JAZA TELECOM COMPANY CUSTOMER CHURN PREDICTION MODEL

NAME: PETER KINYANJUI

INSTITUTION: MORINGA SCHOOL

INSTRUCTOR: FAITH ROTICH

QUOTE:
GOOD CUSTOMER SERVICE COSTS LESS THAN CUSTOMER CHURN

- ANONYMOUS

OVERVIEW

1. **Customer Centrality:**

-Success in telecom hinges on customer satisfaction and loyalty.

2. **Churn Signals:**

-If customers switch due to fees, better offers, or poor service, it reflects telecom performance.

3. **Performance Indicator:**

-Customer churn indirectly gauges telecom effectiveness and competitiveness.

4. **Critical Metric:**

-Monitoring churn is vital for assessing telecom companies' market standing and operational efficiency.

BUSINESS PROBLEM

1. **Customer Retention Focus:**

   - Jaza Telecom historically prioritized getting new customers over keeping existing ones.

2. **Cost Dynamics:**

   - Acquiring new customers can cost up to five times more than retaining current ones.

3. **Profit Potential:**

   -Research shows a 5% increase in retention can boost profits by 25% to 95%.

4. **Churn Impact:**

   -Failure to identify and address customer churn signals can harm the company's long-term performance.

5. **Data Insights:**

   - Customer churn data offers valuable insights and fuels predictive models for proactive actions.

DATA UNDERSTANDING

1. Source of Information:

- Data comes from https://www.kaggle.com/datasets/becksddf/churn-in-telecoms-dataset , providing details on customer interactions.

2. Dataset Overview:

- Contains 3333 entries and 21 columns with information like account length, usage metrics, and customer service calls.

