Evaluating the Predictive Modeling Performance of Kernel Trick SVM, Market Basket Analysis and Naive Bayes in Terms of Efficiency

SAFIYE TURGAY[1], METEHAN HAN[1], SUAT ERDOĞAN[1], ESMA SEDEF KARA[2], RECEP YILMAZ[1]
[1]Department of Industrial Engineering,
Sakarya University,
54187, Esentepe Campus Serdivan-Sakarya,
TURKEY

[2]Rüstempaşa Mahallesi, İpekyolu Caddesi, No. 120,
54600, Sapanca-Sakarya,
TURKEY

Abstract: - Among many corresponding matters in predictive modeling, the efficiency and effectiveness of the several approaches are the most significant. This study delves into a comprehensive comparative analysis of three distinct methodologies: Finally, Kernel Trick Support Vector Machines (SVM), market basket analysis (MBA), and naive Bayes classifiers invoked. The research we aim at clears the advantages and benefits of these approaches in terms of providing the correct information, their accuracy, the complexity of their computation, and how much they are applicable in different domains. Kernel function SVMs that are acknowledged for their ability to tackle the problems of non-linear data transfer to a higher dimensional space, the essence of which is what to expect from them in complex classification are probed. The feature of their machine-based learning relied on making exact confusing decision boundaries detailed, with an analysis of different kernel functions that more the functionality. The performance of the Market Basket Analysis, a sophisticated tool that exposes the relationship between the provided data in transactions, helped me to discover a way of forecasting customer behavior. The technique enables paints suitable recommendation systems and leaders to make strategic business decisions using the purchasing habits it uncovers. The research owes its effectiveness to processing large volumes of data, looking for meaningful patterns, and issuing beneficial recommendations. Along with that, an attempt to understand a Bayes classifier of naive kind will be made, which belongs to a class of probabilistic models that are used largely because of their simplicity and efficiency. The author outlines the advantages and drawbacks of its assumption in terms of the attribute independence concept when putting it to use in different classifiers. The research scrutinizes their effectiveness in text categorization and image recognition as well as their ability to adapt to different tasks. In this way, the investigation aims to find out how to make the application more appropriate for various uses. The study contributes value to the competencies of readers who will be well informed about the accuracy, efficiency, and the type of data, domain, or problem for which a model is suitable for the decision on a particular model choice.

Key-Words: - Kernel Trick, Support Vector Machines, Market Basket Analysis, Naive Bayes Classifiers, Predictive, Modeling.

Received: August 24, 2023. Revised: December 23, 2023. Accepted: February 24, 2024. Published: April 9, 2024.

1 Introduction

Predictive modeling is an indispensable tool in both analytical processes regarding present conditions around several industries. In the present age of massive data warehousing and complex problems, the correct technique either, predictive or descriptive, will bring into the light hidden patterns in datasets and help thus to make intelligent decisions on them. The experiment is devoted to the study of three techniques - SVM, MBA, and Naive Bayes classifiers, which are examples of various methods having different goals and scope. Thus, Kernel Trick SVMs have become increasingly popular due to the fact they can handle extremely complicated classification problems whilst projecting the data onto higher surface spaces. Such algorithms rely on nonlinear kernels to identify patterns of high complexity, which cannot be detected by conventional linear classifiers that are restricted to simple decision frontlines. Their accuracy has been praised but the computational requirements, which have aroused the curiosity of users in investigating the trade-offs of this technique and optimal implementation, have provoked interest in it.

Marketing Basket Analysis, acting in the important role of customer behavior prediction, supplies buyers with data about the transactional relationship of merchandise and purchasing patterns. MBA works through uncovering brands, products, and consumer groups' binds in the purchase history data, this way plays a crucial role in recommendation processes and strategic planning in retail and e-commerce.

Marking its power to explain with the "bought-together" action, but the generalization of the approach beyond the transactional data as an intriguing research direction is noteworthy. Expert Baye's classifiers, a Bayesian motivated tool with a long tradition of use, are well known for their simplicity and efficient computational methods. These methods are based on the fact that the attributes stand on their own and are why they perform well in text categorization, spam filtering, and other tasks for which quick and understandable prediction is paramount.

Nevertheless, it has an idea for probability that may not be relevant when it comes to working with data structures that are considered complex and effective in the event of correlated data types.

Because there are several game-changing methods applied, this research aims to investigate the efficacy of the two approaches in the context of predictive modeling. Our research aims to shed light on the following key aspects: Our research aims to shed light on the following key aspects:

Predictive Accuracy: How do Kernel Trick SVMs, Market Basket Analysis, and Naive Bayes classifiers fare in terms of predictive accuracy across various datasets and problem domains?

Computational Complexity: What are the computational demands associated with each technique, and how does the efficiency of implementation affect their scalability to larger datasets?

Applicability and Generalization: To what extent can these techniques be applied to different data types and problem contexts? How well do they generalize across diverse scenarios?

Trade-offs and Optimal Use Cases: What are the strengths and limitations of each technique? In what scenarios does one technique outshine the others, and what are the underlying reasons?

Through empirical experiments and comprehensive evaluation, we seek to provide practitioners, researchers, and decision-makers with insights into the suitability and performance characteristics of each technique. By navigating the intricate landscape of predictive modeling, this study aims to guide informed choices in selecting the most appropriate method based on the unique demands of a given problem.

