

## Prediction of Bank Customer Potential Using Creative Marketing Based on Exploratory Data Analysis and Decision Tree Algorithm

Fadhil M. Basysyar\*, Arif R. Dikananda, Dian A. Kurnia

STMIK IKMI Cirebon, Cirebon 45131, Indonesia

Corresponding Author Email: [fadhil.m.basysyar@gmail.com](mailto:fadhil.m.basysyar@gmail.com)



<https://doi.org/10.18280/isi.270409>

### ABSTRACT

**Received:** 3 June 2022

**Accepted:** 6 August 2022

**Keywords:**

*decision tree, best trashold determination, classification, creative marketing, Python 3*

Today's bank marketing uses traditional methods, using standard and boring ads to target prospects. Of course, this can impact the banking process, reducing the number of potential banking customers. Banks need to think outside the box and apply creative marketing ideas to drive their profitable marketing development and marketing success potential. Most consumers consider banking a daily necessity, and it is best to avoid it. If they can take a creative approach to bank marketing, that idea can change. Especially when bank marketing integrates creative bank marketing ideas such as gamification, automation, chatbots, and rewards to encourage potential customers to use banking services, therefore; this study uses a decision tree algorithm with the best trash old decisions to perform a classification process on kaggle.com's bank marketing dataset. The classification process uses Python 3 to find the accuracy value of the decision tree algorithm calculation using K-Fold and scale data. This survey achieved classification results with 82% accuracy, 84% recognition, and 85% accuracy with recommendations and creative marketing solutions.

## 1. INTRODUCTION

In recent years, sales have increased due to the economic impact of advertising cooperatives on the practice of small businesses that gives presentations. Introducing cooperatives faces the general belief that defeating smallholders can be an effective system that can be effective system that fosters constraints caused by their limited size and fundamental changes in the value chain of business. This consideration has resurfaced [1-3]. Today, customers can better choose from various banking options [4]. The bench is not an efficient device in the mood of this chief. In addition, banks are an important requirement to support associations to incentive their customers and gain manageable advantages [5]. The expansion of banks has created an interesting quirk for advertisers. Customers have ample options to switch rather than focus on a particular bank [6]. In this way, tagging is a technique that allows associations to promote useful long-term relationships with clients. Therefore, this mark provides a special incentive to help customers perform services and catch up with them [7]. The main resource of many clubs is the club bank [8].

Today advertising regularly relies upon prediction fashions to make, deliver esteem, and oversee institutions with customers [9]. A prediction has empowered advertisers to shape critical showcasing picks utilizing adequate facts that companies information on customer behavior attributes and characteristics, for example, customer spending designs, eyeball developments, pix and comments shared through digital entertainment, customer object reviews, leisure content, and meals and workout propensities [10]. Using a linear hyperplane, the fundamental prediction splits data into two classes: data test and data training. In other circumstances, it is unable to properly classify some data [11, 12]. Prediction-

primarily based prescient exam and concept frameworks dashed up the complete selling process, allowing advertisers to apprehend areas of expertise customers and speak to them via particular multi-channel crusades [13]. In this manner, Prediction assists companies with fostering a sensible top hand via way of means of maintaining far from emblem redundancies, saving cash on unnecessary notices, and staring at post-purchase behavior via form means of dissecting non-stop facts [4]. In any case, marketing and marketing researchers have forewarned that many agencies are endeavoring to increment offers earning via way of controlling Prediction frameworks via the development of incorrect calculations utilizing non-delegate facts to make an uncalled-for gain via selling [14]. A precise customer accumulation is restrained from unbiased admittance to selling contributions [15].

Despite the inequality, inequity, and unfair impact of predictive marketing models, research in this area is mainly anecdotal and fragmented and has not yet evolved into an integrated conceptualization [16]. This study identifies the source of the predictive marketing model to determine the dimension of creative marketing based on the ability of decision tree algorithms in creative marketing, namely Irregularity to know potential clients, Crusade Calls the strategy should be executed, the Age category of potential clients, Occupation of potential clients, House Loans and Balances, and Target of the potential clients with a higher span.

Conducts consciously written research and top-to-bottom meetings using expectations models to facilitate a rational structure of algorithmic predisposition to leaders' abilities. And distinguish between different aspects. The review makes some commitments. Primarily, we use creative marketing based on the decision tree algorithm for the power of the predicted advertising model. Second, the findings show that

clever marketing is interconnected based on decision-making tree algorithms, and understanding these sources can foster the unique management capabilities of the model. It will help. Finally, from a functional point of view, our findings raise concerns about the irrational impact of decision tree algorithms on a particular assembly of a customer in terms of direction, gender, class, age, religion, and identity. I will deal with it. It is as follows. It focuses on written audits of creative marketing based on the decision tree algorithm model. Techniques for researched writing, subjective experience, and relationship triangulation. Applied innovative marketing model based on the capacity of the decision tree algorithm and related recommendations. Discuss research proposals and future testing directions.

The following is how the paper is organized: Section 2 Literature Review explaining the research that is the reference and basis for this research, Section 3 Research Methodology describes the methodology in detail, Section 4 Modeling and Result, Section 5 Conclusion concludes the research, and recommends additional prospective research topics.