3. Data Types:

- Various data types include object (e.g., state, phone number), int64, float64, and bool (for churn).
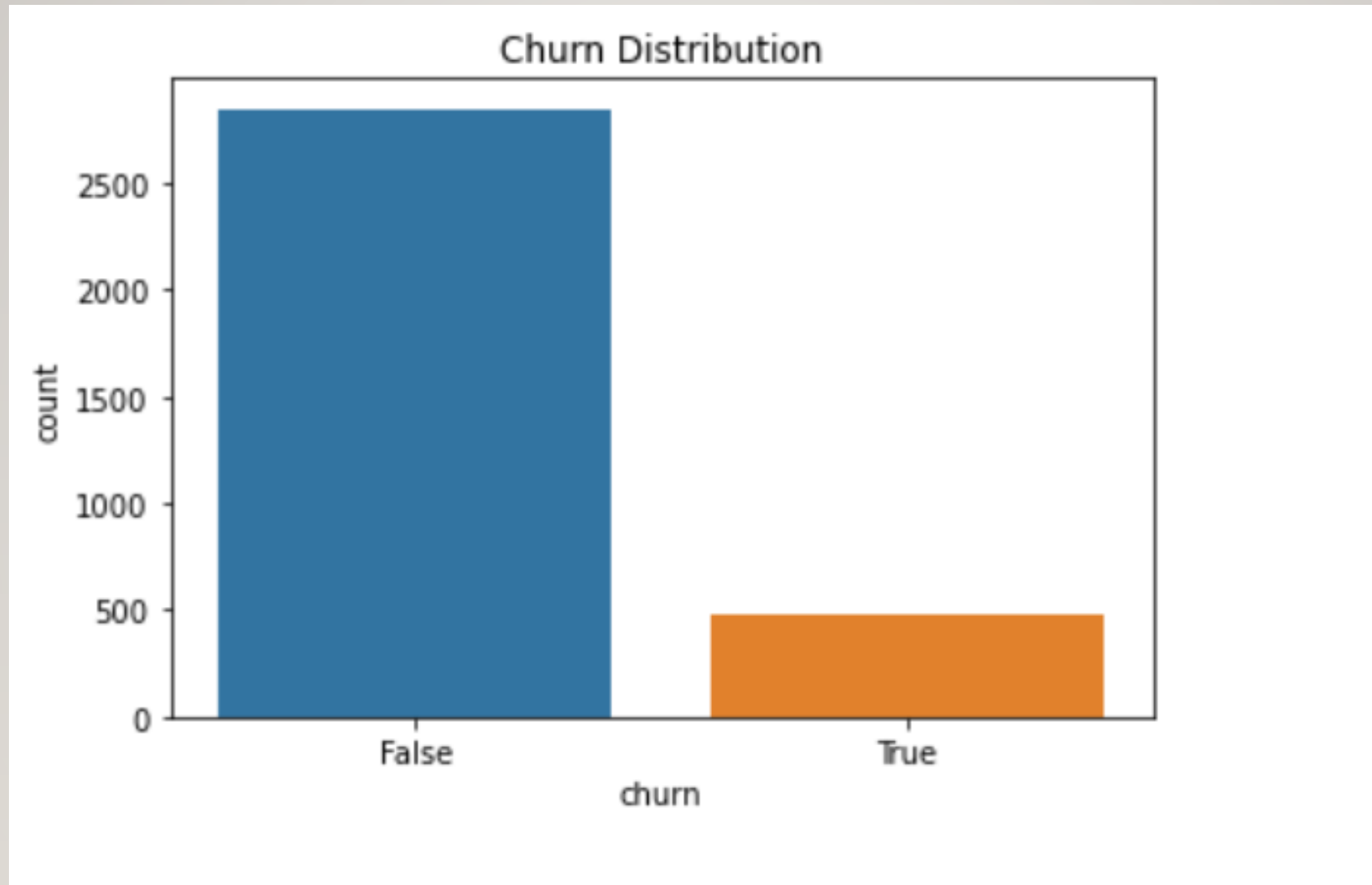
CONTINUATION…
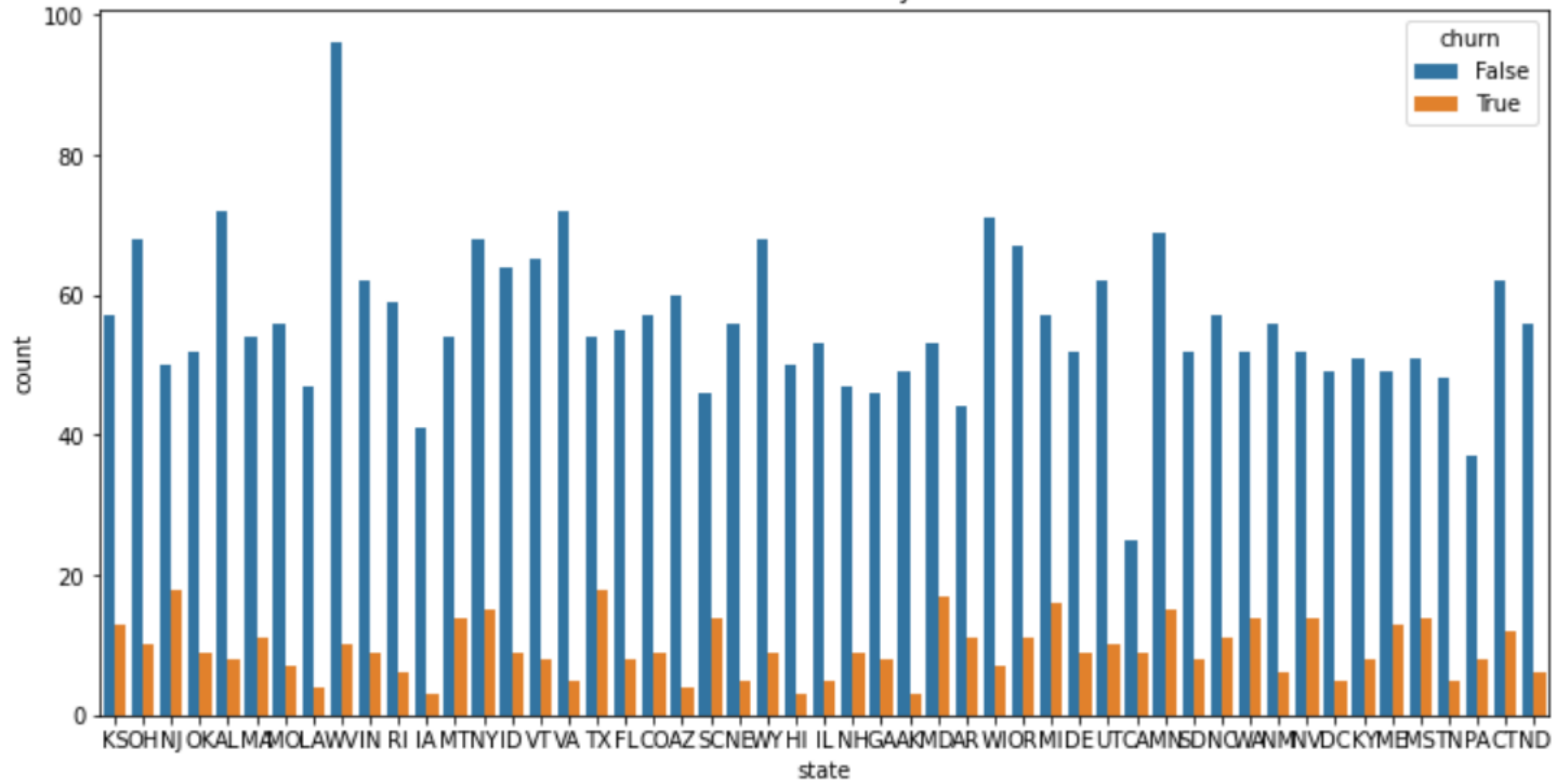
4. **Data Quality:**