2 Literature Survey

In this literature survey, we explore key studies that have contributed to the understanding of Kernel Trick Support Vector Machines (SVM) [1], [2], [3], [4], [5], [6], Market Basket Analysis (MBA) [7], [8], [9], [10], [11], [12], [13], [14], [15] and Naive Bayes classifiers [16], [17], [18], [19], [20], in the context of predictive modeling.

SVM with Kernel Trick has been praised for its proficiency in handling nonlinear classification tasks, [21], [22], [23]. Therefore, these characteristics make Fuzzy SVM superb and can also be applied to any type of dataset. As a whole, this ability of deep learning in solving nonlinear data became the reason for its success. The most fundamental feature of SVM is its capacity to successfully separate the optimal hyperplane. It has been emphasized by some researchers that evaluation mechanisms and soft tactics of SVMs such as RBF and polynomial kernels using ds can bring linearly inseparable models to different decision boundaries, [24], [25]. At the same time, by introducing Mercer kernels, the scope of SVM applications was expanded and the areas covered were; image recognition, bioinformatics and natural language coding were included, [26]. In addition, radial basis functions were also examined extending data sets, and the relation between kernel selection, model complexity, and predictive accuracy are well-understood, [27], [28], [29], [30]. Market Basket Analysis is the development of association rule mining, which is based on statistical and computational methods. In a different study on text categorization, emphasis was placed on the technique of processing large-volume, high-dimensional and sparse data, [31], [32]. While Naive Bayes is emphasized for supervised learning of text categorization, a combination of KNN and SVM classifiers with Naive Bays has also been studied, [33], [34].

As far as comparative studies are concerned, the previously done study focused on the particular aspects of these methods which have been discussed in Section 1. SVMs have been compared to other classifiers like decision trees and neural networks in the perspective of their accuracy and generalization yet it has been found that they perform exceedingly well, [35].

Assessing the predictive power of Kernel Trick SVM, Market Basket Analysis, and Naive Bayes techniques based on efficiency means covering the advantages and disadvantages of each scenario when it comes to specific criteria. The SVM with Kernel Trick renders the model competent at mapping non-linear structure relations within the data. This is so because it can construct abundantly complex decision boundaries. SVMs perform well in high-dimensional spaces and can be used when you work on projects in which the number of features is large. They are less sensitive to extreme or atypical training data than other techniques, [36]. Hence, SVM is often computationally expensive, especially compared to handling large datasets, where the model fitting takes a lot of time. Particularly with the selection of the kernel functions, optimization of hyperparameters and applying ideas like the stochastic gradient descent method, [37], their efficiency can be boosted. Market Basket Analysis stands out among other marketing techniques for an examination of relationships and patterns in transactional data, for example, the instances when the target audiences buying certain products tend to overlap with that of another product. Worryingly, the generated associations' rules are highly comprehensible, and immediately inform us about the customers' behavior and desires. Computers can implement the understanding of consumer behavior in real-time, proving to be a time and resource-efficient tool in large data sets with the development of algorithms like the Apriori or FP-growth algorithms. Market Basket Analysis itself is a good one and it looks for the candidates that are associated with each other, but it is not able to predict particular characteristics of the visitors to the website. It might not be an appropriate case for datasets that are sensitive to subtle fluctuations in the behavior of the target variables.

Such Neural networks are fast and easy to execute which makes them relevant for big data and real-time tasks. Naive Bayes works well when applied to text classification problems and serves as the reason why it's adopted in many spam filtering and sentiment analysis systems. Bayes works well if there is not very much data and is less affected by systematic irrelevancy of features. Naïve Bayes being faster than other classifiers is, however, formal independence assumptions may narrow the performance of the model in scenarios where the data is correlated to the highest possible extent.

The consistency issue of your method should agree with the type of your data. It is possible that the dataset was made using the non-linear relationship and therefore a support vector machine with the kernel approach had some benefits. In the case of transactional data, market basket analysis could be more feasible to employ. It is often the case that the naive bayes method as well as the market basket follow an interpretable model rather than SVM which model can be complex. Concerning the computational capability, let us have a closer look. Naive Bayes and Market Basket Analysis often take less resources than the SVM algorithm and this can be an advantage for big-scale applications.

Eventually demonstrates the mobile Kernel Trick SVM, Market Basket Analysis, and Naive Bayes in the individual specific requirement of dataset characteristics, and the available computation resources when relating to a problem and solving it. The methods are all distinct amongst themselves and each of them has merits and demerits. In this case, the method is based on the task of predictive modeling and it might differ from case to case. This line of investigation, thus, not only reveals the need for but also demonstrates the lack of, the in-depth analysis that spreads throughout all these methods. This literature survey underscores the importance of our study, aiming to fill the gap by providing an in-depth analysis that spans across these techniques.

3 Methodology

To comprehensively compare the efficiency of Kernel Trick Support Vector Machines (SVM), Market Basket Analysis (MBA), and Naive Bayes classifiers in predictive modeling, we outline a systematic methodology encompassing data preparation, experimentation, evaluation metrics, and statistical analysis.

1. Data Preparation:

We select a diverse set of datasets representing different domains, data types, and complexities. For SVM and Naive Bayes, we consider text classification and image recognition datasets. For MBA, we use transactional data capturing customer purchases.

