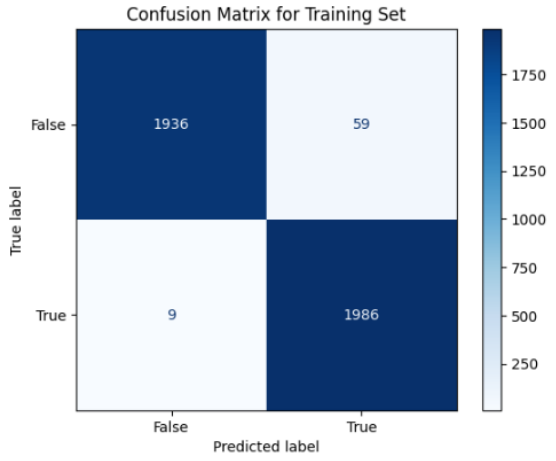
Jivko Jeff Uzoff

Naïve Bayes – Predictice Modeling Comparison

Naïve Bayes – All attributes:

The Naive Bayes algorithm was applied to the training and test datasets using all available attributes. The performance of the model was evaluated based on key metrics such as accuracy, sensitivity (recall), specificity, precision, and the F-measure. These metrics provide insights into the model's effectiveness in classifying instances correctly.

Training Set Results:



Confusion Matrix Output:

True Negatives (TN): 1,936

- The model correctly predicted 'False' for 1,936 instances.

False Positives (FP): 59

- The model incorrectly predicted 'True' for 59 instances where the actual response was 'False'.

False Negatives (FN): 9

- The model incorrectly predicted 'False' for 9 instances where the actual response was 'True'.

True Positives (TP): 1,986

- The model correctly predicted 'True' for 1,986 instances.

Performance Metrics:

Accuracy: 98.30%

- The model was able to accurately predict whether a given instance belonged to the 'True' or 'False' class 98.54% of the time.
- (1,936 + 1,986) / 3,990 = 0.9830

Jivko Jeff Uzoff

Sensitivity (Recall/TPR): 99.55%

- Shows that 99.55% of actual 'True' instances were correctly identified by the model.
- 1,986 / (9 + 1,986) = 0.9955

Specificity (TNR): 97.04%

- Indicates that 97.04% of actual 'False' instances were correctly identified by the model.
- 1,936 / (1,936 + 59) = 0.9704

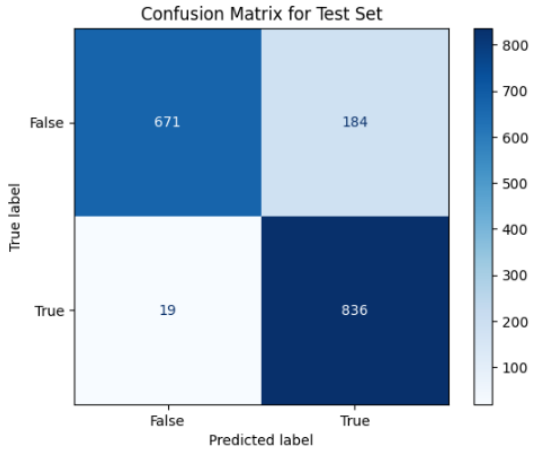
Precision: 97.11%

- Reflects that 97.12% of the instances predicted as 'True' were actually 'True'.
- 1,986 / (59 + 1,986) = 0.9711

F-Measure: 98.31%

- Represents the harmonic mean of precision and recall, indicating a balanced performance.
- 2 * ((0.9712 * 0.9955) / (0.9712 + 0.9955)) = 0.9831

Test Set:



Confusion Matrix Output:

True Negatives (TN): 671

- The model correctly predicted 'False' for 671 instances.

False Positives (FP): 184

- The model incorrectly predicted 'True' for 184 instances where the actual response was 'False'.

False Negatives (FN): 19

- The model incorrectly predicted 'False' for 19 instances where the actual response was 'True'.

True Positives (TP): 836

- The model correctly predicted 'True' for 836 instances.

