1. **Discussion**:
   1. **Interpretation of Results and their significance**

The following section outlines various methods utilized in our project, providing an interpretation of results along with an explanation of their significance for each method.

* **Logistic Regression**:

The results of our Logistic Regression suggest that the model has moderate predictive performance. The balance between precision and recall, as indicated by the F1 score, suggests that the model is reasonably good at identifying the states but may not be highly precise or sensitive.

1. Accuracy (0.597):

- The accuracy is a measure of overall correctness in predicting the states.

- In this case, the model achieved an accuracy of approximately 59.7%, indicating that about 59.7% of the predictions were correct.

2. Precision (0.592):

- Precision is a measure of how many correctly predicted instances belong to the predicted class.

- A precision of around 59.2% suggests that when the model predicts a particular state, it is correct about 59.2% of the time.

3. Recall (0.597):

- Recall (or Sensitivity) is a measure of how many actual instances of a class were correctly predicted by the model.

- The recall of approximately 59.7% indicates that the model is able to capture about 59.7% of the instances of the states.

4. F1 Score (0.591):

- The F1 score is the harmonic mean of precision and recall and provides a balanced measure between the two.

- A F1 score of approximately 59.1% suggests a balance between precision and recall, taking into account both false positives and false negatives.

* **SVM one-versus-one:**

In comparison to the Logistic Regression, the SVM Classifier with a one-vs-one strategy appears to perform better across all metrics, with higher accuracy, precision, recall, and F1 score. These results suggest that the SVM model is more effective in predicting the states, demonstrating a higher overall performance.

1. Accuracy (0.849):

- The accuracy is approximately 84.9%, indicating that the SVM Classifier, with a one-vs-one strategy, correctly predicted the states about 84.9% of the time.

2. Precision (0.850):

- Precision is around 85%, suggesting that when the SVM Classifier predicts a particular state, it is correct about 85% of the time.

3. Recall (0.849):

- The recall is approximately 84.9%, indicating that the SVM Classifier captures about 84.9% of the instances of the states.

4. F1 Score (0.848):

- The F1 score is around 84.8%, providing a balanced measure between precision and recall. It suggests a good balance between avoiding false positives and false negatives.

* **SVM one-versus-all**

Our SVM Model, demonstrates excellent performance with our dataset, utilizing both one-vs-all and one-vs-one strategies. This robust performance may indicate the presence of outliers in the data, which SVMs are known to be more resistant to. Furthermore, the consistent and strong predictive performance across accuracy, precision, recall, and F1 score metrics for both one-vs-all and one-vs-one strategies suggests that the choice of strategy does not significantly impact the SVM Classifier's effectiveness in this particular case. This observation is indicative of a well-balanced dataset that includes examples for all classes and features that contain the necessary information for constructing a reliable decision boundary.

1. Accuracy (0.849):
   * The accuracy is approximately 84.9%, indicating that the SVM Classifier, with a one-vs-all strategy, correctly predicted the states about 84.9% of the time.
2. Precision (0.850):
   * Precision is around 85%, suggesting that when the SVM Classifier predicts a particular state, it is correct about 85% of the time.
3. Recall (0.849):
   * The recall is approximately 84.9%, indicating that the SVM Classifier captures about 84.9% of the instances of the states.
4. F1 Score (0.848):
   * The F1 score is around 84.8%, providing a balanced measure between precision and recall. It suggests a good balance between avoiding false positives and false negatives.

* **Naive Bayes Classifier**

Our Naive Bayes Model performs the worst so far, with it not being able to cross the 50% threshold. This might indicate that the datapoints are not conditionally independent, leading to many wrong predictions by the model. The Naive Bayes model, appears to have a lower overall performance compared to the SVM model discussed earlier. The lower accuracy, precision, recall, and F1 score indicate that the Naive Bayes model might not be well-suited for the task or that the data may not align well with the assumptions of the Naive Bayes algorithm.

1. Accuracy (0.258):

- The accuracy is approximately 25.8%, indicating that the Naive Bayes model correctly predicted the states only about 25.8% of the time.

2. Precision (0.302):

- Precision is around 30.2%, suggesting that when the Naive Bayes model predicts a particular state, it is correct about 30.2% of the time.

3. Recall (0.258):

- The recall is approximately 25.8%, indicating that the Naive Bayes model captures about 25.8% of the instances of the states.

4. F1 Score (0.249):

- The F1 score is around 24.9%, providing a balanced measure between precision and recall. It suggests a relatively low balance between avoiding false positives and false negatives.

