SYRIATEL CUSTOMER CHURN PREDICTION

BERYL AGAI

BUSINESS UNDERSTANDING

- Syriatel is a giant Telco but experiencing a recent spike in customer churn.
- The stakeholders want to know why and how to mitigate this issue to ensure positive business outcome.
- We will build a predictive model to analyze the trends and predict churn rate

POSSIBLE CHALLENGES LEADING TO CHURN

1. **Product quality:**

- Slow internet speeds
- Poor network coverage
- system downtimes

2. Price implications:

3. Customer service experience:

- Long wait times on the IVR
- High abandonment rate

STAKEHOLDER DEFINITION

- The stakeholders vested in this project are:

1. Syriatel executive leadership
2. Customer retention team
3. Customer sales & marketing department

MAIN OBJECTIVE

To identify the factors driving customer churn at Syriatel and build a predictive model with at least 80% precision to enable the implementation of targeted strategies for retention and improved business outcomes.

.

DATA UNDERSTANDING

Our dataset is sourced from Kaggle

It has 3333 labels and 21 features.

It does not have any missing values

The dataset has no duplicated values.

There are 17 continuos features and 4 categorical features

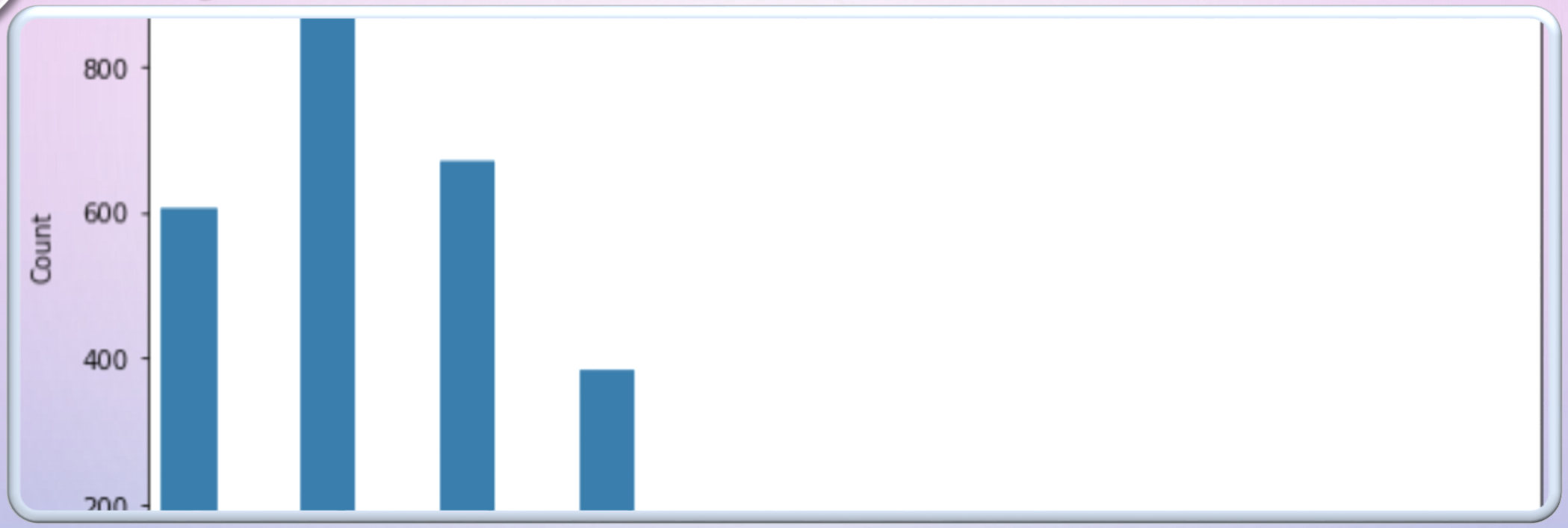
DATA PREPARATION

- We dropped the phone number feature due to ethical reasons.
- One hot encoded the categorical variables.
- Removed outliers.
- Dropped highly correlated variables.

EXPLORATORY DATA ANALYSIS (EDA)

- Most churners are from area code 1&2
- Customers that make more customer service calls are the highest churners.
- Customers with an international plan churn more than those without.
- Churners have higher day-time usage and charges.
- Churn customers mostly have zero voice mail messages, while their counter parts have more.

Effect of area code on churn



Effect of customer sevice calls on churn

MODELING

We built five models and improved each of them to see which was the best performance based on

- Accuracy
- Precision
- Recall
- F1 score
- Roc
- Auc

5) LOGISTIC REGRESSION CLASSIFIER

Performance

- Accuracy: 0.760
- Precision: 0.96
- Recall: 0.76
- F1-score: 0.85
- AUC score: 0.8097
- ROC score: 0.8097
- **Observations**: logistic regression shows the lowest accuracy and f1-score but performs relatively well in AUC and ROC score. It may benefit from more complex models for improved performance.

Logistic Regression feature importance

4)K-NEAREST NEIGHBORS (KNN)

Performance

- Accuracy: 0.893
- Precision: 0.91
- Recall: 0.98
- F1-score: 0.94
- AUC score: 0.765
- ROC score: 0.7650
- **Observations**: KNN demonstrates moderate accuracy but struggles with identifying churn instances accurately, leading to a lower f1-score. It requires further optimization for better performance.