Performance Metrics:

Accuracy: 88.13%

- The model was able to accurately predict whether a given instance belonged to the 'True' or 'False' class 78.50% of the time.
- (671 + 836) / (671 + 184 + 19 + 836) = 0.8813

Sensitivity (Recall/TPR): 97.77%

- Shows that 97.77% of actual 'True' instances were correctly identified by the model.
- 836 / (19 + 836) = 0.9777

Specificity (TNR): 78.48%

- Indicates that 78.44% of actual 'False' instances were correctly identified by the model.
- 671 / (671 + 184) = 0.7848

Precision: 81.96%

- Reflects that 81.81% of the instances predicted as 'True' were actually 'True'.
- 836 / (184 + 836) = 0.8196

F-Measure: 89.17%

- Represents the harmonic mean of precision and recall, indicating a balanced performance.
- 2 * ((0.8181 * 0.9777) / (0.8181 + 0.9777)) = 0.8917

Overall Takeaway from Naïve Bayes Model (All Attributes)

The Naive Bayes model trained using all attributes demonstrates strong performance on the training set with an accuracy of 98.30% and maintains high sensitivity and specificity. However, on the test set, while the accuracy is still respectable at 88.13%, there is a noticeable drop in both specificity and precision, suggesting potential overfitting. The model is particularly adept at identifying positive instances ('True') with very high sensitivity, but it shows decreased effectiveness in correctly classifying negative instances ('False') on the test set. This drop in specificity and precision indicates that the model may be less robust when applied to new, unseen data.

Comparison between Training and Test Sets

- **Accuracy:** The model performs significantly better on the training set (98.30%) compared to the test set (88.13%), which is indicative of potential overfitting. The model achieves near-perfect accuracy on the training set but is less accurate on new data.
- **Sensitivity:** The model maintains high sensitivity on both the training set (99.55%) and the test set (97.77%), showing consistent performance in identifying positive instances.
- **Specificity:** Specificity drops from 97.04% on the training set to 78.48% on the test set, suggesting a decrease in the model's ability to correctly identify negative instances when applied to new data.
- **Precision:** Precision decreases from 97.11% on the training set to 81.96% on the test set, indicating an increase in false positives on new data, which further supports the observation of overfitting.

Jivko Jeff Uzoff

Naïve Bayes – Selected attributes:

The Naive Bayes algorithm was applied to the training and test datasets using a selected subset of attributes. The model's performance was evaluated based on key metrics, including accuracy, sensitivity (recall), specificity, precision, and the F-measure. These metrics help assess the model's effectiveness in classifying instances correctly.

Strategy for Selecting Attributes for the Selected Attributes Model

Attribute 1	Attribute 2	Correlated	P-Value (Signficance Level of 0.05)
Vmail_Message	State	No	0.9946
Intl_Calls	Intl_Mins	No	0.986
State	Vmail_Plan	No	0.9819
Vmail_Message	Day_Charge	No	0.9643
Vmail_Message	Day_Mins	No	0.9642
Intl_Charge	Vmail_Plan	No	0.9413
Intl_Mins	Vmail_Plan	No	0.9394
Eve_Mins	Area_Code	No	0.9334
Eve_Charge	Area_Code	No	0.933
Day_Mins	Vmail_Plan	No	0.9226
Day_Charge	Vmail_Plan	No	0.9225
Area_Code	CLASS ATTRIBUTE - Churn - Whether client has churned or not	No	0.9151
Intl_Calls	State	No	0.9147
Account_Length	State	No	0.9131
Account_Length	CLASS ATTRIBUTE - Churn - Whether client has churned or not	No	0.3398

Filter Attributes by P-Value (Highest > Lowest) and showcase p-values that are 0.9 and above:

- The attributes were filtered based on their p-values in descending order to prioritize removing attributes with the least predictive power. Attributes with higher p-values indicate weaker associations with the target variable and with other attributes.

1) Remove first attribute with highest p-value with class attribute (Churn):

- The attribute "Area_Code" was the first to be excluded from the model, as it had the highest P-value with the class attribute (Churn). This high P-value suggests that "Area_Code" is not significantly correlated with the class attribute 'Churn'. Therefore, its removal is justified to avoid introducing irrelevant information into the model.

2) Remove any attributes with p-values from 0.9 and above that present no correlation with class:

- Next, attributes with p-values at or above 0.9 were evaluated to see their correlation with the class attribute ‘Churn’. "Account_Length" was excluded from the model, as it was present amonst attributes with p-values above 0.9 as well as having no correlation with the class attribute ‘Churn’.