## 2. LITERATURE REVIEW

An outline of a few explores that have been performed up to now connecting with this study is as per the following:

(1) Customer obligation to extravagance brands, Antecedents, and outcomes [15].

(2) The impact of web-based entertainment correspondence on purchaser view of brands [16].

(3) Forerunners and the results of arranged brand character in advanced education [17].

(4) An intermediate task in the brand image is to explore the impact of brand evidence and gossip on the effects of healthcare service management on customer decision-making [18].

### 2.1 Creative marketing ideas for banks

Bank presentation efforts are almost conventional, but traditional bank advertising is exhausted for the two advertisers and their customers. Adopting an innovative bank promotion strategy may change, especially when coordinating creative bank presentation ideas such as gamification, robotization, chatbots, and rewards to convince shoppers to use their controls.

#### 2.1.1 Location-based advertising for banks

Regardless of the size of the bank, the actual space is an integral part of the bank's advertising and publicity activities. Budgets, innovation capabilities, and crowds can make decisions, but with an essential presence, you can accomplish great things [19, 20].

#### 2.1.2 Benefits of gamification

Current buyers expect, try and demand applications and computerized encounters to determine the ideal freedom to present their ideas at the bank without any particular reason. Gamification is a simple and innovative way to attract buyers. Attract buyers using game standards and mechanics. Most banks coordinate gamification to drive desirable behaviors such as savings and planning. Gamification models in money management are ubiquitous. Wells Fargo includes an application as a banking component for making money-saving

decisions like lunch at home. When customers take the test, they accumulate reserve funds and then transfer them directly to their bank account [21].

#### 2.1.3 Further, develop customer service at the bank

Most customer support requests relate to basic questions such as B. Use highlights, cash demands, or data when opening files or new controls. In contrast to customer care, additional connectivity options can broaden the customer's attention and get the individual to talk [20].

#### 2.1.4 Support bank marketing campaigns

The more excellent your financial institution, the greater probably it is far. You have an abundance of examples of overcoming adversity to tug from as a function of your financial institution's showcasing efforts. In particular, you ought to. Whenever you credit score cash, deliver a domestic loan, or format CDs or exceptional reserve budget and securities structures, they have to assist the patron. Taking one's bills and providing them as a thing of a marketing attempt tries to boost your financial institution, assemble trust, and display that you care approximately purchasers. Allstate's "Worth Telling" attempt featured character patron examples of overcoming adversity, sharing meetings, patron stories, and different statistics approximately true folks that had made it in mild of credit or assistance from their financial institution [21].

#### 2.1.5 Social media personality

Web-primarily based leisure is putting down deep roots. Also, you, in all likelihood, have intense regions of power for a media presence as a thing of your financial institution's showcasing efforts. Nonetheless, leading banks use web-primarily based leisure as really standard and even "exhausting" routes in that they live conventionally. Simultaneously, distinctive forms of full-size commercial enterprise are power regions for using media characters to power consumer dedication and assemble patron devotion [21].

#### 2.1.6 Foster partnerships and leverage

Whether with faculties and schools, land places of work, or automobile income facilities, developing institutions and using them to provide well worth to the partner and the patron will help you wear out the modern selling mind on your financial institution. For instance, you could cooperate with instructive foundations to provide financial education and courses, using monetary scalability for understudies, the whole thing being equal. Additionally, you can collaborate with automobile income facilities or land places of work to allow customers to unexpectedly and successfully get credit score endorsement so customers can start a utility via the showroom with the statistics you need at the automobile [22].

#### 2.1.7 Reward users for engagement

Engagement is essential to maintaining clients involved and aware of your financial institution, whether or not online, in-person, on occasions, or at financial institution branches. Taking the time to construct praise packages for non-stop engagement is a worthwhile financial institution advertising concept and could assist you in still powering that engagement. For example, you could praise customers for checking into your financial institution on Facebook or Foursquare, create praise software for dealing with virtual banking and budgeting every day or weekly, and create rewards packages to sell the adoption of the latest packages. For example, if you're

launching a brand new cell app, you could praise folks that undertake it and use it consistently [23].

## 2.2 Decision tree algorithm

Overfitting and extensive growing bushes, pruning might be the satisfactory machine-implemented at the same time as constructing a preference tree or developing a tree [24, 25]. Choice tree pruning disposes of one subtree from a preference tree as a minimum. It continues a strategic distance from overfitting bushes, and pruned bushes are unique in grouping and checking out information. Various techniques have been proposed for preference tree pruning [26-28].

There are several hubs within the preference tree. A particular quantity of seats copes with the check on unique traits, which has unfolded to the instance length within the maximum-minimum class of leaf hubs. There are several styles of preference bushes. The brilliant maximum calculation within the enterprise is the calculation created using Rosquin Institute, mainly used to supply preference bushes. Each hub within the preference tree could have a particular check-in in the computation cycle from beginning to finish. Different branches will appear while the water check outcomes of every touch are unique, ultimately merging to a right leaf hub. The whole exam system is the research and extrudes cycle of the selection tree, which uses many elements to skip judgment on its trait [28-32].

(1) Assume there are  $n$  messages with a similar likelihood, then the average likelihood  $p$  of each letter is  $1/n$ , and how much data sent by notice is  $-\log_2(1/n)$ .

