Targeted Marketing Model: Feature Engineering Report
- FNU Andria Grace

Introduction

The objective of this report is to detail the feature engineering process undertaken to enhance a targeted marketing dataset. The dataset comprises subscriber information, campaign details, and response indicators. The goal is to create new features that capture meaningful patterns and relationships to predict subscriber responses to marketing campaigns more accurately.

Feature Engineering

1. response_rate

Description: Average response rate for each subscriber.
Usefulness: This feature provides insight into individual subscriber engagement. Subscribers with higher response rates are likely more engaged with marketing efforts and may be more receptive to future campaigns.

2. campaign_success_rate

Description: Average success rate of each campaign.
Usefulness: Helps identify which campaigns have historically been more effective in eliciting responses. Campaigns with higher success rates can guide future campaign strategies.

3. response_last_month

Description: Average response rate of each subscriber over the last four weeks.
Usefulness: Captures recent engagement trends, indicating whether subscribers' responsiveness is increasing, decreasing, or stable over time.

4. total_campaigns_received

Description: Total number of campaigns each subscriber has received.
Usefulness: Indicates the level of exposure of each subscriber to marketing campaigns. Higher counts may suggest saturation or fatigue.

5. total_campaigns_responded

Description: Total number of campaigns each subscriber has responded to.
Usefulness: Reflects the overall engagement level of subscribers. Subscribers who respond to multiple campaigns are likely more receptive to marketing messages.

6. average_response_time
Description: Average time (in weeks) between responses for each subscriber.
Usefulness: Provides insights into subscriber behavior patterns, such as regularity or variability in response times. This can inform timing strategies for future campaigns.

7. state_response_rate
Description: Average response rate for each state.
Usefulness: Identifies geographic variations in subscriber engagement. Understanding regional response rates can help tailor campaigns to specific state demographics.

8. sex_response_rate
Description: Average response rate for each gender.
Usefulness: Highlights any gender-based differences in campaign effectiveness. This insight can guide gender-specific marketing strategies.

9. attribute_response_rate
Description: Average response rate for each user category.
Usefulness: Indicates which user categories are more responsive to marketing campaigns. This segmentation helps in targeting specific demographic groups effectively.

10. week_response_rate
Description: Average response rate for each week.
Usefulness: Tracks weekly trends in subscriber engagement. Understanding weekly variations can optimize campaign scheduling and resource allocation.

11. state_campaign_success_rate
Description: Average success rate of campaigns in each state.
Usefulness: Combines geographic and campaign effectiveness data to identify which campaigns perform best in different states. This can inform localized marketing strategies.

12. sex_campaign_success_rate
Description: Average success rate of campaigns for each gender.
Usefulness: Provides insights into gender-specific campaign performance. Tailoring campaigns based on gender responses can enhance overall campaign effectiveness.

13. attribute_campaign_success_rate
Description: Average success rate of campaigns for each user category.
Usefulness: Evaluates campaign performance across user categories, highlighting which demographics respond best to specific campaigns. This segmentation guides targeted marketing efforts.

14. week_campaign_success_rate
Description: Average success rate of campaigns for each week.
Usefulness: Temporal analysis of campaign success rates helps in understanding weekly variations in campaign effectiveness. This insight can optimize campaign scheduling.

15. subscriber_state_frequency
Description: Number of different states a subscriber is associated with.
Usefulness: Indicates subscriber mobility or multi-state presence. Understanding this can help in targeting campaigns that account for geographical mobility.

16. subscriber_sex_frequency
Description: Count of subscriber's gender records.
Usefulness: Ensures data consistency and helps identify any discrepancies in gender reporting. Consistent gender data is crucial for accurate gender-based campaign analysis.

17. subscriber_attribute_frequency
Description: Number of different user categories a subscriber is associated with.
Usefulness: Provides insights into the diversity of subscriber attributes. This information helps in segmenting subscribers based on demographic characteristics.