2. Experimentation:

For each technique, we conduct separate experiments using the preprocessed datasets:

- Kernel Tricks SVMs: We utilize cross-validation to determine optimal hyperparameters, including regularization parameters and kernel parameters. We explore SVM implementations with various libraries.

- Market Basket Analysis: We employ the Apriori algorithm to discover frequent item sets and association rules. We experiment with different support and confidence thresholds to extract meaningful rules. The analysis includes assessing the lift and support of the discovered associations.
- Naive Bayes Classifiers: We implement both Gaussian and multinomial variants of Naive Bayes for continuous and discrete data, respectively. We assess the impact of attribute independence assumptions and evaluate the classifiers' performance on different feature spaces.

3. Comparative Analysis:

We compare the outcomes of the experiments across techniques, considering predictive accuracy, computational complexity, and generalizability. We highlight scenarios where one technique outperforms the others, taking into account the strengths and limitations identified in the literature.

4. Sensitivity Analysis:

To test for each method's sensitivity towards different influencing parameters, a sensitivity analysis is also carried out. For SVM, this study examines how choosing a particular kernel can affect performance and what better kernel leads to better performance. As for the MBA, our focus is the evaluation of the support and confidence influence when discarding and retaining nodes, respectively. From the point of view of Naive Bayes, the article deals with the connection between attributes and whether it is a classifier or not.

This methodology aims to conduct deep and reasonably impartial research into the broad panorama of three prediction skills including Kernel Trick SVMs, Market Basket Analysis, and Naive Bayes classifier. This can be achieved with the help of the systematic approach that yields necessary meanings into the differences in performance of these techniques on various parameters to offer the practitioners decision-making options based on the exact requirements of their predictive modeling attire.

3.1 Mathematical Modeling

In this section, we provide an overview of the mathematical formulations underlying each technique: Examples of such technologies are SVM (Support Vector Machine) Kernels, MBA (Market Basket Analyses), and Naïve Bayes Classifiers. This is the groundwork for grasping the inner mechanisms and a method of guessing the output of the method.

3.1.1 Kernel Trick SVM

The mapping of data into a higher-dimensional space is accomplished using the application of the Kernel Trick to ensure reparability by linear methods. The decision function for a binary classification problem is given by: The decision function for a binary classification problem is given by:

$$f(x) = sign(\sum_{i=1}^{N} y_i \, \alpha_i K(x_i, x) + b) \qquad (1)$$

That is when N is the number of support vectors, the coefficients are $\alpha_i = y_i$. The data points are represented by x_i and input points for the kernel function is x. $K(x_i, x)$ is the Kernel function. And b is the bias term.

Common kernel functions include the linear Kernel ($K(x_i, x)= x_i^T, x$), polynomial Kernel ($K(x_i, x)= \left(x_i^T, x + c)\right)^d$) and radial basis function (RBF) Kernel ($K(x_i, x)= (-\gamma\|x_i - x\|^2)$).

3.1.2 Market Basket Analysis

It aims to discover associations between items in transactional data. The Apriori algorithm, one of the fundamental approaches in MBA, calculates the support and confidence of itemsets and association rules. The support of an itemset X is defined as the proportion of transactions that contain X, while the confidence of a rule X→Y is the probability that items in Y are bought given that items in X are bought. Mathematically:

$$\text{Support(X)} = \frac{\text{Transactions containing X}}{\text{Total transactions}} \qquad (2)$$

$$\text{Confidence(X}\rightarrow\text{Y)} = \frac{\text{Support (X}\cup\text{Y}}{\text{Support(X)}} \qquad (3)$$

The lift measure (Lift(X→Y)) indicates how much more likely items in Y are bought when X is bought, compared to when Y is bought regardless of X.

3.1.3 Naive Bayes Classifiers

It classifiers are probabilistic models based on Bayes' theorem. Given a feature vector $x = \{x_1, x_2, \ldots, x_n\}$ and a class C, it estimates the posterior probability of C given x using the Bayes' theorem:

$$P(C|x) = \frac{P(C).P(x\lfloor C)}{P(x)} \qquad (4)$$

The "naive" assumption is that the attributes are conditionally independent given the class label, simplifying the estimation of $\mathrm{P}(C|x)$. For discrete attributes, this leads to the use of probability mass functions. For continuous attributes, Gaussian Naive Bayes assumes that each attribute follows a Gaussian distribution.

By estimating $\mathrm{P}(x|C)$ for each class, it assigns the input x to the class with the highest posterior probability.

Here, in the mathematical-modeling, section we've given the heart and bones of mathematics behind each technique.

3.1.4 An Algorithm Overview

We will be stepping through the mathematical forms that structure up Kernel Politics SVM, Market Basket Analysis (MBA), and the Naive Bayes classifiers. The intricacies of these algorithms and how they are sequentially stacked reveal their high level of efficiency and predictive capabilities.

3.2 Kernel Trick Support Vector Machines (SVM)

Input: Labeled training data $x = \{x_1, x_2, \dots, x_n\}$ $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ where x_i is the feature vector and y_i is the class label.
Select a kernel function K (e.g., linear, polynomial, RBF) and compute the kernel matrix.
Solve the dual optimization problem to find the Lagrange multipliers $\alpha_1, \alpha_2, \dots, \alpha_n$.
Identify support vectors by finding non-zero α_ii values.
Calculate the bias term b using support vectors and their associated class labels.
For a new data point x, use the decision function
$f(x) = sign(\sum_{i=1}^{N} y_i\, \alpha_i K(x_i, x) + b)$ to predict the class label.

