Group 6

PRN: 2019BTECS00105 & 2019BTECS00111

Batch: T7

SET LAB Assignment 2

1. Study of Frameworks/IDEs/Softwares

Anaconda

The Python & R Distribution

- Original Author : Peter Wang and Travis Oliphant
- Developers : Anaconda Inc. previously Continuum analytics.
- Initial Release : 0.8.0 / 17July 2012
- Stable Release : 2021.11 / 17 November 2021
- Repository : https://repo.anaconda.com/pkgs.
  Anaconda Cloud is a package management service by Anaconda where users can find, access, store and share public and private notebooks, environments, and conda and PyPI packages.
- Written In : Written in Python with some modules written in C
- Operating System Support : Supported for Mac , Linux, Windows.
- Platform and Portability : Supports portability by combining all data science assets and provides Anaconda Navigator.
- Type : Distribution of Python and R programming lang.
- Website : https://www.anaconda.com/
- Features : Helps in managing data science project environments, having strong data science library support.
- Size : 300-400MB
- Privacy and Security : Provides security for projects by detecting errors in an optimal way. Anaconda collects and processes your personal data to help us provide you the product and services.
- Type of software : Open Source.
- Licences : Freemium (Miniconda and the Individual Edition are free software, but the other editions are software as a service).
- Latest Version : 2021.05 release.
- Cloud Support : Anaconda for cluster management can launch and bootstrap clusters on a variety of cloud services.currently supports Amazon EC2.
- Applicability : Used widely for data science projects, using Scipy, Pandas, Numpy packages & adjusts virtual environments using conda navigator.
- Drawbacks : Requires little bit more time for loading. It is very heavy software.

2. Implement linear regression problem using Google colab (Perform preprocessing, training and testing)

Dataset : Weather of Mumbai from 2nd to 9th June,2018

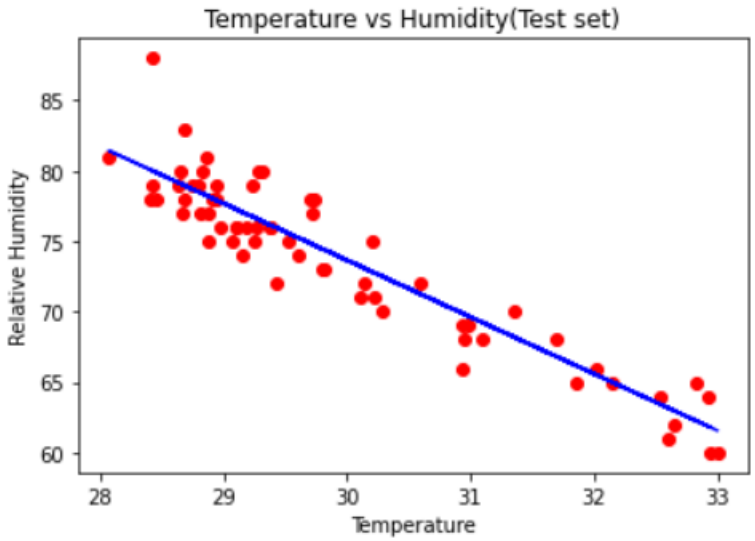
Relative humidity and temperature relation is shown using linear regression.

Google collab is used for implementing the code.

Link for google collab: https://colab.research.google.com/drive/1g6RIISoLk6dR2OekT7x4K-ISAG5Xrwbc?usp=sharing

Result:

```
Coefficients:
 [-4.02872657]
Intercept:
 194.5221290669885
Mean squared error: 4.62
Variance score: 0.87
```



Implementation screenshot: