



Important Predictors for Covid-19 Vaccine Hesitation

Mireille Fangueng^(✉), Mamadou Thiongane, Idrissa Sarr,
and Bitsha-kitime D. Kabkia

Department of Mathematics and Computer Science, University Cheikh Anta Diop,
Dakar, Senegal

{mireille.fangueng,mamadou.thiongane,idrissa.sarr}@ucad.edu.sn,
dieudone.kabkia@eismv.org

Abstract. Hesitation to take the Covid-19 vaccine is one of the main obstacles to the establishment of a general vaccination program that would quickly achieve mass immunity. Identifying the human and societal factors that lead to hesitancy toward the Covid-19 vaccine can be very useful in raising awareness about vaccine acceptance. In this work, we are interested in finding these factors for the African universities population (students and professors). Surveys are conducted in several universities and some information that we believe may influence vaccine hesitancy, vaccine acceptance, and vaccine rejection are collected from individuals in this community. Three classes of people are observed in these data: the vaccinated, the non-vaccinated, and the hesitant. We propose a Bernoulli Mixture Model with conditional class dependency that can estimate the importance candidate predictor variables for a class. We used this model and determined the most important variables to predict Covid-19 vaccine hesitancy in the study population.

Keywords: Bernoulli Mixture Model · Covid-19 · Vaccine Hesitancy

1 Introduction

1.1 Context and Problem

According to the Strategic Advisory Group of Experts on Vaccination (SAGE), vaccine hesitancy is the term used to describe the delay in acceptance or refusal of vaccination despite the availability of vaccination services [10]. Complacency, confidence, and convenience are conducive to this attitude [17]. Complacency denotes a low perception of disease risk, so vaccination is deemed unnecessary. Confidence or lack of confidence refers to the assurance of safety produced by vaccination and generally to the competence of health systems. Convenience involves the availability, affordability and delivery of vaccines in a comfortable setting. Vaccines developed against Covid-19 are particularly in the spotlight for vaccine hesitancy. This disease discovered in Wuhan, China in December 2019 was officially declared by the World Health Organization as a public health

emergency of international concern in January 2020. The high infectivity of this disease has led to various measures to limit its spread: social distancing, wearing a face mask in public places, regular hand washing, etc. At the same time, several research teams around the world have started to develop vaccines against the disease. Although the social measures helped to reduce transmission, they were not without consequences for the social structure of the population, their psychological balance, and the world economy. In addition, there is little evidence to suggest that the spread of Covid-19 will be halted by natural immunity of the population. A general vaccination is therefore an option for overcoming the health crisis. Indeed, the higher the number of people vaccinated, the lower the exposure to the virus. This artificially leads to collective immunity. Hesitation about the Covid-19 vaccine is one of the main obstacles to the establishment of a general vaccination program. This hesitancy is due, in part, to the fact that the development of these vaccines has been particularly rapid compared to the duration of standard protocols. Identifying the behavioral and societal factors leading to hesitancy about the Covid-19 vaccine can be used to understand this attitude and to assess its dangerousness in relation to the overall plan to contain or at least control this pandemic. Our study is conducted on the population of some African universities. Students and professors are the targets of this study.

1.2 Literature Review

Vaccine hesitancy in the case of Covid-19 has been of particular interest to the community in recent months. Several approaches based on statistical and/or machine learning models have already been proposed in the literature to study the elements (predictors) likely to influence the decision in favour or not of a Covid-19 vaccine. A predictor or predictive variable or explanatory variable designates an input used to predict the values of an output variable, or target, or explained variable. In Riad and al [15], decision trees are used to model Covid-19 vaccine hesitancy, with 23 demographic and psychological predictors on a population of dental students from around the world. The results highlight five important predictors of willingness to receive the Covid-19 vaccine: the economic level of the country where the student lives and studies, the individual's trust in the pharmaceutical industry, the individual's misconception of natural immunity, the individual's belief in the risk-benefit ratio of vaccines, and the individual's attitude towards new vaccines. Figueiredo et al. [6] propose a Bayesian ordinal logistic regression model to identify socio-demographic determinants of Covid-19 vaccine acceptance, on a population of 26,759 individuals over 18 years of age across 32 countries worldwide. Ruiz et al. [16] use a community detection and semantic network analysis (SNA) approach for identifying pro- and anti-vaccine influencers through the social network Twitter, and also identify online communities formed around Covid-19 vaccine hesitancy. Lyu et al. [13] use a multinomial logistic regression model and counterfactual analysis to identify sociological factors of importance for vaccine acceptance through tweets as well. In Lombo et al. [12], a Bayesian ordinal logistic regression model is used to establish whether exposure to misinformation has a differential impact on subjects' intention to

