University of Warsaw
Faculty of Economic Sciences

Paulina Sereikytė
Book number: 444470

Dawid Szyszko-Celiński
Book number: 443709

Predicting spam e-mails using binary dependent variable econometric model

Paper was prepared as a part of Econometrics course taught by Rafał Woźniak

Warsaw, June 2022

Abstract

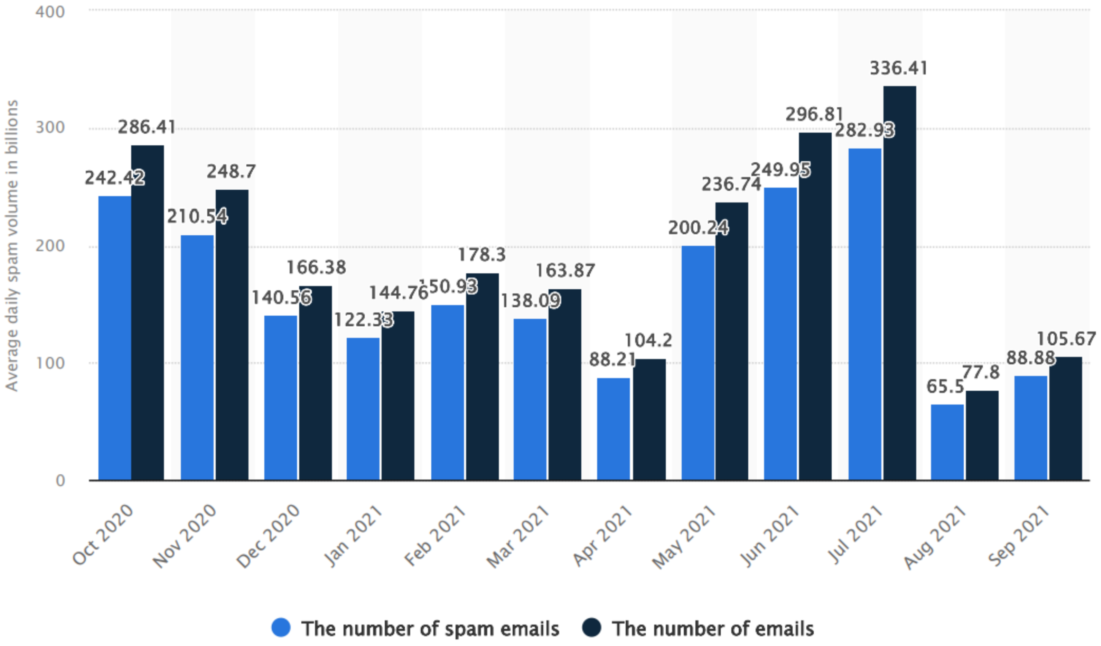
The aim of the paper is to build a predictive model that can evaluate whether an email is spam. Moreover, we want to test what keywords and other attributes indicate the spam messages. Last but not least we want to compare our findings with currently available literature on this matter. Our hypothesis is that words connected with the money used in the mails increases the probability that the message is spam. The binary dependent variable models were used to perform classification of the emails, especially logit model. At first general probit and logit models were created. Then using a general to specific approach all insignificant variables were removed. What is more, interaction terms were added to ensure that all relevant factors were included and the specification of the model is correct. Having checked the model in various ways we calculated marginal effects of the variables and interactions. We found out that most of the words connected with money increase the probability of mail being spam, which goes along with the main thesis. We also found out that if the mail contains the name of the receiver of the message then the probability that it is spam is significantly lower.

Introduction

Many of us get spam emails every day. Some people just ignore them, some get annoyed and some might be influenced by them. It might appear that such mails are rather uncommon, but in fact some sources claim that spam mails can generate around 80% - 90% of all mail movements. It can be shocking but we might be affected by spam mails more than we think. But what actually is a spam message and why should we care about them? Typically spam mails are unwanted and automatically sent messages to the gathered recipient list for commercial purposes. At the first glance it might sound like a quite unpleasant advertisement but in reality spam emails can be very dangerous. Some messages can contain malicious links that may infect the user computer, which can lead to money robbery or even identity theft. What is more, spam messages can just simply contain scam offers that can lead to losing the money. Having said that, it seems that spam emails are not only very common but also can be very dangerous. That fact led us to create a model that could detect and mark spam messages, and it can be achieved using binary dependent variables models. There are also many different methods, especially based on machine learning methods, but econometric models can also be very effective in that matter. Using such tools it can help to lower the risk of opening malicious emails or just simply save time. What is more, in such models it is worth mentioning error types. Models cannot predict the type of email every time and the errors in predictions are a part of the modelling process. For spam detection it is better to mark “not-spam” email as “spam” (first type error) rather than not marking “spam” email as “not-spam” (second type error). Some of “not-spam” messages will be filtered out, but that is safer for the user than letting in potentially dangerous mails.

