Predicting Divorce Likelihood using SVM and Platt Scaling

Jonathan Henin

DSC680

Bellevue University

April 05, 2020

jhenin@my365.bellevue.edu

Instructor: Catherine Williams

Abstract

With the divorce rate still significantly high, this study aims to see if the features that predict divorce are possible to identify in hopes that a tool could be built that helps couples self-assess their marital health early on. This study takes marriage survey data from 170 individuals, both divorced and married, and analyzes using a Support Vector Machine and Platt Scaling. It determines that predicting the likelihood of divorce is possible through self-reporting methods and opens the doors for a tool to be built.

Introduction

The U.S. divorce rate has seen steady declines from its highs in the 90s of 50%, to 44% in 2018 as reported by the CDC/NCHS National Vital Statistics. While this is certainly a positive trend, there are other factors that are playing a role in this decline.

First, the average age of first-time marriages has been steadily increasing over the years, with current rates at 29.8 for men and 27.8 for women, an increase of 7 years for both over the last half century. This has resulted in a drop of the percent married among 18 to 34 year-olds from 59% to 29% over the same time period; the lowest recorded rate on record. There are many speculative reasons for this, such as financial stability, fall in religious adherence, the increase of education and income for women, rising student debt and housing costs, the increase in cohabitation and those who believe marriage is an outdated institution (Zagorsky, 2016).

The second reason is the increase of those choosing to go to pre-marital counselling. A meta-analysis study by Carroll & Doherty (2003) found that couples who went to premarital counselling had 30% higher marital success rates than couples who did not. The study also claims that around 44% of couples who get married today go to a premarital counseling session. Millennials lead with the highest rate at 51%. Early counselling can help provide couples the tools they need to deal with problems before they become insurmountable obstacles. But there are hurdles that can prevent people from seeking counselling.

Marriage counselling can be a challenge for many to attend. The cost of therapy per session can range from $25 up to $250, depending on the type of counsellor; with multiple sessions being required, a 12-week program at the average rates can cost couples between $1300 and $1800. But cost alone isn't the only challenge. Finding time to go as a couple can often mean sacrifices; either in the form of taking time off work, finding a counsellor that is available after work, or going during weekends. Since each counselor has their own personal style, finding the right one that works for both partners mean couples might go through a few different ones before they find one that matches their needs. Stigmas can also play a big role as a study by Baptista et al. (2017) found that self- and social stigmas can be the biggest hurdle that prevents people from seeking professional help. Despite the challenges, delaying marriage counselling can make it harder.

On average, those seeking marriage counselling are 6 years into their marriage (Gottman & J.M., 1994). Given that according to the CDC the average first marriage lasts 8 years (6 years across all marriages), counselling is usually considered a last resort for a failing marriage; colloquially considered divorced counselling. According to Gottman, of those seeking counseling at this point, only between 11% and 18% make any meaningful gains that last more than a year. The longer couples wait to seek help, the harder it is for them to reconcile.

Leaning on the benefits of pre-marital counselling, having the ability to assess a relationship's health early on is an important step to further lowering the divorce rate, and perhaps extending benefits to long term non-married couples. Thus, if a resource, such as an app, was available to provide couples with individualized assessments, that did away with the barriers to entry to counselling, there could be a significant benefit to relationships. To start, we want to see if we can build out a survey that will assess the likelihood of divorce to see if we can identify some of the major warning signs in a relationship preemptively.

Background

This study is based on prior work done by Yöntem, M. K., Adem, K., İlhan, T. ve Kılıçarslan, S. (2019) in their paper *Divorce Prediction Using Correlation Based Feature Selection and Artificial Neural Networks.* In this study, a group of participants filled out a survey answering questions intended to gauge their marital health. These questions were based on previous research by Gottmon who has done over 30 scientific journals and leads the Gottman Institute.

Gottman is perhaps best known for his work in predicting divorce based on indicators in a conversation. In a 1992 study, he interviewed over 52 heterosexual, married couples, each having a 15-minute discussion about an area of conflict in their marriage; the purpose being to study how the couple communicated and handled conflict resolution. They looked at variables such as affection, negativity, expressiveness, gender stereotyping, volatility and martial disillusionment. They were able to predict with 93.6% accuracy which couples would eventually divorce. As negativity was the prevailing factor, in later work, they were able to fine tune their indicators, leading to a prediction accuracy of 98.2% over the course of just a 3-minute conversation (Carrère & Gottman, 1999). These negative factors were named the Four Horsemen of the Apocalypse and are: Criticism, Contempt, Defensiveness, and Stonewalling. With contempt being the biggest indicator of divorce.