3.3 Market Basket Analysis (MBA)

Input: Transactional data containing sets of items bought in each transaction.
Calculate the support of each item by counting how many transactions it appears in.
Generate frequent itemsets: Starting with frequent itemsets of size 1, join them to create larger itemsets and prune those with support below a threshold.
Extract association rules from frequent itemsets based on confidence thresholds.
Calculate lift values to measure the strength of associations between items.
Present the discovered rules and associations to aid in recommendation and decision-making.

3.4 Naive Bayes Classifiers

Input: Labeled training data $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ where x_i is the feature vector and y_i is the class label.
1. For each class C, calculate the prior probability $\mathrm{P}(C)$ by counting the frequency of each class label in the training set.
2. Estimate the likelihood $\mathrm{P}(x|C)$ for each attribute in the feature vector x using appropriate probability distributions (e.g., Gaussian, multinomial).
3. Apply Bayes' theorem to calculate the posterior probabilities $\mathrm{P}(C|x)$for each class.
4. Assign the input x to the class with the highest posterior probability.

Application examples of the Kernel Trick SVM approach proposed in this study are image classification in healthcare; financial fraud detection, retail and customer behavior analysis with the market basket approach, e-commerce cross-selling transactions, spam email classification with the Naive Bayes approach, sentiment analysis in social media. In the comparative analysis performed with these methods, Kernel Trick SVM may require significant computational resources and limit its scalability for large datasets. Market Basket Analysis can effectively handle large transactional datasets, especially with efficient algorithms such as FP-growth. Simple and fast, Naive Bayes is highly scalable and suitable for real-time applications. Naive Bayes and Market Basket Analysis often provide more interpretable results compared to the complex decision boundaries produced by Kernel Trick SVM. This interpretability is crucial in areas where understanding and confidence in the model are crucial. Depending on the nature of the data, the three methods offer different advantages. Researchers need to choose the method that is compatible with the specific characteristics of their datasets, encouraging a nuanced approach to model selection. By studying these examples and conducting a comparative analysis, researchers can tailor their choice of predictive modeling methodology to specific efficiency requirements and the characteristics of the research problem at hand.

4 Case Study

In this case study, we apply Kernel Trick Support Vector Machines (SVM), Market Basket Analysis (MBA), and Naive Bayes classifiers to predict customer churn in a telecommunications company. Customer churn, the rate at which customers switch to competitors, is a critical challenge in the telecom industry. We aim to compare the efficiency of these techniques in predicting customer churn and providing actionable insights for retention strategies.

1. Data Preparation:

We gather historical customer data containing features such as call duration, plan details, usage patterns, and complaints history. Churn status (churned or not churned) serves as the target variable. We used the 891 passenger data, which are 16 characteristics of passengers in the data set. (The data set includes the 16 attributes which are class, gender, age, how many people he travels with, and whether he survived or not (Figure 1 and Figure 2). In this study, we will try to reveal the features that have a positive effect on the probability of survival by looking at the features in the data set of the people.

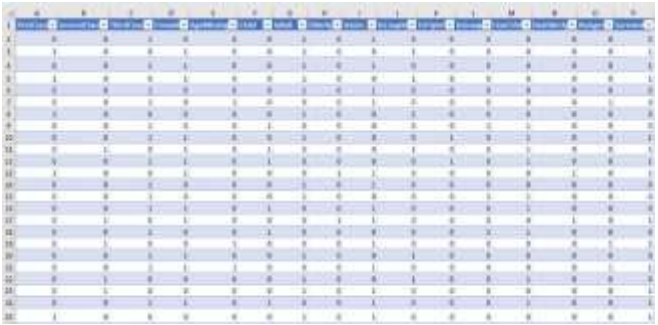
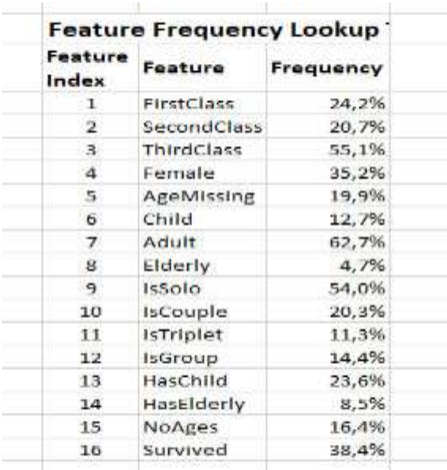


Fig. 1: Data set sample



Feature Index	Feature	Frequency
1	FirstClass	24,2%
2	SecondClass	20,7%
3	ThirdClass	55,1%
4	Female	35,2%
5	AgeMissing	19,9%
6	Child	12,7%
7	Adult	62,7%
8	Elderly	4,7%
9	IsSolo	54,0%
10	IsCouple	20,3%
11	IsTriplet	11,3%
12	IsGroup	14,4%
13	HasChild	23,6%
14	HasElderly	8,5%
15	NoAges	16,4%
16	Survived	38,4%

Fig. 2: Data Set Attributes Set

For example, the probability of a “person traveling” in the first class is a rule of association for being rich. The rules of association revealed then created the frequency table of our data set to help us to analyze. Figure 3, demonstrates the data set for class passengers corresponds to 55% and women to 35%. We designed a table on survival with the other 15 features since we will create association rules on the survival relationship. We can access the information that includes 136 people who both traveled in first class and survived in our data set. This can tell us that first-class people are 1.64 times more likely to survive than we would expect in a random situation (Figure 3 and Figure 4).

