Prediction Divorce Prospect using Machine Learning Methods: A Case Study on Bangladesh

BY

Sherin Tasnim Mukta
ID: 221-25-112

This Report Presented in Partial Fulfillment of the Requirements for

The Degree of Master of Science in Computer Science and Engineering

Supervised By

Abdus Sattar

Assistant Professor & Coordinator MSc.

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This Project/Thesis titled **"Prediction Divorce Prospect using Machine Learning Methods: A Case Study on Bangladesh"**, submitted by **Sherin Tasnim Mukta**, ID No: **221-25-112** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 17-01-2023.

BOARD OF EXAMINERS

Chairman

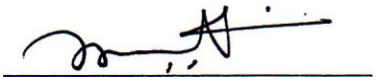
Dr. Sheak Rashed Haider Noori, PhD
Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Ms. Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Mr. Md. Sadekur Rahman
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



External Examiner

Dr. Mohammad Shorif Uddin, PhD
Professor
Department of Computer Science and Engineering
Jahangirnagar University

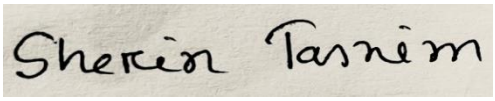
DECLARATION

We hereby declare that, this thesis has been done by us under the supervision of **Abdus Sattar, Assistant Professor, Department of CSE** Daffodil International University. We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:



Sherin Tasnim Mukta
ID: 221-25-112
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final thesis successfully.

We really grateful and wish our profound our indebtedness to **Abdus Sattar**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning" to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this thesis.

We would like to express our heartiest gratitude to **Pro. Dr. Touhid Bhuiyan, Professor, and Head,** Department of CSE, for his kind help to finish our thesis and also to other faculty members and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and passion of our parents.

ABSTRACT

In this paper, we suggest a system for predicting divorce and use the Divorce Predictors Scale to evaluate it (DPS). The DPS is a 40-item self-report questionnaire based on Gottman couple's therapy, which could be used as features or characteristics in a machine learning algorithm. Besides from "Divorce Predictors Scale," a "personal information form" was utilized to collect personal information from participants in a more traditional and disciplined manner. There were 179 (71.6%) married people and 71 (28.4%) divorced people among the 250 participants (N=250). The algorithms of K-nearest neighbors, Naive Bayes, and Decision Tree, SVM, Random Forest were used. We attempted to restrict the field using feature selection. a list of the most important or noteworthy characteristics based on the selection of features based on correlation as a result, characteristics/items were discovered. In the context of Bangladeshi divorce prediction, data. Different algorithms are applied directly to the data throughout this process. The dataset with the greatest accuracy rate for divorce prediction is the SVM method produced an accuracy of 100%. The two algorithm Decision Tree, Naïve Bayes given the same accuracy 94.0%. However, the Random Forest which give accuracy 96.0%. DPS can predict divorce, as evidenced by the results. This scale can be used by family advocates and family experts to help with case characterization and mediation strategy. Furthermore, the results of the study show that the Machine Learning Algorithm treatment verified in the Bangladeshi sample.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGRES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Divorces have increased in Bangladesh in recent years. One of the key reasons of social unrest is rapid growth unbalancing. For any civilization, it is a source of concern. The most important consideration the goal of this study was to determine the divorce rate and the number of divorces in Bangladesh. Why is the divorce rate so high in this society, and who is it? Marriage in today's society is more secure. We made an attempt to gather information from a diverse group of individual we were able to gather 250 pieces of information from them. Bangladesh's entire territory. There are 18 % from joint families and 82 % from single households among them. We received data from 71.6% of married people and 28.4% of divorced people. Males made up 65 % of the participants, while females made up 35%. On the data, we used a variety of algorithms. The family relationship therapist can use it during counseling.

As a result, if couples are aware of their own inadequacies in the relationship, as well as their obligation to one another, they will be more cautious in their interactions. As a result, it will aid in the reduction of divorce rates. In this study, we used multiple machine learning algorithms and compared them to predict divorce among Bangladeshi couples.

1.2 Motivation

According to the Bangladesh Bureau of Statistics (BBS), divorce has been on the rise for several years. Divorce is increasing among educated husbands and wives. According to a BBS report titled The Situation of Vital Statistics published in June 2019, the number of divorces in 2019 has increased by 16 % as compared to 2016. Last year, there were an average of 1.4 divorces per 1,000 men and women over the age of 15. In the petitions, almost all the reasons for divorce were 'not getting married' between the husband and wife. The reasons given by the wife in the application include non-payment of alimony, torture in demanding dowry, suspicious attitude of the husband, relationship with another woman, drug addiction, impotence, conflict of personality. The husband's petitions mention various reasons including wife's bad temper, indifference towards family, not having children, disobedience, not following Islamic law. Many people's income has been reduced due to coronation. But the impact of the crisis does not stop there. Economic misery is breaking down in many worlds as well. In the capital Dhaka alone, 36 divorces are happening

daily. As a result, every 36 minutes a couple grows a beard. Compared to last year, this month has seen an increase of 99 separations per month. According to the data of two cities of Dhaka, 75 percent of divorces are given by women. In the four months from January to April this year, 4,565 divorce applications were received, meaning 1,141 per month. Last year this number was one thousand 42. As a result, the number of divorces has increased by 99 per month this year. Last year too, more divorces were given on behalf of women, 70 %. The trend of divorce has increased not only in Dhaka, but in the whole country. According to the Bangladesh Bureau of Statistics (BBS), the number of divorces has increased by 17 % in 2019 as compared to 2018. According to the data of the two city corporations of Dhaka, the main reasons given by the men for the separation are the suspicion of the wife towards them, the indifference of the wife towards the world, bad temper, infertility etc. Women have shown as reasons for extramarital affairs, torture for dowry, drug addiction etc. There were 12,513 divorces in the two cities last year. Of these, 8,471 were filed by women and the remaining 4,032 were sought by men. There were 12,513 divorces in the two cities last year. Of these, 8,471 were filed by women and the remaining 4,032 were sought by men.

