**Naive Bayes Model - Report**

**1. Objective**

**The goal of this project is to:**

* **Build a Gaussian Naive Bayes model.**
* **Analyze the impact of the var\_smoothing hyperparameter on model performance.**
* **Perform RandomizedSearchCV tuning to find the best var\_smoothing.**

**2. Methodology**

* **Data was split into train/test sets (80% training, 20% testing) with random\_state=42.**
* **Models were trained under three conditions:**
  + **Default setting: var\_smoothing=1e-9**
  + **Manual setting: var\_smoothing=1e-7**
  + **Tuned setting: Using RandomizedSearchCV to find the best var\_smoothing.**
* **Evaluation Metric: Accuracy Score on the test set.**

**3. Experiments and Results**

| **Model Setting** | **Accuracy (%)** |
| --- | --- |
| **Default (var\_smoothing=1e-9)** | **82.07%** |
| **Manual (var\_smoothing=1e-7)** | **79.69%** |
| **RandomizedSearchCV Best Model (1e-9)** | **82.07%** |

**RandomizedSearchCV confirmed that the default value (var\_smoothing=1e-9) is optimal for this dataset.**

**A higher smoothing value (1e-7) decreased accuracy — showing slight underfitting.**

**4. Conclusion**

* **GaussianNB with the default var\_smoothing=1e-9 achieved the best performance.**
* **Increasing var\_smoothing led to lower accuracy, suggesting that too much smoothing harms the model’s ability to fit the data properly.**
* **RandomizedSearchCV tuning confirmed that the default hyperparameter already gives the best results.**

**5. Files**

* **Naive\_Bayes.ipynb: Full Jupyter Notebook with all code (training, evaluation, tuning).**