    - No missing values in the dataset.

    - Categorical variables may need encoding for model training.

5. **Target Variable:**

    - 'Churn' column, a boolean, indicates if a customer has discontinued services.

    - A rich dataset offering insights into customer behavior, serving as a foundation for predicting and addressing churn.

state Distribution by Churn

MODELING

1. Simple model - Logistic regression

2. Complex Model- Random forest

3. Hyperparameter tuned for Random forest model
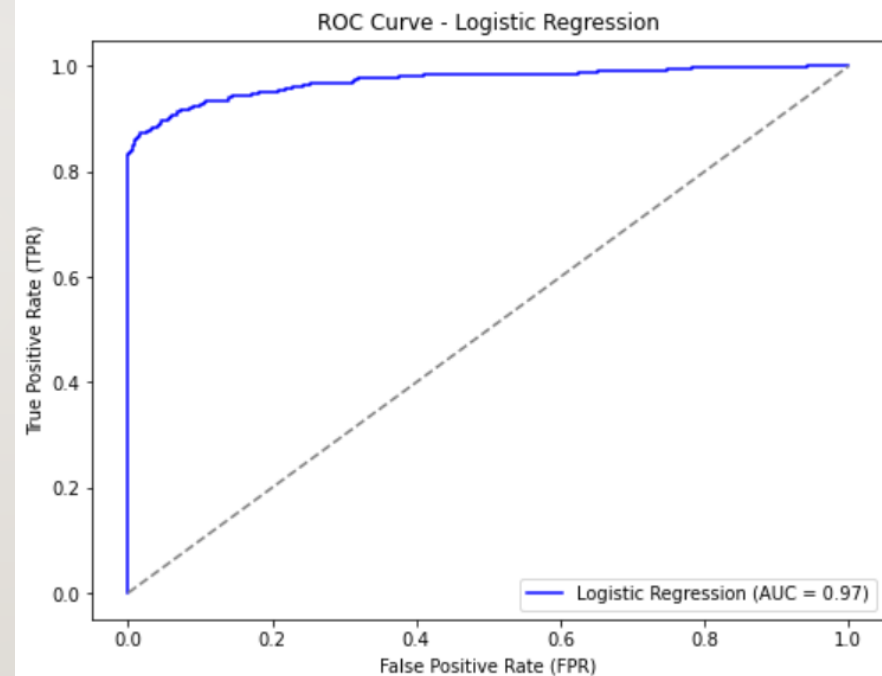
EVALUATION- LINEAR REGRESSION MODEL

```
Logistic Regression Accuracy: 0.92
Confusion Matrix:
[[560  25]
 [ 63 492]]
Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.96      0.93       585
           1       0.95      0.89      0.92       555

    accuracy                           0.92      1140
   macro avg       0.93      0.92      0.92      1140
weighted avg       0.92      0.92      0.92      1140
```