Due to above the problem we can see that the divorce rate in Bangladesh has increased year by year and it is increasing dramatically It is growing at a much higher rate each year than its predecessor. So we want to do some research on this and find out the proper reasons for divorce.

1.3 Problem Definition

Let's also discuss how to avoid divorce. Divorce is becoming more common, which implies more interruptions in family unity, which is an unpleasant experience for children and a cause of acute trauma that can have long-term consequences. Existing public awareness campaigns, such as those against domestic violence and dowry, two major causes of women seeking divorce in Bangladesh, could use these findings to reinforce their message. Extensive research that considers the social complexities of Bangladesh, as well as the resulting disaggregated data, could be very useful in determining the deeper causes of the rise in divorce, as well as ways to prevent it.

1.4 Research Questions

Research Questions are given here for the research purposes.

1. What is the current situation of relationships in Bangladesh?
2. How we can predict a couple get divorced or not?
3. How to reduce this problem using machine learning and AI?

4.What will the quality of our original data be?

5. Is it necessary to teach the machine learning model our original data?

6. How much and where do we collect data?

7. Is our data going to be compatible with machine learning?

8. Should we employ well-known machine learning approaches or develop our own?

1.5 Methodology and Steps

This component of our study paper also covers the Experiment Data Set, Data Pre-processing, Model Architecture, Learning Rate and Optimizer, Data, and Model Training. The performance of the suggested model will be reviewed at the conclusion of this chapter.

1.6 Objectives

There are some benefits of using AI in predicting divorce

a) Proposed an efficient model to predict divorce.

b) To improve result accuracy of predicting.

c) How we can reduce this rate in future.

1.7 Research Outcome

We hope that our findings will assist individuals in making predictions. People can easily and swiftly learn about the divorce and happily married pair utilizing this strategy. Machine learning can also help people understand more about prediction. New or existing machine learning algorithms for divorce prediction have been successfully deployed. We occasionally go to various locations and environments in search of job. We don't know if somebody is happily married or divorced, and if so, to what extent do I have a proclivity to divorce. When we entering a new environment, this research would help us in predicting the likelihood of separation. People are capable of caring for their husbands or wives. People with whom they can make acquaintances, who they are relocating to, and their everyday life, using this information to prevent these situations. It will protect us and our country against the negative consequences of divorce. Furthermore, divorced and happily married people can work together with our model, which will help them in identifying whether or not it is divorced for the right reasons. In addition, the creation of a data set for divorce prediction in Bangladesh.

1.8 Report Layout

Chapter 1 discussed about introduction, objectives, some research question, and outcome of the research.

Chapter 2 discussed about literature review of the existing work and study , the tried to compare with their study and summarize with their study.

Chapter 3 discuss about Research Methodology, Data Collection Process, Research subject and Instrument, Proposed Methodology, Data Preprocessing, Statistical Analysis, Feature Selection and Implementation Requirements.

Chapter 4 discuss about Experimental Results and Discussion

Chapter 5: Here Limitation, Conclusion, Further Study and all the references we used for this research.

CHAPTER 2

BACKGROUND

2.1 Introduction

In Bangladesh there are a little work or research was done which divorce cannot predict perfectly. So the background is the current situation of divorce and the use of machine learning algorithm in this area of Bangladesh.

2.2 Related Works

This section of the study will present all of the known literature on divorce prediction. Until date, Artificial Neural Networks have been used in a variety of divorce prediction research. To achieve the goal, we'll use a combination of network, machine learning, and data mining the most accurate forecast rate.

To the best of our knowledge, we are the first in Bangladesh to collect such a dataset for machine learning analysis. We developed the Divorce Prediction Scale [1-5] based on J. M. Gottman's beliefs. [1] presents a scientific marital therapy, and [2] presents the determination of a correlation between the marital process and results as predictors. These theories emphasized the concepts of making marriage work [5] and resolving trust concerns [4].

In the paper examined Divorce Prediction from a Turkish perspective [6-7]. Divorce Prediction Scale (DPS) was created based on Gottman's relationship theory [1-5]. They used Artificial Neural Networks and relationship-based component determination. The RBF neural network achieved the highest forecast rate of 98.23 percent, followed by ANN at 97.64 percent, and Random Forest at 97.64 percent. They have 98.82 percent with ANN when it comes to relationship-based highlights. Furthermore, the achievement is enhanced by RBF and Random Forest. The percentage was 97.64 percent. As a result, they achieved the greatest results by combining the ANN model with relationship-based analysis. determining the elements. A good married life is essential for the enhancement of society's psychosocial balance. Researchers are attempting to spread recognized methods and counseling to married couples in order to promote healthy marital solutions. Predictive elements of social functioning [8, identifying the rigorous components for psycho-

educational couples [9], and even studies on patient therapy who are hospitalized after suicide attempts [10] have all been brought to the forefront in the quest to find divorce predictors.

