Performance Marketing Project Report

By Benjamin Wong

In the carefully measured, data-driven world of modern marketing, the lure of performance marketing, with its focus on quick and easily quantifiable metrics such as clicks and conversions – has tempted many brands to sacrifice long-term brand equity in favour of short-term gains. Although this method provides measurable outcomes, it tends to ignore the powerful effect of emotional brand story and brand resonance. As pointed out by Zachary Tindall in his article, “Performance Marketing is Killing Your Brand,” the brands that last forever are the ones that create an emotional bond with the audience that creates links at the neurological level that make stories memorable and powerful.

This project aims at investigating the balance between performance marketing metrics and the intangible but crucial aspects of brand storytelling. The question that informs this analysis will be: How can brands incorporate facts-based gestures without sacrificing the emotional stories that create long term bonds with customers?

The results of this research can be especially significant for the marketing specialists, brand strategists, and business leaders who are interested in sustainable growth. Through the dynamics of performance indicators that can be measured with an understanding of the science of storytelling, stakeholders can make informed decisions to affect immediate results that will create long term brand value.

With the drive for rapid, quantifiable growth, performance orientated modern approach to marketing has become overbearing by far. Percentages on click-through rates, conversion rates, and efficiency in ad spend come to power decision-making dashboards, blowing out of perspective the more intangible parts of the brand health—trust, loyalty, emotional engagement, and resonance of a narrative. Although performance marketing offers instant feedback and scope for optimization, there is a danger that marketers get into a habit of focusing on short term returns rather than steering the brand’s value for the long-term.

This change has caused a pressing and relevant question: How much do over-rely on performance marketing tactics affect customer loyalty, emotional connection and perception to the brand – can data mining be of some help to understand how not to overdo it?\

This analytical dig is significant because it touches on a dilemma that exists in many modern companies this day. how to continue being data-driven, without losing the emotional element of their brand. The impact of the lack of attention to this balance is devastating – from the brand equity dilution and the customer loyalty erosion to the lifetime value decrease and the churn increase. Given the number of brands that now use these ever finer metrics to justify marketing investments, it is necessary to check if such indicators really reflect health and future of a brand, or just provide the illusion of its success.

To examine this problem, I will use a combination of data mining techniques such as regression analysis to assess the interconnection between marketing undertakings and conversion behavior, and classification models that define what features predominantly suggest loyalty or conversion outcomes.

Besides, the exploratory data analysis (EDA) will help an analytic to explore patterns of the customers engagement which is beyond the metrics of the performance marketing. This mixed approach offers a greater perspective of how marketing strategy does not only impact transactions but also relationships.

At the end of the day, this project is an attempt to re-interpret the concept of the success of marketing by data. Instead of applying analytics only to optimize for short-term wins, this work will look at how we can leverage data to make it integral in our nurturing, protection, and growth of brand identity in a manner that sustains the long-term growth and connection with the customer.

The central object of this analysis is target variable Conversion - a binary variable of classification that shows whether a customer converted (1) or not (0) post being exposed to a marketing campaign. This target will allow us to formulate the problem in terms of a supervised classification problem, to work with predictive modeling, to find out which of the customer and campaign attributes are the most important in explaining conversion outcomes.

#### **Dataset Overview**

The dataset that was applied in this research provides a detailed picture of customer interaction with digital marketing campaigns. It has 8,000 instances i.e. unique customers. This dataset has 20 attributes covering a number of key dimensions of digital marketing analytics.

* **Demographic Information**: Includes features such as Age, Gender, and Income, which provide context on who the customer is.
* **Marketing-Specific Variables**: These features, such as CampaignChannel, CampaignType, AdSpend, ClickThroughRate, and ConversionRate, capture the nature and performance of different marketing efforts.
* **Customer Engagement Metrics**: Variables like WebsiteVisits, PagesPerVisit, TimeOnSite, EmailOpens, EmailClicks, and SocialShares reflect how users interacted with the brand’s digital presence.
* **Historical Purchase Behavior**: Includes PreviousPurchases and LoyaltyPoints, offering insight into past customer-brand interactions.
* **Platform-Specific Attributes**: Fields such as AdvertisingPlatform and AdvertisingTool, though anonymized, indicate the technologies or ecosystems involved in campaign delivery.

