

Module 4 - Bonus Challenges

§M4: Support Vector Machines

- Below are open-ended bonus challenges; solving them is not required but can help you better understand ML/AI in the context of engineering, and how to use them in practical cases.
- Bonus points earned in all homework assignments will be averaged (6 bonus points for each assignment) and then directly added to your final score to calculate your final letter grade.

Challenge 1.1. (Adapted from [1]) As shown in Figure 1a, consider a heterogeneous column subjected to a fixed level of applied displacement. The column is segmented into 16 parts, each capable of being composed of one of two materials (Figure 1b), with Young's Modulus (E) values of 1 and 4, respectively. Certain distributions of these materials can cause the column to become unstable (buckle).

The Buckling Instability Classification (BIC) dataset [1] contains results from finite element simulations for 65,536 columns with various material distributions. Each input file represents a 16-column vector entry. For each 16x1 vector input, there is a corresponding output that indicates whether the column was stable or unstable under the fixed displacement. An output value of "0" indicates stability, while "1" indicates instability. These data were split into training (`'train_input_data.txt'`, `'train_output_data.txt'`) and testing data (`'test_input_data.txt'`, `'test_output_data.txt'`) .

The goal of this task is to predict the stability of the column (either stable or unstable) based on the material choices of its segments. Please complete this task according to the following requirements: (6pts)

1. Please use the built-in functions `LogisticRegression` and `SVC` in `sklearn`.
2. Compare the performance of Logistic Regression and Support Vector Machine in this specific problem, considering factors such as efficiency, accuracy, and interpretability.
3. Submit your Jupyter notebook file with necessary comments, reporting the score for the classification.

Below are some useful tips:

1. `pandas` library is a useful tool to read the `.txt` file.
2. Refer to the documentation of `LogisticRegression` and `SVC` for guidance. Go over each parameter of the built-in function and assess their relevance to this specific case.
3. Built-in function `Pipeline` and preprocessing module `sklearn.preprocessing` could be useful in this task.

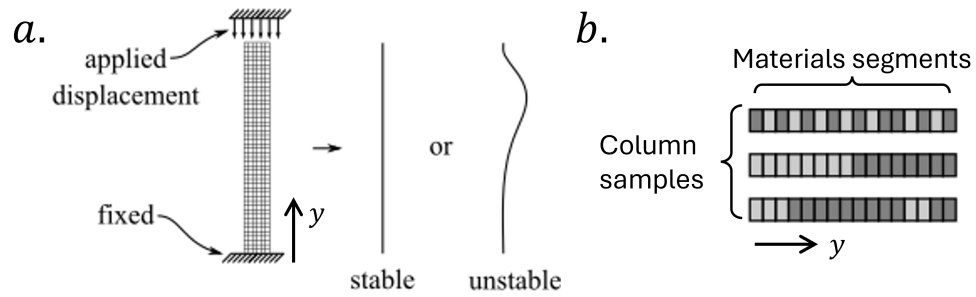


Figure 1: Buckling instability classification of heterogeneous columns, adapted from [1]

References

- [1] Emma Lejeune. *Buckling Instability Classification (BIC)*. 2020. URL: <https://open.bu.edu/handle/2144/40085>.