On our Divorce Prediction dataset, we used the Artificial Neural Network (Multilayer Perceptron), Naive Bayes, and Random Forest methods to determine our prosperity rate. With Nave Bayes, 81. 42 percent, and Random Forest, the most notable forecast precision is 87.14 percent. However, following feature selection, ANN has 84.29 percent, Nave Bayes has 85.71 percent, and Random Forest has 81.42 percent. As a result, our best forecast complements Nave Bayes by identifying options.

2.3 Bangladesh Perspective

During the last few decades, divorce rates have increased not only in our country but also all over the world. The rise in divorce has created concerns among researchers, policymakers and members of the public about the consequences of changes in family structure. Divorce should be seen and discussed in a sophisticated manner. True, long-held beliefs about divorce are gradually being dispelled. One of the main reasons is that women have more decision-making power as a result of their economic independence. But let's also start talking about methods to avoid divorce. Divorce is becoming more common, which means there are more interruptions in family harmony—a unpleasant experience for children and a cause of acute stress that can have long-term consequences. Existing public awareness programs, such as those against domestic abuse and dowry, two major causes for women seeking divorce in Bangladesh, might use these data to reinforce their message. Extensive study that considers Bangladesh's social intricacies and the ensuing disaggregated data might be highly beneficial in determining the fundamental causes of the rise in divorce and measures to prevent it.

2.4 Comparison with related study

With the machine learning algorithm, some work has already been done on prediction and detection. With the usage of Divorce Prediction and various detectors, the application of machine learning technology has expanded in recent years. This section contains a comparison of these connected works. Table 2.0 shows a comparison of several research studies, including their subject, methods, and results.

2.5Scope of the Problem

This forecast will have a major impact on society. It is possible to avoid divorce and live a happy married life. This model will be usable by machine learning for a variety of purposes. It is harmful to divorce since it might have a negative impact on their lives and their children. As a result, this concept could help ordinary people and conscientious people avoid divorce. People are too concerned about their children and the future, which may impede a generation's growth and reflect its negative repercussions on society. Machine learning algorithms have recently been utilized for various object detection and divorce, as well as several types of disease predictions, with promising results. As a result, we decided to construct a divorce prediction model using machine learning.

2.6Challenges and Limitations

People who have recently divorced find it difficult to talk and disclose their feelings. Furthermore, ordinary folks and divorcees are difficult to identify. We read a lot of publications and spoke with a variety of people, including neighbors, but no one would offer us any information regarding divorced people. It was extremely difficult to gather information about divorced people from bus stops, train stations, and other unfamiliar locations. After that, we were able to obtain our information from a domestic staff in Dhaka. We looked for other divorce community Facebook groups and judge court information but declined to assist her with any information. We needed to speak with Parsons to Parsons for data gathering, but we couldn't since the employees refused to answer our questions.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The focus of this research is to create a method for predicting the release of a divorce prediction. The Prediction Model is built using people's personal and everyday life data, as well as some other related data. We used a variety of machine-learning methods to generate this model. Different machine-learning algorithms have been used. kNN, SVM, Nave Bayes, decision tree, and random forest were all employed. The classification algorithms employed in the model. We looked at forty important variables that were all linked to divorce. We looked into some of the factors that contributed to the final result. Pre - implantation, we analyzed the information. We compared and computed several types. To choose the best method for the analysis, we analyzed and evaluated its accuracy, sensitivity, specificity, and precision. Support vector machine was determined to be the most accurate and appropriate for our suggested model.

3.2 Data Collection Process

The data set contains a massive collection of useful and related coordinates that can be easily accessed and modified. We start by looking for divorced persons in our neighborhood and other areas. However, we witnessed someone close to us divorcing and having an unlawful relationship, but it was kept a secret, and the judge court and municipal corporation refused to assist. Then we go out and questioned random people, including such rikshaw drivers, about our issues, and gathered information. We also get information from domestic servants in Dhaka. We have been unable to conveniently gather data when we talked to the marriage people due to privacy concerns. As a consequence, we generated a form and distributed it to them, and the participants assisted us by providing information on their cause of divorce. We were able to get information from 250 persons based on 40 variables. We have information about 71 divorces and 179 happily married couples. We got all of our information from random strangers, a Google form, and a home servant, among other places. The following factors influenced our data collection question:

1. Do you have any other say about what a married relationship should or should not be?

What about the role of the spouse you are now married to?

3. Are you economically self-dependent?

4. Are you both educated?

5. Is the age difference between the two of you more than 5-7 years?

6. Are you childless couple?

7. Are you drug addict?
8. Are you physically incapable?
9. Are you engaged in any post marital affair?
10. Do you have any demand of hyper sexuality?
11. Do you repeatedly give birth to daughter children?
12. Have you ever being pressurized for taking/giving dowry?
13. Are you abused by your husbands/wife?
14. Are you being tortured at your in law's house?
15. Do you do all household chores together?
16. Do you trust each other equally?
17. Do you know about each other likes and dislikes?
18. Do you live in a joint family?
19. Do you have any boyfriend/girlfriend before marriage?
20. If so then, did you marry her/him?
21. Are you married with the consent of the family?
22. Do you know what your spouse likes?

Each of these criteria must be considered while determining the likelihood of divorce. We learn about these factors by speaking with a variety of people and conducting past research.