accept a vaccine for themselves, depending on their socio-demographic background. Lange et al. [11] combine a multivariate regression model and a random forest model to quantify the factors contributing to vaccine hesitancy at the county level in the continental United States. The joint effects of several variables (race/ethnicity, politics, age, etc.) are considered simultaneously to capture and quantify the factors that affect vaccination rates. Bouguila et al. [2] address the problem of unsupervised selection of binary features using finite mixture models of multivariate Bernoulli distributions. Pires [1] does a review of works that identify and analyse the predictors of Covid-19 vaccine acceptance and/or hesitancy in different studied populations. In these studied groups as in ours, the most predominant predictors of vaccine hesitancy were a lower perceived risk of getting infected, a lower level of institutional trust, not being vaccinated against influenza, lower levels of perceived severity of Covid-19, or stronger beliefs that the vaccination would cause side effects or be unsafe. Syed et al. [14] examine state-level features and policies that are most important in achieving a threshold level vaccination rate to curve the effects of the Covid-19 pandemic.

1.3 Motivation and Objective

Most of the works presented in the previous section, attempt to develop models capable of capturing the importance of each predictor (selected a priori in an intuitive way) in the choice of vaccine hesitancy. The main objective of these works is not to predict whether an individual will be vaccinated or not, but rather to identify the factors (the predictors) that would most or least guide the choices. A wide range of relevant factors to predicting Covid-19 vaccine hesitancy for several population categories have been identified in these work. However, the above methods look at the importance of predictors for classification in general. They do not look at the importance for discrimination of specific classes. Indeed, looking specifically at the importance of a predictor in relation to a class provides more information, and therefore understanding.

In this paper, we propose a method that captures the importance of predictors variables conditional on classes for any classification dataset. It is a mixture model for supervised classification which estimates the importance of predictors as a function (conditional) of classes. We assume a Bernoulli distribution for our predictors that are independent of each other. The activation of a predictor is modeled as a binary latent variable, whose distribution parameter provides information on its importance in the classification. We use a fully frequentist approach, unlike Bouguila et al. [2] where a Bayesian approach to clustering is used. We estimate a conditional dependence model of class importances that allows us to obtain the predictor importances for specific class by using the Expectation-Maximisation (EM) algorithm [4]. This model can capture variables importance for any classification dataset. Thus, the validation of our model is done with a well known dataset from the literature and information given by our model confirms well known knowledge that are not given by existing models [2].

We conduct surveys in several African universities and collect some information (from students and professors) that we believe may influence Covid-19

vaccine hesitancy, acceptance, or rejection. We used our proposed method to determine the most important predictors for Covid-19 vaccine hesitancy for the african universities population. Indeed, academics have a strong voice among other segments of the population and can influence others to accept the Covid-19 vaccine. Knowing the factors that cause professors and students to be reluctant to take the Covid-19 vaccine can help in their awareness of vaccine acceptance.

1.4 Structure of the Paper

The remainder is organized as follows. Section 2 presents our Bernoulli Mixture Model Method and its validation with a well known dataset. Section 3 describes the collected dataset with surveys. Numerical results for variables importance in the Covid-19 dataset are reported in Sect. 4. Finally, concluding remarks are given in Sect. 5.

2 The Bernoulli Mixture Model Method and Validation

2.1 Bernoulli Mixture Model

On an independent and identically distributed sample (i.i.d.) of n individuals from a population structured in K classes, we observe D binary variables supposed to determine the class to which each individual belongs. The dataset is then presented in the form:

$$\mathcal{D}_n = \{(x_i, y_i)\}_{i=1}^n \tag{1}$$

$x_i = (x_{i,d})_{d=1:D}$ is the vector of D binary predictors of individual i ; $(y_i)_{d=1:D}$ is the vector of D binary predictors of individual i ; $(y_i)_{d=1:D}$ the observed class (output variable) for individual i , such that: $y_i^K, \sum_{k=1}^K y_{i,k} = 1$. This is a common representation of classes as K -lists where each position in the list represents a class, the value 1 indicates the activation of a specific class while the values 0 obviously indicate the disactivation of all other classes, since an individual belongs to only one class. We note respectively $(\omega_k)_{1:K}$ and $(\theta_{k.})_{1:K}$ the proportions of the classes and the parameters of the distributions of the variables within the classes.

