

Supervised Machine Learning:

Supervised learning uses a training set to teach models to yield the desired output. This training dataset includes inputs and correct outputs, which allow the model to learn over time. The algorithm measures its accuracy through the loss function, adjusting until the error has been sufficiently minimized.

One of the most popular supervised machine learning models is :

Regression Model :

- Machine Learning Regression is a technique for investigating the relationship between independent variables or features and a dependent variable or outcome. It's used as a method for predictive modeling in machine learning, in which an algorithm is used to predict continuous outcomes.

What are the types of regression?

- Simple Linear Regression
- Multiple linear regression
- Logistic regression

What is simple linear regression?

Simple Linear regression is a linear regression technique that plots a straight line within data points to minimize error between the line and the data points. The relationship between the independent and dependent variables is assumed to be linear in this case.

What is multiple linear regression?

Multiple linear regression is a technique used when more than one independent variable is used. Polynomial regression is an example of a multiple linear regression technique. It is a type of multiple linear regression, used when there is more than one independent variable. It achieves a better fit in comparison to simple linear regression when multiple independent variables are involved. The result when plotted on two dimensions would be a curved line fitted to the data points.

What is logistic regression?

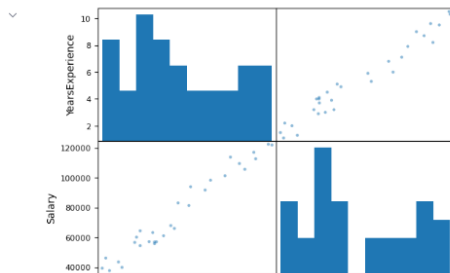
Logistic regression is used when the dependent variable can have one of two values, such as true or false, or success or failure. Logistic regression models can be used to predict the probability of a dependent variable occurring. Generally, the output values must be binary. A sigmoid curve can be used to map the relationship between the dependent variable and independent variables.

→ in this lab we had the chance to work with simple regression (Experience, Salary), multiple regression (Insurance dataset), linear polynomial regression (China GDP)

Part 1: Simple regression (Experience Salary)

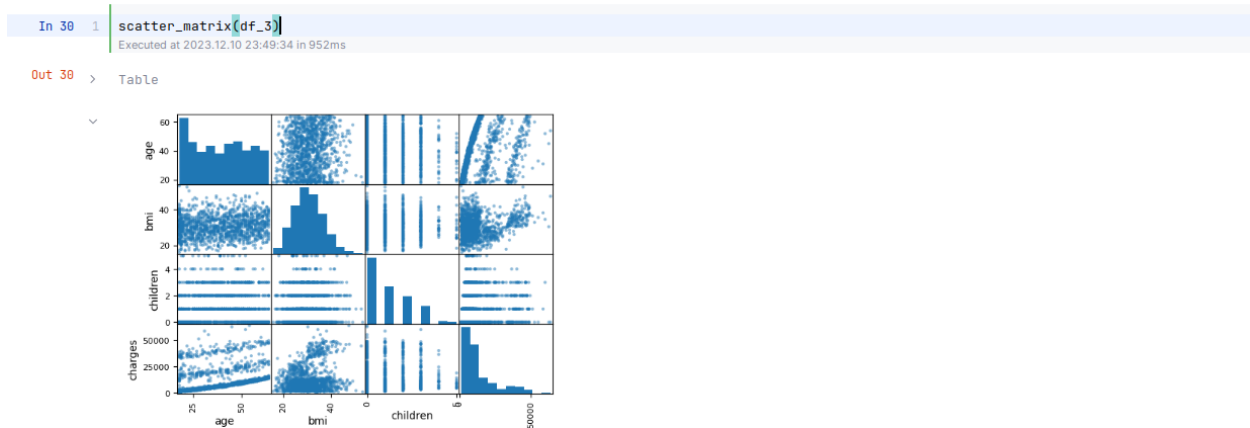
```
In 34 1 from pandas.plotting import scatter_matrix
      2
      3 # scatter matrix of the first data set
      4 scatter_matrix(df)
      5 # strong positive correlation between the years of experience and the salary
      6 # the more the years of experience the more the salary
      Executed at 2023.12.09 12:02:19 in 334ms
```

```
Out 34  array([[<Axes: xlabel='YearsExperience', ylabel='YearsExperience'>,
               <Axes: xlabel='Salary', ylabel='YearsExperience'>],
               [<Axes: xlabel='YearsExperience', ylabel='Salary'>,
               <Axes: xlabel='Salary', ylabel='Salary'>]], dtype=object)
```