3.3 Research Subject and Instrumentation

We'll use our data to test different algorithms to evaluate which ones will work best for our model. kNN, support vector machine (SVM), Nave Bayes, Decision tree, and Random Forest are some of the machine-learning techniques we utilize. In our research, we employed Python as a programming language, Google Collaboratory as a data mining tool, and Microsoft Excel as our dataset

3.4 Proposed Methodology

In figure 3.0 shown a diagram of our proposed methodology.

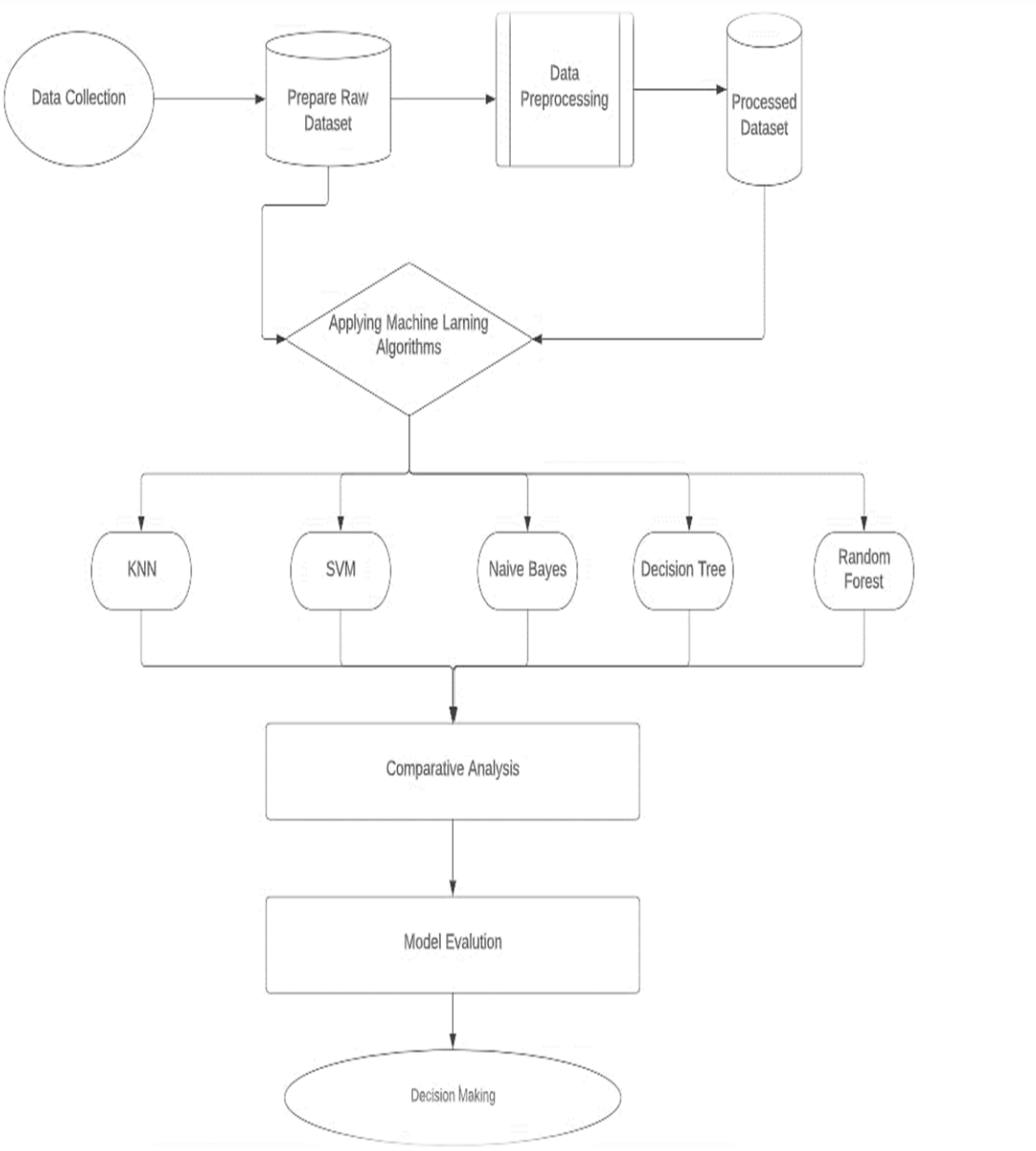


Figure 3.0 Proposed methodology flowchart

3.5 Data Preprocessing

We got some missing data, numerical and text data after collecting the data. Then we decide that by Processing the data, we can make it acceptable for algorithms. Data processing is the capacity to convert data into a useful format once it has been collected. Information or data is processed in a certain format to facilitate output.

Figure 3.1 depicts our data preparation procedure.

