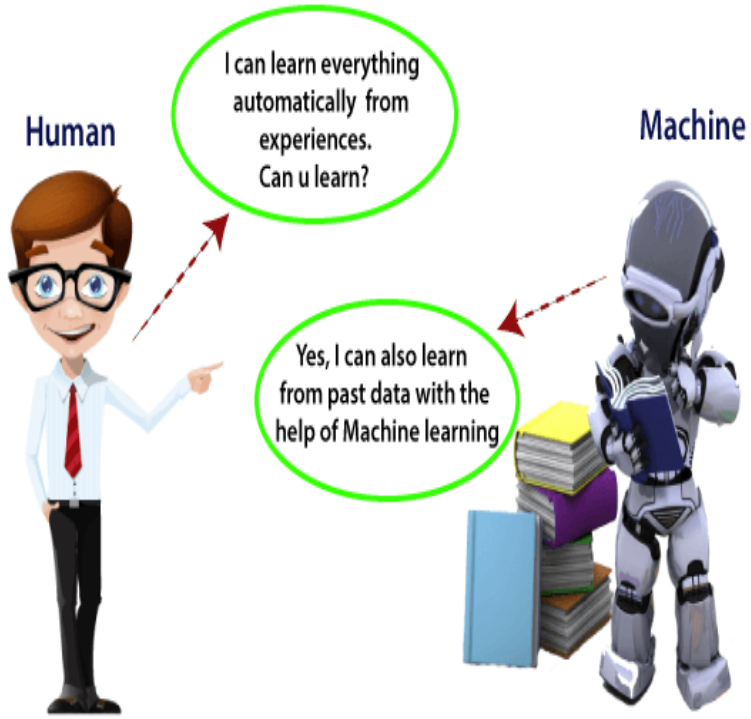
Connecting Dreams

Contents

- Machine Learning
- Relation between Machine learning , Deep Learning and AI
- Types of Machine Learning
- Concept of Regression and Types
- Linear Regression
- Evaluation methods
- Tools
- Hands-on

