

PySpark Assignment – 2

Task:

Each Company publishes the Company Annual Report, it is mandatory to stick with one of the UNESCO Sustainable Development. The task is to create a Machine Learning Model that classifies the Goal to which the company is working on given the company report as input.

PS: Ofcourse, you should use the DataFrames and ML-Lib from PySpark to create the model

For more information check the following links:

SGD: <https://sdgs.un.org/goals>

Dataset: <https://zenodo.org/records/5550238/files/osdg-community-dataset-v21-09-30.csv>