Our data preprocessing method is shown below in Figure 3.1

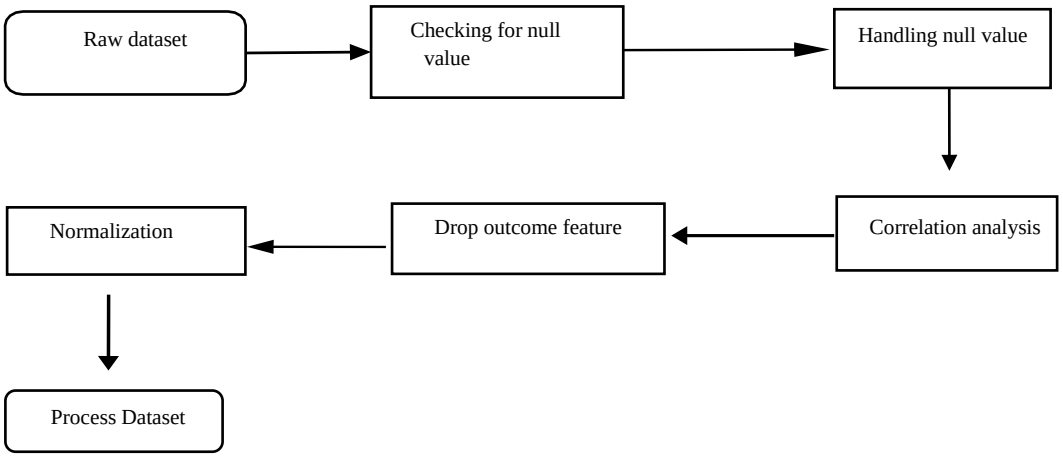


Figure 3.1: Steps of data preprocessing.

First, we completed the data cleaning process. We examine the data set for any null values. We used imputer and median to tackle the missing value problem. The correlation matrix is then examined as part of the information implementation phase. The ratio of each data connected to each data is shown in this matrix. A positive value shows that data is highly connected, a negative value suggests that data is negatively related, and zero implies that data does not connect to itself. Then we remove the predict column, which was our outcome feature. In feature engineering, we generate a separate histogram for each feature to aid data reduction and visualization. The data processing was completed by normalization.

As a result, we now have the processed data set in our hands. The Google Collaboratory was used throughout the data processing process.

3.6 Statistical Analysis

We were able to get information from 250 persons. We gathered information from people of various jobs, ages, and locations. Figure 3.2 depicts the number of divorced and married people in our data set. Our model was built using data from 71 divorced people and 179 married people.

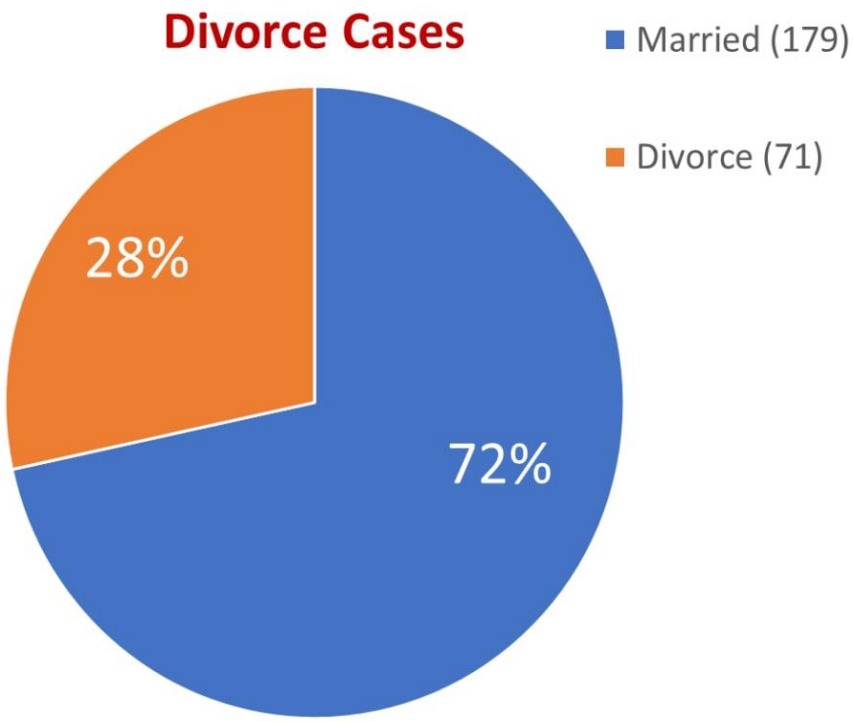


Figure 3.2: Divorce and married cases.

3.7 Feature Selection

It is possible to pick the features in our data which are most effective or important for the issue we are working on automatically. Selection of features is the name for this method. We'll discuss about feature selection and the different types of strategies we used here study in this section. Variable selection or attribute selection are other terminology for selecting features. It's the task of automatically recognizing the qualities in our data (including such columns in data tables) which are most significant to the predictive modeling task at hand. Dimensionality decrease is not the

same as selecting features. Most methods are aimed to reduce the number of attributes in a dataset, but dimensionality reduction methods do by creating new combinations of characteristics, whereas techniques simply include and omit existing attributes in the data. Feature selection strategies help us fulfill our idea of producing a precise predictive model. They enable us by feature extraction that will provide good or excellent accuracy with much less information. Methodologies is used to recognize and remove unnecessary, irrelevant, and redundant attributes from data that often do not add to the accuracy of a predictive model or may reduce overall the model's accuracy. Fewer characteristics are preferable since they decrease the model's complexity, and a simplified model is better to understand and explain. We used two feature selection methods in this study f classif and chi square. After using this feature, we can see which critical feature has the most influence and which has the least. Figure 3.3 depicts how many questions have a greater or lesser impact on divorce prediction. This figure is f_classif feature selection.

