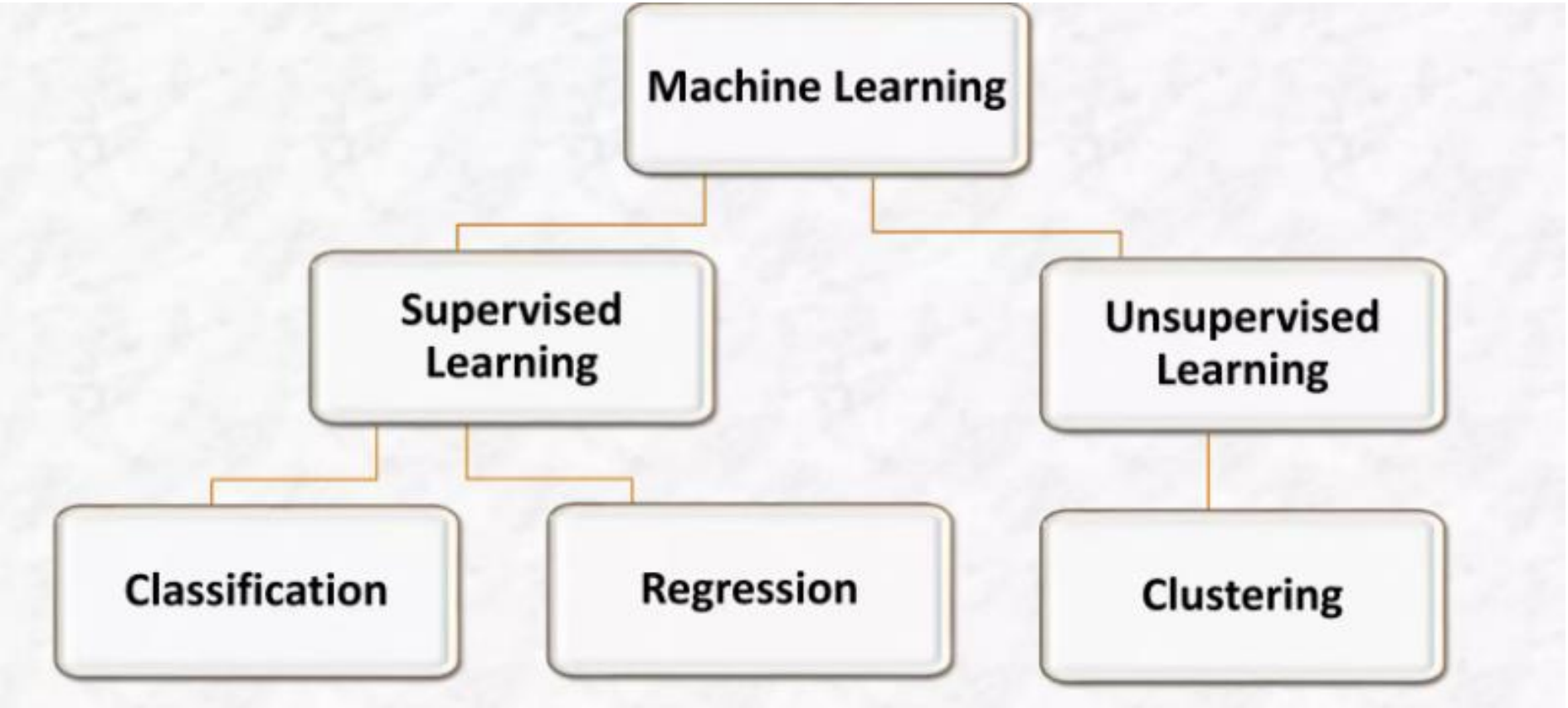
