

HOW CAN WE OPTIMIZE THE EFFECTIVENESS OF ONLINE ADVERTISING THROUGH PRECISE SEGMENTATION OF TARGET AUDIENCES ?



Soumia ZARKAN
MSC AIBA
August 2023

Table of contents

01 AI in the digital marketing era

Definitions and context.

02 Methodology

Methods used to answer the question.

03 Results

Results and comparisons of the algorithms used.

04 Recommendations

Managerial recommendations and open discussions about AI in programmatic advertising,



01

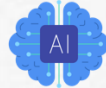
AI IN THE DIGITAL MARKETING ERA



AI definition

Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.

- Encyclopædia Britannica



Digital Marketing and AI

AI leverages customer data to identify patterns and preferences, enabling personalized recommendations and improved engagement. It also aids in customer segmentation for targeted campaigns based on demographics and behavior, enhancing overall marketing effectiveness.

MARKETING SEGMENTATION



Definition

The practice of dividing a broader consumer or market base into distinct and homogeneous groups or segments based on shared characteristics, behaviors, preferences, or needs.

These segments are created to enable more targeted and effective marketing strategies, allowing businesses to tailor their products, messages, and campaigns to better resonate with the specific traits and desires of each segment.



AI and segmentation

Market segmentation has evolved into a more comprehensive strategy that considers psychographics, behavior, and personalization. AI and data analytics have played a significant role in enhancing the effectiveness and precision of market segmentation.

PROGRAMMATIC ADVERTISING



Definition

Programmatic advertising can be defined as “the automated serving of digital ads in real-time based on individual ad impression opportunities”

- Busch, 2016.



AI and programmatic advertising

Bidding on Ads : To know how much to bid on an ad is fundamental in programmatic advertising. If the bid is too low, the ad is less likely to be shown. Whereas if the bid is too high, there is a chance to spend more money than necessary.

Predictive Analytics : And above all, machine learning is used to improve the results of programmatic campaigns. Through analyzing metrics from past campaigns, machine learning can identify which targeting strategies, pieces of ad content and campaign targets are working and which aren't.

Identifying Patterns : Analyze data and identify patterns in user behavior. This data is then used to make predictions about what users are likely to do in the future.

Predicting Outcomes : Another important use of machine learning algorithm technology in programmatic is to forecast the likelihood of potential customers taking desired actions, such as clicking on specific Google Ads or making a purchase.



02

METHODOLOGY

Data exploration

Database composed of 18 columns and 12330 rows. Cleaning of the data, check of the missing values and correlations.

Comparison of algorithms

- Decision Trees
- KNN Models
- Naive Bayesian
(Gaussian/Multinomial)



Decision Trees

A machine learning model that makes predictions by recursively partitioning data into subsets based on the values of input features, forming a tree-like structure of decisions.

KNN Models

KNN (K-Nearest Neighbors) is a machine learning model that makes predictions by looking at the most similar past examples in the data.

Naive Bayesian

Naive Bayes is a machine learning model that predicts outcomes by using the probability of events and assuming that features are independent.



03

RESULTS

Decision Tree:

Accuracy: 0.8249
Precision: 0.8282
Recall: 0.8249
F1 Score: 0.8247

KNN model:

Accuracy: 0.6761
Precision: 0.6928
Recall: 0.6761
F1 Score: 0.6667

Gaussian Naive Bayes (improved):

Accuracy: 0.8428
Precision: 0.8436
Recall: 0.8428
F1 Score: 0.8426

Multinomial Naive Bayes:

Accuracy: 0.6939
Precision: 0.7119
Recall: 0.6939
F1 Score: 0.6854

Explanation of the scores

- **The Accuracy** is the proportion of correct predictions (positive and negative) in the sample. The higher the accuracy means the better the model predicts efficiently. Here the Gaussian Naive Bayes has the highest accuracy with 0.84.

- **Precision** is the proportion of correct "positive" predictions in the sample. A high precision value indicates that the model has a great rate of true positives and a low rate of false positives. The Gaussian Naive Bayes has the highest precision with 0.84.

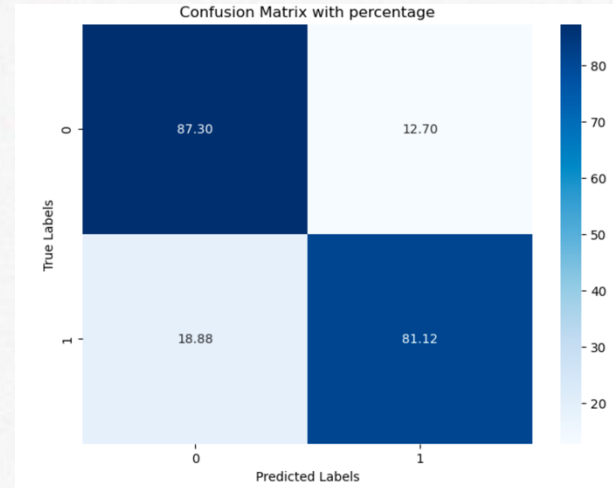
- **The recall** is the proportion of "positive" actual records correctly predicted as "positive". It identifies the model's ability to identify positive instances. If the recall is high it means that there are low rates of false negatives. The Gaussian Naive Bayes has the highest recall with 0.84.

- **F1-score** is the harmonic mean between precision and recall. More informative than Accuracy for unbalanced datasets. It is particularly interesting here because the target variable was unbalanced at the beginning. Like for the rest of the metrics the Gaussian Naive Bayes has the best F1 Score with 0.84.

Confusion Matrix's interpretation of the Gaussian model

We can observe a rate of 87.30 of true negatives which is very great. Same thing for the true positive with a rate of 81.12. It seems very balanced. We can say that the model is quite good. However it can be optimized by using more sophisticated sampling techniques.

With the cross validation we can see a certain consistency in the prediction. However there is an improvement the score starts at 0.82 and increase to 0.86. Nevertheless we need to check at the parameters because there's something that could improved.



Cross-Validation Scores:

Fold 1: 0.8220
Fold 2: 0.8639
Fold 3: 0.8479
Fold 4: 0.8409
Fold 5: 0.8182



04

RECOMMENDATIONS

MY RECOMMENDATIONS

Implementation of the Gaussian Naive Bayes to have a better segmentation.

Make sure that the dataset is clean and balanced before testing the model to avoid any error

Use Session-Level Clustering or User-Level Clustering.

Implementation of a system that would monitor and analyze the campaigns performance.

**THANK YOU FOR
YOUR ATTENTION !**

