**AI ML Internship Log**

# Day 7 - Implementing SVM Model and Comparison with Logistic Regression

# Date - 21 June 2025

# Team Role - Member

# Project Title - Personality Prediction from Social Media

**🔷 Objective of the Day:**

* To implement **Support Vector Machine (SVM)** as a second ML model, evaluate its performance, and compare it with **Logistic Regression**. Understand the results deeply and reflect on model behavior.

**✅ Tasks Completed:**

* ✔️ Revised Day 6 model and TF-IDF feature generation
* ✔️ Implemented LinearSVC model in Colab
* ✔️ Made predictions and generated classification report
* ✔️ Compared SVM results with Logistic Regression
* ✔️ Understood the meaning of macro vs weighted average F1-scores
* ✔️ Reflected on why SVM didn’t beat Logistic Regression in all metrics
* ✔️ Explored how SVM handles rare classes better
* ✔️ Planned next steps for trying advanced SVM variants

📊 **SVM Model Evaluation Results:**

|  |  |  |
| --- | --- | --- |
| **Metric** | **Logistic Regression** | **Linear SVM** |
| Accuracy | 64% | 61.56% |
| Macro Avg F1 | 0.43 | 0.47 |
| Weighted Avg F1 | 0.62 | 0.61 |

**🧠 Learnings & Understanding:**

* SVM is a margin-based classifier that works well on high-dimensional text data like TF-IDF
* It is known for handling imbalanced class distributions better than Logistic Regression
* Macro F1 gives equal importance to rare classes, and SVM showed an improvement here
* Weighted F1 and accuracy slightly dropped because SVM prioritized fairness across all classes rather than just the majority ones
* Logistic Regression had better overall accuracy, but SVM was more balanced

❓ **Confusions and Doubts Faced:**

* Expected SVM to improve all metrics — but it only improved **macro F1**
* Felt confused why accuracy dropped despite better handling of some classes
* Took time to understand how margin-maximization relates to classification results in real, messy data
* Needed clarity on when to prefer macro avg vs weighted avg

🧪 **Final Result Interpretation:**

SVM did **not outperform** Logistic Regression on accuracy, but it:

* ✅ Showed **better balance across classes**
* ✅ Handled **minority classes** more fairly

Logistic Regression remains a strong **baseline model**

SVM gives value in cases where **equal treatment of all personality types** matters

🌟 **Self-Reflection:**

* Today was an exciting and meaningful day. I didn’t just run code — I did real analysis. I saw how SVM works, why it behaves differently than Logistic Regression, and how to compare models using multiple metrics.
* I also faced some confusion — but I didn’t give up. I stayed curious, learned the theory behind SVM, and analyzed results with confidence. I now feel much more confident in interpreting model outputs and understanding why some models perform better on some tasks.