3) Remove attribute that appears most frequently with p-values of 0.9 and above:

- Finally, "VMail_Plan" was removed from the model. This attribute appeared most frequently among the attributes with P-values of 0.9 and above, suggesting it consistently failed to show a significant relationship with the target variable and other attributes. Removing this attribute helped further reduce noise and enhance the model's performance.

From this strategy, the following attributes were selected to go into the selected attributes Naïve Bayes model:

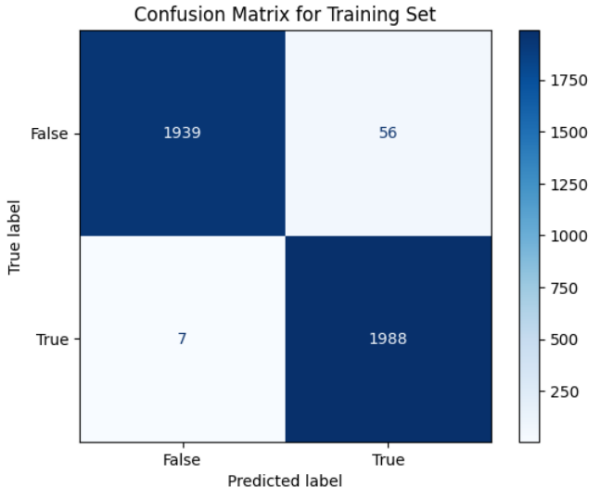
- State

Jivko Jeff Uzoff

- Phone
- Intl_Plan
- Vmail_Message
- Day_Mins
- Day_Calls
- Day_Charge
- Eve_Mins
- Eve_Calls
- Eve_Charge
- Night_Mins
- Night_Calls
- Night_Charge
- Intl_Mins
- Intl_Calls
- Intl_Charge
- CustServ_Calls
- Churn

Jivko Jeff Uzoff

Training Set Results:



Confusion Matrix Output:

True Negatives (TN): 1,939

- The model correctly predicted 'no' for 1,939 instances.

False Positives (FP): 56

- The model incorrectly predicted 'yes' for 56 instances where the actual response was 'no'.

False Negatives (FN): 7

- The model incorrectly predicted 'no' for 7 instances where the actual response was 'yes'.

True Positives (TP): 1,988

- The model correctly predicted 'yes' for 1,988 instances.

Performance Metrics:

Accuracy: 98.42%

- The model accurately predicted whether a given instance belonged to the 'yes' or 'no' class 98.48% of the time.
- (1939 + 1988) / 3990 = 0.9842

Sensitivity (Recall/TPR): 99.65%

- The model correctly identified 99.65% of actual 'yes' instances.
- 1988 / 1995 = 0.9965

Specificity (TNR): 97.19%

- The model correctly identified 97.20% of actual 'no' instances.
- 1939 / 1995 = 0.9719

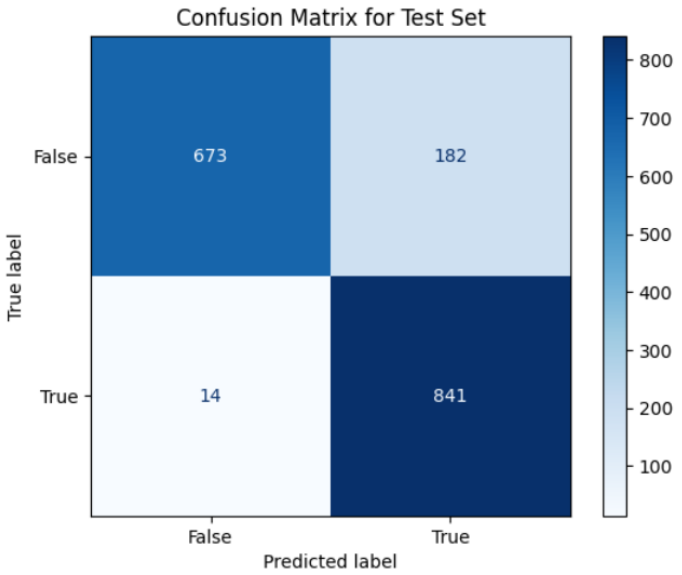
Precision: 97.26%

- The model correctly predicted 'yes' for 97.26% of the instances.
- 1988 / 2044 = 0.9726

F-Measure: 98.44%

- The harmonic mean of precision and recall, indicating a balanced performance.
- 2 × (0.9726 × 0.9965) / (0.9726 + 0.9965) = 0.9844

Test Set:



Confusion Matrix Output:

True Negatives (TN): 673

- The model correctly predicted 'no' for 673 instances.