Methods

Data

The dataset used for this study originated from Yöntem et al. (2019) and made public through UCI Machine Learning Repository. In this study, 170 participants from regions in Turkey, 84 of them being divorced and 86 of them being married (who are not considering divorce) filled out a survey asking them about their marriage. Unfortunately, none of the demographic information was provided in the dataset so we cannot test any correlations to the outcome. However, the information is rather interesting so here is what the study describes. The ages of the participants ranged from 20 to 63 with the mean age being 36.04 with a standard deviation of 9.34. Additional statistics from the study:

- 74 (43.5%) were married for love vs 96 (56.5%) were married in an arranged marriage.
- 127 (74.7%) had children vs 43 (25.3%) had no children.
- 18 (10.58%) were primary school graduate, 15 (8.8%) were secondary school graduate, 33 (19.41%) were high school graduate, 88 (51.76%) were college graduate, and 15 (8.8%) had master's degree.

The data consists of 54 feature scores, ranging in values between 0 and 4, and a Classification (1 = Divorced, 0 = Married). The 54 scores correspond to 54 questions in the survey, which can be seen in the compendium below. The survey questions are based Yöntem and İlhan (2017, 2018) work which developed the questions and the scoring methodology, named Divorce Predictors Scale (DPS). Using this method, each question is answered on a 5-point scale i.e., (0 = Never, 1 = Seldom, 2 = Averagely, 3 = Frequently, 4 = Always). The questions are re-weighted, such that negative responses are weighted higher, e.g., a response of Always on question 9 will result in a 0 value rather than a 4.

Model Evaluation

With the objective being to provide a percentage of likelihood, a probabilistic model was determined to be the best candidate. During the Exploratory Data Analysis the classes revealed that they were well defined such that a Linear Classification model could perform well. With those two considerations, the models that were considered were a Logistic Regression model and a Support Vector Machine. Logistic Regression can have issues with multicollinearity, that is when independent variables are strongly correlated with each other. Thus, it is necessary to determine the level of correlation. As seen below in Figure 1, many of the questions do have correlation between them. Performing a Logistic Regression model would require pruning of the questions just to get it to run. As well, reducing features removes questions from contributing to the likelihood result, making those questions irrelevant (the score would not change no matter the answer). Due to those restrictions, it was decided not to include a Linear Regression.