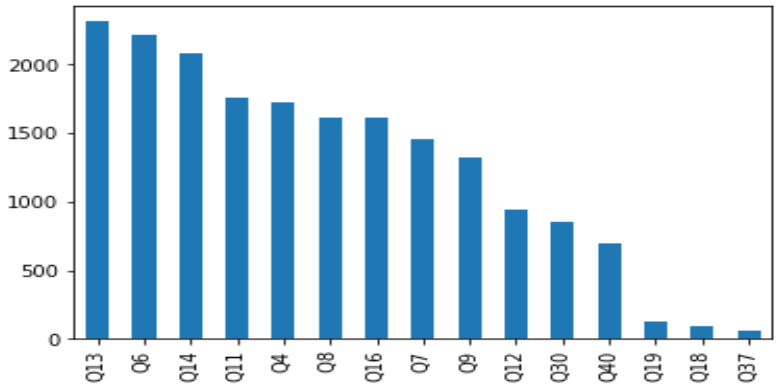


Figure 3.3: More impact question in f classif feature selection.

Figure 3.4 for each input feature, a bar chart of feature relevance scores is constructed.

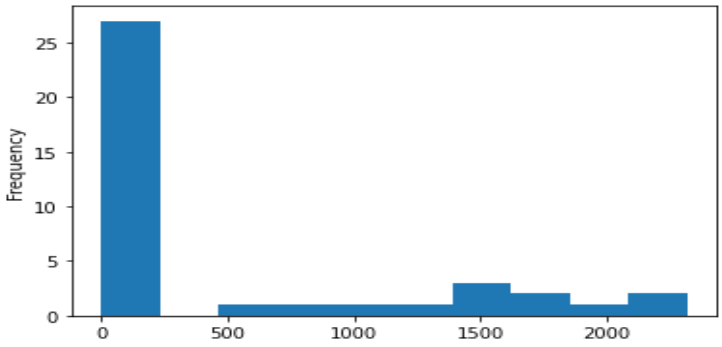


Figure 3.4: Bar chart of Score in f classif feature selection.

We also choose features using the chi-square method. It test and how it makes a difference in machine learning By studying the connection between the features, the chi-square test aids us in fixing the issues of feature selection. Figure 3.5 depicts how many questions have a greater or lesser impact on divorce prediction. This figure is chi-square feature selection

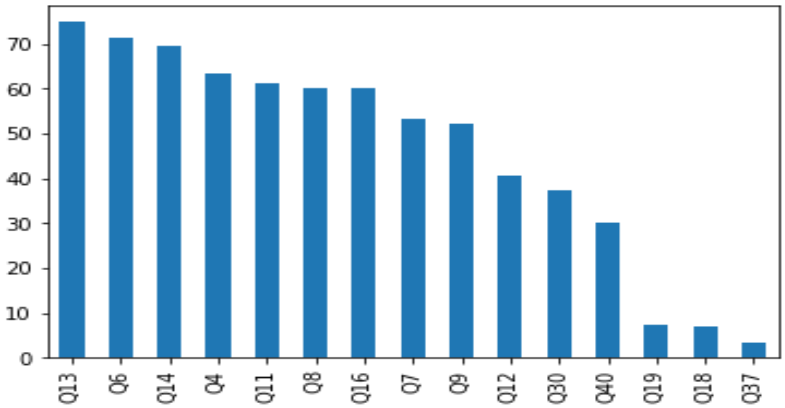


Figure 3.5: More impact question in chi-square feature selection.

Figure 3.6 for each input feature, a bar chart of feature relevance scores is constructed in chi-square feature selection.

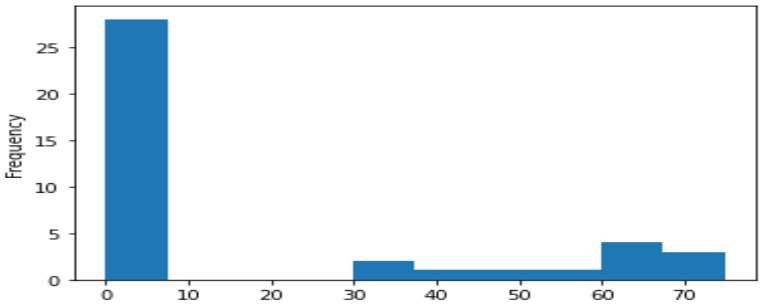


Figure 3.6: Bar chart of Score in chi-square feature selection.

Following feature selection, we obtain a figure 3.7 depicting the number of divorced and happily married couples in the data sheet.

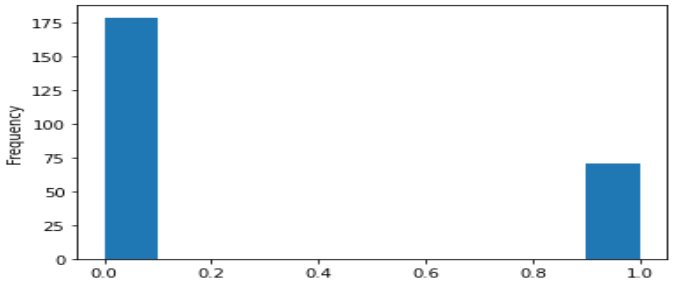


Figure 3.7: Bar chart of divorce and happily married couple.

The correlation between the features is shown in Figure 3.8, and the result is shown using a box plot.