(2) Assume there are  $n$  messages whose given likelihood conveyance is  $p=(p_1, p_2, \dots, p_n)$ , then at that point, how much data is sent by this dissemination is known as the entropy of  $p$ .

(3) On the off chance that a recordset is separated into free classes of  $C_1, C_2, \dots, C_k$  as per the worth of the class trait, how much data is expected to distinguish which type a component has a place with is  $Info(T)=I(p)$ , in which  $P$  is the likelihood dissemination.

$$p = \frac{|C_1|}{|T|}, \dots, \frac{|C_k|}{|T|} \quad (1)$$

(4) Assuming the initial gap  $T$  sets as  $T_1, T_2, \dots, T_n$  per the worth of non-class quality  $x$ , the data of a component class is not set in stone by the decided weighted average value of  $T_i$ , so the weighted average worth  $Info(T_i)$  of is:

$$Info(X, T) = \sum_{i=1}^n \left( \frac{|T_i|}{|T|} I_{T_i}(T_i) \right) \quad (2)$$

(5) Data gain is the contrast between two data amounts, one of which is the data amount of a component not set in stone. The other is that the data amount of an element of still up in the air after the acquired worth of quality  $X$ . The data gain equation is:

$$Gain(X, T) = Info(T) - Info(X, T) \quad (3)$$

## 3. RESEARCH METHODOLOGY

A reasonable model of research, given speculations and the issues examined in the segment, the calculated model of examination is planned as shown in Figure 1.

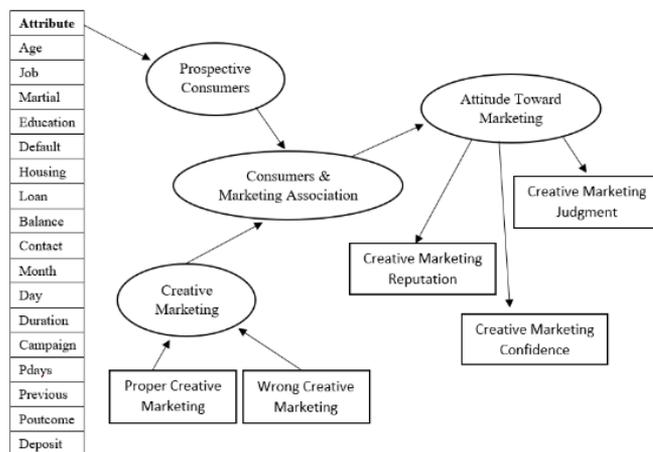


Figure 1. Reasonable model of the research

### 3.1 Exploratory data analysis (EDA)

Exploratory Data Analysis was mentioned as an option [33] instead of CDA. As mentioned earlier, EDA is an attitude or theory of how information reviews are completed, rather than the placement of appropriate methods that EDA often associates with investigator work. At EDA, the job of scientists is to look up information in as many ways as expected under a given situation until a possible "story" of the data emerges. Therefore, the "information detective" must have serious doubts about the "wrong value" of information and the prospect of accepting unexpected results when secret examples are revealed. The use of logical and factual expressions is central to EDA. Information recognition focuses on only one perspective of information presentation (design), whereas EDA covers a broader focus, as outlined in the last three components, so we do not compare them [34]. As NIST Semantech states, EDA involves several processes to fulfill its associated obligations.

#### 3.1.1 Checking assumptions on dataset

Attempt another strategy rather than a conventional examination and make another class for this. The motivation behind the course is to examine every one of the information accordingly. The class will comprehend which information type the section has a place with and picture in like manner. At the same time, it will provide appropriate information for segments whose information type is number and naturally show us the mean, standard deviation, and most minor qualities. When there is complete information, it will outwardly show us how much from which classification. The outline of the Exploratory Data Analysis is depicted in Figure 2.



Figure 2. The outline of the exploratory data analysis

Based on the results of Exploratory Data Analysis by deposits then the dataset has 12 unique values in the Job column, 3 unique values in the Marital column, 4 unique values in the Education column, 2 unique values in Has credit

in the Default column, 2 unique values in the Housing column, 2 unique values in Loan column, 3 unique values in Contact column, 12 unique values in Month column, 4 unique values in outcome column, 2 unique values in Deposit column, column age value of 18, a maximum weight of 95 with a mean of 41.2 and a standard deviation of 11.91, a minimum value of -6847 for a balance column, a maximum weight of 81204 with a mean of 1529 and a standard deviation of 3225, a minimum day column value of 1, a maximum weight of 31 with a mean 15.66 and standard deviation 8.42, duration column minimum value 2, maximum value 3881 with mean 372 and standard deviation 347, campaign column minimum value 1, maximum value 63 with mean 2,508 and standard deviation 2,722, pdays column minimum value -1, maximum value 854 with a mean of 51.3 and a standard deviation of 108.8, the previous column has a minimum value of 0, a maximum value of 58 with a mean of 0.833 and a standard deviation of 2.292 is displayed in Figure 3 and the matrix is shown in Figure 4.