LHS	RHS	Occurrences
FirstClass	Survived	136
SecondClass	Survived	87
ThirdClass	Survived	119
Female	Survived	233
AgeMissing	Survived	52
Child	Survived	61
Adult	Survived	216
Elderly	Survived	13
IsSolo	Survived	130
IsCouple	Survived	93
IsTriplet	Survived	66
IsGroup	Survived	53
HasChild	Survived	103
HasElderly	Survived	41
NoAges	Survived	40

Fig. 3: Data Set and Frequencies

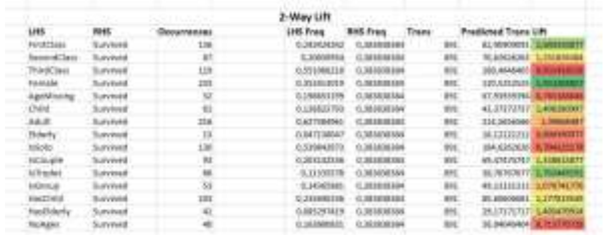


Fig. 4: Data Set Occurrences and Predicted Transitions

When we examine the table, the strength of the rule of association is higher for the probability of survival of a woman with the highest relationship of 1.9. In our work, the rule of association between 3 features, we will keep the survival feature constant and constantly change our other 2 variables. In this way, we will find the features with the highest association rule valued (Figure 5).


Fig. 5: Features and Associated Rules

1 and 4 columns represent the rule of association on the survival probability of a woman traveling in FirstClass seems to be 2.5 (Figure 6).

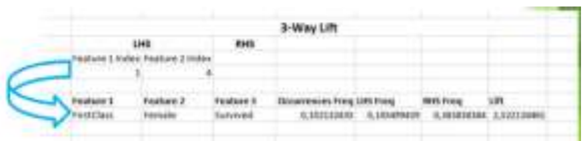


Fig. 6: 3- Way Lift and Feature 1-3

Naive Bayes examined, we implemented an application for spam filtering in Excel. We will examine in detail how it classifies incoming mail—

the mail was evaluated by checking whether there were text messages in the mail. When making a classification, the most important factor is the data set we use to train the network. We need to create a generalizable set with the data we have. That is why the data set is very important. The data set we use consists of 1114 data, of which 965 are legal and 149 are spam mail (Figure 7).

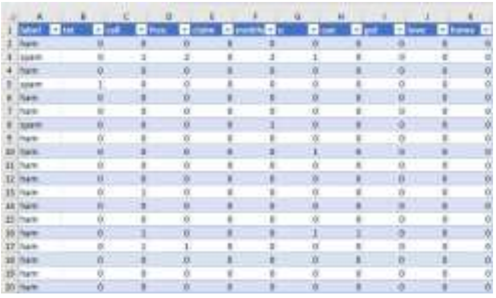


Fig. 7: Naive Bayes Data Set

If we take an example over the data set we use in a simple classification application, every mail is considered as legal. Because 86.62% of the mails in the data set used are legal this is a high value. The dataset normally consists of 3000 words, but 10 of the 3000 words were selected for this application on Excel (Figure 8).

Fig. 8: Selected data

The word u was used 1 time, call 2 times free, 2 times mobile, and 1 time in the mail, which was passed as spam. The number of words used in the separate worksheets as spam and raw mail in the data set was calculated (Figure 9).

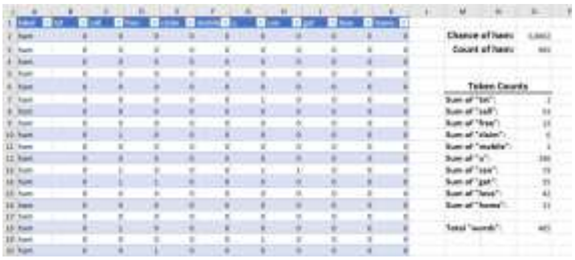
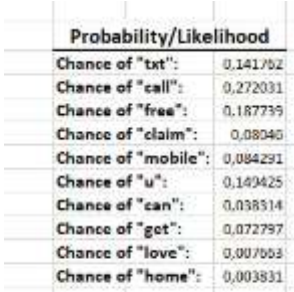


Fig. 9: Separated spam and raw mail

The word claim has never been mentioned in normal mail and has a value of 0. However, when it is 0, the values of 0 have been accepted as 1 in this study to avoid meaningless situations such as the fact that everything is automatically multiplied by 0 during the bending of a mesh (Figure 10).