18. subscriber_campaign_frequency
Description: Number of campaigns each subscriber has received.
Usefulness: Similar to total_campaigns_received, but focused on individual subscribers. Helps understand the distribution of campaign exposure among subscribers.

19. subscriber_response_frequency
Description: Number of responses each subscriber has made.
Usefulness: Provides a direct count of subscriber engagements with campaigns. Subscribers with higher response frequencies are more actively engaged with marketing efforts.

20. average_weekly_responses
Description: Average number of responses per week across all subscribers.
Usefulness: Provides an overall view of weekly response trends, helping to identify peak engagement periods and plan campaigns accordingly.

21. subscriber_weekly_response_rate
Description: Response rate for each subscriber on a weekly basis.
Usefulness: Granular analysis of subscriber engagement patterns over time. Helps in understanding weekly response consistency or variability.

22. campaign_duration
Description: Duration of each campaign in weeks.
Usefulness: Provides insights into campaign longevity and duration effects on response rates. Longer campaigns may have different engagement patterns compared to shorter ones.

23. state_gender_interaction
Description: Interaction between state and gender.
Usefulness: Captures combined effects of geographic location and gender on campaign responses. Helps in targeting campaigns that consider both demographic factors simultaneously.

24. state_attribute_interaction
Description: Interaction between state and user attribute.
Usefulness: Combines geographic and user attribute data to understand localized demographic responses. Essential for geographically targeted marketing strategies.

25. gender_attribute_interaction
Description: Interaction between gender and user attribute.
Usefulness: Evaluates how gender influences responses across different user attributes. Helps in gender-specific campaign customization.

26. average_state_response_rate
Description: Average response rate for each state.
Usefulness: Provides a consolidated view of state-wise response trends. Essential for regional campaign planning and resource allocation.

27. average_gender_response_rate
Description: Average response rate for each gender.
Usefulness: Highlights gender-based response patterns, aiding in gender-specific campaign strategies.

28. average_attribute_response_rate
Description: Average response rate for each user attribute.
Usefulness: Evaluates response rates across different user categories. Essential for targeted marketing campaigns based on demographic attributes.

29. average_week_response_rate
Description: Average response rate for each week.
Usefulness: Tracks weekly response trends, indicating variations in subscriber engagement over time. Essential for optimizing campaign timing.

30. total_state_campaigns
Description: Total number of campaigns delivered in each state.
Usefulness: Quantifies campaign activities by state, aiding in regional campaign planning and analysis.

31. total_gender_campaigns
Description: Total number of campaigns delivered for each gender.
Usefulness: Provides a gender-based view of campaign distribution and engagement. Essential for gender-specific campaign strategies.

32. total_attribute_campaigns
Description: Total number of campaigns delivered for each user attribute.
Usefulness: Quantifies campaign reach across user categories, aiding in demographic-specific campaign analysis.

33. total_week_campaigns
Description: Total number of campaigns delivered in each week.
Usefulness: Quantifies weekly campaign activities, identifying periods of high or low campaign intensity.

34. state_response_rate_std
Description: Standard deviation of response rates for each state.
Usefulness: Measures variability in response rates across states. Highlights states with consistent or fluctuating engagement levels.

35. gender_response_rate_std
Description: Standard deviation of response rates for each gender.
Usefulness: Measures response rate variability by gender. Helps in understanding gender-specific response consistency.

36. attribute_response_rate_std
Description: Standard deviation of response rates for each user attribute.
Usefulness: Evaluates response rate variability across user categories. Essential for assessing demographic response consistency.

37. week_response_rate_std
Description: Standard deviation of response rates for each week.
Usefulness: Measures weekly response rate variability, indicating periods of stable or fluctuating subscriber engagement.

38. state_response_rate_median
Description: Median response rate for each state.
Usefulness: Represents the central tendency of response rates across states. Complements average response rate insights with distribution characteristics.

39. gender_response_rate_median
Description: Median response rate for each gender.
Usefulness: Provides a central measure of response rates by gender. Useful for understanding gender-specific response trends.