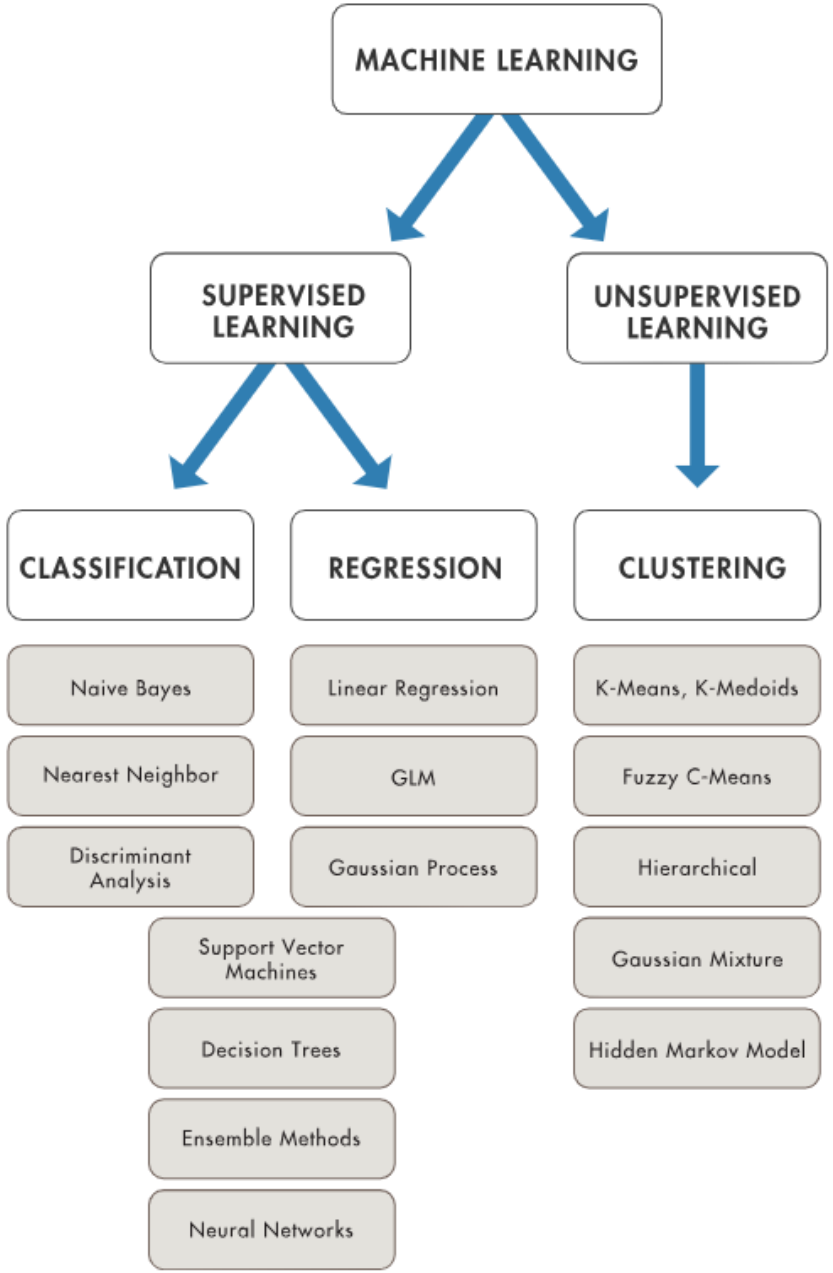
Artificial intelligence