We begin by making the basic assumptions without which it would be impossible to infer the laws of a sample on the global population [3]. Two individuals of “close” variables belong to the same class. The classification boundary runs through the low density areas. We also make the assumptions that the data come from a Bernoulli mixture distribution. Indeed, according to the maximum entropy principle, the multivariate Bernoulli distribution is the best model for binary vectors [7–9]. Thus $y_i = (y_{i,k})_{k=1}^K \sim \mathcal{M}(1; \omega_{1:K}) \sim \prod_{k=1}^K \omega_k^{y_{i,k}}$; multinomial distribution with one trial and K possible outcomes of probability ω_k each; and

x is Bernoullian conditional on y and s ; i.e. $x_{id}|(y_{ik} = 1, s_{id} = 0) \sim \mathcal{B}(\beta_d)$ et $x_{id}|(y_{ik} = 1, s_{id} = 1) \sim \mathcal{B}(\alpha_{kd})$;

We add here an assumption of heterogeneity of the variables with respect to the classes because of the conditional dependence with respect to the classes; ie, $s_{id}|y_{ik} = 1 \sim s_{ikd} \sim \mathcal{B}(\gamma_{kd}) \sim \gamma_{kd}^{s_{ikd}}(1 - \gamma_{kd})^{1-s_{ikd}}$; $s_{ikd} = 1$ (i.e. $s_{id} = 1|y_{ik} = 1$) if the variable x_d is discriminating for the classification of the individual i with respect to the class k , $s_{ikd} = 0$ otherwise.

1. Likelihood of observations

The likelihood of the observations in this case is given by:

$$P(D_n, \theta) = \prod_{i=1}^n \prod_{k=1}^K \left\{ w_k \prod_{d=1}^D \left[(1 - \gamma_{kd})p(x_{id}; \beta_d) + (\gamma_{kd})p(x_{id}; \alpha_{kd}) \right] \right\}^{y_{ik}} \quad (2)$$

2. The complete log-likelihood

The full likelihood is found in:

$$\begin{aligned} p(D_n^c; \theta) &= \prod_{i=1}^n \prod_{k=1}^K w_k^{y_{ik}} \times \prod_{i=1}^n \prod_{k=1}^K \prod_{d=1}^D \gamma_{kd}^{s_{ikd}y_{ik}} (1 - \gamma_{kd})^{(1-s_{ikd})y_{ik}} \\ &\times \prod_{i=1}^n \prod_{k=1}^K \prod_{d=1}^D p(x_{id}; \beta_d)^{(1-s_{ikd})y_{ik}} \\ &\times \prod_{i=1}^n \prod_{k=1}^K \prod_{d=1}^D p(x_{ik}; \alpha_{kd})^{s_{ikd}y_{ik}} \end{aligned} \quad (3)$$

The full log-likelihood is thus deduced:

$$\begin{aligned} \ln p(D_n^c; \theta) &= \sum_{k=1}^K \left(\sum_{i=1}^n y_{ik} \right) \ln w_k + \sum_{k=1}^K \sum_{d=1}^D \left(\sum_{i=1}^n y_{ik} s_{ikd} \right) \ln \gamma_{kd} \\ &+ \sum_{k=1}^K \sum_{d=1}^D \left(\sum_{i=1}^n y_{ik} (1 - s_{ikd}) \right) \ln (1 - \gamma_{kd}) \\ &+ \sum_{i=1}^n \sum_{d=1}^D \left(\sum_{k=1}^K (1 - s_{ikd}) y_{ik} \right) \ln p(x_{id}; \beta_d) \\ &+ \sum_{i=1}^n \sum_{d=1}^D \sum_{k=1}^K s_{ikd} y_{ik} \ln p(x_{ik}; \alpha_{kd}) \end{aligned} \quad (4)$$

This time, we have $x|y$ as the marginal distribution of $x, s|y$. We are still looking for the maximum likelihood estimator, and the idea remains to use the Log of the likelihood obtained in Eq. (4). But the sum brought by the marginal distribution makes it difficult to calculate by the classical differentiable optimisation methods. A way of getting around this difficulty is proposed in Dempster et al. [4] and known as the Expectation-Maximisation algorithm, or EM algorithm.