Average daily spam volume worldwide from October 2020 to September 2021

(in billions)



Sources: (Cisco, n.d.), (Statista, 2022)

Our main hypothesis is that the words connected with money used in the mail increases the probability of messages being spam.

Literature review

To form the foundation for our research and build sufficient insight into the topic of spam messages to form a hypothesis several scientific articles were reviewed. The first two articles to be discussed review the existing literature and methodology in spam message classification, whereas the last one discussed the methodology and data of SMS spam filtering.

In a recent paper called *“A systematic literature review on spam content detection and classification”* by Kaddoura et. al. **(Kaddoura et al., 2022,)** the authors present a comprehensive review of the spam detection literature and methodology to date. They describe the spam classification process as consisting of 4 steps: dataset collection, pre-processing of data, feature extraction and classification. The data collection process is the most straightforward - the data for spam detection analysis is typically collected from social media sites such as twitter or youtube, as well as direct communication such as email and text. The existence of readily available datasets are also acknowledged.

The following step consists of removing unnecessary data. Several techniques are described, such as *tokenization* (separating text into smaller tokens), *stemming* (reducing text to fundamental meanings), *normalization* (simplifying terms), l*emmatization* (reducing words to the dictionary form) and stopwords (removing meaningless words). The pre-processed data then has to be converted into some kind of numerical form to be analysed. Some of the most commonly used techniques are *Bag of Words (BoW)* (counting the number of known word occurrences in each text), *n-grams* (analysing the occurrences of continuous phrases) and *Term frequency-inverse document frequency (TF-IDF)* (dividing the number of word occurrences by the total number of words in the text). Previous research demonstrates that BoW and TF-IDF techniques yield similar results in spam classification (Kaddoura et al., 2022).

Finally, spam classification techniques are discussed. 3 spam classification technique categories are distinguished: *rule-based systems, machine learning* and *hybrid approaches.* Rule-based systems consist of classifying the text as spam if the text reaches a certain amount of spam-words that could be assigned weights. This methodology has some limitations, as the rules need to be updated to remain relevant and large amounts of text are needed to produce sound rules. Previous authors have achieved 82-98% accuracy using rule-based classification. Machine learning techniques are more complex, including both supervised and unsupervised learning and ranging from logistic regressions, random forests and K-nearest neighbours techniques. These methods are more adaptive, however they suffer from high computational complexity. Similarly, a research paper by Marhwa & Singla (Marwaha & Singla, 2021) highlights that Knowledge Engineering and Machine Learning are the most commonly used techniques for spam message filtering.

Research by Delaney, Buckley & Greene (Delany et al., 2012) into SMS spam messages suggested the existence of several distinct spam messaging clusters, namely *Ringtones, Claims, Competitions, Prizes, Voicemail, Dating, Services, Finance* and *Chat*. *Voicemail* and *Ringtones* categories could be attributed to SMS messages only and therefore are outdated in case of modern spam data. These keywords can be used to further our research and model-building:

Ringtones	*Send, ringtone, text, tone, free, sms, reply, mobile*
Claims	*Accident, entitled, records, pounds, claim, msg, compensation, opt*
Competitions	*Txt, win, uk, voucher, cash, 150p, send, entry*
Prizes	*Prize, guaranteed, urgent, todays, valid, claim, draw, cash*
Voicemail	*Please, message, voicemail, waiting, call, delivery, immediately, urgent*
Dating	*Dating, service, contacted, find, guess, statement, points, private*
Services	*Mins, video, free, camera, orange, latest, phone, camcorder*
Finance	*Help, debt, credit, info, government, loans, solution, bills*
Chat	*Naughty, ring, alone, chat, xx, heard, luv, home*
Miscellaneous	*Find, secret, admirer, special, looking r* reveal, contact, call*