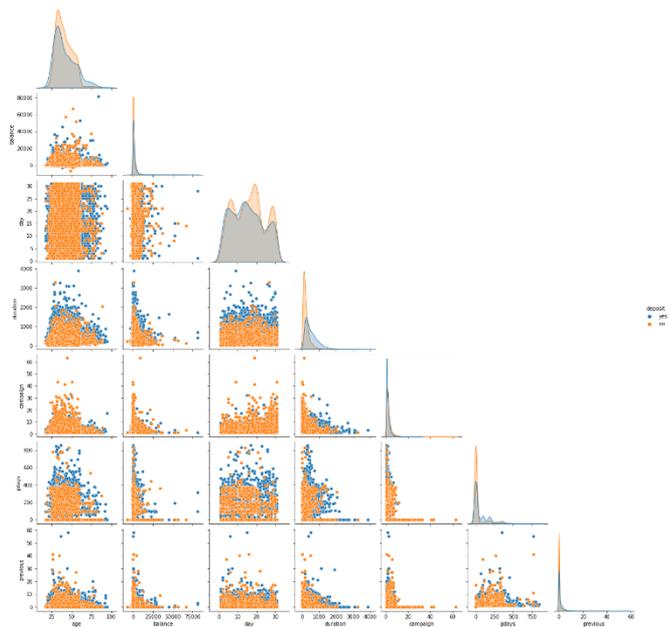


Figure 3. The results of exploratory data analysis by deposits

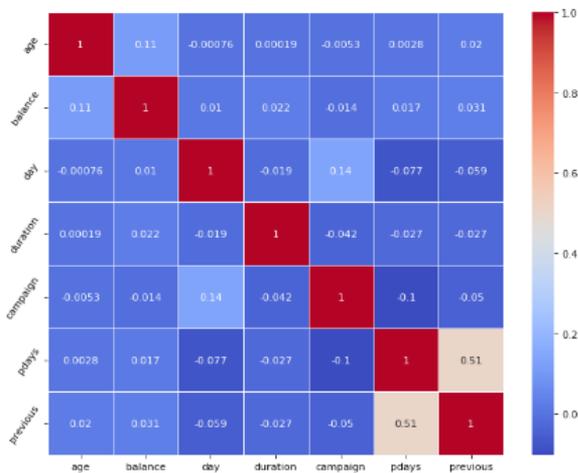


Figure 4. Matrix of exploratory data analysis by deposits

### 3.1.2 Spotting outliers on dataset

Exploration of data by determining Target Data and Spotting Outliers to assess the relationship between attributes

in the dataset in determining predictions in carrying out creative promotions. Then get the result Management has more likelihood to put a deposit. That most dynamic balance and most deposit accounts are mostly management is displayed in Figure 5.



Figure 5. Links the relationship between work and deposit

Determine the relationship between marital status and job as management impacts the deposit held, and the result is displayed in Figure 6.

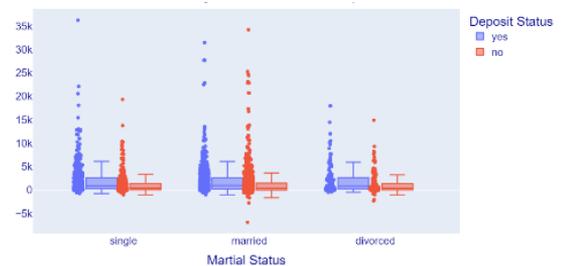


Figure 6. Relationship between marital status and management

Determine the relationship between marital status and job as the technician impacts the deposit held, and the result is displayed in Figure 7.

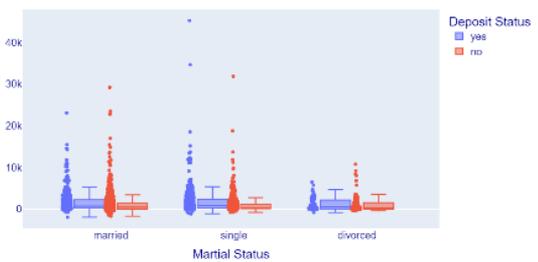


Figure 7. Relationship between marital status and technician

Determine the relationship between marital status and job as blue-collar impacts the deposit held, and the result is displayed in Figure 8.

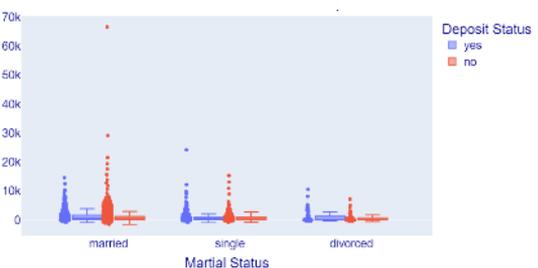


Figure 8. Relationship between marital status and blue-collar

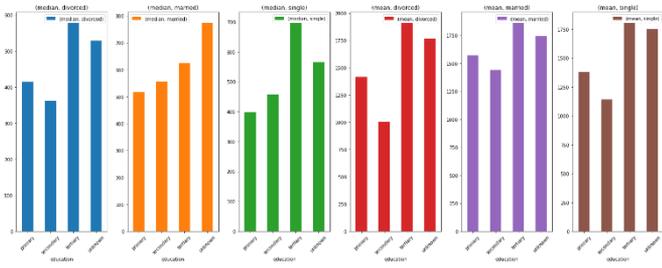


Figure 9. Educational level and marital status

After obtaining the relationship between employment and marital status in the impact on the deposit held, the next step is to determine the relationship between marital status and the last education held. From the calculations, it is found that educational level and marital status are pretty influential on the number of deposits held, as shown in Figure 9.