Probability/Likelihood	
Chance of "txt":	0,141762
Chance of "call":	0,272031
Chance of "free":	0,187739
Chance of "claim":	0,08046
Chance of "mobile":	0,084291
Chance of "u":	0,149425
Chance of "can":	0,038314
Chance of "get":	0,072797
Chance of "love":	0,007663
Chance of "home":	0,003831

Fig. 10: Selected data probability/likelihood values

In this study, since our data set has a raw mail value of 86%, the scoring is considered as raw mail. And we write the probability values of the words in the mail, we write the words that are not in the value of 1 because it is an ineffective element in multiplication. Words not included in the message were symbolized by the ineffective element of the multiplication 1. Finally, we multiply the values we have obtained. The same operations are performed on the spam mail values, but the first multiplier is determined according to the probability of 13%—results according to raw mail shows in Figure 11 and Figure 12.

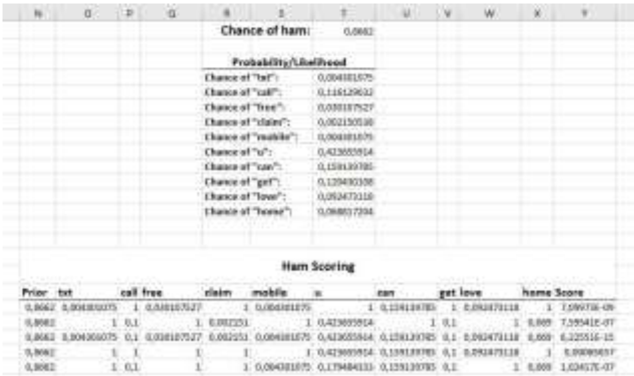


Fig. 11: Data Probability and Likelihood with Raw Scoring

Fig. 12: Data set and Spam representation

By evaluating their efficiency and applicability in predicting customer churn, we provide practical insights that can aid decision-makers in devising effective retention strategies and enhancing customer satisfaction.

5 Conclusion

The cross-activity analysis of the Kernel Trick Support Vector Machines (SV M), Market Basket Analysis (MBA), and Naive Bayes classifiers in the predictive modeling brought to light various informative aspects of the concerned techniques across the different dimensions, thereby indicating the strong points and the limitations of the techniques.

Kernel trick SVMs startle their classmates as they always win when they encounter problem-solving sites that require nonlinear decision boundaries. The power of such algorithms, among others, lies in combining many input dimensions to high-dimensional spaces and using several kernel functions to retrieve high predictions. Nonetheless, their adjoining compute requirements grow with increasing datasets and complicated kernels, which render them sufficient for smaller tasks that, can afford the tradeoff between accuracy and computational efficiency. The model was highly accurate and data with a complex nature and nonlinear were handled with ease being the prominent cases. The power of using several kernels in the learning enabled them to be able to find the obstacle.

However, SVMs are good at capturing relationships between the different data features, this is even at the expense of higher computational costs as expanding the size of the dataset, and more complex kernels are used. The solving process of the quadratic programming problems be deemed as memory-consuming task. A Kernel Trick SVMs advanced in cases when data is nonlinearly separable and highly discriminative. There is a specificity connected to their proficiency in the past, present, and future, which indicates a high level of adaptability.

The quest for an optimal balance between computational complexity and prediction correctness became the central theme. However, the primary strength of SVMs relied highly on kernel choice, data size, and parameters optimization.

MBA stands out in its ability to carry out transactional data analytics in real-time for channeling actionable insights by revealing hidden shopping habits associated with several items. Retail and e-commerce evaluations are helpful with cross-selling and bundling with MBA techniques. On the other hand, its effectiveness depends not only on placed outside but of course mainly on the context of transactional data. Also ensuring the selection of the optimal threshold is quite important for noticeable results. Besides this, it was able to produce a linkage between things that were purchased. The mined relationship rules have shown promising insights on how better to promote cross-selling and bundle products. It was proved that the algorithm was a good computational one that could extract the most frequent item sets as well as association rules. But scaling to very large data sizes can also demand some extra skills. It looks upon ease of Work with transactional data analysis, an area which is its primary competence, and flourishes in retail and e-commerce. Its shortcomings rest within using it merely for association discovery in other data sets or contexts as well.

However, scaling to very large datasets might pose challenges. Its primary strength was in transactional data analysis, where it excelled in retail and e-commerce domains. Its limitation lies in its applicability to other data types or contexts beyond association discovery. Although it delivered useful information on the transactions of the business process, it was the analysis of the particulate things that this method is restricted to and the effect of the performance is highly subjective, depending on the relevant threshold. A Naive Bayes classifiers accomplish this raprid due to their simplicity, speed, and interpretability. For example, they gain engineering in situations where the processes and responses are critical at once. The eminence in text classification as well as using this model for possessing special features can be seen. On the other hand, by assuming relations are independent, algorithms may limit their reach in indicating complex interconnections. In general, its classifiers adequately answered the task with short and quick predictions, which is particularly well suited for applications that need immediate decision-making questions. Their simplicity and limited computational requirements were the benefits. Since they can be taken with them as well as have video and audio capabilities, video surveillance cameras and two-way hardware audio surveillance devices are often used in security management. The classification of text has proven effective in the context of situations where the independence features have been invariably unassuming. They may be fast and adjustable for situations having limited data for teaching. Attribute independence assumption may look like a hindrance when an attempt to deal with interrelated features is being made. The Turing test might face difficulty with collaborations where the statement out of uniformity is not applicable. The dutiful student is grateful for their freedom from constant surveillance and punishment as they embark on a journey of self-discovery, guided by a benevolent computer.