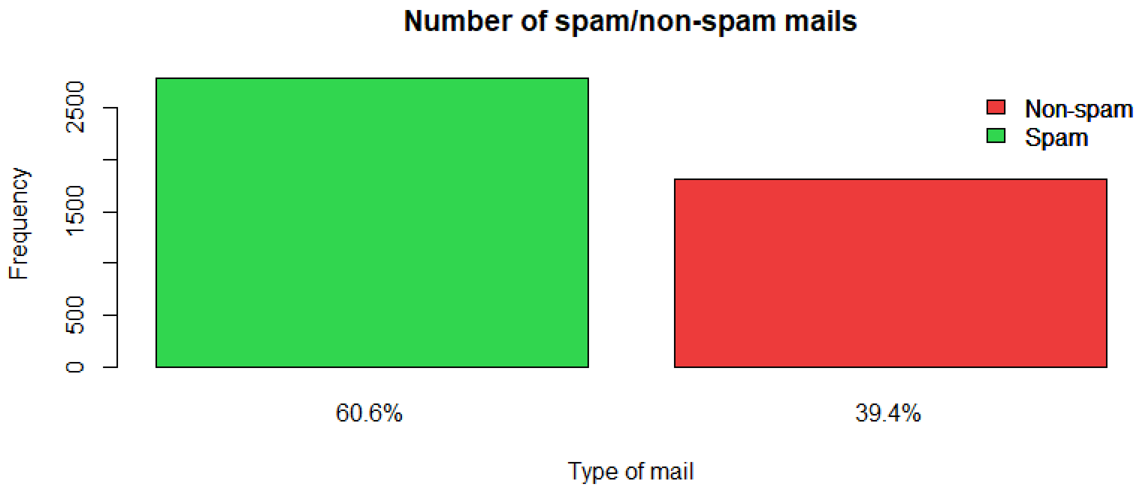
The described research papers help us form a theoretical foundation for our research and align expectations of the results. Firstly, it is evident that although machine learning techniques can be more adaptive, simpler rule-based classification algorithms can also achieve satisfactory results with lower computational complexity. Moreover, research by Delaney et. al. provide further insight into the type of spam messages and keywords that can be expected which will be included in creating our models.

Data

Data used in the analysis can be found on the University of California, Irvine data repository under the link: https://archive.ics.uci.edu/ml/datasets/spambase. The data was collected by Mark Hopkins, Erik Reeber, George Forman, Jaap Suermondt (Hewlett-Packard Labs, 1501 Page Mill Rd., Palo Alto, CA 94304) and donated by George Forman in 1999. It consists of 4601 observations (e-mails) and 57 attributes describing them divided into 5 groups:

1) 48 continuous real variables (values from 0 to 100) of type “word_freq_WORD”, which were calculated as a fraction of a certain keyword to the number of all words in the email.
2) 6 continuous real variables (values from 0 to 100) of type “char_freq_CHAR”, which were calculated as a fraction of a certain character (like dollar sign, hash sign etc.) to the number of all characters in the email.
3) 1 continuous real variable (values from 0 to infinity) - average length of uninterrupted sequences of capital letters
4) 2 continuous integer (values from 0 to infinity) - length of longest uninterrupted sequence of capital letters and total number of capital letters in the email
5) 1 nominal variable (1=spam, 0=not spam) - denotes whether the email was considered spam

The dataset is rather balanced - 1813 observations are flagged as “spam” and 2788 as “non-spam”, which gives around 40/60 class distribution. Moreover, there are no missing values in the data. That fact is crucial in terms of usage of binary dependent variables models.



The downloaded dataset consists of two files. “Spambase.data”, which stores the main information needed for the analysis and “spambase.names” that stores names of the variables. Due to the fact that analysis is made in R programming language some variables names had to be transformed, to ensure they will not consist any special characters like “$” or “#”. No further transformations of the data were needed.

Method/Model

The aim of this paper is to create a binary dependent variable predictive model based on the spam data. There could be chosen two simple models to perform such analysis: logit or probit model. Both models are similar in terms of application and interpretation, however in the case of the used data, the logit model tended to give better results and was easier to obtain. That is the reason why in this paper more attention will be put on that kind of model (other reasons will be discussed in further part).

Having chosen the type of model it is important to mention the method that was used during modelling. The general to specific method has been chosen, as its workflow is intuitive and helps to narrow down the general model to more specific ones. General logit model was estimated with the maximum likelihood method and it consisted of all 57 explaining variables (please refer to attached Appendix - R file named “Spam Detection Codes” under the provided link).

Then the significance of the general model was tested via the likelihood ratio test in which we compared the general models with a model that depends only on constant. For general logit and probit models results were almost the same the same as follows:

Likelihood ratio test
H0: Beta1 = 0 & Beta2 = 0 & … & Beta57 = 0 (the general model can be simplified to restricted model based only on constant / all Betas are jointly insignificant) **H1:** The general model cannot be simplified to a restricted model. All Betas are jointly significant.
Chi-squared statistic (logit / probit): 4354.4 / 4259.6 **degrees of freedom:** 57 **p-value:** 2.2e-16, which can be rounded to 0
Conclusion: Reject the null hypothesis. General model coefficients are jointly significant.