A distinction is made between the observations \mathcal{D} and the so-called complete data \mathcal{D}^c made up of observations and the latent variable s .

The classical EM algorithm is as follows:

```

Input: Dataset
Output: Parameter and latent variable estimates
1 Initialization ;  $t \leftarrow 0$ 
2 while no-convergence do
3   Step E:
      $Q(\theta|\theta^{(t)}) = \mathbb{E}_s[\ln p(D^c; \theta) | D = \mathcal{D}; \theta = \theta^{(t)}] = \mathbb{E}_{s|\mathcal{D}; \theta^{(t)}}[\ln p(D^c; \theta)]$ 
4   Step M:  $\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}} Q(\theta|\theta^{(t)})$ 
5 return

```

Thus,

Step E of the EM algorithm

Using the expectation of the complete log-likelihood following the distribution of the latent variable s conditional on the observations and parameters, we have:

$$\gamma_{ikd} = \frac{\left[\gamma_{kd} p(x_{id}, \alpha_{kd}) \right]^{y_{ik}}}{\left[\gamma_{kd} p(x_{id}, \alpha_{kd}) \right]^{y_{ik}} + \left[(1 - \gamma_{kd}) p(x_{id}, \beta_d) \right]^{y_{ik}}} \tag{5}$$

Step M of the EM algorithm

Using the full log-likelihood expectation:

Estimation of $w_k \forall k = 1 \dots K$:

$$w_k = \frac{1}{n} \left(\sum_{i=1}^n y_{ik} \right) \tag{6}$$

Estimation of $\gamma_{kd} \forall k = 1 \dots K$ and $\forall d = 1 \dots D$:

$$\gamma_{kd} = \frac{\sum_{i=1}^n \gamma_{ikd} y_{ik}}{\sum_{i=1}^n y_{ik}} \tag{7}$$

Estimation of α_{kd}
 $\forall k = 1 \dots K$ and $\forall d = 1 \dots D$:

$$\alpha_{kd} = \frac{\sum_{i=1}^n \gamma_{ikd} y_{ik} x_{id}}{\sum_{i=1}^n \gamma_{ikd} y_{ik}} \tag{8}$$

Estimation of β_d
 $\forall d = 1 \dots D$:

$$\beta_d = \frac{\sum_{i=1}^n \left(\sum_{k=1}^K (1 - \gamma_{ikd}) y_{ik} \right) x_{id}}{\sum_{i=1}^n \sum_{k=1}^K (1 - \gamma_{ikd}) y_{ik}} \tag{9}$$

2.2 Validation of the Model

The validation of our model was done with the “Zoo dataset” [5]. The later is constituted of 101 animals of a zoo described by 16 traits or attributes supposed to allow to classify them according to the following 7 classes: mammals, birds, reptiles, fish, amphibians, insects and invertebrates. The common use is to predict the class of animals on the basis of the variables. The observed attributes and their types are: name (unique for each instance), hair (boolean), feathers (boolean), eggs (boolean), milk (boolean), aerial (boolean), aquatic (boolean), predator (boolean), toothed (boolean), spine (boolean), breathes (boolean), venomous (boolean), fins (boolean), legs (set of values: 0,2,4,5,6,8), tail (boolean), domestic (boolean), cat size (boolean), class (integer values in the range [1–7]). Except for the “legs” variable, all predictors are binary. To make the dataset suitable for our model, we have eliminated the “legs” column as in Bougila et al. [2]. The dataset we use is thus made up of 101 individuals described by 15 binary predictors and structured in 7 classes.

The average success rate for the classification is 0.85 ± 0.04 for the classical model [2], 0.91 ± 0.02 for the model with class independent importance [2], and 0.90 ± 0.04 for the model with class dependent importance (our model) on this

dataset. For the variable importance, the model with class independent importance show that “tooth” possession or “egg” laying are discriminating criteria for the different classes of animals, without further clarification. However, our model with class-dependent importance goes further and indicates that “milk” production is more discriminating for the mammalian class; “tooth” possession is strongly discriminating for insects, invertebrates and birds; and “egg” laying is strongly discriminating for mammals. This is perfectly consistent with our knowledge of these categories of animals. Let us also note, for example, that the “venomous” character is weakly discriminating for any class; this is also intuitive, since venomous animals are found in each of these categories. Figure 1 and 2 show the result for the importance of the predictors, and the probabilities of membership for each class according to the discriminant variables respectively.

