Spectrum Club,

College of Engineering and Technology

Bhubaneswar

www.spectrumcet.com

Dear Intern,

Welcome to Data **Science and Machine Learning** internship drive. This internship will be basically divided into 3 phases, each with its own task, which will all be preceded by a pre-task phase. This main aim of this task is to guide you through the initial setup and the prerequisite knowledge required before you finally dive into the start of your main internship phases. So go through the details minutely, and prepar yourselves well in advance with the resources, because at the end of the internship, you will be a full fledged data scientist ready to change the world!

**PRE-TASK: -**

Welcome to the pre-task phase of the internship. This phase is all about getting you familiar with the environment you are going to work in, as well as the language you are going to work with. Data Science and Machine Learning techniques can be applied using any language such as R, C++, Java, etc. but in this internship we will be using **Python 3.6** for it.

Python is well preferred and widely used in the field of Data Science mainly due to two foremost reasons, its **simplicity** and its arsenal of available libraries **and frameworks**. Before the beginning of the internship, we would like the interns to be familiar with python basics, but in case you are not, we have provided you with YouTube links and resources to prepare well in advance before the beginning of the actual internship phase.

There are multiple platforms you can use for the purpose of data science, but for this internship we would prefer if you used **Jupyter Notebook, or Spyder IDE**(both are part of Anaconda Distribution) for this purpose. There is also the option for you to work on the cloud, such as **Google Colab**, which would help you work in a really powerful environment built mainly for data analysis and artificial intelligence purpose. The resources and links to set up each of the environment has been provided with this document.

**ALL INTERNS ARE REQUIRED TO GET ACQUAINTED WITH GITHUB IN THE PRE-TASK PERIOD AS THEY WILL HAVE TO SUBMIT THEIR TASKS THROUGH GITHUB.**

**LINK TO TUTORIALS FOR GITHUB AND GITHUB DESKTOP ARE GIVEN IN RESOURCES SECTION. YOU ARE REQUESTED TO FIRST UNDERSTAND THE BASICS OF GITHUB AND THEN PROCEED WITH THE GITHUB DESKTOP TUTORIAL.**

**RESOURCES: -**

**Set up-**

* **GitHub Desktop: -** <https://desktop.github.com/>
* **Anaconda Platform: -** <https://www.anaconda.com/distribution/>
* **Google Colab Notebook:-** <https://colab.research.google.com/notebooks/intro.ipynb>

**Tutorials-**

* **GitHub: -** <https://www.youtube.com/watch?v=SWYqp7iY_Tc>
* **GitHub Desktop: -** <https://www.youtube.com/watch?v=kFix7UDJ7LA>
* **Anaconda Installation with Jupyter Notebook and Spyder: -** <https://www.youtube.com/watch?v=5mDYijMfSzs>
* **Google Colab:-** <https://www.youtube.com/watch?v=i-HnvsehuSw>
* **Jupyter Notebook tutorial:** - <https://www.youtube.com/watch?v=HW29067qVWk>
* **Python Tutorial -** <https://www.youtube.com/playlist?list=PLBZBJbE_rGRWeh5mIBhD-hhDwSEDxogDg>

**Note:-** If you are new to python, or even if you know the basics of python, it is advisable to go through the concepts of **loops, dictionaries, lists and slicing** in python. We will be using these concepts quite thoroughly throughout these internship tasks.

WARM REGARDS,

**SPECTRUM, CET-B**