#### **Data Types and Structure**

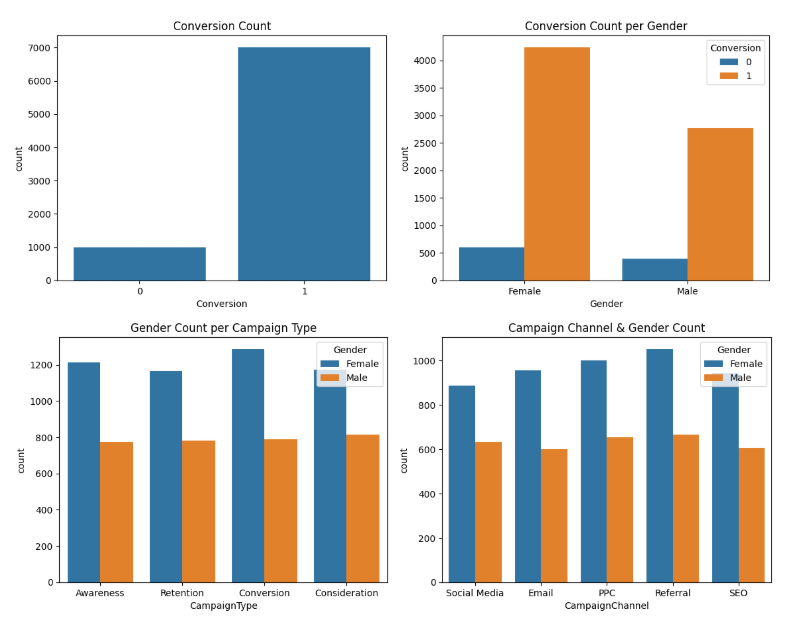
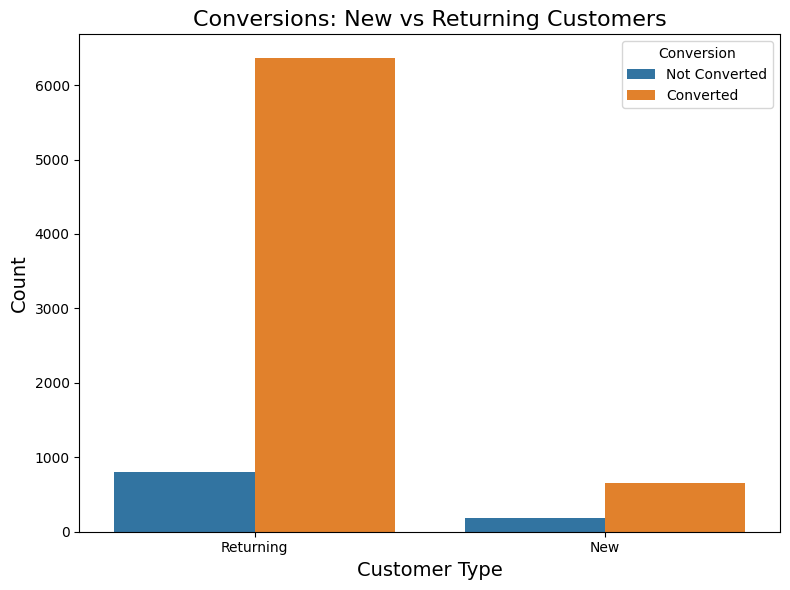
The dataset constitutes both numerical (integer and float) as well as categorical (objects) types of variable. Specifically:

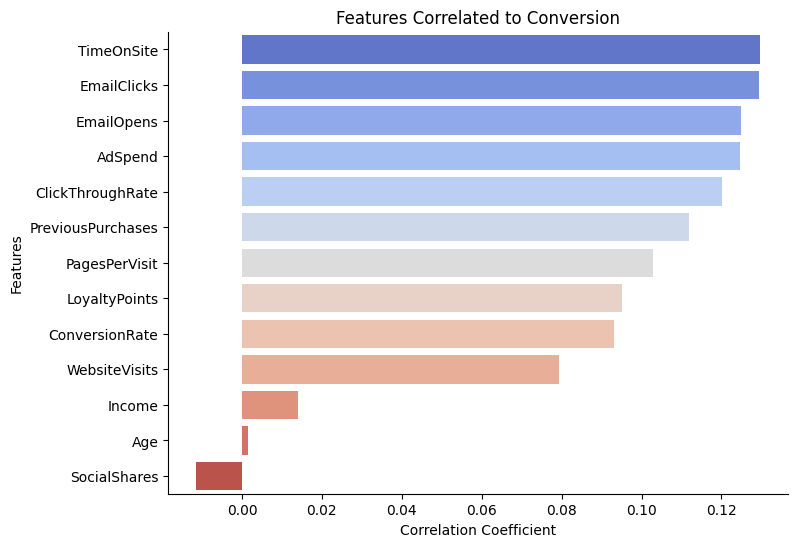
* **Numerical Attributes**: 15 columns including integers (Age, Income, etc.) and floats (AdSpend, ClickThroughRate, etc.).
* **Categorical Attributes**: 5 columns representing classification features such as Gender, CampaignChannel, and CampaignType.

The structure of the data is pristine and fluent.:

* There are no missing values, which makes it convenient to go ahead into modeling and analysis..
* Some of the numeric features like ClickThroughRate, ConversionRate occur naturally in a normalised range 0 to 1 and don’t require further scaling.
* Other numeric features with large domains of values (such as AdSpend, Income, LoyaltyPoints) were defined for scaling during preprocessing, in order to make sure that the model will be balanced**.**

Although the dataset was clean, scaling of features, and interpretation of results, were still a problem. Thoughtful pre-processing was necessary in order to make sure that attributes are handled with care, particularly when some of the values (such as rates and counts) would differ by several orders of magnitude. In addition, some categorical variables like CampaignChannel and CampaignType are associated with subtlety information regarding the marketing strategy that can inform the need to encode and chastely feature engineer for them to be preserved in a modeling process. Nevertheless, the dataset underpins both the predictive modeling and enhances deeper insights into marketing effectiveness – which is crucial to answer the overall question of how brands can optimize performance, but without losing strategic and emotional impact.

