|  |
| --- |
| DSP Module II  Python Final Assignment |
| 01 / 08 / 2020 |

Seven Sins Restobar

A new restaurant startup has born in German Village, and three people are working very extensive hours to make things happen. The entry barriers are high, as there is such a huge contest against big companies to attract people that could buy their products/services.

As an EVS new initiative, the idea is to support new local businesses to achieve prosperity in the region, and you were selected to develop a study that could lead them to improve their results.

The success is vital for these people, as their growth will bring new people and increase employment.

With a bunch of effort and energy, they have been collecting reviews from friends, family and new customers, and they have stored them in a tiny dataset, which consists in two features:

-The restaurant review

-The thumb of nail critic: They did or did not like the service received.

Right now, they don’t have enough time to conduct any sort type of analysis coming from them. So they don’t know what to do in terms of strategy.

As a data science analyst, you are requested to perform an analysis on the data that they have and give them some feedback. Your critical thinking is your best friend for developing a case, and support it with data driven conclusions.

But you’re not alone in this process: You are part of a talented people team, and all of you are using this opportunity to escalate and generate more impact in our community. The person in charge only expects from you a Python code (.py file), with the possibility to include comments and all sort type of data driven analysis (Markdown is your friend). Due the lack of time and resources available: Your opinion and recommendations will only be considered if they are clear and follow a convincing and clear path and objective.

There have 500 labeled reviews (Liked 1; Disliked: 0).

They have other 500 reviews, but unlabeled: A model is required for classifying them.

Your boss requires a python (.py) file that includes:

* Dataset review (descriptive)
* Pre-processing steps and methods.
* Define a main objective: What is the object of your study? To analyze the relation of certain type of products/services with the feedback? Detect things that lead people to write positive/negative reviews? A mixture? To classify new upcoming reviews? Why?
* The feature engineering process (techniques used, commented).
* A model selection process (parameters, candidates, selected (why?) and evaluated metrics).
* Analysis based in selected model.
* Final conclusion and recommendation(s) (be concise)

Evalueserve Disclaimer

The information contained in this report has been obtained from reliable sources. The output is in accordance with the information available on such sources and has been carried out to the best of our knowledge with utmost care and precision. While Evalueserve has no reason to believe that there is any inaccuracy or defect in such information, Evalueserve disclaims all warranties, expressed or implied, including warranties of accuracy, completeness, correctness, adequacy, merchantability and / or fitness of the information.