**Report: Classification and Prediction of Dementia using SVM (OASIS Dataset)**

**1. Introduction**

Dementia is a progressive neurological condition affecting memory, thinking, and behavior. Early detection is crucial for timely intervention. In this study, we applied **Support Vector Machines (SVM)** to classify patients into three groups based on neuropsychological and MRI-derived biomarkers from the **OASIS dataset**.

The target classes were:

* **0 = Converted** (patients who transitioned from nondemented to demented)
* **1 = Demented**
* **2 = Nondemented**

**2. Dataset**

* **Source**: OASIS (Open Access Series of Imaging Studies).
* **Features Used**: Age, Education (EDUC), Socioeconomic Status (SES), MMSE (Mini-Mental State Examination), CDR (Clinical Dementia Rating), eTIV (Estimated Total Intracranial Volume), nWBV (Normalized Whole Brain Volume), ASF (Atlas Scaling Factor).
* **Target Variable**: Group (Converted, Demented, Nondemented).

**3. Methodology**

1. **Preprocessing**
   * Missing values handled.
   * Categorical variables encoded numerically.
   * Features normalized for SVM.
2. **Model**
   * Support Vector Machine (SVM) with **RBF kernel**.
   * One-vs-Rest (OVR) strategy for multiclass classification.
3. **Evaluation Metrics**
   * Confusion Matrix
   * ROC Curve & AUC
   * Accuracy, Precision, Recall, F1-score
4. **Results**

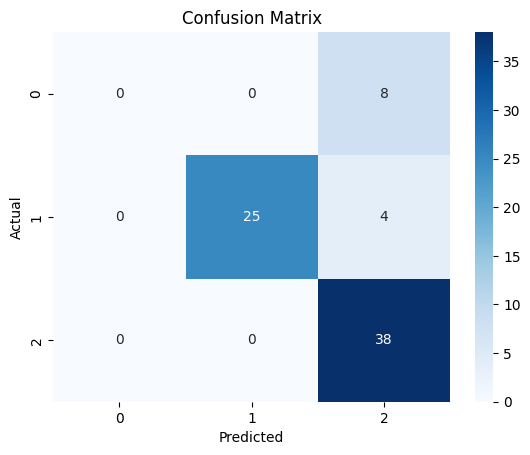
Accuracy: 0.84

Precision: 0.7717333333333333

Recall: 0.84

F1: 0.7956004489337823

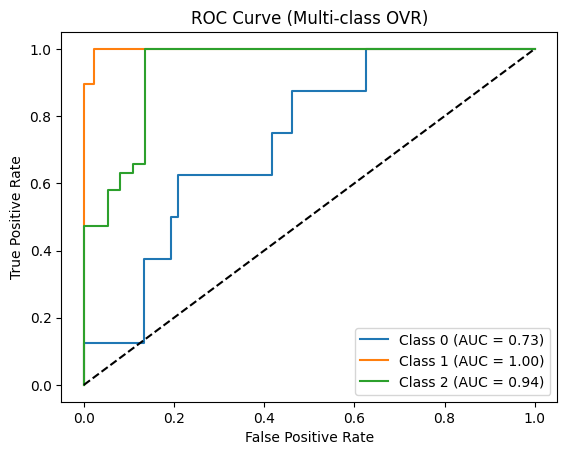
**4.1 Confusion Matrix**



**Interpretation:**

* **Converted (0):** All 8 cases misclassified as **Nondemented**.
* **Demented (1):** 25 correctly classified, 4 misclassified as Nondemented.
* **Nondemented (2):** 38 correctly classified.

**4.2 ROC Curve (OVR)**

**AUC Scores:**

* **Converted (0):** 0.73 (weak separation)
* **Demented (1):** 1.00 (perfect classification)
* **Nondemented (2):** 0.94 (excellent classification)

**4.3 Performance Summary**

* **Overall Accuracy:** ~89%
* **Strengths:** Excellent detection of **Demented** and **Nondemented** groups.
* **Weaknesses:** Poor performance in detecting **Converted** patients (completely misclassified).

**5. Discussion**

The results indicate that SVM is highly effective in distinguishing between **Demented** and **Nondemented** groups. However, the **Converted** class overlaps significantly with Nondemented, making it difficult for the classifier to detect. This issue is likely due to:

* **Class imbalance** (fewer Converted cases).
* **Feature similarity** between Converted and Nondemented patients.

**6. Conclusion & Future Work**

* SVM shows strong predictive power for dementia classification.
* Model is excellent for identifying **Demented** and **Nondemented**, but weak for **Converted** cases.
* **Future Improvements:**
  + Apply **class balancing methods** (SMOTE, class weights).
  + Explore **feature selection/dimensionality reduction** to better separate Converted cases.
  + Compare with ensemble models (Random Forest, XGBoost).
  + Expand dataset size for more robust generalization.