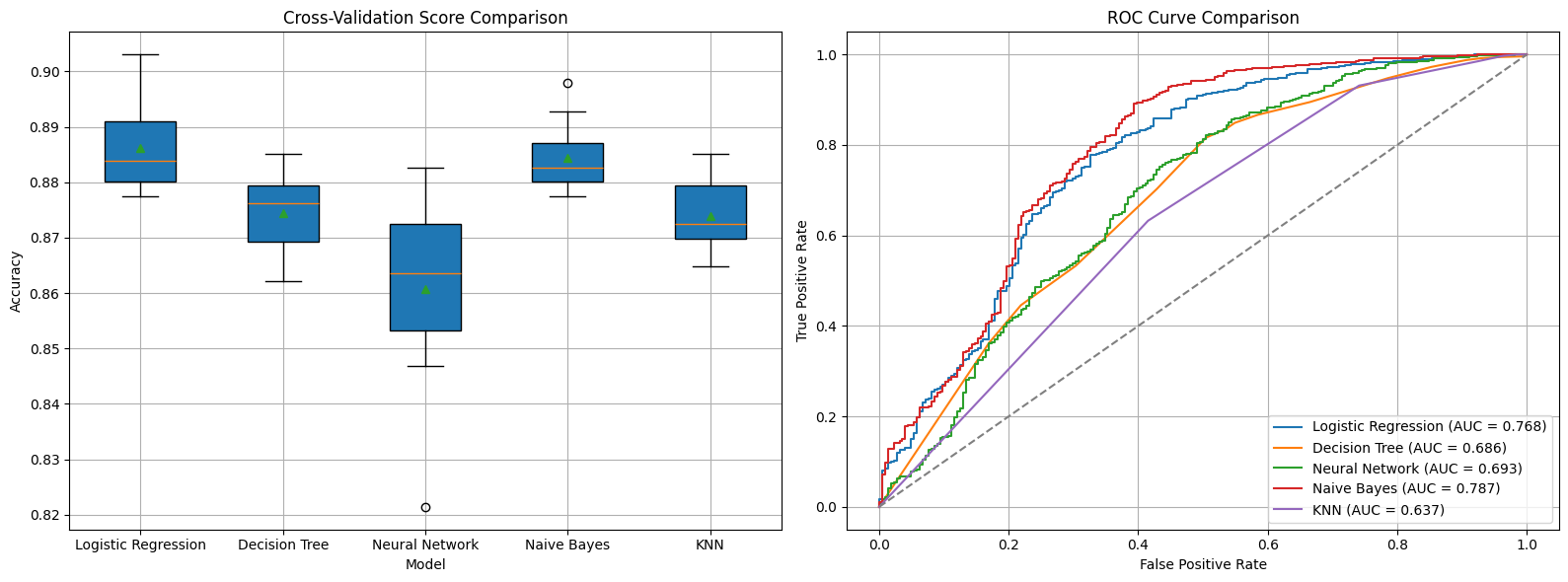
* The first graph above is a portrayal of the distribution in values of our target variable, Conversion, with 0 when a user did not convert to a customer and 1 when a user converted. After looking at the findings from the graph it can be said as there is imbalance on our target.
* The three graphs post conversion distribution, are the count plots for distribution of categorical feature Gender in relation to other categorical variables after conversion. With the help of these visualizations, one can make crucial observations on the target population of the company. It is worth mentioning that data reveals that a rather sizeable share of the customer base is female, which means that women could be the main target for the company’s products or services.
* Based on this graph, we get more idea about the target demographic of the company. As depicted most of the campaigns have largely convinced returning customers as compared to new customers. remarkably, these campaigns have been successful in reconverting these users to be repeat buyers.

After undertaking an assessment of the categorical variables, focus was shifted towards the numerical attributes in the dataset. To know their relationship with the target variable, a correlation matrix was created. This enabled determination of numerical features and the degree of linear correlation of such features with the target variable, Conversion. This step gave an insight into numerical features that potentially could be the most influential in predicting the conversion outcomes.

* A bar plot was created with the help of Matplotlib to visualize correlation between numerical features and target variable in an ascending order. For all features, the correlation coefficients were not too high, none was above 0.15. Although this implies weak linear relationships with the target variable, there are some observations that can be made. For example, the campaign-specific features demonstrated a slightly lower correlation as opposed to the features related to email use, such as email clicks and email opens. Although the values of the correlations are low, these features will be kept as input variables in the modeling phase to determine the possible predicative power of the additional variables.
* Such a finding may indicate that the selection of performance metrics or marketing-oriented variables, like click-through rate, ad spend, and conversion rate, may not be as decisive as usually underlined in the performance marketing strategies.

After the assessment of the correlation of features with the target variable, the following step was to scale the numerical features so as to minimize the variance among instances. Variable in the feature values can degrade performance and accuracy of model. To compensate for this, we used StandardScaler of scikit-learn to standardize the numerical features. All of the numerical features were scaled except for ClickThroughRate and ConversionRate. These two characteristics were already within reasonable limits (between about 0.01 to 0.03), meaning that there was little variation hence did not need rescaling. This discriminate methodology made sure that features that play an important part in training of the model were normalized to a uniform scale without uncalled for modification of data that was already distributed in an optimal manner.