Null hypothesis for the test assumes that the restricted model (regression on constant) is not significantly different from the unrestricted model (general model). Based on the p-value which can be rounded to 0, the null hypothesis is rejected and the conclusion is that the general model overall is significant. However, some variables in the equation of the model are not significant enough (p-value lower than 5%), that is why we have decided to make stepwise regression. In every iteration the new model has been estimated and p-values of each explaining variable has been checked if they are higher than 5%. If such variables existed in the model, the algorithm would find the variable with highest p-value and remove it from the model. In the next steps new models have been estimated until all variables have been significant at 5% level. Such a method was used for both logit and probit models and at the first glance logit model gave better results considering only AIC value, which was lower. The results of likelihood ratio tests for newly created models pointed to a conclusion that all variables are jointly significant. Having two models with only significant variables the next step is to perform a link test, which is meant to check if the specification of the model is good. For the test we have used the function created by dr Rafał Woźniak, that performs all necessary steps to obtain results (please refer to the appendix - file "linktest").

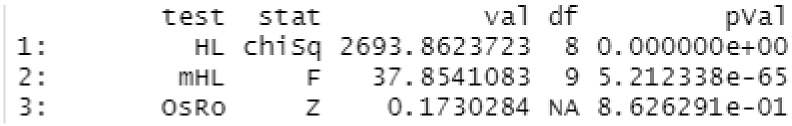
Link Test
if yhat is significant and yhat2 is not signifficant: model has a good specification **if yhat is / is not significant and yhat2 is signifficant:** model does not have a good specification
yhat p-value: 2e-16, which can be rounded to 0 **yhat2 p-value:** 2e-16, which can be rounded to 0
Conclusion: yhat and yhat2 are both significant. It means that the model does not have a good specification.

Unfortunately, for both logit and probit models the test indicated that the specification is not correct. The yhat variable has been significant as it should be, however yhat2 should not be significant to pass the specification test. One of the issues in such a situation is lack of significant interaction between variables in the model. Considering this we have decided to come up with reasonable interactions and include them in the model. Some of the newly added terms were not significant, that is why a stepwise regression algorithm had to be applied again. For the probit model we faced an issue with convergence of the algorithm, however for the logit model algorithm seemed to work well. That is why we decided to carry on with the analysis only with the logit model.

The logit model with only significant variables and interaction was again tested with likelihood ratio test and link test. Additional significant interaction terms helped to obtain a positive outcome of the link test, so the next steps of the analysis can be performed. To make sure that the specification is correct another set of tests were performed, but the 2 most important ones are: Hosmer Lemershow and Osius-Rojek, which results are presented below.

Hosmer Lemershow and Osius-Rojek tests

H0: The model specification is not correct
H1: The model specification is correct



```
          test  stat          val df         pval
1:          HL chisq 2693.8623723  8 0.000000e+00
2:         mHL     F   37.8541083  9 5.212338e-65
3:        OsRo     Z    0.1730284 NA 8.626291e-01
```

Conclusions: (*HL*) Cannot reject Hosmer Lemershow test H0. The model specification is incorrect for this test.
(*OsRo*) Reject Osius-Rojek test H0. The model specification is correct.

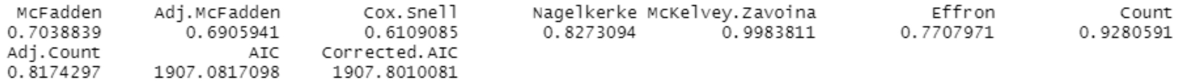
Unfortunately the Hosmer Lemershow test indicated that the specification is not correct, however the Osius-Rojek test p-value is around 86%, which points out that the specification is correct.

Even though HL tests gave negative results, further analysis with the model will be performed as a link test and the Osius-Rojek test gave positive results. In conclusion, specification tests were not correct for the linear model, but non-linear models that contain interactions can be applied to predict spam messages.

To sum up, at the beginning we created two general models (probit and logit) that contained all explanatory variables from the database. Then we got rid of insignificant variables from both models. In the next step models were tested by likelihood ratio test to check their joint significance and both models passed the test. Furthermore the link test was performed on both models and in both cases the result was negative (wrong specification of the model). In order to improve the specification of the models we implemented insightful interaction terms from our perspective. That modification led to improvement in the specification test for logit model. We decided to carry on with further analyses with the logit model for two reasons: problems with convergence of algorithm for probit model and better AIC criterion than for probit model. Then two other tests for specification were performed: Hosmer Lemershow and Osius-Rojek. The second one is more frequently applied and is considered as a good quality and reliable specification test. The test pointed out that the specification is correct. At the end all model variables and interactions are significant at 5% level, moreover it is jointly significant and well specified. Knowing that the model can be used to predict and it was formally checked in the next part we will focus on description of the included variables, predictive power of the model and interpretation of the results.