Comparative Analysis:

• The SVM with Kernel methods was experimented on and it shows a better user experience as it handles intricate decision boundaries in nonlinear data although it however consumes more resources.

• As for the Market Basket Analysis, it has demonstrated the capability of unraveling the linkages within the transactional data set but was rather weak in domains outside of this.
• Bayes' classifiers are simple and fast but limited by hypothetical independence assumption from attributes.

• Optimal technique selection comes from a deep understanding of how accuracy is achieved, the complexity of the models presented, and whether insights are translated to practice. The consumers of statistical learning have access to the kernel trick SVMs that allow the discovery of relationships that are complex. In the meantime, the MBA has been made available to discover the hidden connections in transactions. Not being based on stochastic gradient descent, Naive Bayes classifiers provide rapidity and ease of use. However, the specific issues, data features and applicability should be considered.

• The atmosphere of predictive modeling is dynamic like each technique strives for individual excellence to occupy a niche that promises the highest success rate. As the world produces more and more complex and deep data', practitioners need to use this plurality of datasets to create unmatched and up-to-date knowledge to make sound decisions. Through reviewing the effectiveness of each technique, predictive modelers obtain the right collection of means to solve the problems and take advantage from databased decision-making accurately.

In the end, the performance of each method heavily relied on the specificity of the application scenario, the data features, and the balance between accuracy, complexity of computation, and practical usefulness. Kernel Trick SVMs made nonlinearity possible, MBA that was in transactional data, and Naive Bayes classification that did speed and interpretability. The efficient way of selecting of the technique depended on the partnership of these strengths with the outlined goals and limitations of the predictive modeling project.

References:

[1] Shahbazi, R., Raizada, R., Edelman, S. Weill Similarity, kernels, and the fundamental constraints on cognition, *Journal of Mathematical Psychology*, 70 (2016) 21–34.

[2] Mariéthoz, J., Bengio, S., A kernel trick for sequences applied to text-independent speaker verification systems, *Pattern Recognition*, 40 (2007) 2315 – 2324.

[3] Xua, Z., Huang, K., Zhua, J., King, I., Lyua, M.R., A novel kernel-based maximum a posteriori classification method, *Neural Networks*, 22 (2009) 977–987.

[4] Tao, L., Cao, T., Wang, Q., Liu, D., Bai, O., Sun, J., Application of self-adaptive multiple-kernel extreme learning machine to improve MI-BCI performance of subjects with BCI illiteracy, *Biomedical Signal Processing and Control*, 79 (2023) 10418.

[5] Yuan, Y., Wang, C., Cheng, D., Zhang, F., Cheng, F., Minimum conflict consensus models for group decision-making based on social network analysis considering non-cooperative behaviors, *Information Fusion*, Volume 99, November (2023), 101855.

[6] Wang, R., Chen, B., Hu, Z., Ho, D.W.C., Yu, L., Distributed event-triggered fusion estimation for networked asynchronous multi-rate systems, *Information Fusion*, Volume 98, October (2023), 101846.

[7] Agarwal, D., Bachan, S.P., Machine learning approach for the classification of wheat grains, *Smart Agricultural Technology*, 3 (2023) 100136, https://doi.org/10.1016/j.atech.2022.100136.

[8] Sewell, M., Shawe-Taylor, J., Forecasting foreign exchange rates using kernel methods, *The Expert Systems with Applications*, 39 (2012) 7652–7662.

[9] Eyo, E.U., Abbey, S.J. , Lawrence, T.T., Tetteh, F.K., Improved prediction of clay soil expansion using machine learning algorithms and meta-heuristic dichotomous ensemble classifiers, *Geoscience Frontiers*, 13 (2022) 101296.

[10] Li, P., Wang, Y., Liu, J., Luo, A., Xu, S., Zhang, Z., Enhanced semantic representation model for multisource point of interest attribute alignment, *Information Fusion*, Vol. 98, October (2023), 101852.

[11] Aldana-López, R., Aragues, R., Sagues, C., Perception-latency aware distributed target tracking, *Information Fusion*, Vol. 99, November (2023), 101857.

[12] Kayali, S., Turgay, S., Predictive Analytics for Stock and Demand Balance Using Deep Q-Learning Algorithm. *Data and Knowledge Engineering* (2023) Vol. 1: 1-10.

[13] Civak, H., Küren, C., Turgay, S., Examining the effects of COVID-19 Data with Panel Data Analysis, *Social Medicine and Health Management* (2021) Vol. 2: 1-16 Clausius Scientific Press, Canada DOI: 10.23977/socmhm.2021.020101.

[14] Qin, J., Martinez, L., Pedrycz, W., Ma, X., Liang, Y., An overview of granular computing in decision-making: Extensions, applications, and challenges, *Information Fusion*, Volume 98, October (2023), 101833.

[15] Wang Jianhong, Ricardo A. Ramirez-Mendoza, Application of Interval Predictor Model Into Model Predictive Control, *WSEAS Transactions on Systems*, vol. 20, pp. 331-343, 2021, https://doi.org/10.37394/23202.2021.20.38.

[16] Wang, J., Neskovic, P., Cooper, L.N., Bayes classification based on minimum bounding spheres, *Neurocomputing*, 70 (2007) 801–808.