OUTLINE

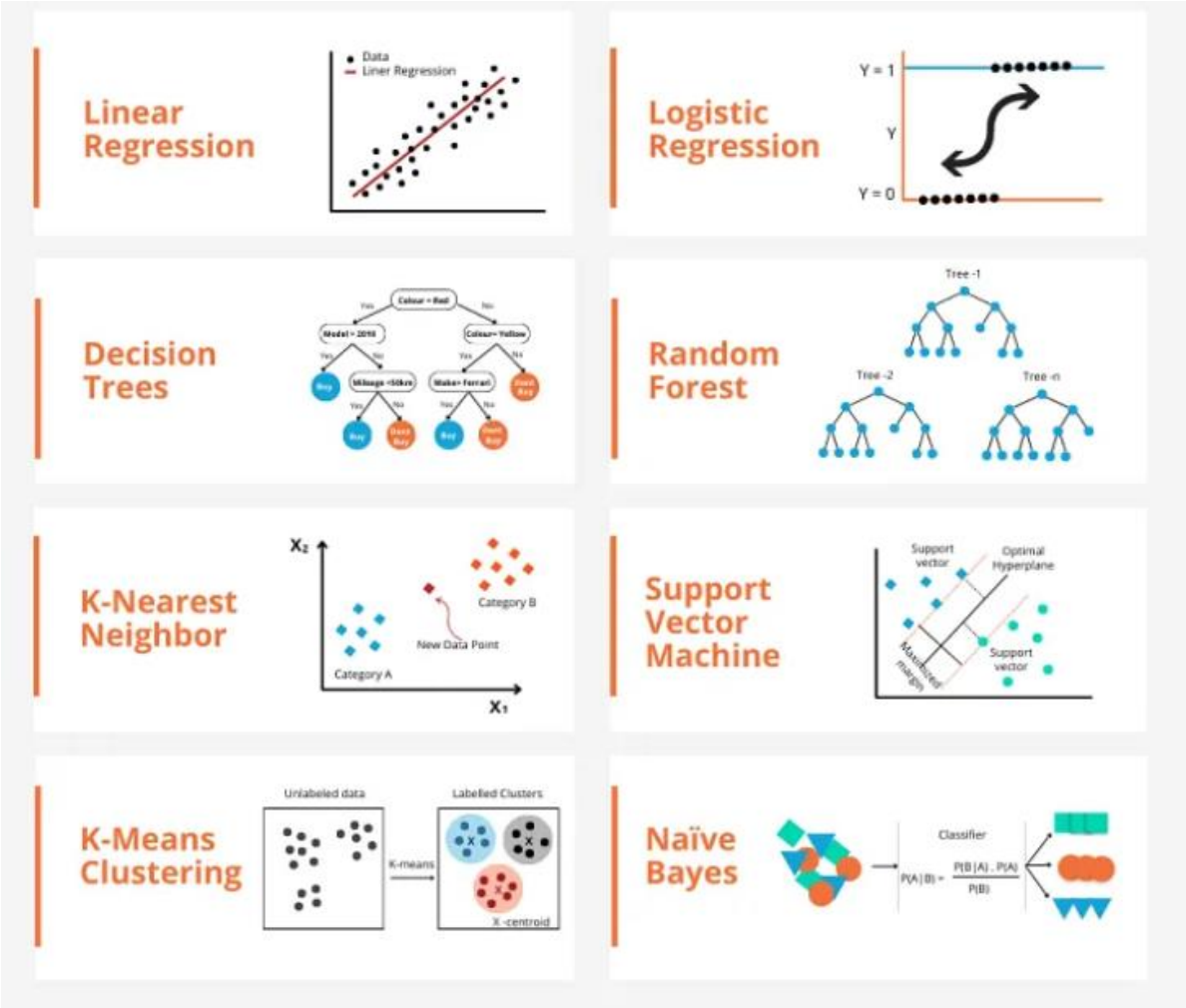
- **Machine learning**
- **Regression**
- **Regression vs classification**
- **Linear Regression**
- **Linear Regression vs Logistic Regression**

Machine learning




MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
CLASSIFICATION
REGRESSION
CLUSTERING
Naive Bayes
Nearest Neighbor
Discriminant Analysis
Support Vector Machines
Decision Trees
Ensemble Methods
Neural Networks
Linear Regression
GLM
Gaussian Process
K-Means, K-Medoids
Fuzzy C-Means
Hierarchical
Gaussian Mixture
Hidden Markov Model