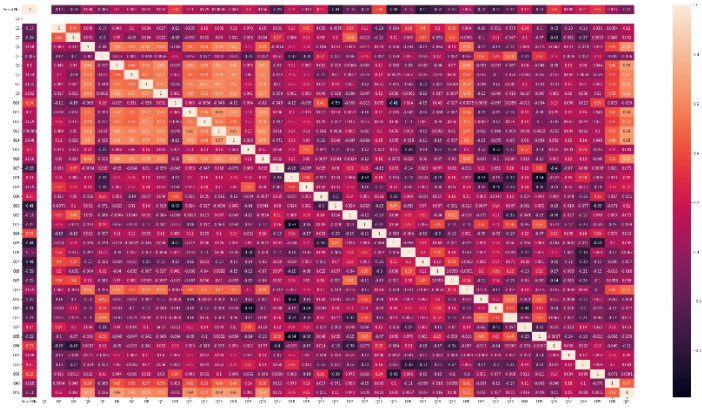


Figure 3.8: Correlation Matrix.

A correlation matrix now describes the features' interrelation. So according statistics, 71 individuals divorce each year, whereas 179 people remain blissfully married

CHAPTER 4

RESULTS AND ANALYSIS

4.1 Introduction

The dataset processing techniques were addressed in the preceding section. Some algorithms use the processed data, and the outcomes of the algorithm will be discussed in this section. All of these algorithms are used, and the results are compared to see which approach delivers the best accuracy: kNN, support vector machine (SVM), nave Bayes, decision tree, and random forest. We acquire 250 data points from divorced and non-divorced people, of which 80% is used as training data and 20% as test data. Divorce Data is the name of our dataset.

4.2 Experimental Results & Analysis

We employed five machine-learning algorithms and compared them by measuring the accuracy and confusion matrix of each technique.

4.3 Experimental Evaluation

On generated datasets, we run five machine-learning algorithms. The accuracy of five algorithms is shown in Figure 4.0. kNN has an accuracy of 88.0 %, SVM has an accuracy of 100 %, Nave Bayes has an accuracy of 94.0 %, random forest does have an accuracy of 96.0 %, and decision tree has an accuracy of 94 %.

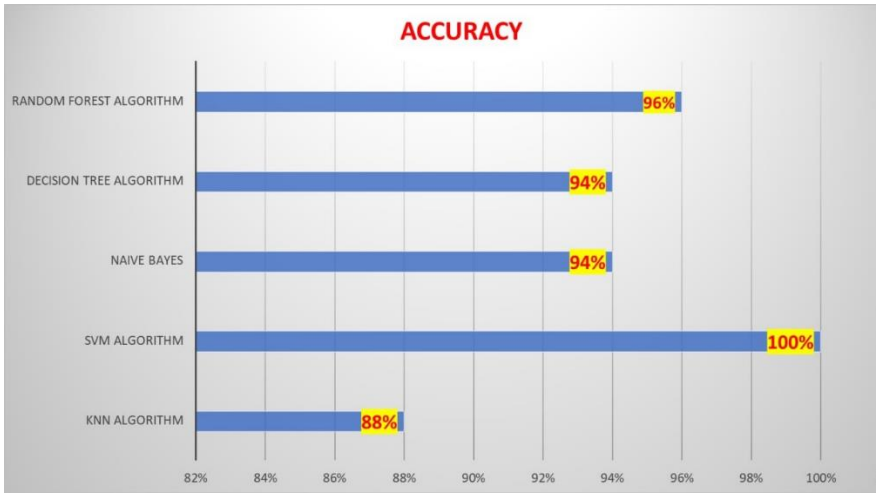


Figure 4.0: Accuracy Chart

Table 4.1 depicts the table where all test data accuracy is listed. Only the SVM algorithm gives the best results. The second best algorithm, random forest, provides excellent accuracy. After that, give the best accuracy Nave Bayes, decision tree, and KNN algorithms.

Algorithms	**Test data usage rate (%)**			
	20%	**30%**	**40%**	**50%**
*k*NN	88.0	86.0	90.0	91.0
SVM	100	95.0	96.0	96.8
Naïve Bayes	94.0	92.0	93.0	93.0
Random forest	98.0	95.0	93.0	95.0
Decision tree	90.0	91.0	85.0	82.0

Figure 4.1: Accuracy of test Data

4.4 Descriptive Analysis

We calculated not only the accuracy of multiple mechanisms, but also the confusion matrix for each algorithm. Any developed model should include an evaluate of that approach. Certain classifications must be checked in the case of model evolution. For better measurement, classifications are measured using the test data set. Amongst the most significant performance evaluation approaches for machine learning classification is the confusion matrix. This will execute the classification techniques against by the test data and produce the true positive, true negative, false-positive, and false-negative values in a tabular manner. The Confusion Matrix is critical for analyzing any classifier's performance.

The confusion matrix of all methods utilized in our model is shown in Table 4.2. In the following table, the model evaluation of each classifier is described with a value.

TABLE 4.2: ALL CLASSIFIER CONFUSION MATRIX

Random Forest

True Class		No	Yes
	No	90	1
	Yes	5	29
	Predicted Class		

Decision Tree

True Class		No	Yes
	No	73	18
	Yes	4	30
	Predicted Class		

Naïve Bayes

True Class		No	Yes
	No	86	5
	Yes	4	30
	Predicted Class		

SVM

True Class		No	Yes
	No	91	0
	Yes	4	30
	Predicted Class		

KNN

True Class		No	Yes
	No	83	8
	Yes	3	31
	Predicted Class		

4.5 Discussion

This section examines algorithm performance, including accuracy, sensitivity, and specificity. The role and equations of evolution models are also examined. As seen, the SVM method has the best accuracy (100%). 96.0 % accuracy was attained by the Random Forest method, 94.0 % by Decision Tree, 94.0 % by Nave Bayes, and 88.0 % by KNN. Finally, we observe that our divorce predictive algorithm performs best when implementing the SVM method.