Then the next step is to determine the cause of the number of deposits consumers have and the relationship between what has been discussed previously with the reason for having at least a small warranty. Stating that the result of the cause of the deposit amount owned is Marital Status because the impact of a divorce has a significant effect on the balance of the individual. Education because the level of education also significantly impacts the amount of credit a prospect has. Loans because whether the candidate has a previous loan substantially affects the amount of balance they have, as shown in Figure 10.

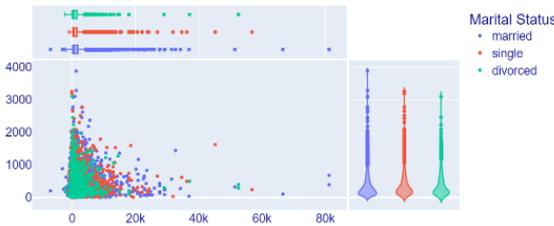


Figure 10. Determine the cause of the number of deposits consumers

### 3.1.3 Transformation data on dataset

Before the Transforming dataset, the examination objective is to have practically the same measure of information for each class by using the counterplot value\_counts. So can be seen that the analysis results stated that as many as 5873 data did not have a deposit, or the promotions carried out were less successful. As many as 5289 had guarantees or promotions carried out quite successfully in Figure 11 and Figure 12.

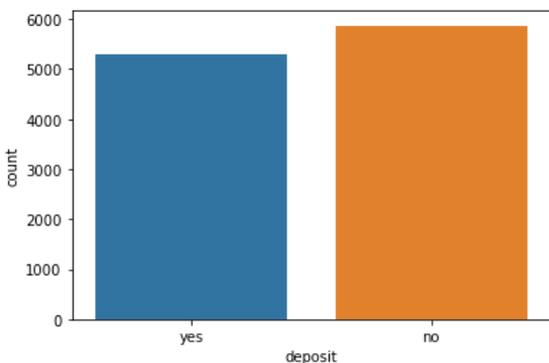


Figure 11. Class by using the countplot value\_counts

```
Out[14]: no    5873
         yes    5289
         Name: deposit, dtype: int64
```

Figure 12. Analysis results stated

Divide the information needed to encode the unmitigated section. For this reason, we are going to utilize two distinct techniques. Mark Encoder is the picture result for the name encoder. Label Encoding is a famous encoding strategy for dealing with downright factors. Each name is relegated to a unique number in light of sequential requests. One-Hot Encoder is one more well-known procedure for treating aspects. It bakes extra highlights in light of the number of exceptional qualities in the downright element.

## 4. MODELING AND RESULT

Information perception intends to develop the straightforwardness of the insightful interaction further. While speculation analyzers present the information to confounded calculations without understanding how Wilk's Lambda and the p-esteem are registered, information visualizers could straightforwardly "see" the example on the diagram. Not in the least do information experts like the straightforwardness and interpretability that outcomes from representation; however, most educators and speakers likewise prefer to utilize diagramming methods to introduce individual results and complex information structures substantially and engagingly.

```
In [21]: param_grid = {'criterion': ['gini', 'entropy'], 'max_depth': [5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100], 'min_samples_split': [2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100], 'min_samples_leaf': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100], 'max_features': ['sqrt', 'log2', 'best'], 'random_state': [0]}

In [22]: from sklearn import GridSearchCV
param_grid = param_grid
scoring = 'roc_auc'
grid_search = GridSearchCV(estimator=DecisionTreeClassifier(),
                           param_grid=param_grid,
                           scoring=scoring,
                           cv=5,
                           n_jobs=-1,
                           refit=True,
                           verbose=0)
grid_search.fit(X_train, y_train)

In [23]: print(grid_search.best_estimator_.get_params())
Decision Tree Parameters: {'criterion': 'entropy', 'max_depth': None, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 2, 'random_state': 0, 'splitter': 'best'}

In [24]: dt_tuned = DecisionTreeClassifier(criterion='gini',
                                         min_samples_leaf=4,
                                         min_samples_split=4,
                                         max_features='best',
                                         random_state=0)

In [25]: dt_tuned.fit(X_train, y_train)

In [26]: dt_tuned.score(X_test, y_test)
Out[26]: 0.822044889426
```

Figure 13. Decision tree modeling

```
In [31]: print(classification_report(y_test, pred))
              precision    recall  f1-score   support
0               0.85         0.81         0.83         1471
1               0.80         0.84         0.82         1319

 accuracy          0.82         0.82         2790
 macro avg         0.82         0.82         0.82         2790
 weighted avg      0.82         0.82         0.82         2790

In [32]: fig, ax = plt.subplots(figsize=(10, 10))
plot_confusion_matrix(dt_tuned, scaled_test_y, test_ax=ax)
Out[32]: <sklearn.metrics._plot_confusion_matrix.ConfusionMatrixDisplay at 0x2648e94148d>
```