ROC Curve - Logistic Regression

1.Accuracy:
Achieved a commendable 92% accuracy in predicting customer churn.

2.Confusion Matrix:
Balanced performance with 25 false positives and 63 false negatives.

3.Precision-Recall:
Precision of 95% for non-churn and 90% for churn.
Balanced recall of 89% for churn and 96% for non-churn.
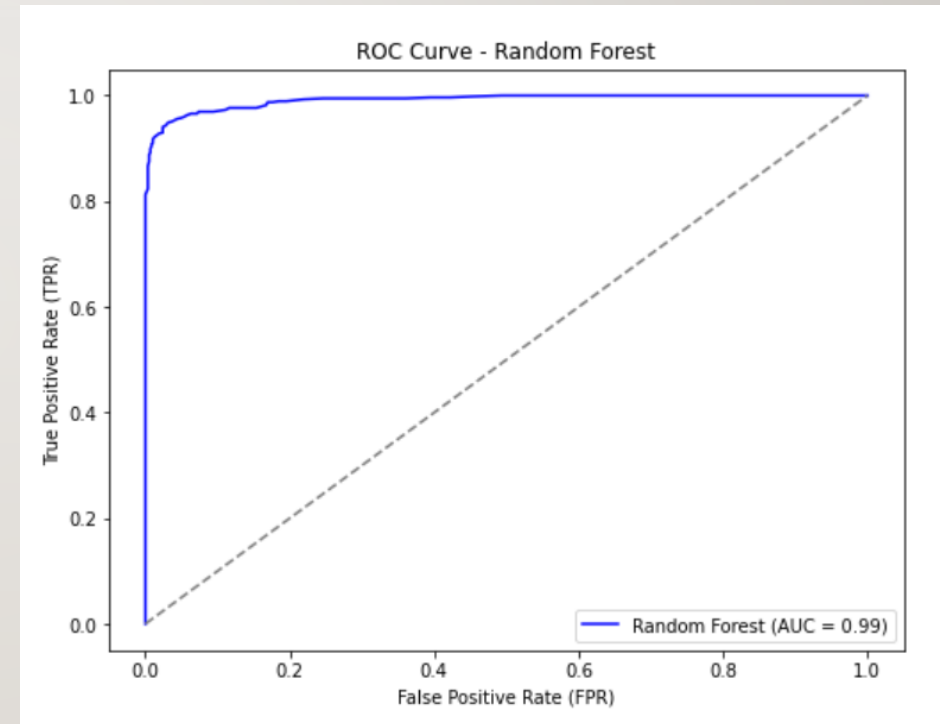
EVALUATION- RANDOM FOREST MODEL

1. **Accuracy Boost:**
   Demonstrated superior performance with a higher accuracy of 96%.
2. **Confusion Matrix Improvement:**
   Significantly reduced false positives (20) and false negatives (28).
3. **Precision-Recall Enhancement:**
   Maintained high precision and recall for both churn and non-churn classes.