40. attribute_response_rate_median
Description: Median response rate for each user attribute.
Usefulness: Central measure of response rates across user categories. Complements average response rate insights with distribution characteristics.

41. week_response_rate_median
Description: Median response rate for each week.
Usefulness: Central measure of response rates by week. Useful for understanding weekly response rate distribution.

42. state_response_rate_min
Description: Minimum response rate for each state.
Usefulness: Identifies states with the lowest response rates. Helps in targeting underperforming regions for improvement strategies.

43. gender_response_rate_min
Description: Minimum response rate for each gender.
Usefulness: Identifies gender groups with the lowest response rates. Essential for gender-specific engagement enhancement strategies.

44. attribute_response_rate_min
Description: Minimum response rate for each user attribute.
Usefulness: Identifies user categories with the lowest response rates. Essential for demographic-specific engagement enhancement strategies.

45. week_response_rate_min
Description: Minimum response rate for each week.
Usefulness: Identifies weeks with the lowest response rates. Essential for identifying periods of low engagement for campaign planning.

46. state_response_rate_max
Description: Maximum response rate for each state.
Usefulness: Identifies states with the highest response rates. Highlights regions with strong engagement for campaign prioritization.

47. gender_response_rate_max
Description: Maximum response rate for each gender.
Usefulness: Identifies gender groups with the highest response rates. Essential for gender-specific campaign prioritization strategies.

48. attribute_response_rate_max
Description: Maximum response rate for each user attribute.
Usefulness: Identifies user categories with the highest response rates. Essential for demographic-specific campaign prioritization strategies.

```
Feature Engineering completed. The enhanced dataset is saved.
```

```
In [6]: # Print all columns
        print("All columns after feature engineering:")
        print(df.columns)
```

```
All columns after feature engineering:
Index(['week', 'id.no', 'attribute', 'state', 'sex', 'campaign', 'response',
       'response_rate', 'campaign_success_rate', 'response_last_month',
       'total_campaigns_received', 'total_campaigns_responded',
       'average_response_time', 'state_response_rate', 'sex_response_rate',
       'attribute_response_rate', 'week_response_rate',
       'state_campaign_success_rate', 'sex_campaign_success_rate',
       'attribute_campaign_success_rate', 'week_campaign_success_rate',
       'subscriber_state_frequency', 'subscriber_sex_frequency',
       'subscriber_attribute_frequency', 'subscriber_campaign_frequency',
       'subscriber_response_frequency', 'average_weekly_responses',
       'subscriber_weekly_response_rate', 'campaign_duration',
       'state_gender_interaction', 'state_attribute_interaction',
       'gender_attribute_interaction', 'subscriber_state_gender_interaction',
       'subscriber_state_attribute_interaction',
       'subscriber_gender_attribute_interaction',
       'average_state_response_rate', 'average_gender_response_rate',
       'average_attribute_response_rate', 'average_week_response_rate',
       'total_state_campaigns', 'total_gender_campaigns',
       'total_attribute_campaigns', 'total_week_campaigns',
       'state_response_rate_std', 'gender_response_rate_std',
       'attribute_response_rate_std', 'week_response_rate_std',
       'state_response_rate_median', 'gender_response_rate_median',
       'attribute_response_rate_median', 'week_response_rate_median',
       'state_response_rate_min', 'gender_response_rate_min',
       'attribute_response_rate_min', 'week_response_rate_min',
       'state_response_rate_max', 'gender_response_rate_max',
       'attribute_response_rate_max', 'week_response_rate_max'],
      dtype='object')
```

Conclusion

The feature engineering process has significantly enriched the targeted marketing dataset by incorporating new features. These features capture diverse aspects of subscriber behavior, campaign effectiveness, and demographic interactions. The insights derived from these features are crucial for optimizing campaign strategies, enhancing subscriber engagement, and improving the overall effectiveness of marketing efforts. Moving forward, leveraging these features in

machine learning models will enable more accurate predictions of subscriber responses, thereby driving more targeted and successful marketing campaigns.

This concludes the feature engineering report for the targeted marketing dataset.