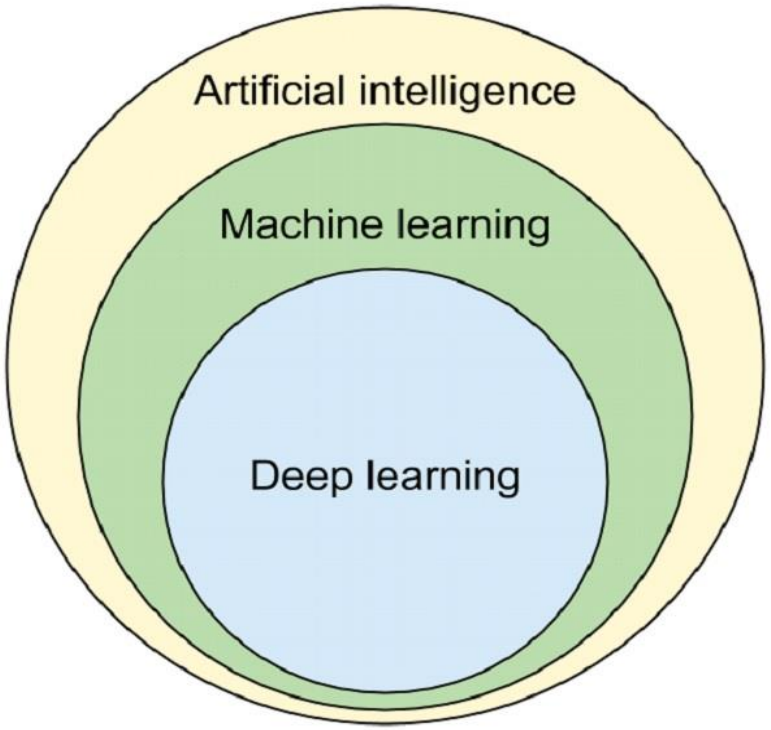
Machine Learning

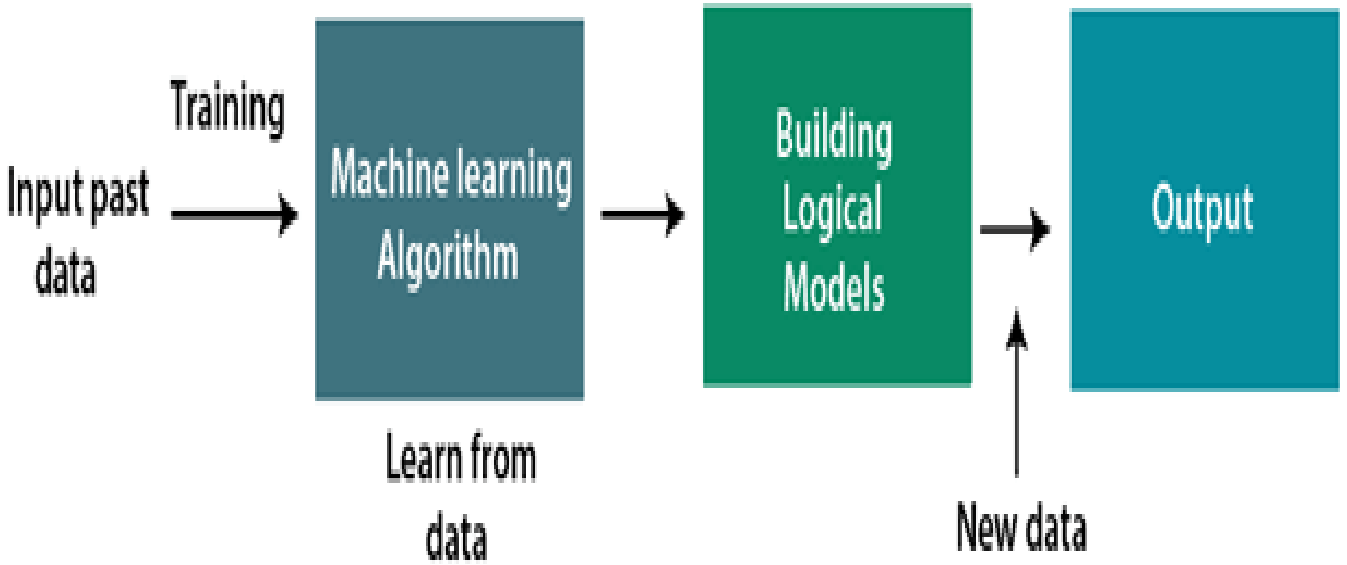
Machine learning (**ML**) is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so.[2]



Source - [1]