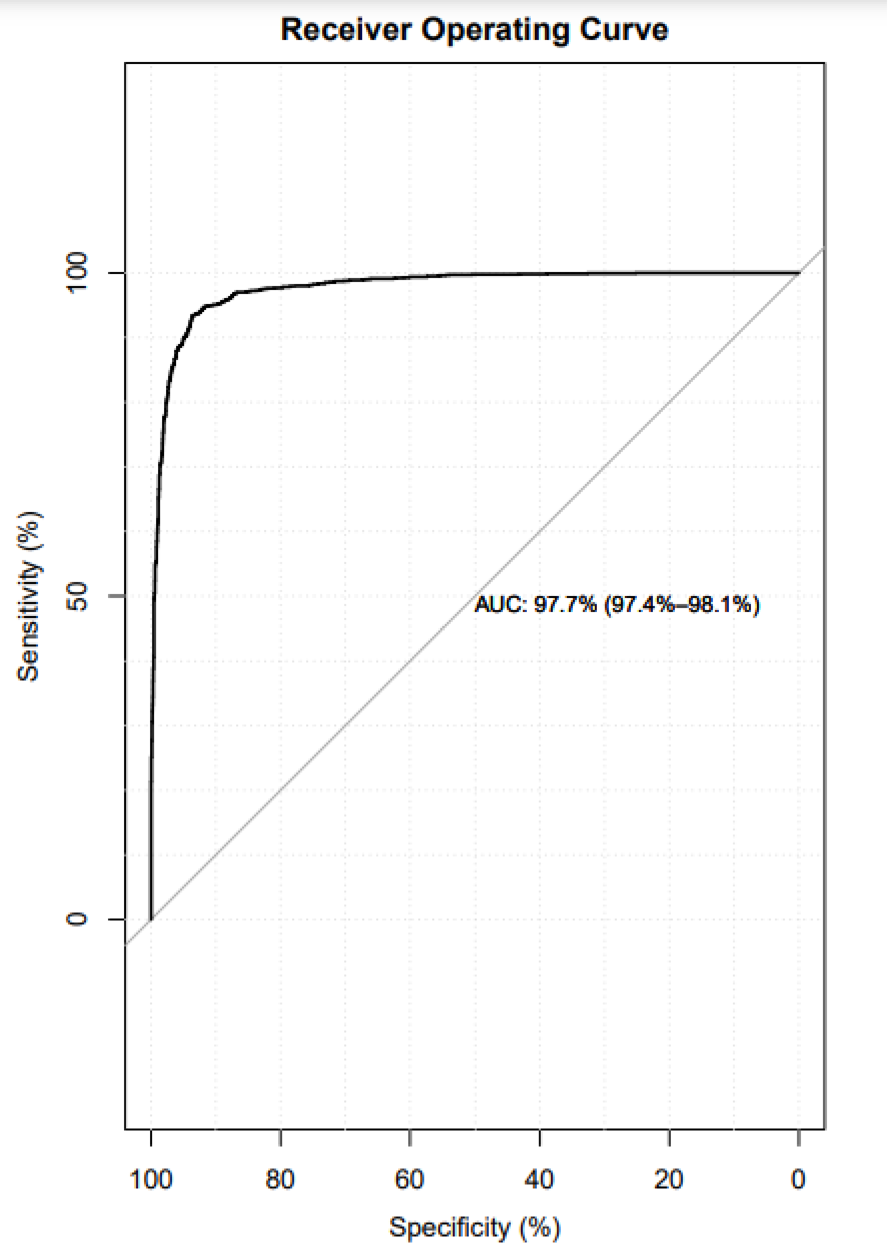
Results

The predictive power of the model can be obtained using a few different indicators: McKelvey - Zavoina, count, adjusted count.



```
   McFadden     Adj.McFadden        Cox.Snell       Nagelkerke McKelvey.Zavoina           Effron            Count
  0.7038839        0.6905941        0.6109085        0.8273094        0.9983811        0.7707971        0.9280591
  Adj.Count              AIC    Corrected.AIC
  0.8174297     1907.0817098     1907.8010081
```

1. McKelvey - Zavoina statistic can be interpreted similarly to R squared statistic for standard regression model. For the estimated model the statistics is equal to 99.84%. It means that if the latent (hidden) variable is observed as a dependent variable the model would explain around 99.8% of variance.
2. Count statistics can be interpreted that 92.8% observations are well predicted.
3. Adjusted count statistics can be interpreted that 81.74% correct predictions of observations were because of variation of explaining variables excluding effect of p star (threshold level) compared to standard count statistics that takes into consideration p star level.



