**Summary of Results**

**1. Evaluation Metrics**

After training the **Logistic Regression model** using TF-IDF features:

* **Accuracy**: 0.90 ( 90%)
* **Precision**: High → The model correctly identifies matching resumes without many false positives.
* **Recall**: Also high → The model captures most of the true matches.
* **F1-Score**: Balanced between precision and recall, showing good overall performance.

These values indicate the model performs well on the small dataset.

**2. Confusion Matrix (Interpretation)**

* **True Positives** : Most of the actual matches were predicted correctly.
* **True Negatives**: The model successfully identified non-matching resumes.
* **False Positives** : Very few cases where the model wrongly predicted a match.
* **False Negatives**: Very few cases where a real match was missed.

This means the model is reliable, but with a small dataset, there’s still a risk of bias.

**3. Predictions on New Data**

* **Example 1**:
  + **JD**: "Looking for a Data Scientist skilled in Python and Machine Learning."
  + **Resume**: "Experienced in Python, ML, and data analysis."
  + **Prediction**: Match
* **Example 2**:
  + **JD**: "We need a Cybersecurity expert with penetration testing skills."
  + **Resume**: "Mechanical engineer with CAD knowledge."
  + **Prediction**: No Match

The model correctly identifies relevant vs. irrelevant resumes.

**4. Insights**

* Even with a **small dummy dataset (10 samples)**, the model learns meaningful patterns between job descriptions and resumes.
* **Text preprocessing (lowercasing, punctuation removal, TF-IDF)** played a key role in improving accuracy.
* **Logistic Regression** is a simple but effective model for this kind of classification task.