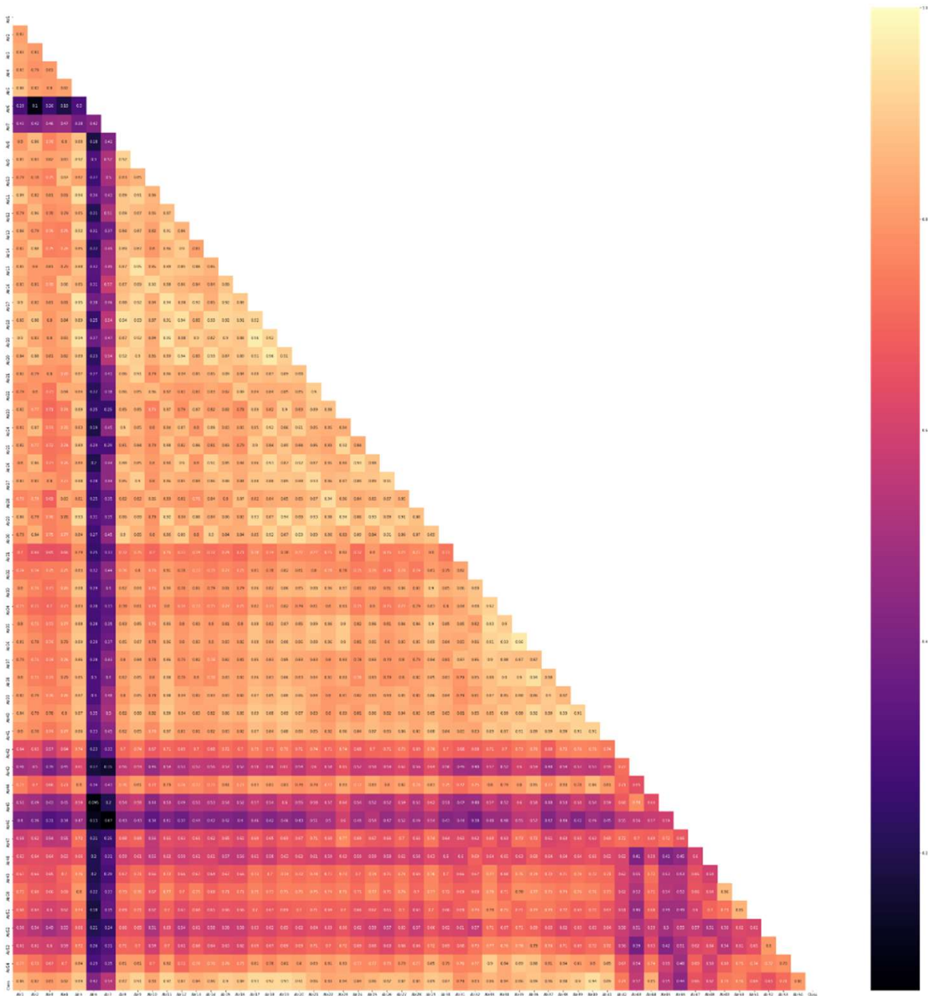

Figure 1. Correlation of questions. Highly Correlated = Top of Scale. No Correlation = Bottom of Scale.

With a Support Vector Machine model, depending on the kernel being used (particularly a linear one), multicollinearity can have an impact on the accuracy and performance of the model as well, but it is not necessary to get the model to run. Rather, feature reduction is done to help improve a model. In this analysis, improving the accuracy of the model was not necessary as will be noted in the results. The benefit of a SVM is that it treats every feature as an independent variable and has no problem running models with many features (except performance run time). However, SVMs are not probabilistic by default, in order to achieve that, they need to be calibrated. The two most common methods for calibrating probabilities are Platt Scaling and Isotonic Regression. Platt Scaling is simpler and because the predicted probabilities are sigmoid-shaped, due to squashing the predicted values between 0 and 1, this method also performs the best.

In summary, we passed every feature through an SVC model, and then calibrated it using Platt Scaling to transform the analysis into a sigmoid probability. The model was fitted using a 10-fold cross validation method with the following parameters:

C=1.0, break_ties=False, cache_size=200, class_weight=None, coef0=0.0, decision_function_shape='ovr', degree=3, gamma='scale', kernel='linear', max_iter=-1, probability=False, random_state=None, shrinking=True, tol=0.001, verbose=False), cv=10, method='sigmoid’

Technology

The model was coded in Python, using scikit-learn’s SVC and CalibratedClassifierCV methods for the SVM model and calibration. The original survey was also reproduced in Python. A Docker file was created which contains the model and is hosted using Google Cloud’s Kubernetes engine. An API call was developed using Flask. The API takes a json file, which are key:value pairs of the questions and scores, and then processes it against the model; providing a likelihood of divorce.

Survey

The original survey was recreated in Python and is presented to the user through the command line interface. In its current state, a user can take the survey and upon completion is presented with a likelihood of divorce percentage. Given more time a better methodology would be to hook up to Survey Monkey and pass their API through to my API and then present the score, though it is not known if that workflow is supported.

Results

Model performance was assessed using a Receiver Operating Characteristics (ROC) curve, with an Area Under the Curve (AUC) score of 1.00. Additional metrics for model performance were run with the following results: Accuracy = 1.0, Precision = 1.0, Recall = 1.0, F1 Score = 1.0. A preliminary correlation analysis of the features versus the target class was performed and are presented in Table 1.

Table 1. Initial Correlation Analysis

Features	Correlation
x40	.939
x17	.929
x19	.928
x18	.923
x11	.918
x9	.912

To identify the feature importance of the model, the exponential of the average of the coefficients from the 10 SVM models was calculated for each feature. The results are presented in Table 2. This is a method described by Guyon et al. (2002) in which they identified gene selection for cancer classification using Support Vector Machines. Validation of the weights was performed by getting a baseline prediction (7.71%), by passing a survey set of all 0 scores, and then testing the impact of setting each question independently to a value of 4.