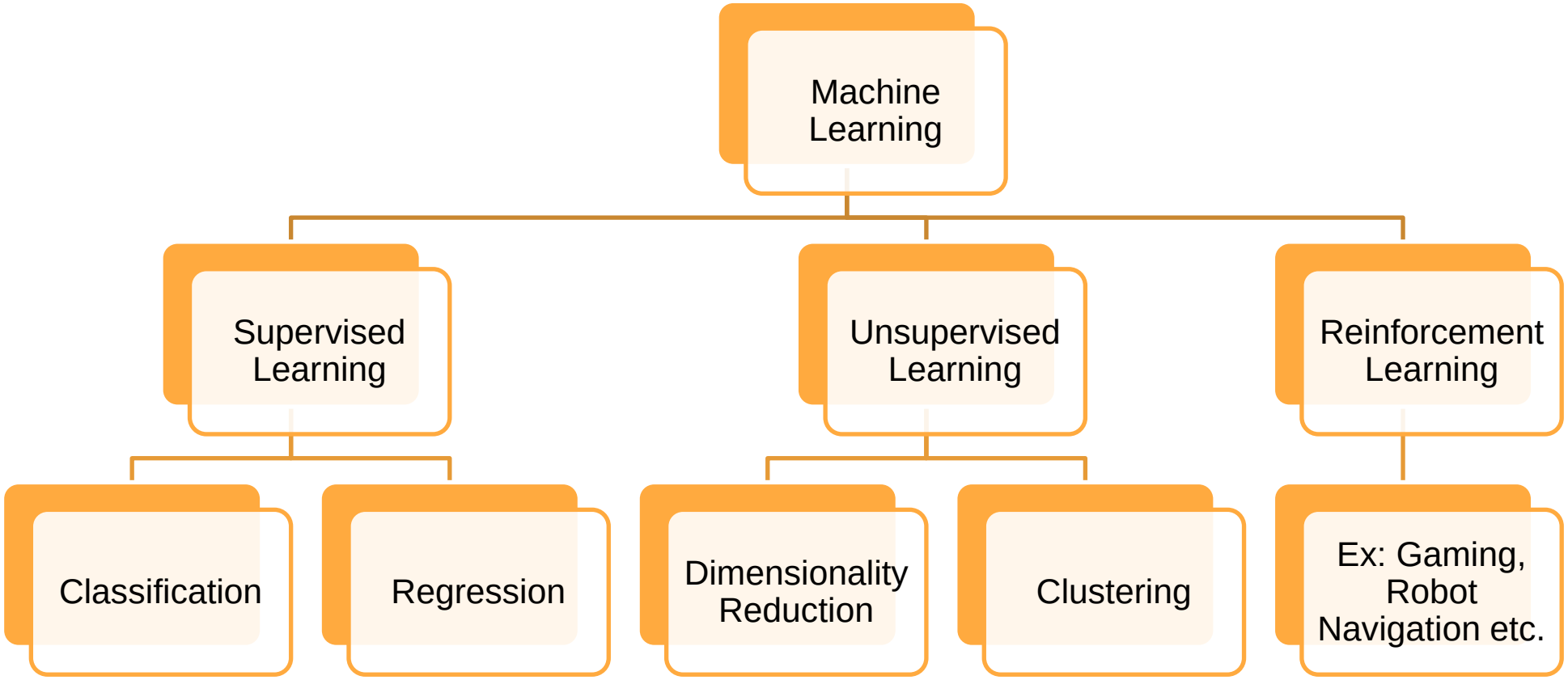
Relation between ML , DL & AI





Source - [1,3]

Types of Machine Learning



Vodafone Idea Foundation

_VOIS

Connecting Dreams

Supervised

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

Unsupervised

Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labelled responses.

Reinforcement

Reinforcement Learning is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.

Vodafone Idea Foundation

_VOIS

Connecting Dreams

Supervised Learning(Task Driven)

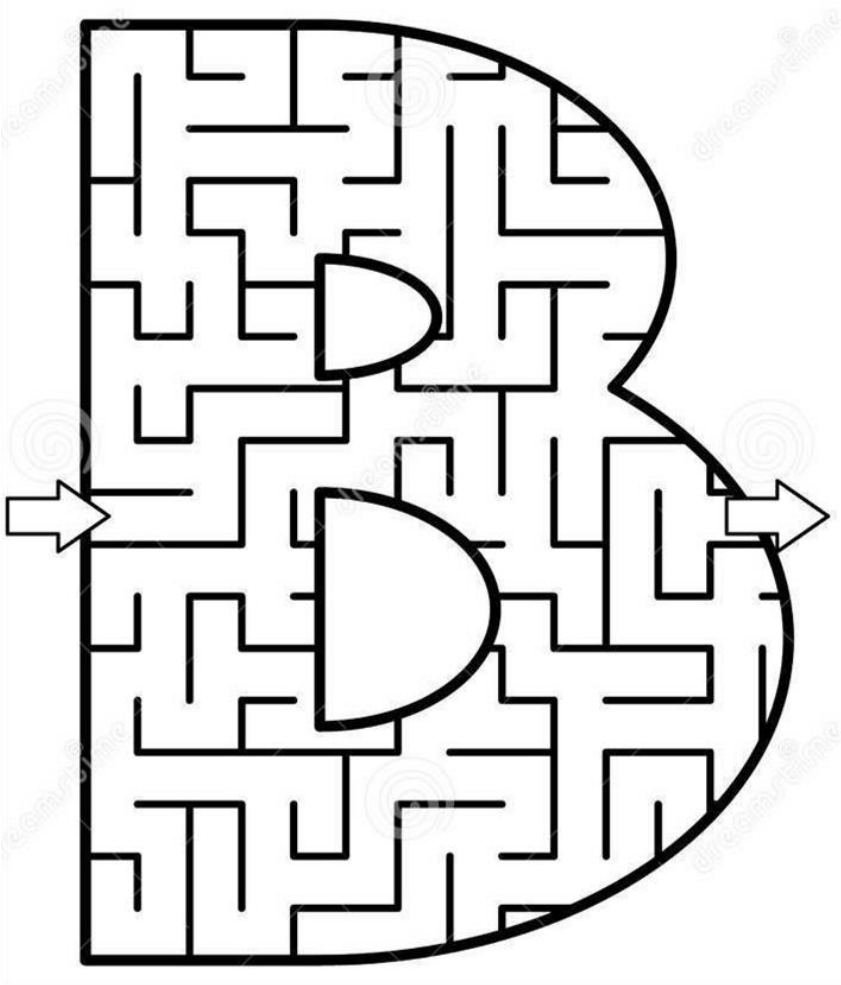
Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