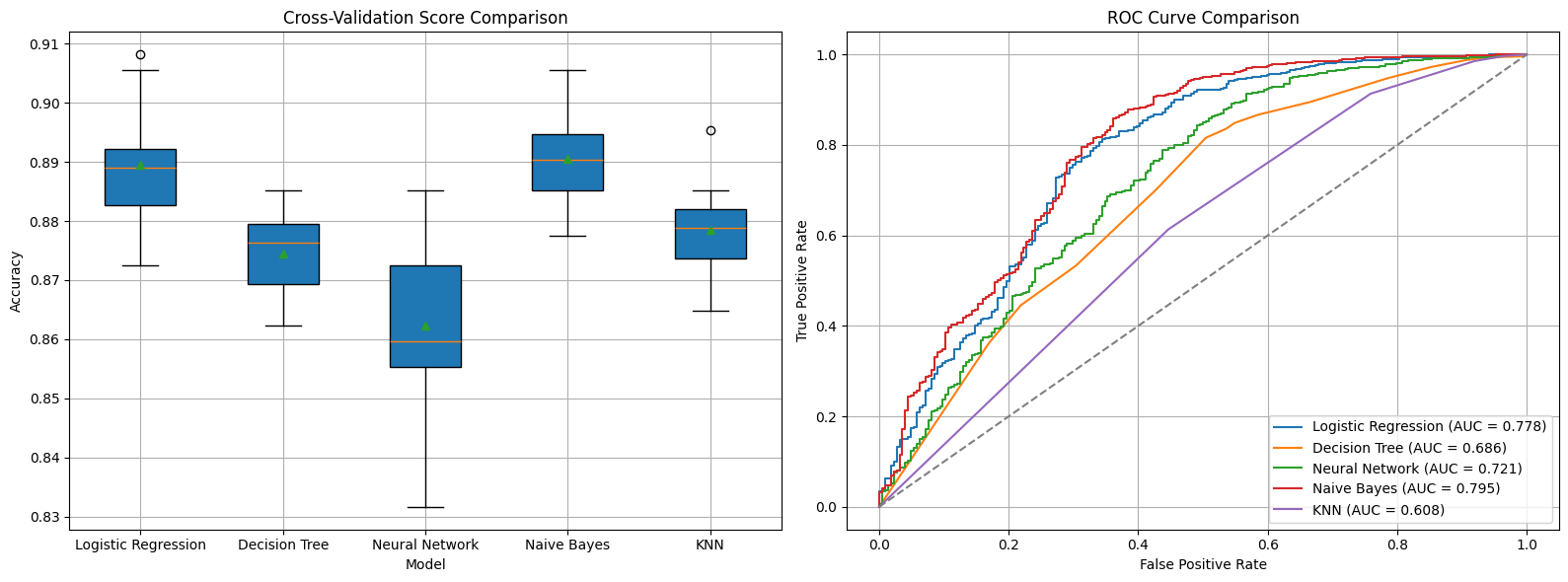
For the assessment of performance through various classification algorithms, 5 models were taken into consideration as the basis of comparison; Logistic Regression, Decision Tree, Neural Networks, Naive Bayes, and K-Nearest Neighbors (KNN). In order to study the effect of distinct feature representations over model performance, three individual training datasets were formed. one with only numerical features, another with one-hot encoding for categorical variables, and the third one using target encoding. This approach made it possible to understand more broadly how data preprocessing methods affect model’s outcomes. We need to consider the results of the first training set.



| Model | Mean CV Score | Std Dev | AUC Score |
| --- | --- | --- | --- |
| Logistic Regression | 0.8862 | 0.0084 | 0.7683 |
| Decision Tree | 0.8745 | 0.0076 | 0.6865 |
| Neural Network | 0.8607 | 0.0168 | 0.6929 |
| Naive Bayes | 0.8844 | 0.0064 | 0.7874 |
| KNN | 0.8740 | 0.0062 | 0.6370 |