Table 2. Calibrated Model Feature Significance

Features	Exponent of Coefficient
x40	1.184
x49	1.177
x26	1.155
x17	1.127
x28	1.116
x3	1.106

In contrast, the original study by Yöntem et al. (2019) used an Artificial Neural Network (ANN). Their model's top 6 most effective features were x2, x6, x11, x18, x28, and x40 with an accuracy of .9882.

Discussion

The developed model performed with perfect accuracy among every metric of performance tested. These results are generally in line with the original studies 98.82% accuracy and Gottman's couples therapy (which the survey is based on) which claims an accuracy of 98.23%. Looking at the high correlation of the features to the target class, it might not be that unusual. However, the perfect accuracy of the model is more a statement on the data than a real-world reflection of the accuracy of the model. The data collected are from two highly different groups of individuals, happily married and ones who are divorced. Data was not collected from anyone about to go through a divorce or questioning. It stands to reason that due to the data being from two polar opposite groups, a clear delimitation boundary would be observed that would make a SVM model perform very well. While the two groups were relatively balanced, the low number of participants (n = 170) also contributes to an inflated model performance. In an ideal situation, we would include individuals across the spectrum of marital health (self-reported) as well as increase the number of participants. Then track them over the course of time to see how their survey responses change. This is essential to see if a couple always had issues or if they developed over time.

Given that there are clear differences in responses between the two groups, it is worth it to examine the questions which are high / low predictors of divorce. Categorizing the questions, the questions that have the highest predictive power deal with Issue Resolution, Communication, Views on what makes a Happy Marriage, and Familiarity with their spouse. While things that scored low in predictive power are categories such as Common Goals, Common Interests and Stonewalling.

Again, without knowing any baselines of these individuals, it is hard to draw any conclusion whether these were always issues.

According to the Gottman study in 1999, contempt was classified as the most destructive form of communication. Gottman defines and says this of contempt, “When we communicate with contempt, we are truly mean. Treating others with disrespect and mocking them with sarcasm and condescension are forms of contempt. So are hostile humor, name-calling, mimicking, and body language such as eye-rolling and sneering. In whatever form, contempt is poisonous to a relationship because it conveys disgust and superiority, especially moral, ethical, or characterological.” In his book, he notes, “When contempt begins to overwhelm your relationship you tend to forget entirely your partner’s positive qualities, at least while you’re feeling upset.” Contempt questions are 32-36 and 52-54. The majority of these questions fall in the top 20 predictors in our model. While the physical signs of contempt might be easier to spot during a recorded conversation with a partner, they are harder to determine in self-reporting surveys. However, they could manifest in questions referring to listing positive qualities about their partner.

The goal of developing this model and survey were to see if it was possible to develop a self-reporting questionnaire that could predict eventual divorce between a married couple. This initial model is very promising and gives credence that an app or tool could be built that would help couples monitor their marital health. As well, considering that the number of married individuals is falling, while cohabiting partners is rising, we see potential in expanding the scope to cover premarital relationships in future work. Identification of issues at any stage during a relationship could prove beneficial, though we note that there may be challenges in getting accurate unbiased self-reporting results in early stages of a relationship. Still, an application that could help in a relationship should not be limited to just married individuals, as seen from the benefits of premarital counselling. Continuation of this work could prove beneficial to anyone in a relationship, and our hope that such an application can be built.

Acknowledgements

Thank you to the peers and professors in DSC680. This research was made possible with the resources provided by Bellevue University.

Compendium

List of Questions:

1 If one of us apologizes when our discussion deteriorates, the discussion ends
2 I know we can ignore our differences, even if things get hard sometimes
3 When we need it, we can take our discussions with my spouse from the beginning and correct it
4 When I discuss with my spouse, to contact him will eventually work
5 The time I spent with my wife is special for us
6 We don't have time at home as partners
7 We are like two strangers who share the same environment at home rather than family
8 I enjoy our holidays with my wife
9 I enjoy traveling with my wife
10 Most of our goals are common to my spouse
11 I think that one day in the future, when I look back, I see that my spouse and I have been in harmony with each other
12 My spouse and I have similar values in terms of personal freedom
13 My spouse and I have similar sense of entertainment
14 Most of our goals for people (children, friends, etc.) are the same
15 Our dreams with my spouse are similar and harmonious
16 We're compatible with my spouse about what love should be
17 We share the same views about being happy in our life with my spouse
18 My spouse and I have similar ideas about how marriage should be
19 My spouse and I have similar ideas about how roles should be in marriage
20 My spouse and I have similar values in trust
21 I know exactly what my wife likes
22 I know how my spouse wants to be taken care of when she/he sick
23 I know my spouse's favorite food
24 I can tell you what kind of stress my spouse is facing in her/his life
25 I have knowledge of my spouse's inner world
26 I know my spouse's basic anxieties
27 I know what my spouse's current sources of stress are
28 I know my spouse's hopes and wishes
29 I know my spouse very well
30 I know my spouse's friends and their social relationships
31 I feel aggressive when I argue with my spouse