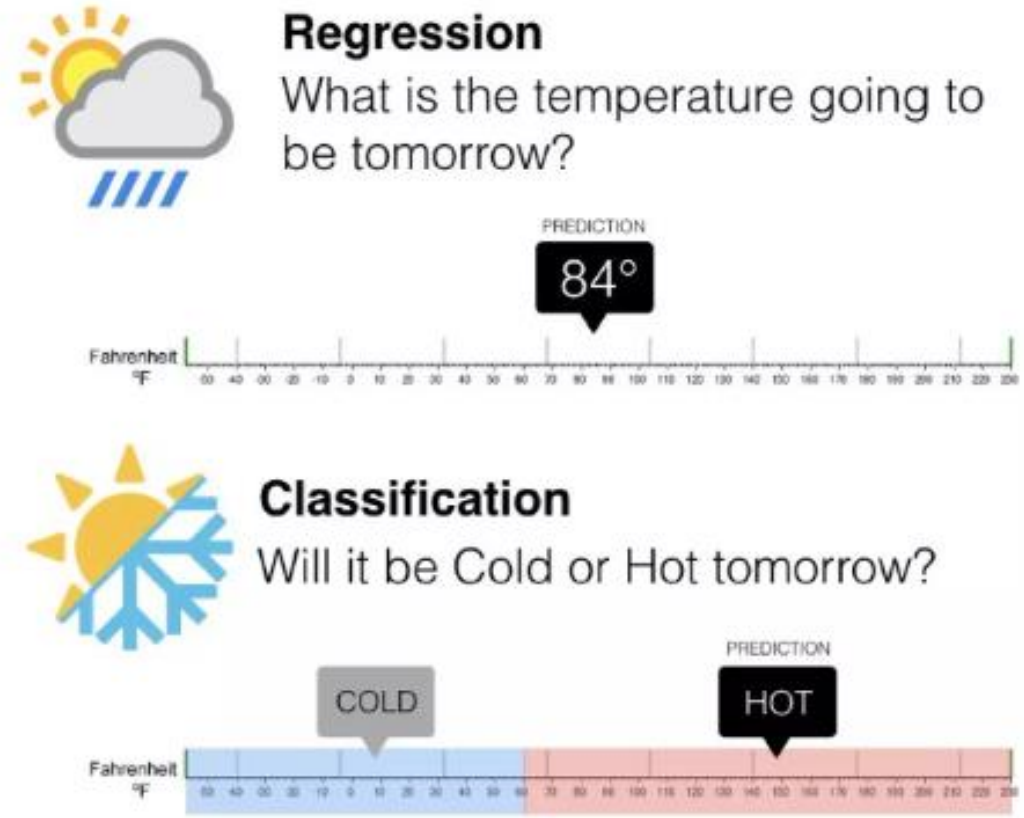
TOP 8 ML MODELS



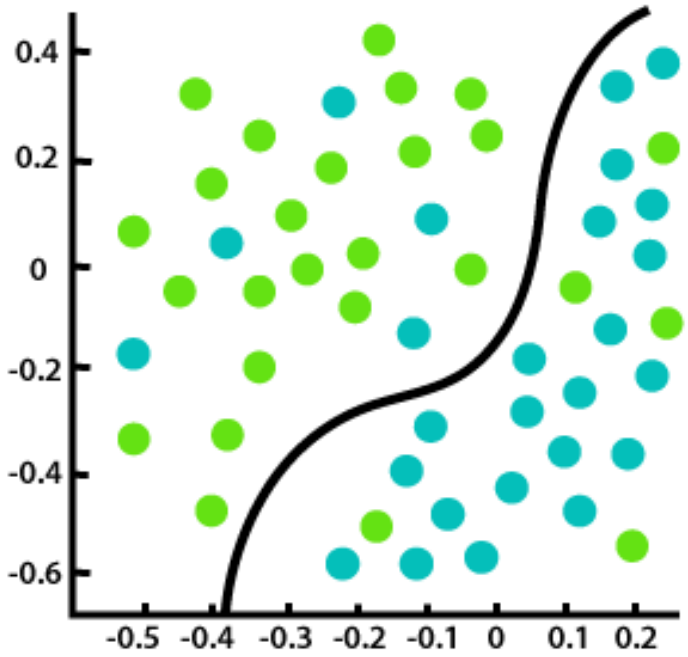
Regression

- **Regression analysis** is a set of statistical methods used for the estimation of relationships between a dependent variable(often called outcome variables) and one or more independent variables (often called features).
- It can be utilized to assess the strength of the relationship between variables and for modeling the future relationship between them.
- Regression analysis includes several variations, such as linear, multiple linear, and nonlinear. The most common models are simple linear and multiple linear. Nonlinear regression analysis is commonly used for more complicated data sets in which the dependent and independent variables show a nonlinear relationship.