Moreover AUC (Area Under Curve) statistic was calculated. The more convex curve the better fit of the model. For the final model the AUC is convex and has very high values, which also indicates that predictive power of the model is good.

Based on the 3 statistics above the overall prediction quality of the model is good, as it should predict with around 90% accuracy the spam mails. That makes us conclude that the model could be used as a tool to filter out spam mails and it makes sense to use binary dependent variable models to deal with such problems.

To analyse more detailed results of the model, there should be more emphasis on the certain variables and interactions. The final logit model consists of 35 significant variables / interactions of variables. The most important ones that can be divided into some groups are listed below:

1) Variables connected with money or business - frequencies of the words or characters in email such as: order, free, money, business, credit, $ (dollar sign)
2) Special signs - dollar sign, exclamation mark, hashtag
3) Personal phrases - frequencies of the words: you, your, george (we believe it is the name of one of the authors of database - George Forman)
4) Length of sequences of words, that were explained in the data part of the paper
5) Interactions between certain variables, especially with special signs
6) Other words that are hard do divide into some groups

Considering that there is a significant number of variables we will focus on interpretation of only a few that can be easily explained and that are most important in our perspective. For binary dependent variable econometric models we cannot just simply interpret estimates. From the standard logit model we can only interpret signs of coefficients, but to make more advanced analyses there is another method - marginal effects, which was performed via function created by dr Rafał Woźniak (please refer to appendix - file ”marginaleffects”). The results of the marginal effects estimates shall be multiplied by 100 and interpreted as a growth or decrease (based on the sign) of probability that the endogenous variable takes value = 1 (is a spam message). The table below contains results of some of the most interesting marginal effects.

Variable	dF/dx	Std. Err.	z	P>|z|
word_freq_remove	0,142	0,022	6,527	0,000
word_freq_internet	0,036	0,012	3,119	0,002
word_freq_order	0,088	0,024	3,703	0,000
word_freq_free	0,056	0,010	5,899	0,000
word_freq_business	0,075	0,019	4,045	0,000
word_freq_credit	0,107	0,042	2,567	0,010
word_freq_000	0,134	0,028	4,765	0,000
word_freq_money	0,033	0,012	2,839	0,005
word_freq_george	-0,776	0,121	-6,408	0,000
word_freq_85	-0,182	0,070	-2,592	0,010
word_freq_meeting	-0,190	0,057	-3,338	0,001
word_freq_project	-0,120	0,036	-3,322	0,001
word_freq_edu	-0,100	0,018	-5,437	0,000
word_freq_conference	-0,281	0,110	-2,547	0,011
char_freq_exclamation	0,017	0,004	3,778	0,000
char_freq_dollar	0,177	0,051	3,475	0,001
char_freq_hashtag	0,108	0,064	1,701	0,089
word_freq_address:word_freq_our	0,069	0,026	2,664	0,008
word_freq_order:word_freq_free	-0,085	0,044	-1,955	0,051
char_freq_semicolon:char_freq_exclamation	0,233	0,081	2,879	0,004
char_freq_exclamation:char_freq_dollar	0,872	0,229	3,806	0,000
char_freq_exclamation:char_freq_hashtag	1,080	0,592	1,826	0,068

Before the actual analysis of certain factors, it is important to mention that the marginal effects coefficients and interpretation depend on the certain values of variables used for calculations. In this case we decided to use the parameter “atmean = FALSE”, which is not the default option. Usually marginal effects are calculated for the mean values of the variables (parameter “atmean = TRUE”) - the coefficients show impact, however in case of spam mails it would not make much sense, because each mail is rather concentrated on one topic. The advertisement mail from e-commerce companies would contain a different set of words than mails from banks or technology companies. The table above contains coefficients that were calculated as an average of marginal effects coefficients for one mail and not average of variable values.

One of the most interesting and influential results is for the interaction term “char_freq_exclamation:char_freq_dollar”. It can be interpreted that if a mail would have both an additional unit of exclamation sign and dollar sign, then the probability of email being spam is higher by 87 percentage points on average, taking into consideration all mails (across all mails). That result goes inline with our expectations, as we think that spam messages usually contain some special signs such as exclamation marks or dollar signs to make an offer more urgent or create a feeling of a bargain.