Unsupervised Learning(Data Driven)

Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labelled responses.

Reinforcement Learning(Learn From error)

Reinforcement Learning is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.



Vodafone Idea Foundation

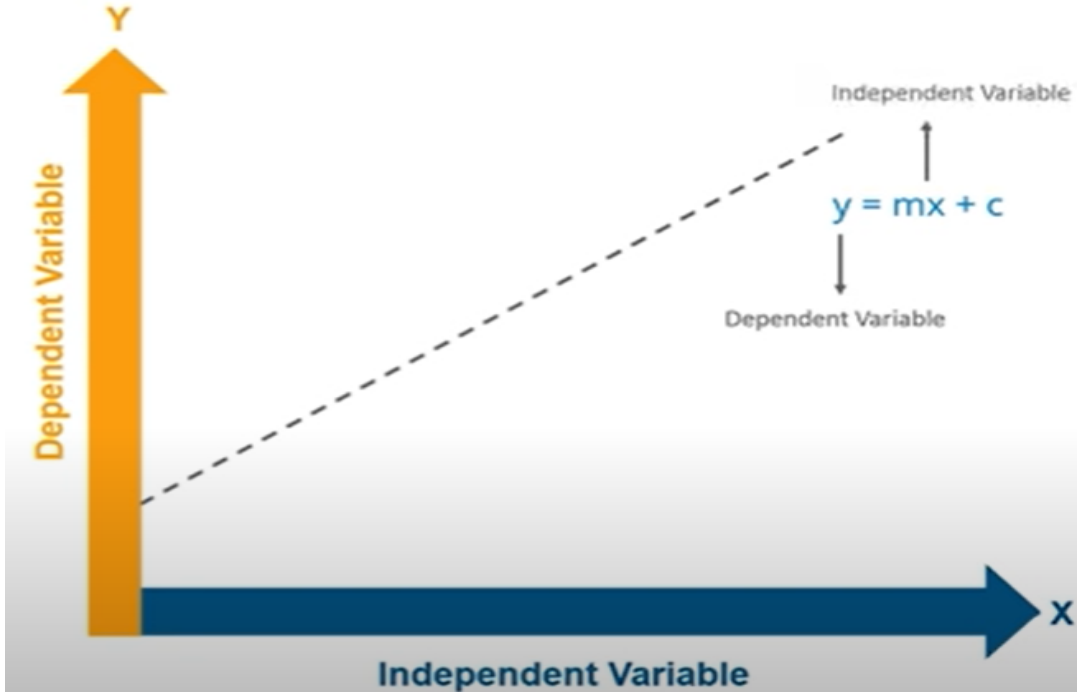
_VOIS

Connecting Dreams

Regression Methods : Linear Regression & Evaluation

Concept of Regression

Statistical process of estimating relation between independent and dependant variable[7]



How to identify problem is a regression problem ?

Dependant variable (outcome) should be continuous value (numeric) not a categorical value.

e.g.

