

DATA SCIENCE PLATFORM

DSP (Data Science Platform) is a platform to do analyse any kind of data. It exposes a number of Machine learning algorithms to do the Analysis and train the model.

The key difference between the algorithms exposed in the DSP vs those exposed in traditional frameworks (like MatLab, R, Python, SAS, Excel, SPSS etc) is that

- ❖ **DSP is built for scalability and distributed computing (for e.g. Running a Linear Regression algorithm on TeraBytes of data)**
- ❖ **DSP exposes Philips proprietary clinically oriented Algorithms (for e.g. Automatic extraction of Anatomies, Findings, Diseases, Procedures etc. from Radiology reports)**
- ❖ **DSP eases prodxcionisng the models**
- ❖ **DSP provides a RESTful set of APIs (behind which the ML algorithms are implemented in a highly distributed nodes)**
- ❖ **DSP provides a number of Deep Learning techniques**
- ❖ **DSP exposes the ML algorithms in various categories (Classification, Regression, Clustering, Dimensionality Reduction, Pattern extraction, etc)**
- ❖ **Even though the algorithms are implemented in different languages/Tools/Technologies, it exposes the capabilities via a set of thin packages in R and Python. So data scientists who are proficient in R and Python can directly use this.**
- ❖ **DSP also allows users to plugin custom algorithms (for e.g. Custom code to pre-process the data using any packages publicly available in R and Python)**
- ❖ **DSP can be deployed on-prem or on the Cloud**
- ❖ **When Deployed on the cloud, DSP can elastically grow (based on the resources required (compute power, memory etc)**
- ❖ **DSP has the capability to decommission the cluster once the model is trained (thereby reducing the cost)**

- ❖ When the user logs on to DSP, he/she can choose either Python NoteBook or RStudio to develop the model .

Note that for every algorithm that is exposed by DSP, a sample code in R/Python, documentation is present so that user can get to know the usage of any functionality