|  |
| --- |
| **Startup prediction system report** |
|  |
| This report contains a brief description of how to deal with the project dataset and machine learning algorithms. |
|  |

|  |  |  |
| --- | --- | --- |
| Dataset explanation: Data set contains approximately 60.000 column of startups name, category, fund, place and status if it closed or still operating. Cleaning dataset: There are two datasets in repo.  First “companies.csv” it’s the main dataset with the original columns and features.  Second “sample\_out.csv” updated dataset with clean columns and rows and filtration to run the Naïve-Bayes algorithm. Dataset accuracy: Calculating dataset with machine learning code using sklearen, pandas, NumPy.  Dataset accuracy result = 81% |  |  |