Categorical data – months of year

Continuous data – Temperature measured for a city

Types of Regression[8]

- Linear Regression
- Logistic Regression
- Ridge Regression
- Lasso Regression
- Polynomial Regression
- Bayesian Linear Regression

Linear Regression

- The predictor/features/ independent variables are in linear relation with dependant variable (outcome).
- If only one independent variable is present then this is called as simple linear regression.

$$Y = \alpha_0 + \alpha_1 . X1 + E$$

Where y = Outcome / Dependant variable

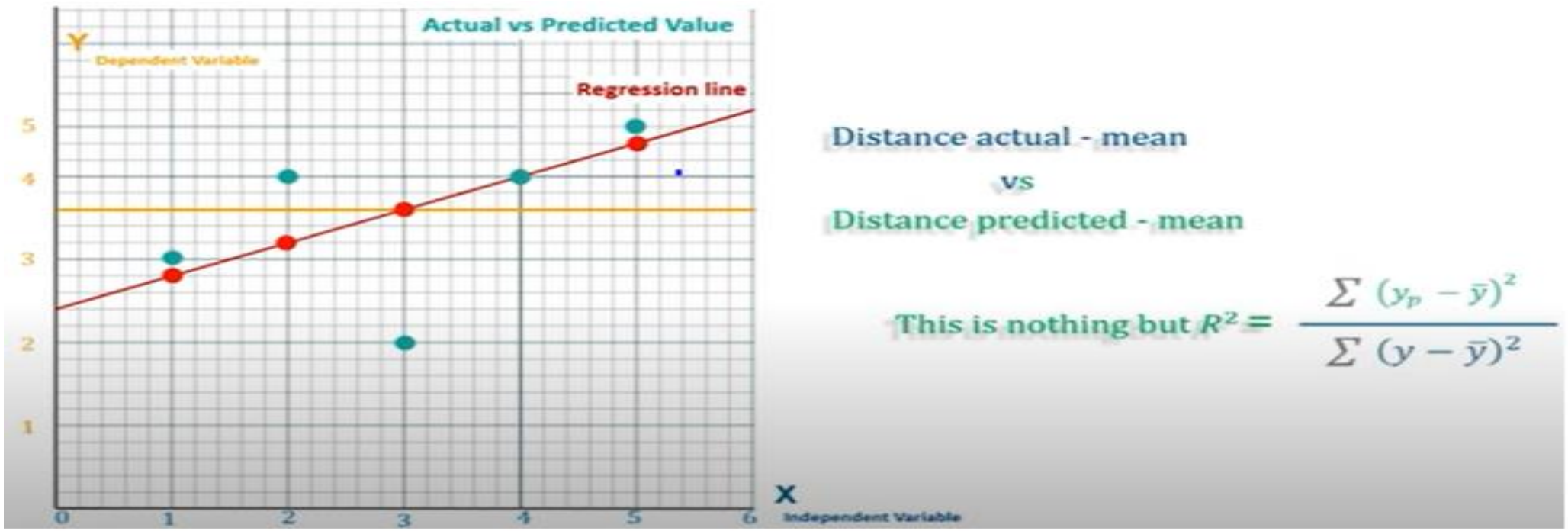
α_0 & α_1 = regression coefficients

X1 = Independent variable / predictor / feature

E = Error between predicted and original value

- If the number of independent variables are more than one then its Multiple linear regression.

$$Y = \alpha_0 + \alpha_1 . X1 + ...+\alpha_n . Xn + E$$



Vodafone Idea Foundation

_VOIS

Connecting Dreams

Evaluation methods

- **Mean Absolute Error:** Summation of the absolute value distance from the points to the line. [10][11][12]
- **Mean Squared Error:** A summation of the square of distances from the points to the line. [10][11][12]

Mean Absolute Error (MAE)	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
Mean Squared Error (MSE)	$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
Root Mean Squared Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$

Where,

$\hat{y}$ – *predicted value of y*

$\bar{y}$ – *mean value of y*

```
# calculate MAE, MSE, RMSE
print(metrics.mean_absolute_error(y_true, y_pred))
print(metrics.mean_squared_error(y_true, y_pred))
print(np.sqrt(metrics.mean_squared_error(y_true, y_pred)))
```