False Positives (FP): 182

- The model incorrectly predicted 'yes' for 182 instances where the actual response was 'no'.

False Negatives (FN): 14

- The model incorrectly predicted 'no' for 14 instances where the actual response was 'yes'.

True Positives (TP): 841

- The model correctly predicted 'yes' for 841 instances.

Performance Metrics:

Accuracy: 88.54%

- The model accurately predicted whether a given instance belonged to the 'yes' or 'no' class 88.02% of the time.
- (673 + 841) / 1710 = 0.8854

Sensitivity (Recall/TPR): 98.36%

- The model correctly identified 98.36% of actual 'yes' instances.
- 841 / 855 = 0.9836

Specificity (TNR): 78.71%

- The model correctly identified 78.71% of actual 'no' instances.
- 673 / 855 = 0.7871

Precision: 82.21%

- The model correctly predicted 'yes' for 82.20% of the instances.
- 841 / 1023 = 0.8221

F-Measure: 89.56%

- The harmonic mean of precision and recall, indicating a balanced performance.
- 2 × (0.8220 × 0.9836) / (0.8220 + 0.9836) = 0.8956

Overall Takeaway from Naive Bayes Model (Selected Attributes):

The Naive Bayes model, trained using selected attributes, shows strong performance on both the training and test sets. With an accuracy of 98.42% on the training set and 88.54% on the test set, the model successfully classifies most instances correctly. The model exhibits high sensitivity (recall) across both sets, proving particularly effective at identifying positive instances ('yes'). However, there is a notable decrease in specificity and precision from the training to the test set, indicating a higher occurrence of false positives when applied to new data.

Comparison between Training and Test Sets:

- **Accuracy**: The model performs better on the training set (98.42%) compared to the test set (88.54%). This gap suggests that the model may be overfitting slightly to the training data but still generalizes reasonably well to new data.
- **Sensitivity**: The model maintains very high sensitivity across both sets, indicating consistent performance in identifying positive instances.

- **Specificity**: The specificity drops significantly from 97.19% (training) to 78.71% (test), highlighting a reduction in the model's ability to correctly identify negative instances in the test set.
- **Precision**: Precision decreases from 97.26% on the training set to 82.21% on the test set, suggesting a higher occurrence of false positives when applied to new data.
- **F-Measure**: The F-Measure also drops from 98.44% on the training set to 89.56% on the test set, indicating a decrease in overall model balance when applied to unseen data.

Jivko Jeff Uzoff

Naïve Bayes Model Comparison – All attributes vs Selected attributes:

This section compares the performance of Naïve Bayes models using all attributes versus a selected subset of attributes. The comparison is based on confusion matrix outputs and performance metrics, visualized in Figure 1 and Figure 2 (in appendix).

The improvements observed with the selected attributes model not only highlight its effectiveness in terms of prediction accuracy and reliability but also suggest that careful feature selection can lead to a more efficient model.

Confusion Matrix Improvements:

Training Set:

- True Negatives (TN): Improved by 3 (from 1,936 to 1,939), showing better accuracy in identifying 'no' instances.
- False Positives (FP): Decreased by 3 (from 59 to 56), reflecting fewer incorrect 'yes' predictions.
- False Negatives (FN): Decreased by 2 (from 9 to 7), indicating fewer missed 'yes' instances.
- True Positives (TP): Increased by 2 (from 1,986 to 1,988), demonstrating improved identification of 'yes' instances.

Test Set:

- True Negatives (TN): Improved by 2 (from 671 to 673), resulting in better identification of 'no' instances.
- False Positives (FP): Decreased by 2 (from 184 to 182), leading to fewer incorrect 'yes' predictions.
- False Negatives (FN): Decreased by 5 (from 19 to 14), showing fewer missed 'yes' instances.
- True Positives (TP): Increased by 5 (from 836 to 841), reflecting improved correct identification of 'yes' instances.