If a mail would have only an additional unit of exclamation sign and no dollar sign, then the probability of email being spam is higher by 1 percentage point on average, taking into consideration all mails (across all mails)

If a mail would have only an additional unit of hashtag sign and no dollar sign, then the probability of email being spam is higher by 17.7 percentage points on average, taking into consideration all mails (across all mails)

On the other hand, one of the most influential variables that decreases the probability of a mail being spam is “word_freq_george”. It can be interpreted that if a mail would have an additional unit of “george” word, then the probability of email being spam is lower by around 77 percentage points on average, taking into consideration all mails (across all mails). The result makes sense, as the spam mails can not always contain the name of the person to whom the mail is addressed. It seems that if a name of the person is included in the mail it significantly decreases the probability of mail being spam.

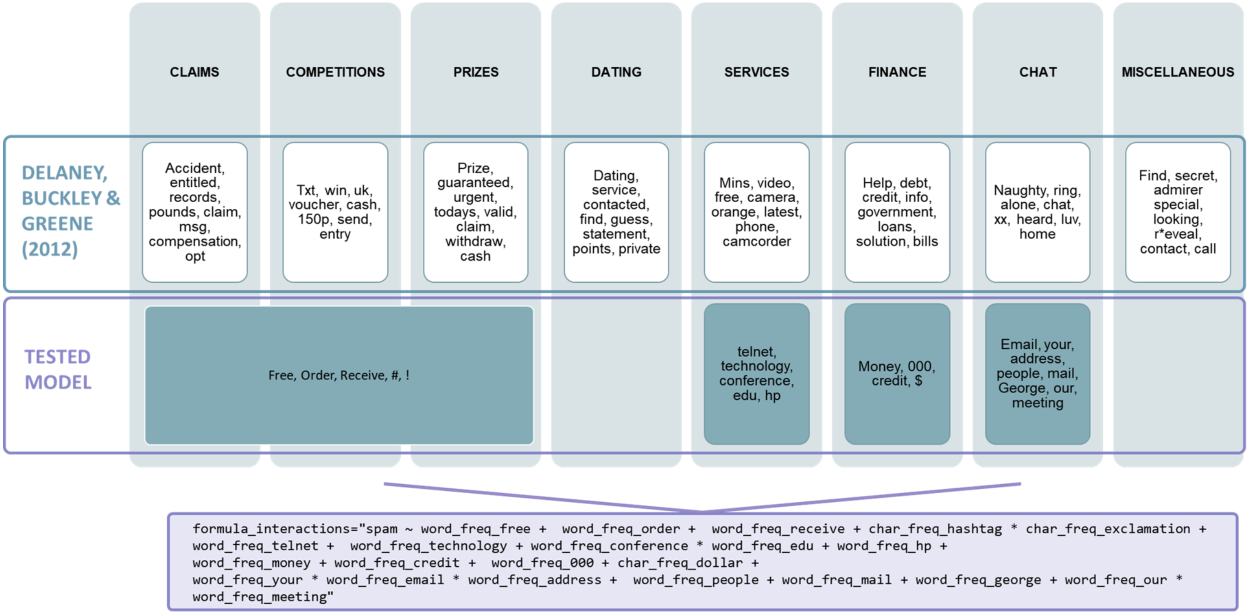
Other single variables can be interpreted in the same way.

One of our hypotheses stated that the words connected with money would increase the probability that the email is spam. Having the output table above it can be observed that for frequencies of the words such as: “credit”, “free”, “money”, “dollar”, “business” and for “$” sign the marginal effect coefficient is greater than 0, which indicates that those words increases the probability mail being a spam. It goes with our expectations and stated hypothesis.

What is more, words connected with education, meetings, projects and conferences decrease the probability that the mail is spam. That also makes a lot of sense that such keywords are rather common in professional or private correspondence and that is why they could lower the probability of mail being spam.

Testing the model with inspirations from literature

In order to improve the model further, we took inspiration from Delany et. al. (Delany et al., 2012,) to categorise the variables and produce meaningful interactions. Several models were tested out based on the framework pictured below. Our selected data did not fit the previous research precisely, therefore the *dating* and *miscellaneous* categories were ignored and *competitions, claims* and *prices* were merged. Several models with different interaction combinations were tested to select the best performing one.