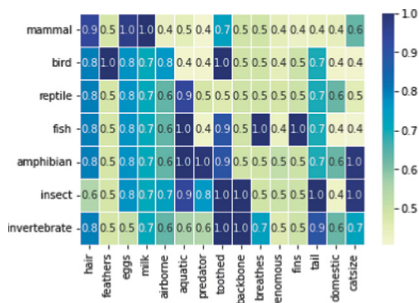


Fig. 1. Importance of the predictors

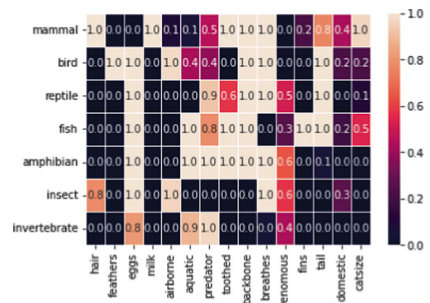


Fig. 2. Class membership probabilities for class-predicting variables

3 The Collected Covid-19 Dataset

We conducted surveys from six African universities to get data on factors that we believe may influence the choice to accept, reject, or be hesitant to take the vaccine against Covid-19. Our candidate variables are: profession of the individu, gender, age, level of education, place of residence, risk or danger associated with Covid-19, number of prevention methods used (the prevention methods are: wearing a mask in public places, use of disinfectant gel to clean hands, respecting barrier measures, regular hand washing), intentions of the vaccine manufacturers (good or bad), and effectiveness of the vaccines. The universities concerned by the surveys are Université Cheikh Anta Diop de Dakar (Sénégal), Université de Thiés (Sénégal), Université de Yaoundé 1 (Cameroon), Université de Yaoundé 2 (Cameroon), Université Alassane Ouattara (Ivory Coast), and Université de Lomé (Togo). As mentioned earlier in the introduction, the target population was professors and students. These surveys are done in two phases, the first one from June 2 to July 31, 2021, and the second one was done between

May 28 and June 15, 2022. Forums and social networks were used to distribute the questionnaires and 1882 people have completed the questionnaires. Table 1 shows the survey questionnaires and the possible answers for each question.

Table 1. Questionnaire and possible answers of the surveys

1. What is your profession ? (a) Professor (b) Student	2. What is your gender ? (a) Female (b) Male
3. Are you under 35 years old ? (a) Yes (b) No	4. Is your level of education below Master 2? (a) Yes (b) No
5. Do you live outside the university campus ? (a) Yes (b) No	6. Do you think there are any risks/dangers associated with Covid-19 ? (a) Yes (b) No
7. Do you practice more than two preventive methods ? (a) Yes (b) No	8. Do you think there are bad intentions in the design of vaccines ? (a) Yes (b) No
9. Is the Covid-19 vaccine a prevention measure ? (a) Yes (b) No	10. Have you take a COVID 19 vaccine? (a) Yes (b) No (c) I am still hesitating

Here are some statistics summary on collected data. We notice that 44,98% of the people who responded to the surveys are not vaccinated, 25,08% of people are vaccinated, and 29,92% of people are hesitant. We observe that 13,52% of the respondents are professors and 86,47% are students, 56,04% are men and 43,95% are women. For the age of the persons, 26,51% of the respondents are over 35 years old and 73,48% are under 35 years old. For intentions on the manufacture and preventability of the Covid-19 vaccine, so 56,93% of people thinks that there is no bad intentions with the vaccines but 43,06% thinks there is a bad intentions with vaccines, and 51,95% of people thinks that vaccines are not preventive and 48,04% thinks that there are preventive.

4 Experiment and Result

Based on the collected and preprocessed dataset, we use our Bernoulli mixture model for classification to determine the importance of each variable in predicting

vaccine hesitancy. We also determine the importance variable in predicting vaccine acceptance or rejection. The evaluation is done in a 5-fold cross-validation context. The dataset is partitioned into 5 subsets, each of which will be used for testing, the remaining 4/5 will be used for training. An average of the inter-fold results allows an overall assessment of the model performance.