Performance Metrics Improvements:

Training Set:

- Accuracy: Improved by 0.12 percentage points (from 98.30% to 98.42%), indicating a more accurate model overall.
- Sensitivity (Recall): Increased by 0.10 percentage points (from 99.55% to 99.65%), demonstrating enhanced ability to identify true 'yes' instances.
- Specificity: Improved by 0.15 percentage points (from 97.04% to 97.19%), reflecting better performance in identifying true 'no' instances.
- Precision: Increased by 0.15 percentage points (from 97.11% to 97.26%), indicating more reliable predictions of 'yes'.

- F-Measure: Rose by 0.13 percentage points (from 98.31% to 98.44%), showcasing a better balance between precision and recall.

Test Set:

- Accuracy: Improved by 0.41 percentage points (from 88.13% to 88.54%), indicating better overall performance on new data.
- Sensitivity (Recall): Increased by 0.59 percentage points (from 97.77% to 98.36%), highlighting improved detection of 'yes' instances.
- Specificity: Improved by 0.23 percentage points (from 78.48% to 78.71%), showing better identification of 'no' instances.
- Precision: Increased by 0.25 percentage points (from 81.96% to 82.21%), meaning predictions of 'yes' are more accurate.
- F-Measure: Rose by 0.39 percentage points (from 89.17% to 89.56%), reflecting an improved balance between precision and recall.

Overall Takeaway from All vs. Selected Attribute Model Comparison:

The selected attributes model demonstrates incremental improvements over the all-attributes model across both confusion matrix components and performance metrics. The model using selected attributes shows enhanced classification accuracy, sensitivity, specificity, precision, and F-measure for both training and test sets. Although the performance gains are subtle, they indicate that focusing on the most relevant features has slightly improved the model's ability to classify instances more accurately and effectively, reducing misclassifications and improving overall performance.

By refining the feature set, the selected attributes model maintains high accuracy while slightly boosting performance in identifying the target class ('yes'). The increase in sensitivity underscores better detection of instances of interest ("yes"), which is critical for the model's effectiveness. Improved specificity and precision suggest that the model is also more adept at correctly predicting 'no' and avoiding false positives.

In summary, targeted feature selection has positively impacted the model's capacity to predict the target class with greater accuracy and reliability, while maintaining a balanced performance between precision and recall.