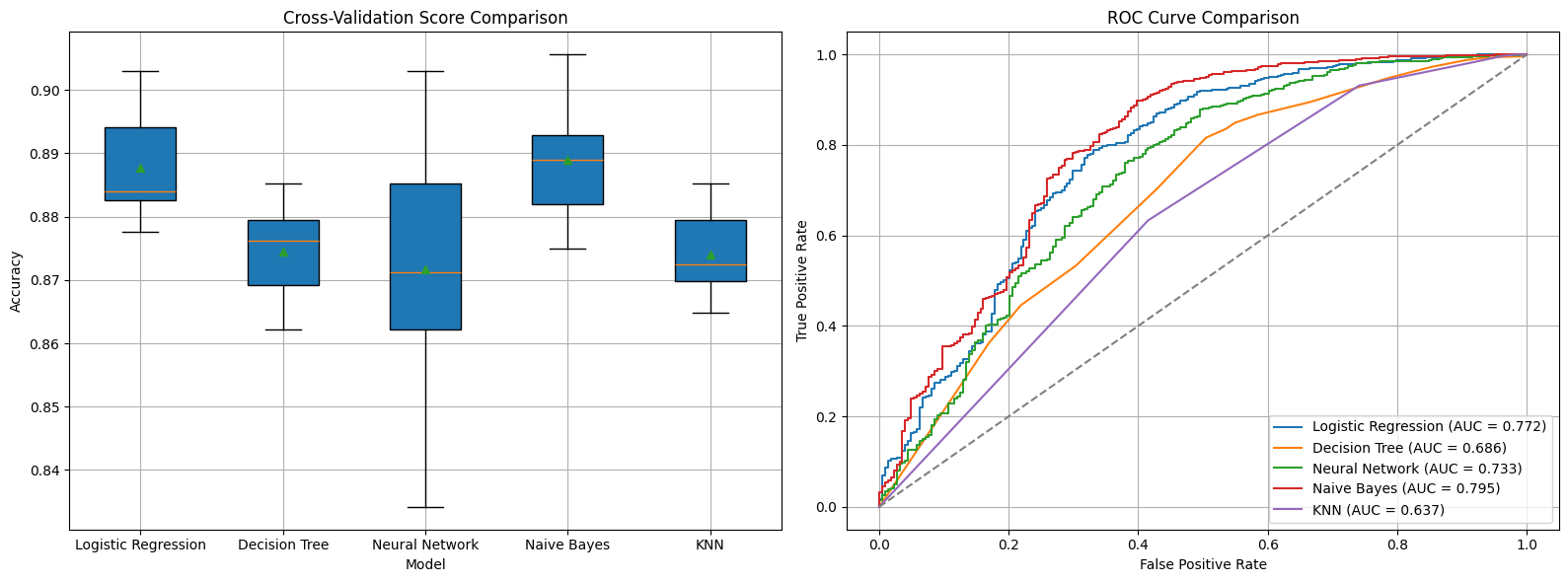
* These are the model’s results on the training set in terms of only scaled numerical data (without con version and click through rates) – as we can see Naive Bayes has the highest Cross validation score with low variance and Naive Bayes has the highest rank on the ROC Curve & AUC score. Quite irrespective of its failing behind logistic regression in cross validation score by 0.0018 as well as having more variation than KNN by 0.0002. It remains to performances excellently and consistently.

The purpose of using this training set is to showcase a model result for a marketing department that was performance driven. Despite scaling and using other numerical features such as user engagement & customer history metrics, this training set doesn’t take into account the categorical features that drive brand longevity. As we can see with the roc curve, the beginning half of the results leans more towards false positives then goes back to true positives. This is possibly due to the fact that this training set was leaving out other important features in our dataset, which could potentially harm brand longevity and the trust of solely performance driven marketing teams. The next few training sets and results will showcase how a marketing department would put both performance metrics & categorical metrics to good use.



| Model | Mean CV Score | Std Dev | AUC Score |
| --- | --- | --- | --- |
| Logistic Regression | 0.8895 | 0.0103 | 0.7784 |
| Decision Tree | 0.8745 | 0.0076 | 0.6865 |
| Neural Network | 0.8564 | 0.0164 | 0.7298 |
| Naive Bayes | 0.8906 | 0.0086 | 0.7954 |
| KNN | 0.8786 | 0.0081 | 0.6080 |

* Here are the results for a training set that included both the scaled numerical features along with categorical features that were one hot encoded. As we can see, the best performing model for this training set would be the Naive Bayes model. Despite being shy of 0.0010 standard deviation score behind the Decision Tree Model it still performed very consistently with this training set. It also excelled in having the highest cross validation and AUC score. Looking at the ROC curve it’s dip in the beginning half isn’t as drastic as the first training set with only numerical scaled features. This could be due to this training set accounting for each time that categorical values are present whereas the previous one didn’t.



