**AI ML Internship Log**

# Day 8 - SVM RBF Implementation, Evaluation & Final Model Selection

# Date - 22 June 2025

# Team Role - Member

# Project Title - Personality Prediction from Social Media

**🎯 Objective of the Day:**

* To implement **SVM with the RBF kernel**, understand the influence of C and gamma, and evaluate it against Logistic Regression and Linear SVM to select the best model for personality prediction.

✅ **Tasks Completed:**

* Implemented SVC(kernel='rbf') on the TF-IDF dataset
* Understood key SVM concepts: C, gamma, margin, and boundary behavior
* Successfully executed and evaluated the RBF model (took ~3 min to train)

Compared RBF SVM results with:

* Logistic Regression (baseline model)
* Linear SVM (previously best macro F1)
* Analyzed detailed classification metrics
* Identified final best-performing model

❓ **Challenges, Confusions, and Learning Points:**

* Confused initially between C and gamma — both seemed similar
* C = how strict the model is about classification errors
* gamma = how much each data point influences the decision boundary
* Struggled with terms like "linearly separable", "curved boundary", and "overlapping classes", but got conceptual clarity with examples
* Faced confusion around why RBF performs better despite lower accuracy than Logistic Regression — resolved after understanding weighted and macro F1.

📊 **Model Comparison Summary:**

| **Metric** | **Logistic Regression** | **Linear SVM** | **SVM (RBF Kernel)** |
| --- | --- | --- | --- |
| Accuracy | **✅ 64.14%** | 61.56% | 64.03% |
| Macro Avg F1 | 0.43 | **✅0.47** | 0.46 |
| Weighted Avg F1 | 0.62 | 0.61 | **✅0.63** |

**🧠 Conclusion:**

While Logistic Regression had the highest accuracy (64.14%), the SVM with RBF kernel achieved a better overall performance:

* Higher **weighted F1-score (0.63)**
* Better balance between common and rare MBTI types
* Flexible boundary learning through gamma
* Stronger generalization due to non-linear kernel
* ✅ Hence, **SVM with RBF kernel was selected as the final model** for the project at this stage.

🌟 **Self-Reflection:**

* Today was one of the most enriching days of the project. I didn’t just run the code but **understood and analyzed it deeply**.
* Though I felt confused multiple times — especially with C, gamma, and boundary logic — I took the time to break it down and reflect.
* This has strengthened my confidence in not just using ML models, but understanding **why and how they work**.
* I feel more aligned with the project now and am ready for the next stage!