3)DECISION TREE CLASSIFIER

Performance

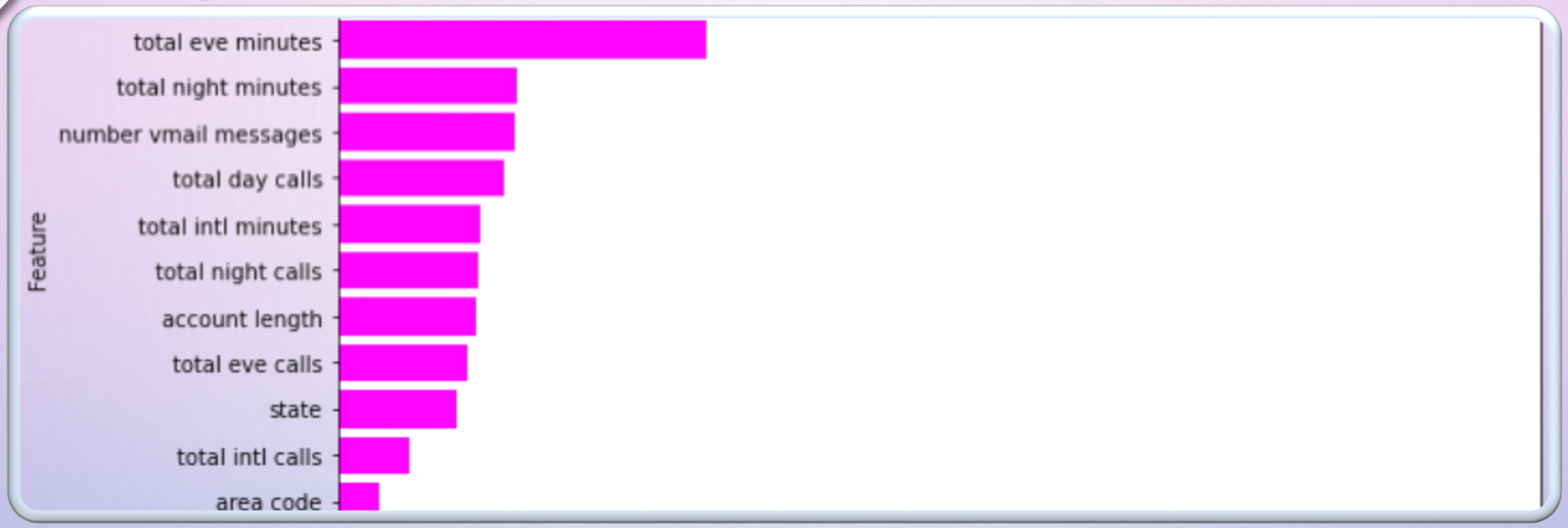
- Accuracy: 0.950
- Precision: 0.95
- Recall: 1.00
- F1-score: 0.97
- AUC score: 0.8130
- ROC score: 0.7773
- **Observations**: the decision tree achieves competitive results similar to random forest but falls marginally behind in AUC and ROC score due to lower recall for churn, indicating slightly less robust performance.

Decision Tree feature importance

2) RANDOM FOREST CLASSIFIER

Performance

- Accuracy: 0.951
- Precision: 0.95
- Recall: 1.00
- F1-score: 0.97
- AUC score: 0.8876
- ROC score: 0.8876
- **Observations**: random forest exhibits robust performance, especially in identifying non-churn instances. However, it slightly lags in recall for churn instances, impacting its overall accuracy.



Random forest feature importance

XGBOOST

- **Performance**
- Accuracy: 0.958
- Precision: 0.96
- Recall: 0.99
- F1-score: 0.98
- AUC score: 0.8966
- ROC score: 0.8966
- Observations: xgboost outperforms other models with the highest accuracy, f1-score, AUC, and ROC score. It effectively predicts non-churn instances and maintains a balance between precision and recall for churn.

Feature Importance as per XGBoost

OBSERVATIONS AND RECOMMENDATIONS

CUSTOMER SERVICE CALLS

Observations

The calls made to customer service affects churn the most.

Recommendations

- Enhance customer service training for the representatives.
- Reduce response times.
- Implement proactive customer satisfaction surveys and follow-up calls to address potential issues before they lead to churn.

TOTAL DAY MINUTES

Observation

- Daytime minutes offered affects churn as the second feature.
- **Recommendations**
- Introduce flexible and affordable day minute plans to cater to heavy daytime users.
- Offer incentives or discounts for plans with higher daytime minutes to retain customers.

NUMBER OF VOICE MAIL MESSAGES

Observation

- This feature contributes notably to the model's predictions and likely represents an important factor

Recommendations

- Promote the use of voicemail services by highlighting their benefits.
- Offer bundled packages with enhanced voicemail features to add value to customer subscriptions.

TOTAL EVENING MINUTES

Observation

While less influential than the top three features, it still holds considerable importance

Recommendations

- Develop evening-specific plans or promotions to attract customers who use services predominantly during these hours.
- Monitor evening usage patterns to tailor future offerings better.

TOTAL NIGHT MINUTES AND EVENING CALLS

Observation

This feature has a moderate impact

Recommendations

- Introduce plans that offer competitive rates for night usage.
- Engage night-time users through targeted marketing campaigns and personalized offers.
- Offer loyalty rewards for consistent usage.
-

TOTAL DAY CALLS, TOTAL INTERNATIONAL CALLS, ACCOUNT LENGTH, TOTAL NIGHT CALLS, TOTAL INTERNATIONAL MINUTES

Observation

These features have no impact on the churn rate.

Recommendations

- Ensure these features are optimized for efficiency and customer satisfaction.
- Periodically review these features to ensure they meet customer needs without significantly impacting operational costs.
-

MITIGATION EXPECTATIONS

With the implementation of the insights listed,we expect:

1. Increased customer retention
2. Increased NPS and customer obsession
3. Increased brand love.
4. Increased revenue
5. Increased market dominance.

THANK YOU!

ANY QUESTIONS?

Name:Beryl Agai

Email: agai.beryl@gmail.com

LinkedIn: Beryl Agai