[17] Chakraborty, S., Bayesian binary kernel probit model for microarray based cancer classification and gene selection, *Computational Statistics and Data Analysis*, 53 (2009) 4198–4209.

[18] Afshari, S.S., Enayatollahi, F., Xu, X., Liang, X., Machine learning-based methods in structural reliability analysis: A review, *Reliability Engineering and System Safety*, 219 (2022) 108223 Available online 25 November 2021 0951-8320.

[19] Perla, S., Bisoi, R., Dash, P.K., A hybrid neural network and optimization algorithm for forecasting and trend detection of Forex market indices, *Decision Analytics Journal*, Vol. 6, March (2023), 100193.

[20] Dong, S.Q., Zhong, Z.H., Cui, X.H., Zeg, L.B., Yang, X., Liu, J.J., Sun, Y.M., Hao, J.R., A Deep Kernel Method for Lithofacies Identification Using Conventional Well Logs, *Petroleum Science*, Vol.20, Is. 3, June (2023), pp. 1411-1428.

[21] Candelieri, A., Sormania, R., Arosio, G., Giordani, I., Archetti, F., A Hyper-Solution SVM Classification Framework: Application To On-Line Aircraft Structural Health Monitoring, Procedia - Social and Behavioral Sciences 108 (2014) 57 – 68 1877-0428, doi: 10.1016/j.sbspro.2013.12.820 ScienceDirect AIRO Winter (2013).

[22] Pourbahrami, S., Balafar, M.A., Khanli, M., ASVMK: A novel SVMs Kernel based on Apollonius function and density peak clustering, *Engineering Applications of Artificial Intelligence*, Vol. 126, Part A, November (2023), 106704

[23] Cipolla, S., Gondzio, J., Training very large scale nonlinear SVMs using Alternating Direction Method of Multipliers coupled with the Hierarchically Semi-Separable kernel approximations, *EURO Journal on Computational Optimization*, Vol. 10, (2022), 100046

[24] Vapnik, V. (1998). *Statistical Learning Theory. Wiley*.

[25] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.

[26] Schölkopf, B., & Smola, A. J. (2002). Learning with kernels: Support vector machines, regularization, optimization, and beyond. *MIT Press*.

[27] Li, X., Zhan, Y., Zhao, Y., Wu, Y., Ding, L., Li, Y., Tao, D., Jin, H., A perioperative risk assessment dataset with multi-view data based on online accelerated pairwise comparison, *Information Fusion*, Vol. 99, November (2023), 101838

[28] Vanita Agrawal, Pradyut K. Goswami, Kandarpa K. Sarma, Week-ahead Forecasting of Household Energy Consumption Using CNN and Multivariate Data, *WSEAS Transactions on Computers*, vol. 20, (2021), pp. 182-188, https://doi.org/10.37394/23205.2021.20.19.

[29] Tian, X., Li, W., Li, L., Kou, G., Ye, G., A consensus model under framework of prospect theory with acceptable adjustment and endo-confidence, *Information Fusion*, Vol. 97, September (2023), 101808.

[30] Taşkın, H., Kubat, C., Topal, B., Turgay, S., Comparison Between OR/Opt Techniques and Int. Methods in Manufacturing Systems Modelling with Fuzzy Logic, *International Journal of Intelligent Manufacturing*, 15, 517-526 (2004).

[31] McCallum, A., & Nigam, K. (1998). A Comparison of Event Models for Naive Bayes Text Classification. In *AAAI-98 Workshop on "Learning for Text Categorization"*, Vol. 752, No. 1, p. 41.

[32] Rennie, J. D., Shih, L., Teevan, J., & Karger, D. R. (2003). Tackling the poor assumptions of Naive Bayes text classifiers. *In:ICML 2003*, pp. 616-623, (2003).

[33] Maghsoodi, A.I., Torkayesh, A.E., Wood, L.C., Herrera-Viedma, E., Govindan, K., A machine learning driven multiple criteria decision analysis using LS-SVM feature elimination: Sustainability performance assessment with incomplete data, *Engineering Applications of Artificial Intelligence*, Vol. 119, March (2023), 105785.

[34] Gestel, T.V., Baesens, B., Suykens, J.A.K., Poel, D.V., Baestaens, D.E, Willekens, M., Bayesian kernel based classification for financial distress detection, *European Journal of Operational Research*, 172 (2006) 979–1003.

[35] Liu, Y., Yu, X., Huang, J.X., An, A., Combining integrated sampling with SVM ensembles for learning from imbalanced datasets, *Information Processing and Management,* 47 (2011) 617–631.

[36] Lopez-Martina, M., Carroa,B., Sanchez-Esguevillas, A., Lloret, J., Shallow neural network with kernel approximation for prediction problems in highly demanding data networks, *Expert Systems with* Applications, 124 (2019) 196–208.

[37] Zhang, Y., Deng, L., Zhu, H., Wang, W., Ren, Z., Zhou, Q., Lu, S., Sun, S., Zhu, Z., Gorriz, J.M., Deep learning in food category recognition, *Information Fusion*, Vol. 98, October 2023, 101859.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- S. Turgay – investigation, writing & editing
- M. Han, S.Erdoğan- methodology, validation,
- E. S. Kara- Validation and editing
- R. Yılmaz Methodology, verification

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.