| Model | Mean CV Score | Std Dev | AUC Score |
| --- | --- | --- | --- |
| Logistic Regression | 0.8872 | 0.0080 | 0.7717 |
| Decision Tree | 0.8745 | 0.0076 | 0.6865 |
| Neural Network | 0.8704 | 0.0133 | 0.7303 |
| Naive Bayes | 0.8888 | 0.0088 | 0.7984 |
| KNN | 0.8740 | 0.0062 | 0.6373 |

* Here are the results for the scaled training set with the categorical features target encoded. Target encoding finds the probability of a user converting for each of the encoded values. For instance if the campaign type social media 4 customers converted out of 10 times that it’s shown up as a campaign type on the dataset then target encoding will give that campaign type a 0.4 probability of the target. So while the previous one hot encoding has provided categorical context for our training set, this further contextualizes the categorical features as each time that a value occurs it takes into account the rest of the times it occurs and not only in a single instance.
* After reviewing the results of the models on the target encoded training set, the best performing model is the Naive Bayes Model. It ranked the highest in the roc auc score, has a relatively low variance & has the highest average cross validation score. Even when looking at the roc curve graph, the naive bayes curve is notably more true positive leaning than the other models and previous training sets.

Strategies that completely rely on performance data like clicks, conversions and efficiency on ad spend may truthfully compromise on a brand’s perception unintentionally because customer behavior is rather complex and laden with nuance. One hot encoding as a method of preserving categorical context treats each category as an isolated entity and disregards the relation with the outcome. However, the target encoding improves model learning by mapping categorical variables according to their historical correlation with target variable. This probabilistic representation provides the additional insight to our models about how various features affect conversion, so the target-encoded, scaled dataset is a more contextually-rich and strategic option – even if other training sets garnered slightly higher cross-validation scores.

From among the evaluated models, Naive Bayes was consistently the best when working with the one-hot and target-encoded training sets. This can be explained by its capacity to work effectively with probabilistic input and thus be very useful in the nature of target encoding. Because Naive Bayes assumes feature independence and can effectively operate with categorical data, which are converted into likelihoods, it is potentially useful to analyze the statistical tendencies in the dataset without overfitting. With its good AUC scores and its consistent cross-validation ability it can be argued that in this particular setting it describes the real behavior of the customer more accurately than more complex models.

This analysis shows that obsession with performance marketing metrics – ranging from ad spend, click-through rates & conversion rates – causes imperfect and misleading customer behavior and brand health visibility. Although these metrics give information that is immediate, our models demonstrated that if one relies on numerical performance indicators, the result is concealed and there is a probability of false positives. This is the pattern of how solely performance drive strategies in marketing can overlook the subtle variables that stimulates long term customer engagement and loyalty.

Using categorical features especially through target encoding that provides more truthful model performance, Naive Bayes is the most consistent in terms of classifier. This means that there is predictive power for features other than conversion provided by brand relevant attributes like campaign channel & types. These features are the proxies for the brand narrative and audience resonance which are many times lost in the performance – driven marketing.

Thus, after all, data mining can give a more balanced approach. Compiling data in the form of performance metrics and strategically encoded categorical data, brands can identify patterns that indicate not only short and long-term relationship building. This explains the fact that proper marketing should not be a choice between storytelling and data, but data, and storytelling.

**Works Cited**

Tindall, Z. (2025, April 23). *Performance marketing is killing your brand. Here’s the science.* Tindall Effect.<https://www.tindalleffect.com/blog-2-1/performance-marketing-is-killing-your-brandnbspheres-the-science>

Wong, B**.** (2025, May 1). *Benjamin\_Wong\_FinalProject* [Colab notebook]. Google Colaboratory.<https://colab.research.google.com/drive/1USn5J-atcX6xTNfaBQXYcaofw8MLIIJC>