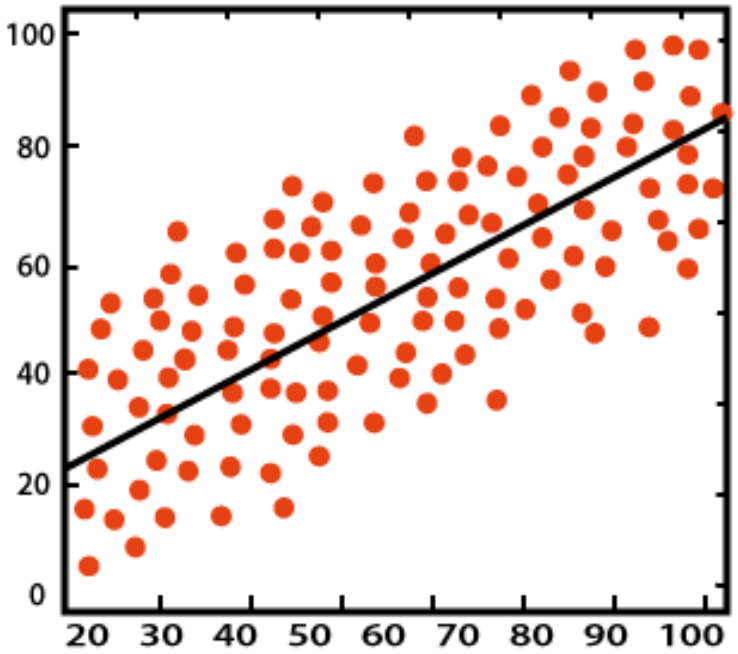
Regression vs classification



Regression vs classification



Classification



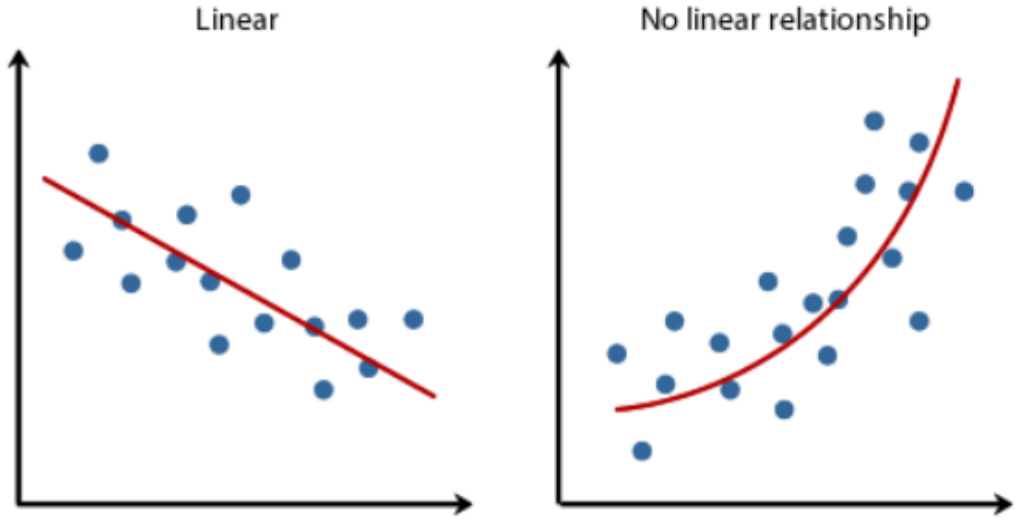
Regression

Regression vs classification

Property	Classification	Regression
Output type	Discrete (class labels)	Continuous (number)
What are you trying to find ?	Decision boundary	Best fit line
Evaluation	Accuracy	Sum of squared error or R-squared
Examples	Images, person who wrote an email	Speed, age, weather, income