Figure 14. Precision value, recall value, and the decision matrix

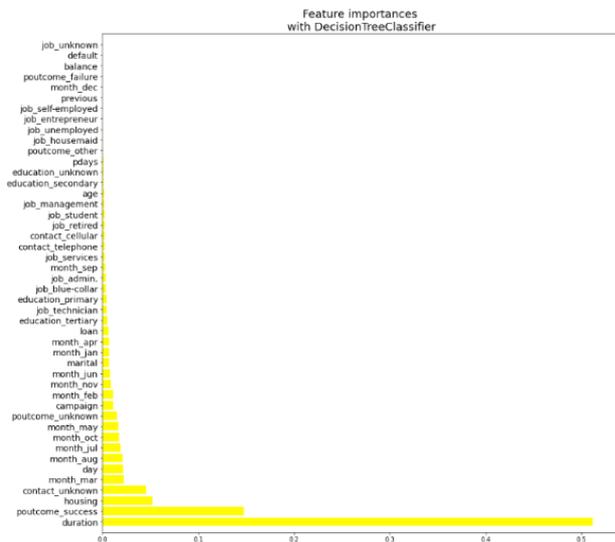


Figure 15. Feature importances

Decision Tree modeling obtained prediction accuracy results of 82%, with a precision value of 85% and a recall value of 84%, which can be seen in Figure 13 and Figure 14.

The feature importances that are the ballast of successful marketing is duration with the most decisive influence, an outcome with the second-highest force, and housing with the third-strongest power, as shown in Figure 15.

Decide Best Threshold use accuracy and review tradeoff strategy. Choose which edge to utilize and return the scores rather than expectations. Precision and Recall Tradeoff with Best Precision and Recall Balance is at the 0.58 threshold in Figure 16.

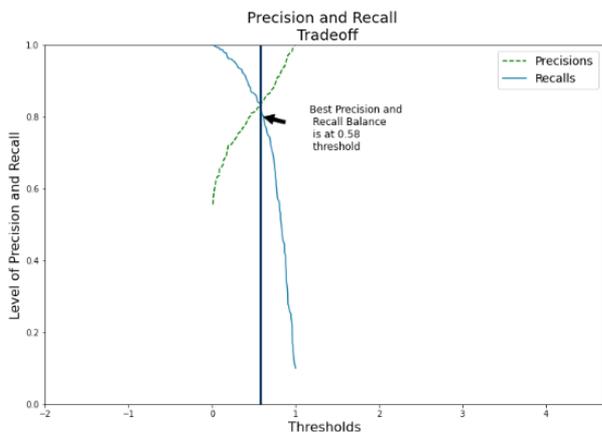


Figure 16. Best precision and recall

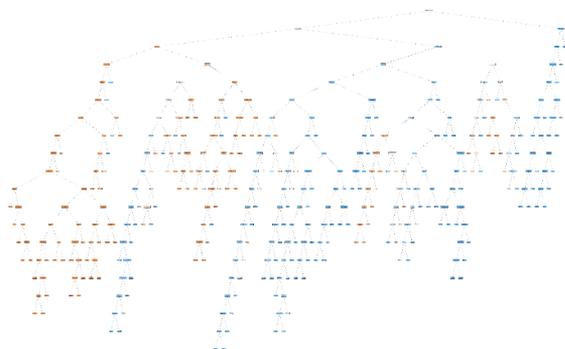


Figure 17. Tree plot

Tree Plot test counts that are shown are weighted with any sample\_weights available. The perception is fit consequently to the size of the pivot. Utilize the fig size or dpi contentions. Figure to control the size of the delivery. A tree was too enormous to investigate. Snap or save the plot and see the hubs and leaf—shows how the choice made by the model can be seen in Figure 17.

## 5. CONCLUSIONS

Long periods of Marketing Activity the period of the most significant level of showcasing movement was May. Nonetheless, this was the month that potential clients would generally offer oddball term stores (Lowest effective rate: -34.49%). It will be insightful for the bank to concentrate the advertising effort during the long periods of March, September, October, and December for the following promoting action. (December ought to be getting looked at because it was the month with the most reduced showcasing movement, there may be a justification for why December is the least).

Irregularity, Potential clients picked to subscribe term stores during the times of fall and winter. The following showcasing effort ought to concentrate its movement throughout these seasons.

Crusade Calls, A strategy should be executed that expresses that something like three calls ought to be applied to a similar possible client to save time and exertion in getting new expected clients. The more we call a similar potential client, the more probable the person will decline to open a term store.

Age Category, the next showcasing effort of the bank ought to target expected clients in their 20s or more youthful and 60s or more established. The most youthful class had a 60% possibility of subscribing to a term store, while the oldest type had an 82% possibility of subscribing to a term store. It would be perfect if the bank tended to these two classes for the following effort and, like this, improve the probability of more term store subscriptions.

Occupation, of course, potential clients that were understudies or resigned were probably going to subscribe to a term store. Left people will generally have more term stores to acquire some money through premium installments. Remember that term stores are momentary credits. The person (for this situation, the retiree) makes a deal to avoid pulling out the money from the bank until a specific date concurs between the individual and the monetary organization. After that time, the individual gets their capital back and its advantage on credit. Resigned people tend not to spend vast amounts of money, so they are more likely to give their cash something to do by loaning it to the monetary establishment. Understudies were the other gathering that used to subscribe term stores.