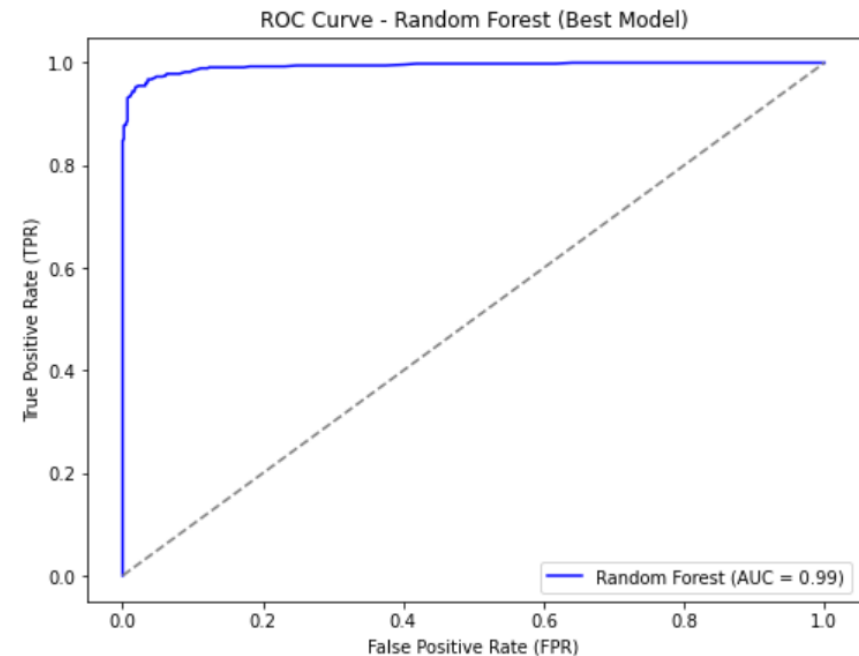# EVALUATION- HYPERPARAMETER TUNED FOR RANDOM FOREST MODEL

1.Optimized Accuracy:
Maintained high accuracy of 96%, matching the untuned model.

2.Further Confusion Matrix Refinement:
Continued reduction in false positives (19) and false negatives (22).

3.Precision-Recall Consistency:
Precision and recall metrics remained consistently high.

4.Visual Validation:
Confusion matrix heatmap and ROC curve illustrated the continued effectiveness of the tuned model.

RECOMMENDATION

- By incorporating these recommendations, Jaza Telecom can effectively integrate predictive modeling into its operational and strategic decision-making processes, ultimately improving customer retention and overall business performance.

1. **Deployment Strategy:**

   Consider deploying the hyperparameter-tuned Random Forest model for its superior accuracy and optimized performance.

2. **Continuous Monitoring:**

   Implement a system for ongoing model monitoring to ensure sustained effectiveness in predicting customer churn.

CONTINUATION…

3. **Strategic Insights:**

Leverage insights from model evaluations to inform strategic decisions related to customer retention and satisfaction.

4. Collaboration with Teams:

Collaborate with marketing and customer service teams to implement targeted strategies based on model predictions.

5. Training and Communication:

Conduct training sessions to familiarize relevant teams with the models, their predictions, and the associated business implications.

NEXT STEP

1. Model Deployment:
Roll out the selected model (hyperparameter-tuned Random Forest) into our production environment for real-time predictions.

2. Integration with Systems:
Integrate the model seamlessly with existing operational systems to ensure a smooth flow of information and decision-making.

3. User Interface Development:
Develop a user-friendly interface for teams to interact with the model predictions, making insights easily accessible and actionable.

4. Automation of Actions:
Implement automated processes triggered by model predictions to streamline and optimize customer retention strategies.

5. KPI Monitoring:
Set up key performance indicators (KPIs) to monitor the effectiveness of the model-driven strategies and regularly evaluate against predefined benchmarks.