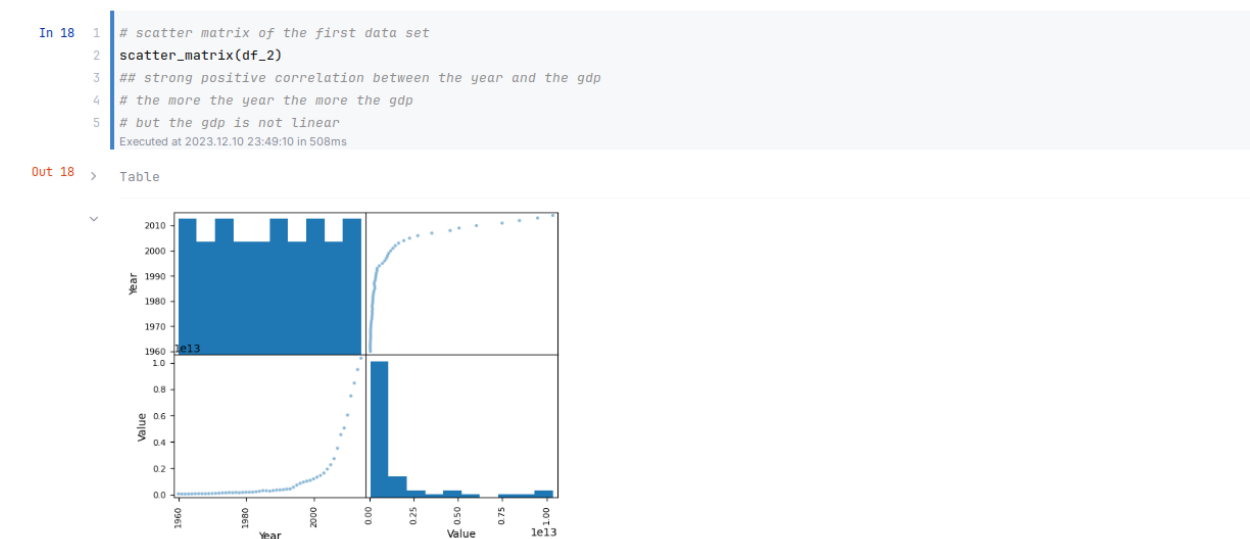
From the data, it is evident that there exists a strong positive linear correlation among the observed variables.

Part 2: Multiple regression (Insurance dataset)



→ bmi, children data not are strongly correlated with the target column

Part 3: Linear polynomial regression (China GDP)



→ the year and value are strongly correlated but not linear, polynomial

Conclusion :

→ Linear regression is a fundamental machine learning algorithm that has been widely used for many years due to its simplicity, interpretability, and efficiency. It is a valuable tool for understanding relationships between variables and making predictions in a

variety of applications. However, it is important to be aware of its limitations, such as its assumption of linearity and sensitivity to multicollinearity. When these limitations are carefully considered, linear regression can be a powerful tool for data analysis and prediction.

References :

<https://www.seldon.io/machine-learning-regression-explained>

<https://www.investopedia.com/terms/r/regression.asp>

<https://www.ibm.com/topics/supervised-learning>

<https://www.geeksforgeeks.org/ml-linear-regression/>