House Loans and Balances, Potential clients in the low equilibrium and no equilibrium classification were bound to have a house credit than individuals in the standard and high equilibrium classification. What's the significance here of having a house credit? Implies that the potential client has financial trade-offs to repay its home credit, and subsequently, there is no money for the individual to subscribe to a term store account. Nonetheless, we see that possible clients with average and high balances are less inclined to have a house credit and hence, bound to open a term store. In conclusion, the following showcasing effort ought to zero in on people of standard and high adjusts to improve the probability of subscribing to a term store.

Foster a Questionnaire during the Calls; since the span of the call is the element that most decidedly corresponds with regardless of whether a potential client will open a term store, by giving a fascinating questionnaire to possible clients during the calls, the discussion length could increment. It doesn't guarantee that the potential client will subscribe to a term store! By the by, we free nothing by carrying out a procedure that will expand the potential client's degree of commitment, prompting an increment likelihood of subscribing to a term store. Like this, an expansion in viability for the following showcasing effort the bank will execute.

Target people with a higher span (over 375). Target the objective gathering that is better than expected in length, there is a profound probability that this target gathering would open a term store account. The likelihood that this gathering would open a term store account is 78%, which is high. Would permit that the achievement pace of the following showcasing effort would find success.

Even though those strategies produced good results, the Decision Tree model has its own set of advantages and disadvantages. Decision Tree model architecture is not too complex, but it is the best solution when accuracy is critical. Even though the Decision Tree model produces satisfactory results, time complexity must be considered because both contain a model with high time complexity. The decision tree model is a mechanism as well, and it has the lowest time complexity. To further study the Decision Tree model, particularly the combinations of multiple models, more research on the following subjects should be performed. For optimal solutions, various Machine Learning approaches are compared and examined.

## REFERENCES

- [1] Barham, J., Chitemi, C. (2009). Collective action initiatives to improve marketing performance: Lessons from farmer groups in Tanzania. *Food Policy*, 34(1): 53-59. <https://doi.org/10.1016/j.foodpol.2008.10.002>
- [2] Das, G., Agarwal, J., Malhotra, N.K., Varshneya, G. (2019). Does brand experience translate into brand commitment? A mediated-moderation model of brand passion and perceived brand ethicality. *Journal of Business Research*, 95: 479-490. <https://doi.org/10.1016/j.jbusres.2018.05.026>
- [3] Haumer, F., Kolo, C., Reiners, S. (2020). The impact of augmented reality experiential marketing on brand equity and buying intention. *Journal of Brand Strategy*, 8(4): 368-387.
- [4] Shukla, P., Banerjee, M., Singh, J. (2016). Customer commitment to luxury brands: Antecedents and consequences. *Journal of Business Research*, 69(1): 323-331. <https://doi.org/10.1016/j.jbusres.2015.08.004>
- [5] De Chernatony, L., Segal-Horn, S. (2001). Building on services' characteristics to develop successful services brands. *Journal of Marketing Management*, 17(7-8): 645-669. <https://doi.org/10.1362/026725701323366773>
- [6] Parayitam, S., Kakumani, L., Muddangala, N.B. (2020). Perceived risk as a moderator in the relationship between perception of celebrity endorsement and buying behavior: evidence from rural consumers of India. *Journal of Marketing Theory and Practice*, 28(4): 521-540. <https://doi.org/10.1080/10696679.2020.1795687>
- [7] Davenport, T., Guha, A., Grewal, D., Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1): 24-42. <https://doi.org/10.1007/s11747-019-00696-0>
- [8] Hagen, L., Uetake, K., Yang, N., et al. (2020). How can machine learning aid behavioral marketing research? *Marketing Letters*, 31(4): 361-370. <https://doi.org/10.1007/s11002-020-09535-7>
- [9] Ma, L., Sun, B. (2020). Machine learning and AI in marketing—Connecting computing power to human insights. *International Journal of Research in Marketing*, 37(3): 481-504. <https://doi.org/10.1016/j.ijresmar.2020.04.005>
- [10] Vermeer, S.A., Araujo, T., Bernritter, S.F., van Noort, G. (2019). Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media. *International Journal of Research in Marketing*, 36(3): 492-508. <https://doi.org/10.1016/j.ijresmar.2019.01.010>
- [11] Rizal, A., Handzah, V.A.P., Kusuma, P.D. (2022). Heart sounds classification using short-time Fourier transform and gray level difference method. *Ingénierie des Systèmes d'Information*, 27(3): 369-376. <https://doi.org/10.18280/isi.270302>
- [12] Gao, J., Ismail, N., Gao, Y.J. (2022). Computer big data analysis and predictive maintenance based on deep learning. *Ingénierie des Systèmes d'Information*, 27(2): 349-355. <https://doi.org/10.18280/isi.270220>
- [13] Kevill, A., Trehan, K., Harrington, S., Kars-Unluoglu, S. (2021). Dynamic managerial capabilities in micro-enterprises: Stability, vulnerability and the role of managerial time allocation. *International Small Business Journal*, 39(6): 507-531. <https://doi.org/10.1177/0266242620970473>
- [14] Hariyadi, Mutira, P., Nguyen, P.T., Iswanto, I., Sudrajat, D. (2020). Traveling salesman problem solution using genetic algorithm. *Journal of Critical Reviews*, 7(1): 56-61. <http://dx.doi.org/10.22159/jcr.07.01.10>
- [15] Mihic, S., Radjenovic, D., Supic, D. (2013). Consumer behaviour-building marketing strategy. *Metalurgia International*, 18(8): 116.
- [16] Foroudi, P., Dinnie, K., Kitchen, P.J., Melewar, T.C., Foroudi, M.M. (2017). IMC antecedents and the consequences of planned brand identity in higher education. *European Journal of Marketing*, 51(3). <https://doi.org/10.1108/EJM-08-2015-0527>
- [17] Sohrabi, M.S., Mahdavi, N. (2019). The effect of participatory ergonomic interventions on the level of satisfaction and the motivating potential score of employees of a medical diagnostic laboratory in Isfahan. *Iranian Journal of Ergonomics*, 7(2): 1-10.
- [18] Shin, W., Lin, T.T.C. (2016). Who avoids location-based advertising and why? Investigating the relationship between user perceptions and advertising avoidance. *Computers in Human Behavior*, 63: 444-452. <https://doi.org/10.1016/j.chb.2016.05.036>
- [19] Sochor, R., Schenk, J., Fink, K., Berger, J. (2021). Gamification in industrial shopfloor—development of a method for classification and selection of suitable game elements in diverse production and logistics environments. *Procedia CIRP*, 100: 157-162. <https://doi.org/10.1016/j.procir.2021.05.024>
- [20] Hanukov, G. (2022). Improving efficiency of service systems by performing a part of the service without the

