University of Essex

Combined coursework reassessment for CE802

Pilot-Study Proposal: Investigating the Effectiveness of Machine Learning for Targeted Email Marketing

Author: 2202026

Publication date: 22/08/2023

Word count: 740

Table of Contents

Introduction

In digital marketing, targeted emails boost engagement and sales. Firms aim for personalized communication to boost conversions. A method is targeted email marketing, adjusting emails using data.

The batch-and-blast method in email marketing is outdated, causing lower engagement, high unsubscribes, and poor brand image. It sends generic content to all, ignoring preferences. According to Hart (2021), this harm returns, relationships, and reputation. On the other hand, targeted email marketing with machine learning and past data is personalized, improving engagement, conversions, and loyalty.

The primary goal of this pilot study is to improve the effectiveness of the online computer store's email advertisement campaigns through targeted email marketing. This involves identifying who are most likely to make a purchase based on historical email campaign data and conversion outcomes. The availability of historical data on previously sent emails and their conversion outcomes provides a unique opportunity to explore these objectives. By combining this data with advanced machine learning techniques, the company can transition from a one-size-fits-all approach to a more personalized and precise marketing strategy, fostering higher customer satisfaction, engagement, and ultimately, business growth.

Classification is a machine learning technique that involves categorizing recipients into distinct classes based on their propensity to convert, allowing the company to allocate its marketing resources more efficiently and tailor its content for maximum impact. It is well-suited for scenarios where the goal is to assign observations to predefined categories. This makes it particularly useful when predicting outcomes with discrete labels which is suitable in this case as the goal is to know whether an email should be sent to a potential client.

Informative Features for Prediction

Classification involves the categorization of data into distinct classes based on specific features or attributes. In targeted email marketing, the specific features for classification involve various attributes and behaviours of recipients. These features may include demographic data such as their age, gender, income level and location. Other key features are past purchase history which reveals patterns vital for understanding consumer behaviour, engagement rate with previous emails and website interactions. These features are crucial as they provide insights into each recipient's interests, preferences, and likelihood to convert. Categorizing recipients using these features enables customized content and offers, boosting engagement and conversions.

Learning Pipeline

Given that the organization has access to historical data of each previously sent email and whether it led to a sale it is crucial for it to be gathered and pre-processed and formatted for analysis. The data needs to be by categorizing data into classes to transform the problem into a classification task.

A pre-processing task that can be performed is feature scaling which involves normalizing the numerical features of a dataset to ensure they are on a similar scale and the handling of missing data as incomplete information would nullify the effectiveness of the model. Classification will then proceed with a two-step procedure that involves the construction of the classification model (training phase) and utilizing the resulting classifier (validation phase).

The predictive task is essentially a binary classification problem, where each contact is classified as either a potential buyer or a non-buyer for a given product campaign. Algorithms that can be used for binary classification include, Logistic Regressions, k-Nearest Neighbours, Decision Trees, Support Vector Machine and Naive Bayes. The performance of the algorithm will then be evaluated by various metrics.

Recommended Learning Method: Decision Tree Classifier

Decision trees are non-parametric supervised learning algorithms used for classification tasks. They have a hierarchical tree-like structure that is based on feature values. This algorithm is good due to their simplicity and ability to handle complex problems. Decision trees can be utilized in targeted email marketing by segmenting the audience based on attributes such as demographics, past behaviour, and engagement. However, to ensure that the best possible choice is made the consideration of other classification algorithms will be done. They will be compared to the decision tree classifier based on various evaluation metrics.

Evaluation Metrics for Success

When evaluating the performance of classification algorithms metrics comes into play. They include accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of predictions, while precision focuses on the proportion of true positive predictions among all positive predictions. Recall highlights the fraction of true positive predictions among actual positive instances. The F1-score combines precision and recall providing a balanced assessment of the model's effectiveness. These metrics collectively paint a comprehensive picture of the algorithm's capabilities and guide the fine-tuning process for optimal results.

Works Cited

Bansal, S. (2023 , January 25). *What is Classification Algorithm in Machine Learning? With Examples*. Retrieved from Analytix Labs: https://www.analytixlabs.co.in/blog/classification-in-machine-learning/#Email_Classification

Hart, K. (2021, July 27). *Why "Batch and Blast" Email is Destroying Your Business (And What to do Instead)*. Retrieved from Quantum Lifecycle Marketing: https://quantumdigitalmarketing.net/batch-and-blast/#:~:text=Not%20only%20does%20batch%20and,can%20negatively%20impact%20your%20bottom.