Vodafone Idea Foundation

_VOIS

Connecting Dreams

Tools

- Python 3
- Anaconda
- Packages – NumPy , Pandas , Matplotlib , Scikit Learn
- Jupyter Notebook or Google Colab

Vodafone Idea Foundation

_VOIS

Connecting Dreams

Hands - on

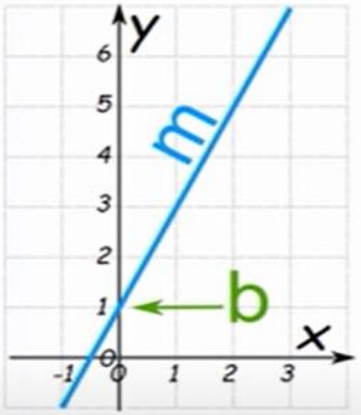
Vodafone Idea Foundation

_VOIS

Connecting Dreams

Case 1: Home Price Prediction (Based on area)

	area	price
0	2600	550000
1	3000	565000
2	3200	610000
3	3600	680000
4	4000	725000



price = m * area + b

y = mx + b

Slope (or Gradient) Y Intercept

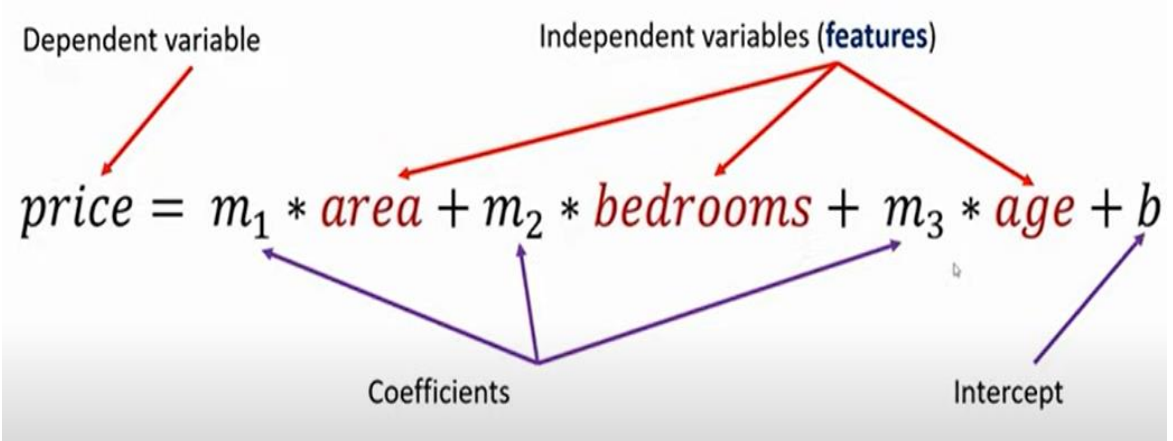
Vodafone Idea Foundation

_VOIS

Connecting Dreams

Case 2: Home Price Prediction (Based on Multiple feature)

	area	bedrooms	age	price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	NaN	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000
5	4100	6.0	8	810000



Vodafone Idea Foundation

_VOIS

Connecting Dreams

References

1. https://www.javatpoint.com/machine-learning
2. https://en.wikipedia.org/wiki/Machine_learning
3. https://en.wikipedia.org/wiki/Machine_learning#/media/File:Fig-X_All_ML_as_a_subfield_of_AI.jpg
4. https://numpy.org/
5. https://pandas.pydata.org/
6. https://matplotlib.org/
7. https://scikit-learn.org/stable/
8. https://en.wikipedia.org/wiki/Regression_analysis
9. https://www.upgrad.com/blog/types-of-regression-models-in-machine-learning/
10. https://commons.wikimedia.org/wiki/File:Residuals_for_Linear_Regression_Fit.png
11. https://medium.com/acing-ai/how-to-evaluate-regression-models-d183b4f5853d
12. https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/regress.htm#CHDBHBDI
13. https://towardsdatascience.com/what-are-the-best-metrics-to-evaluate-your-regression-model-418ca481755b