- customer's presence. *European Journal of Operational Research*, 302(2): 606-620. <https://doi.org/10.1016/j.ejor.2022.01.045>
- [21] Zhang, Y., Tian, Y. (2021). Choice of pricing and marketing strategies in reward-based crowdfunding campaigns. *Decision Support Systems*, 144. <https://doi.org/10.1016/j.dss.2021.113520>
- [22] Sendogdu, A.A., Koyuncuoglu, O. (2022). An analysis of the relationship between university students' views on distance education and their computer self-efficacy. *International Journal of Education in Mathematics, Science and Technology*, 10(1): 113-131. <https://doi.org/10.46328/ijemst.1794>
- [23] Li, F., Lu, H., Hou, M., Cui, K., Darbandi, M. (2021). Customer satisfaction with bank services: The role of cloud services, security, e-learning and service quality. *Technology in Society*, 64: 101487. <https://doi.org/10.1016/j.techsoc.2020.101487>
- [24] McNeish, R., Rigg, K.K., Tran, Q., Hodges, S. (2019). Community-based behavioral health interventions: Developing strong community partnerships. *Evaluation and Program Planning*, 73: 111-115. <https://doi.org/10.1016/j.evalprogplan.2018.12.005>
- [25] Shahbaznezhad, H., Dolan, R., Rashidirad, M. (2021). The role of social media content format and platform in users' engagement behavior. *Journal of Interactive Marketing*, 53: 47-65. <https://doi.org/10.1016/j.intmar.2020.05.001>
- [26] Windeatt, T., Ardeshir, G. (2001). An empirical comparison of pruning methods for ensemble classifiers. In *International Symposium on Intelligent Data Analysis*, pp. 208-217. [https://doi.org/10.1007/3-540-44816-0\\_21](https://doi.org/10.1007/3-540-44816-0_21)
- [27] Esposito, F., Malerba, D., Semeraro, G., Kay, J. (1997). A comparative analysis of methods for pruning decision trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5): 476-491. <https://doi.org/10.1109/34.589207>
- [28] Han, J., Pei, J., Tong, H. (2022). *Data mining: concepts and techniques*. Morgan kaufmann. <https://doi.org/10.1016/C2009-0-61819-5>
- [29] Bertsimas, D., Dunn, J. (2017). Optimal classification trees. *Machine Learning*, 106(7): 1039-1082. <https://doi.org/10.1007/s10994-017-5633-9>
- [30] Maimon, O., Rokach, L. (2005). Introduction to knowledge discovery in databases. In: *Data mining and knowledge discovery handbook*. Springer, Boston, MA, pp. 1-17. [https://doi.org/10.1007/0-387-25465-x\\_1](https://doi.org/10.1007/0-387-25465-x_1)
- [31] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3). <https://doi.org/10.1609/aimag.v17i3.1230>
- [32] Abelairas-Etxebarria, P., Astorkiza, I. (2020). From exploratory data analysis to exploratory spatial data analysis. *Mathematics and Statistics*, 8(2). <https://doi.org/10.13189/ms.2020.080202>
- [33] Watts, B.R., Erickcek, G.A., Duritsky, J., O'Brien, K., Robey, C., Robey, J. (2011). What should EDA fund? Developing a model for preassessment of economic development investments. *Economic Development Quarterly*, 25(1): 65-78. <https://doi.org/10.1177/0891242410377084>
- [34] Yu, C.H., Stockford, S. (2002). Evaluating spatial-and temporal-oriented multi-dimensional visualization techniques. *Practical Assessment, Research, and Evaluation*, 8(1): 17.