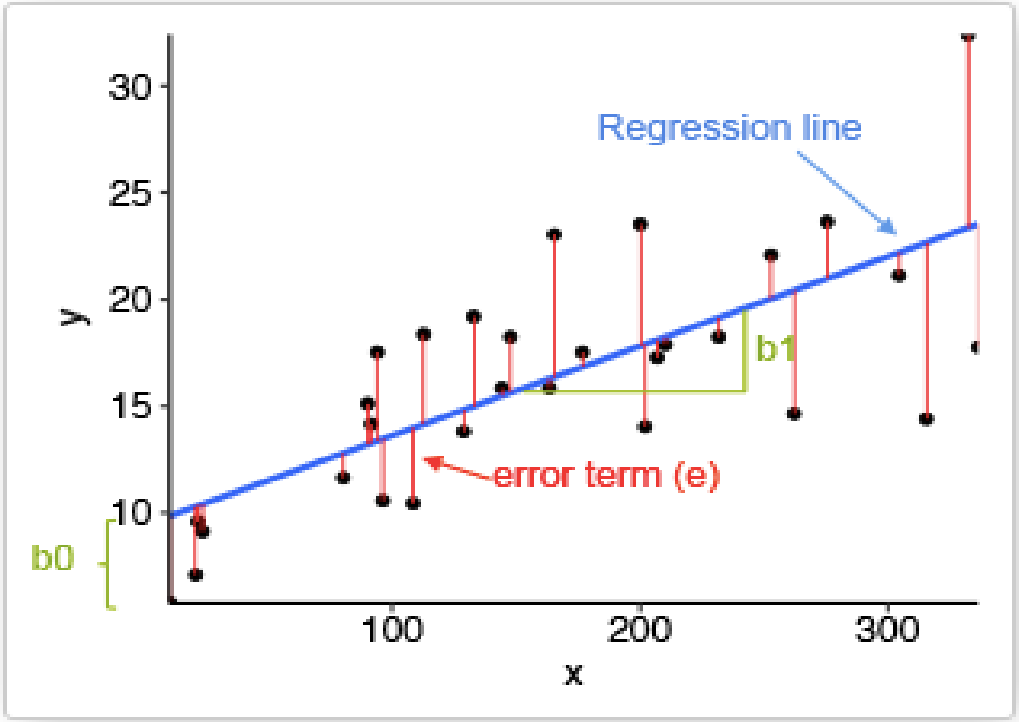
Linear Regression

- Linear regression is a linear model, e.g. a model that assumes a linear relationship between the input variables (x) and single output variable.
- Straight line equation: y = m * x + b
- M: slope
- B: intercept



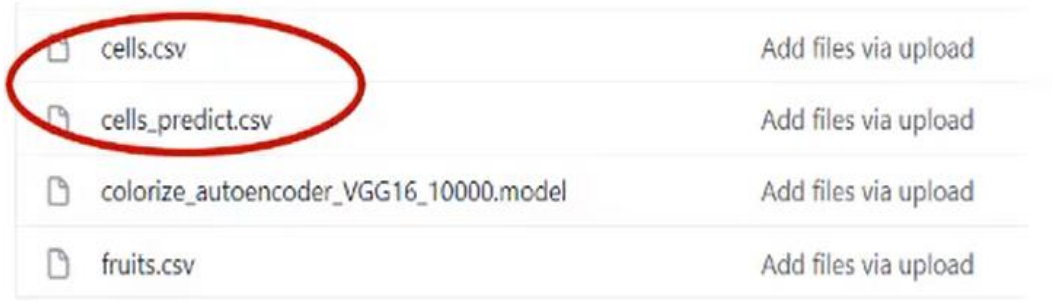
The figure below illustrates the linear regression model, where:

- the best-fit regression line is in blue
- the intercept (b0) and the slope (b1) are shown in green
- the error terms (e) are represented by vertical red lines



Linear regression using sklearn

- Requirements:
  - Data: https://github.com/bnsreenu/python_for_microscopists/tree/master/other_files
  - Pip install numpy
  - Pip install matplotlib
  - Pip install pandas