32 When discussing with my spouse, I usually use expressions such as ‘you always’ or ‘you never’
33 I can use negative statements about my spouse's personality during our discussions
34 I can use offensive expressions during our discussions
35 I can insult my spouse during our discussions
36 I can be humiliating when we discussions
37 My discussion with my spouse is not calm
38 I hate my spouse's way of open a subject
39 Our discussions often occur suddenly
40 We're just starting a discussion before I know what's going on
41 When I talk to my spouse about something, my calm suddenly breaks
42 When I argue with my spouse, ı only go out and I don't say a word
43 I mostly stay silent to calm the environment a little bit
44 Sometimes I think it's good for me to leave home for a while
45 I'd rather stay silent than discuss with my spouse
46 Even if I'm right in the discussion, I stay silent to hurt my spouse
47 When I discuss with my spouse, I stay silent because I am afraid of not being able to control my anger
48 I feel right in our discussions
49 I have nothing to do with what I've been accused of
50 I'm not actually the one who's guilty about what I'm accused of
51 I'm not the one who's wrong about problems at home
52 I wouldn't hesitate to tell my spouse about her/his inadequacy
53 When I discuss, I remind my spouse of her/his inadequacy
54 I'm not afraid to tell my spouse about her/his incompetence

References

Baptista, Makilim Nunes, & Zanon, Cristian. (2017). Why not Seek Therapy? The Role of Stigma and Psychological Symptoms in College Students. Paidéia (Ribeirão Preto), 27(67), 76-83. https://doi.org/10.1590/1982-43272767201709

Carrère, S., & Gottman, J. M. (1999). Predicting divorce among newlyweds from the first three minutes of a marital conflict discussion. Family Process, 38, 293-301. doi:10.1111/j.1545-5300.1999.00293.x

Carroll, J. S., & Doherty, W. J. (2003). Evaluating the Effectiveness of Premarital Prevention Programs: A Meta-Analytic Review of Outcome Research. *Family Relations*, 52(2), 105–118. doi: 10.1111/j.1741-3729.2003.00105.x

Center for Disease Control's National Vital Statistics System (2018). Provisional number of marriages and divorces. Retrieved from https://www.cdc.gov/nchs/data/dvs/national-marriage-divorce-rates-00-18.pdf

Franc, V., Zien, A., & Sch¨olkopf, B. (2011). Support Vector Machines as Probabilistic Models. *International Conference on Machine Learning*. Retrieved from http://www.icml-2011.org/papers/386_icmlpaper.pdf

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63*(2), 221–233. https://doi.org/10.1037/0022-3514.63.2.221

Gottman, J. M. (1994). *What predicts divorce? The relationship between marital processes and marital outcomes*. Lawrence Erlbaum Associates, Inc.

Guyon, I., Weston, J., Barnhill, S. et al. Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning 46, 389–422 (2002). https://doi.org/10.1023/A:1012487302797

Shapiro, A. F., Gottman, J. M., & Carrère, S. (2000). The baby and the marriage: Identifying factors that buffer against decline in marital satisfaction after the first baby arrives. Journal of Family Psychology, 14(1), 59–70. doi: 10.1037/0893-3200.14.1.59

Yöntem, M. K., Adem, K., İlhan, T. ve Kılıçarslan, S. (2019). Divorce Prediction Using Correlation Based Feature Selection and Artificial Neural Networks. Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 9(1), 259-273.

Yöntem, M.K. and İlhan, T. (2018). Boşanma Göstergeleri Ölçeğinin Geliştirilmesi.[Development of the Divorce Predictors Scale]. Sosyal Polika Çalışmaları Dergisi. 41, 339-358.