The top 5 Machine Learning Library

1. Pandas
2. Numpy
3. Sk-learn
4. NLTK
5. Keras

The given table has row columns and values which has given different names in machine learning.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 |
| 4 | 5 | 3 | 1 | 4 | 0 |
| 1 | 1 | 3 | 2 | 7 | 1 |
| 2 | 3 | 1 | 3 | 5 | 1 |
| 6 | 5 | 3 | 3 | 8 | 0 |

Column = **Attributes**

Row = **Observation**

Usually Column 6 is result we need called as **Target Value**

Column 1 to column 5 are called as **Training data**

Total table is called as **Data set.**

**Simple steps:**

10%

80%

Data is collected, and it should to be cleaned. All data obtained are not clean. Cleaning data takes long period. 80%. Then the model is built, and prediction is done, if it gives satisfying output. If not, again changes are made on model and prediction is done again.

**SciKit- Learn:**

* Modeling data
* Not focused on loading, manipulating, summarizing data. Pandas and Numpy is used for manipulation of data.

**NLP (Natural Language Processing):**

* Analysis, or processing where machine would be able to know the meaning or related forms or implies of text.
* Whole paragraph can be tokenized into sentences using sent\_tokenize from nltk.tokenize
* One sentences can be tokenized into single words using word\_tokenize from nltk.tokenize