R-squared

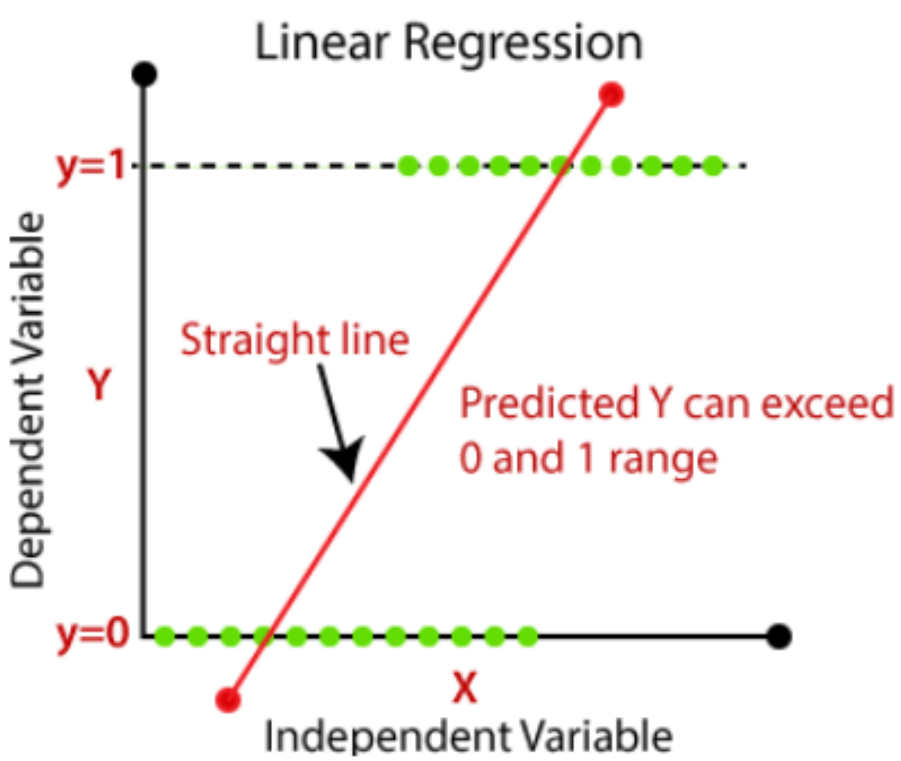
- After fitting a linear regression model, you need to determine how well the model fits the data. Does it do a good job of explaining changes in the dependent variable?
- R-squared measures the strength of the relationship between your model and the dependent variable on a convenient 0 – 100% scale.
- Usually, the larger the R^2, the better the regression model fits your observations.

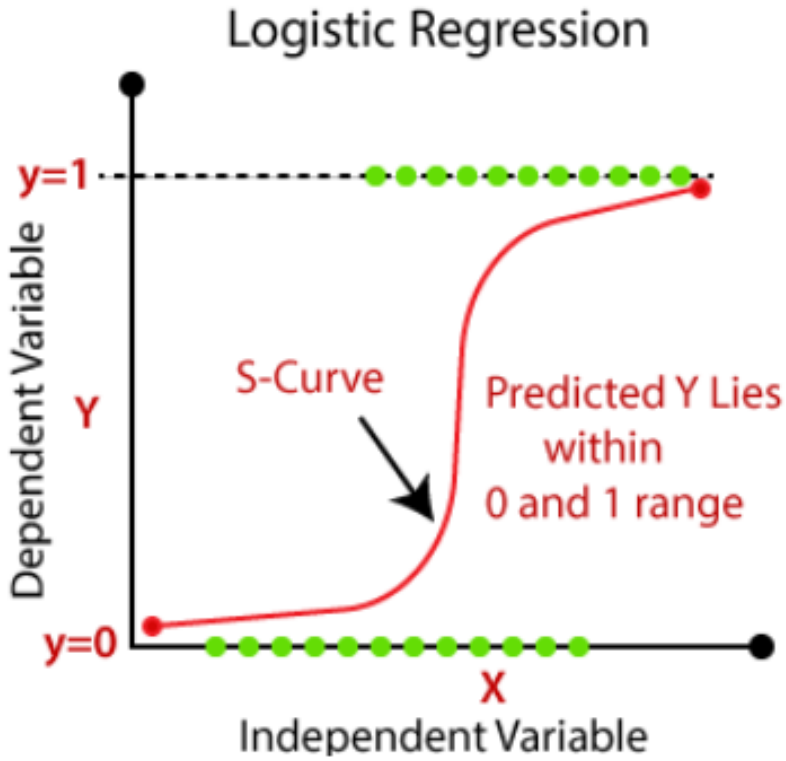
```
print(reg.score(x_df,y_df))  #Prints the R^2 value,
```

Output : 0.9784252641866714

Linear Regression vs Logistic Regression

- Linear Regression and Logistic Regression are the two famous Machine Learning Algorithms which come under supervised learning technique.
- Since both the algorithms are of supervised in nature hence these algorithms use labeled dataset to make the predictions.
- But the main difference between them is how they are being used. The Linear Regression is used for solving Regression problems whereas Logistic Regression is used for solving the Classification problems. The description of both the algorithms is given below along with difference table.


Linear Regression
y=1
y=0
Dependent Variable
Y
Straight line
Predicted Y can exceed
0 and 1 range
X
Independent Variable


Logistic Regression
y=1
y=0
Dependent Variable
Y
S-Curve
Predicted Y Lies
within
0 and 1 range
X
Independent Variable

Thank you