

SVM Parameter Optimization Assignment UCS654

Project Title

Support Vector Machine (SVM) Parameter Optimization Using GridSearchCV

Dataset

The dataset used in this assignment includes a classification problem where the target variable is `quality`. The features are preprocessed using `StandardScaler` for normalization.

Methodology

1. Data Preprocessing

- The dataset is loaded using `pandas`.
- The target variable is separated from the feature set.
- Features are standardized using `StandardScaler` to improve model performance.

2. Model Selection

- A **Support Vector Machine (SVM)** classifier is selected from `sklearn.svm`.

3. Parameter Optimization

- A **GridSearchCV** is used to perform hyperparameter tuning over a specified parameter grid (e.g., `C`, `gamma`, and `kernel`).
- Cross-validation is used to ensure the generalizability of the model.

4. Model Evaluation

- Accuracy score is used as the primary performance metric.
- Additional visualizations (like count plots) are created using `seaborn` to explore class distributions.



Result Table

The results are summarized in a structured format, showing performance metrics for each parameter combination. The best parameter set is selected based on the highest cross-validated accuracy.

Creation of a Result Table



result



| | Sample | Best Accuracy | Best Kernel | Best Nu | Best Epsilon |
|--|--------|---------------|-------------|---------|--------------|
|--|--------|---------------|-------------|---------|--------------|

| | | | | | |
|---|---|------|-----|------|------|
| 0 | 1 | 0.54 | rbf | 7.37 | 6.94 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 1 | 2 | 0.53 | rbf | 8.18 | 7.91 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 2 | 3 | 0.50 | rbf | 9.20 | 6.99 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 3 | 4 | 0.56 | rbf | 8.77 | 5.30 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 4 | 5 | 0.58 | rbf | 2.02 | 0.91 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 5 | 6 | 0.55 | rbf | 2.99 | 8.02 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 6 | 7 | 0.58 | rbf | 3.54 | 4.11 |
|---|---|------|-----|------|------|

| | | | | | |
|---|---|------|-----|------|------|
| 7 | 8 | 0.54 | rbf | 2.10 | 8.21 |
|---|---|------|-----|------|------|

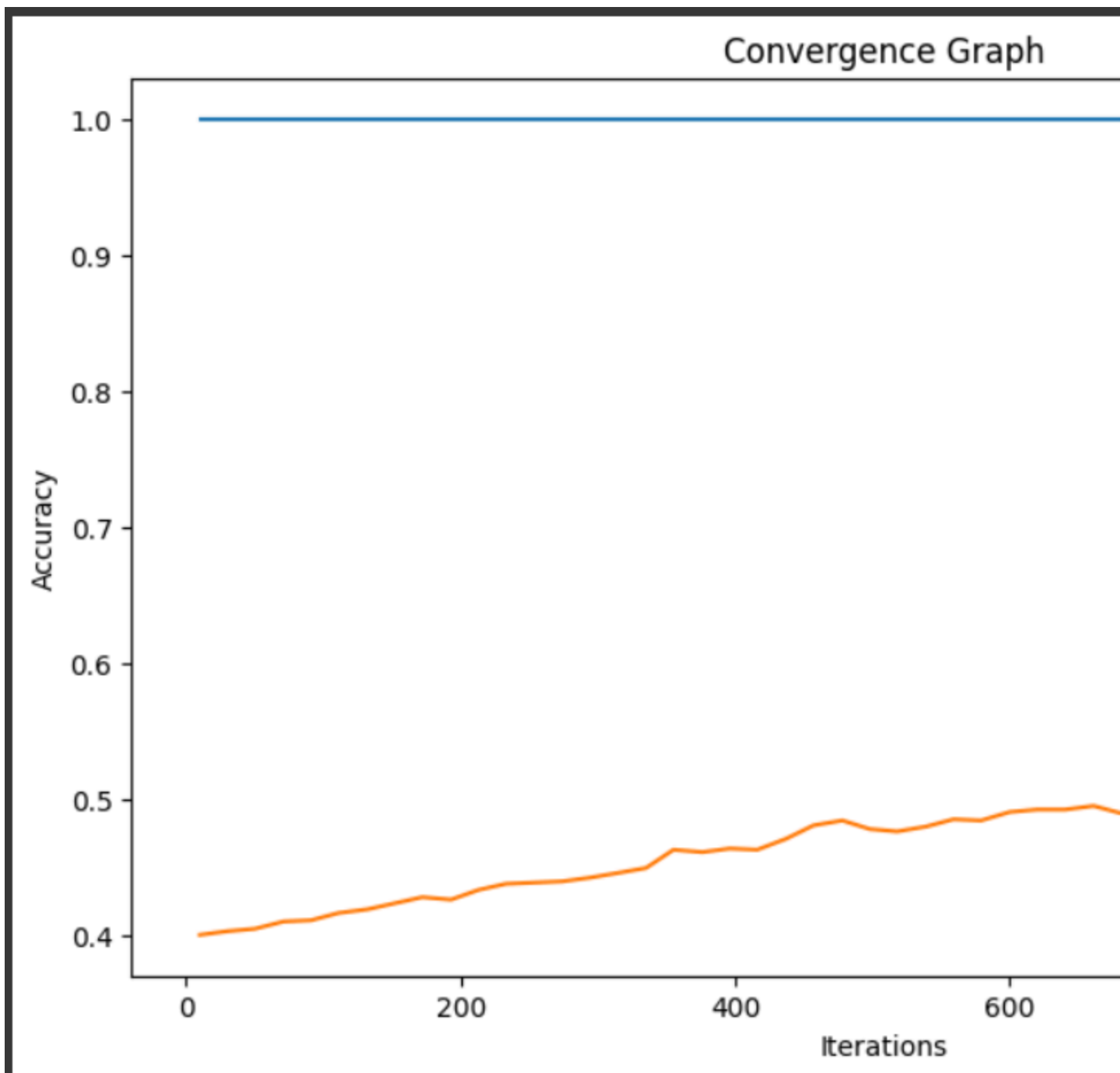
| | | | | | |
|---|---|------|-----|------|------|
| 8 | 9 | 0.46 | rbf | 0.33 | 7.44 |
|---|---|------|-----|------|------|

| | | | | | |
|---|----|------|-----|------|------|
| 9 | 10 | 0.58 | rbf | 3.48 | 7.71 |
|---|----|------|-----|------|------|





Result Graphs



Conclusion

- SVM performs well with properly scaled features.
- Parameter tuning via GridSearchCV significantly improves model accuracy.

- Visualization helps in understanding the dataset distribution and model behavior.
-

Libraries Used

- pandas
- numpy
- matplotlib
- seaborn
- scikit-learn

How to Run

jupyter notebook 102203351_Assignment_ParameterOptimization.ipynb

Ensure all dependencies are installed and the dataset is properly loaded within the notebook.