CHAPTER 5

LIMITATION, CONCLUSION AND FURTHER STUDY

5.1 Limitations and Conclusions

Our research is based about using machine learning techniques to predict divorce. Our work and approach both have some limitations and flaws. We selected a rather small data set; a larger and more diverse data collection would have been preferred. People from various professions, districts, region and social groups were unable to collect data due to certain limitations. Many complex approaches might be employed for data processing, and the model could be nicely presented using many variants in algorithm application. On the dataset used in this study, we used five algorithms. SVM and Random Forest, for instance, have accuracy of 100% and 96.0 %, respectively. Our purpose was to use machine learning to predict divorce rates among Bangladeshi families. If an effective detection method can be developed with the knowledge presented in the study, this will save thousands of families from being broken.

5.2 Implication for Further Study

In future research, the classification performance of various machine learning models can be improved by increasing the quantity of datasets. Feature selection can potentially be employed in the future to help reduce training time and improve our model's accuracy.

REFERENCE

[1] Gottman, J. M., "The Marriage Clinic: A Scientifically-Based Marital Therapy", New York: WW Norton and Company, 1999.

[2] Gottman, J. M., “What Predicts Divorce? The Relationship Between Marital Processes and Marital Outcomes”, New York: Psychology Press, 2014.

[3] Gottman, J. M. and Gottman, J.S. Çiftler Arasında Köprüyü İnşa Etmek: Gottman Çift Terapisi Eğitimi 1. Düzey Kitabı, "Level 1 Clinical Training. Gottman Method Couples Therapy. Bringing to Couple Chasm", İstanbul: Psikoloji İstanbul, 2012.

[4] Gottman, J. ve Silver, N., Aşk Nasıl Sürdürülür. Aşk Laboratuarından Sırlar. (trans. Gül, S.S.) "What Make Love Last. How to Build Trust and Avoid Betrayal", İstanbul: Varlık Yayınları, 2014.

[5] Gottman, J. and Silver, N. Evliliği Sürdürmenin Yedi İlkesi.(trans. Gül, S.S.). "The Seven Principles for Making Marriage Work" İstanbul: Varlık Yayınları, 2015.

[6] Yöntem, M.K. and İlhan, T., "Development of the Divorce Predictors Scale". Sosyal Polika Çalışmaları Dergisi. 41, 339-358, 2018.

[7] Yöntem, M , Adem, K , İlhan, T , Kılıçarslan, S . "Divorce Prediction Using Correlation Based Feature Selection And Artificial Neural Networks", Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi , 9 (1) , 259-273, 2019. Retrieved from https://dergipark.org.tr/en/pub/nevsosbilen/issue/46568/549416

[8] Bae, S. M., Lee, S. H., Park, Y. M., Hyun, M. H., and Yoon, H., "Predictive Factors of Social Functioning in Patients With Schizophrenia: Exploration For The Best Combination of Variables Using Data Mining", Psychiatry Investigation, 7(2), 93-101, 2010.

[9] Babcock, J.C., Gottman, J., Ryan, K. and Gottman, J., "A Component Analysis of a Brief Psycho-Educational Couples' Workshop: One-year Follow-up Results", Journal of Family Therapy, 35(3), 252-280, 2013.

[10] Baca-García, E.,"Using Data Mining to Explore Complex Clinical Decisions: A Study of Hospitalization After Suicide Attempt", Journal of Clinical Psychiatry, 67(7), 1124-1132, 2006.

[11] Ruck, D. W., Rogers, S. K., & Kabrisky, M., "Feature selection using a multilayer perceptron", Journal of Neural Network Computing, 2(2), 40- 48, 1990.

[12] Pal, S. K., & Mitra, S.,"Multilayer perceptron, fuzzy sets, classifiaction", 1992.

[13] Rish, I., "An empirical study of the naive Bayes classifier", In IJCAI 2001 workshop on empirical methods in artificial intelligence, Vol. 3, No. 22, pp. 41-46, 2001.

[14] Murphy, K. P., "Naive bayes classifiers", University of British Columbia, 18(60), 2006.

[15] Oshiro, T. M., Perez, P. S., & Baranauskas, J. A.,"How many trees in a random forest?", In International workshop on machine learning and data mining in pattern recognition, pp. 154-168, Springer, Berlin, Heidelberg, 2012.

[16] Hall, M. A. and Smith, L. A., "Practical Feature Subset Selection for Machine Learning", In Computer Science'98 Proceedings of the 21st Australasian Computer Science Conference ACSC, 98: 181-191, 1998.

[17] Hall, M., "Correlation-Based Feature Selection for Machine Learning", Phd Thesis, Department Of Computer Science, Waikato University, New Zealand, 26-28, 1999.

Prediction Divorce Prospect using Machine Learning Methods

ORIGINALITY REPORT

22%	16%	7%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	12%
2	Md. Ashrafujjaman Tutul, Md. Mehedi Hasan, Shashata Kumar Mondol, Moshaddek Hossain, Ahmed Al Marouf. "Divorce Prediction using Machine Learning Methods-Bangladesh Perspective", 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021 Publication	4%
3	Submitted to Daffodil International University Student Paper	2%
4	www.thedailystar.net Internet Source	2%
5	"Contemporary Issues in Family Studies", Wiley, 2013 Publication	1%
6	"Intelligent Information and Database Systems", Springer Science and Business	1%