Appendix

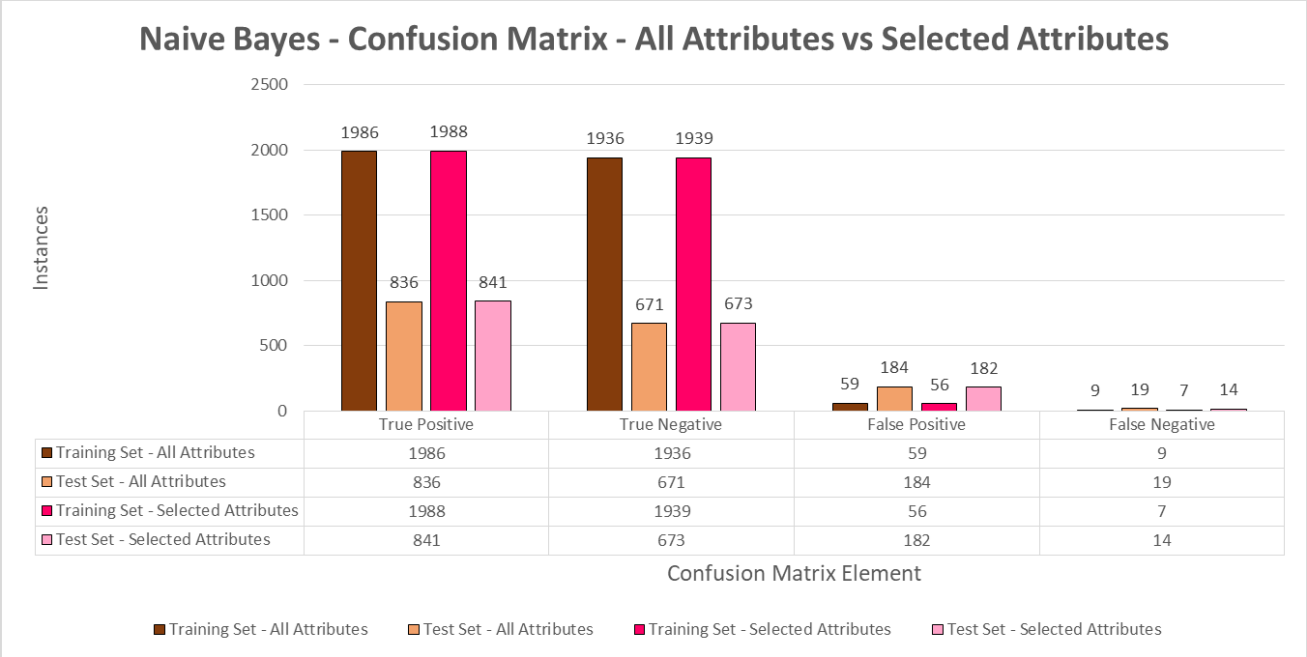


Figure 1: Confusion Matrix – All Attributes vs Selected Attributes

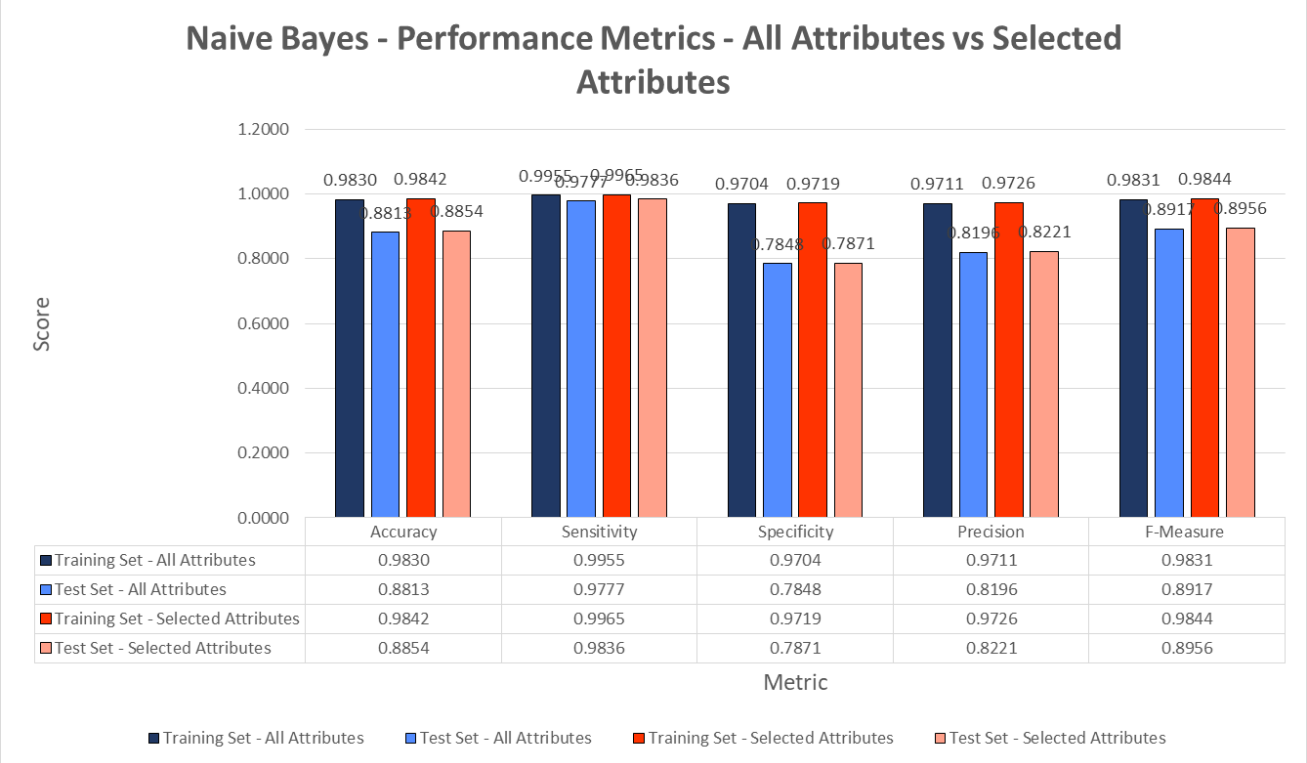


Figure 2: Performance Metrics – All Attributes vs Selected Attributes