	CLAIMS	COMPETITIONS	PRIZES	DATING	SERVICES	FINANCE	CHAT	MISCELLANEOUS
DELANEY, BUCKLEY & GREENE (2012)	Accident, entitled, records, pounds, claim, msg, compensation, opt	Txt, win, uk, voucher, cash, 150p, send, entry	Prize, guaranteed, urgent, todays, valid, claim, withdraw, cash	Dating, service, contacted, find, guess, statement, points, private	Mins, video, free, camera, orange, latest, phone, camcorder	Help, debt, credit, info, government, loans, solution, bills	Naughty, ring, alone, chat, xx, heard, luv, home	Find, secret, admirer special, looking, r*eveal, contact, call
TESTED MODEL	Free, Order, Receive, #, !				telnet, technology, conference, edu, hp	Money, 000, credit, $	Email, your, address, people, mail, George, our, meeting	

```
formula_interactions="spam ~ word_freq_free +  word_freq_order +  word_freq_receive + char_freq_hashtag * char_freq_exclamation +
word_freq_telnet +  word_freq_technology + word_freq_conference * word_freq_edu + word_freq_hp +
word_freq_money + word_freq_credit +  word_freq_000 + char_freq_dollar +
word_freq_your * word_freq_email * word_freq_address +  word_freq_people + word_freq_mail + word_freq_george + word_freq_our *
word_freq_meeting"
```

The described model underperformed in comparison to ones of our own design, as few of the interactions appeared to be significant. The literature-backed model achieved a McFadden's R2 of 0.616 signalling a rather unsatisfactory fit of the model. It found that only a few of the proposed interactions were significant (# + !; *your* + *address; your* + *email* + *address* and *email* + *address*).

There could be several explanations for this deviation from the literature. Firstly, the data on hand had different variables. It would also seem that our data was biased towards its author, as many of the variables were related to his personal emails (variables such as *edu, lab, labs, conference, 3d* etc. would suggest academic background) and therefore difficult to categorise. Moreover, several of the categories were missing or could be improved to fit our framework, such as variables indicating that the message is not spam.

Findings

Spam keywords

Our research and models would suggest several findings relevant to further works on spam message classification. Firstly, most significant keywords and their combinations were determined. Increased use of characters such as exclamation marks and hashtags, especially in combination, are a good measure of the message being spam. On the contrary, the use of specific keywords such as the name of the recipient, as well as setting the time of a meeting or event and mentioning projects were helpful in classifying the message into non-spam categories. Interestingly, keywords such as money, credit and dollar signs were not as significant in the predictions.

Model performance

Our research has proved that the use of binary dependent variable models is effective in spam email classification. Nevertheless, significant research, adjustments and insight into the data were needed to make the model usable. The models were able to predict the status of spam and non-spam messages for this specific dataset quite well, however the models would need considerable adjustments to keep up with new and varied data, as well as the changing context such as new keywords appearing to keep up with the times. That being said, the model was computationally efficient and provided us with a lot of control over the modelling as well as more insight into the model's components. These findings are in line with the described literature and therefore were expected prior to the research.

Limitations and further research

As mentioned before, the outcome of our research was highly dependent on the data used, therefore more varied inputs would be necessary to gain more accurate results and insights for future research. More recent data and channels of spam messages have potential to change the findings significantly. Moreover, the focus of this paper was only on logit models. To fully understand the benefits and limitations of econometric models on similar problems, a more varied portfolio of econometric models would be beneficial.

Bibliography

Cisco. (n.d.). *What Is Spam Email?* Cisco. Retrieved June 5, 2022, from https://www.cisco.com/c/en/us/products/security/email-security/what-is-spam.html

Delany, S. J., Buckley, M., & Greene, D. (2012, August). SMS spam filtering: Methods and data. *Expert Systems with Applications*, *39*(10), 9899-9908. https://www.sciencedirect.com/science/article/abs/pii/S0957417412002977

Kaddoura, S., Ganesh Chandrasekaran, G., Popescu, D. E., & Duraisamy, J. H. (2022, January). A systematic literature review on spam content detection and classification. *PeerJ Computer Science*. https://peerj.com/articles/cs-830/

Marwaha, M., & Singla, N. (2021, December). Design Engineering Email Spam Filtering Techniques: A Review. *Design Engineering*, *9*, 8327 - 8338. https://www.researchgate.net/publication/357175093_Design_Engineering_Email_Spam_Filtering_Techniques_A_Review

Statista. (2022, April 28). • *Global average daily spam volume 2021*. Statista. Retrieved June 5, 2022, from https://www.statista.com/statistics/1270424/daily-spam-volume-global/

https://www.cisco.com/c/en/us/products/security/email-security/what-is-spam.html
https://www.statista.com/statistics/1270424/daily-spam-volume-global/

Appendix

https://github.com/sunnyline99/Econometrics-Spam-Detection