* **Custom Neural Network**

The model's performance on the testing dataset is consistently high, with an average testing accuracy of approximately 92.2% and the latest testing accuracy slightly improved to 93.3%. This suggests that the model is effective in making accurate predictions on unseen data.

Similarly, both average and latest testing precision scores are notably high, with average precision around 92.4% and the latest precision at 93.3%. This indicates that when the model predicts a particular class, it is correct about 92.4% (average) to 93.3% (latest) of the time.

The recall scores for both average (92.2%) and latest (93.3%) testing show that the model is successful in capturing a substantial portion of the instances of the classes.

The F1 scores, which balance precision and recall, are also consistently high, with an average of 92.2% and the latest at 93.3%.

For our custom neural network, we can see the highest scores of any model trained so far. This high accuracy shows that most of its predictions are correct, signifying its overall performance. The high precision and recall show that the model is great at minimizing false positives while being good at predicting actually positive instances of each class. The high F1 score also confirms that the model has a good balance of its precision and recall scores. The training scores being very close to actual test scores also signify that the model has not been overfitted to the training data even with the large number of epochs.

* 1. **limitations and implications**

In the forthcoming section, we delve into an in-depth exploration of the respective limitations and implications associated with each approach.

* **Logistic Regression**:

1. Data Imbalance:

- The metrics are influenced by an unequal distribution of classes, potentially leading to bias and reduced accuracy for minority classes.

2. Generalization Challenges:

- The model's ability to generalize to unseen data, crucial for real-world applications, is not guaranteed.

* **SVM one-versus-one:**

1. Dependence on Data Characteristics:

- The observed superior performance of SVM with a one-vs-one strategy is contingent on the specific characteristics of the dataset used. Its effectiveness may vary with different data distributions.

2. Sensitivity to Hyperparameters:

- Our reported results are based on the specific hyperparameters used during training. The model's performance can change based on different HP.

3. Computational Complexity:

- SVM, especially with one-versus-one classification, tends to have higher computational requirements compared to some other algorithms. This could limit its scalability for large datasets or real-time applications.

* **SVM one-versus-all:**

1. Generalization to Outliers:

- While SVMs are known for their resilience to outliers, the assumption of outlier resistance might not hold universally. The model's performance in the presence of extreme outliers can be lower.

2. Dependency on Balanced Dataset:

- The inference of a well-balanced dataset contributing to the model's success implies potential challenges if the dataset distribution changes. The model's reliance on balanced class representation can limit its effectiveness in imbalanced scenarios.

**Implications**:

1. Robust Outlier Handling:

- The SVM's robust performance suggests its capability to handle outliers effectively. This makes it a favorable choice when dealing with datasets that may contain noise or anomalies.

2. Strategy Insensitivity:

- The consistent performance across one-vs-all and one-vs-one strategies implies that, in this specific case, the choice of strategy doesn't significantly impact the SVM Classifier's effectiveness.

* **Naive Bayes Classifier**

1. Sensitivity to Outliers for Extreme Values:

- Naive Bayes can be sensitive to outliers for extreme values, leading to inaccurate probability estimations.

2. Difficulty Handling Non-Normal Distributions:

- Naive Bayes assumes that features follow a Gaussian (normal) distribution. If the data distribution deviates from this assumption, the model's ability to accurately capture the underlying patterns may be compromised, as evidenced by the lower accuracy, precision, recall, and F1 score.

3. Limited Capability for Capturing Complex Relationships:

- The Naive Bayes model, being a simple probabilistic model, struggles to capture complex relationships between features

4. Sensitive to Feature Scaling:

- Naive Bayes assumes that features are conditionally independent given the class, but it doesn't account for variations in feature scales. Sensitivity to feature scaling could result in certain features dominating the model's decision-making process, especially if their scales differ significantly.

* **Custom Neural Network**

1. Computational Intensity:

- Custom neural networks, especially sophisticated architectures, can be computationally intensive during training and inference for high number of epochs. This can limit their applications requiring real-time predictions when the resource is limited.

2. Interpretability Challenges:

- The inherent complexity of neural networks can result in a lack of interpretability, making it challenging to understand the underlying decision-making process. This limitation could be critical in applications where interpretability is a primary concern. As discussed during our class, higher the performance, lower the interpretability.

3. Data Dependency:

- The exceptional performance observed may be highly dependent on the characteristics of the training data. The model might struggle when faced with novel patterns or datasets with different distributions.