Figure 3 and Fig. 4 report the results of the importance of the variables, and probabilities of membership for each class according to the discriminant variables. For the hesitant class, we observe that the scores importance of the variables varies between 0.5 and 0.68. The variable “prevention” with a score of 0.68 stands out the most in terms of importance compared to the other variables. We notice that hesitant people tend to increase preventive actions other than vaccination (see Fig. 4). The variable “age” with score of 0.52 arrives in second position. The “sex” and “effectiveness” arrive in third position with a score of 0.51. The “profession”, “study level”, “risk” and “intention” variables are the least discriminating. If we combine these importances with probabilities of membership for each class according to the discriminant variables values, we notice that the hesitant individuals are less likely to think that there are bad intentions in the design of the vaccines, believe more that there are risks and dangers associated with Covid-19, and believe a little in the effectiveness of the vaccine. These elements are likely to justify their “hesitant” position.

We can also see that the “effectiveness” of Covid-19 vaccines character is discriminating for the class of the vaccinated. So vaccinate people believe in the effectiveness of the vaccine.

We notice that the character “prevention” is also discriminating for the non-vaccinated, and vaccinated class.

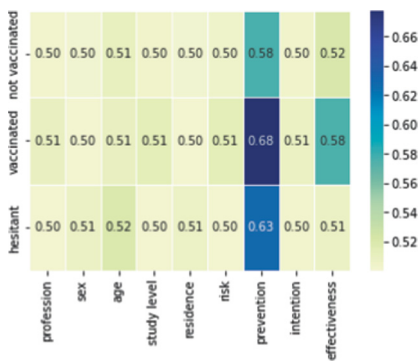


Fig. 3. Importance of the predictors

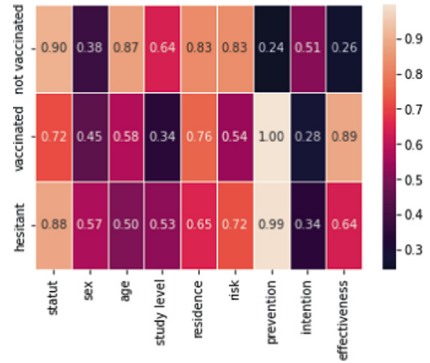


Fig. 4. Class membership probabilities for class-predicting variables

5 Conclusion

We introduced the assumption of class dependencies in a Bernoulli mixture model to estimate the importance of the predictors in a classification model. We proposed to define latent variables to model these class dependencies and we used the Expectation-Maximisation algorithm to compute variables importance. This class dependence hypothesis proved to be effective in providing finer-grained explanations. The method was validated with an animal dataset. The latter was able to provide well known knowledge with these data. This gives confidence in the model. This model has allowed us to identify some factors that drive the African university population to be reluctant to be vaccinated against Covid-19. For example, we learn that the age of the person, and the number of prevention methods used are determining factors in predicting vaccine hesitancy.

References

1. Carla, P.: Global predictors of Covid-19 vaccine hesitancy: a systematic review. *Vaccines* (Basel). MDPI **10**(8), 1349 (2022). <https://www.mdpi.com/2076-393X/10/8/1349>
2. Bouguila, N.: On multivariate binary data clustering and feature weighting. *Comput. Stat. Data Anal.* **54**(1), 120–134 (2010)
3. Chapelle, O., Scholkopf, B., Zien, A.: Semi-supervised learning. *IEEE Trans. Neural Netw.* **20**(3), 542–542 (2009). (Chapelle, O., et al. (eds.) 2006)[book reviews]
4. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* **39**(1), 1–22 (1977)
5. Dua, D., Graff, C.: UCI machine learning repository (2017). <http://archive.ics.uci.edu/ml>
6. de Figueiredo, A., Larson, H.J.: Exploratory study of the global intent to accept Covid-19 vaccinations. *Commun. Med.* **1**(1), 1–10 (2021)
7. Govaert, G., Nadif, M.: Block Bernoulli parsimonious clustering models. In: *Selected Contributions in Data Analysis and Classification*, pp. 203–212. Springer, Cham (2007)
8. Gyllenberg, M., Koski, T.: Numerical taxonomy and the principle of maximum entropy. *J. Classif.* **13**(2), 213–229 (1996)
9. Juan, A., Vidal, E.: On the use of Bernoulli mixture models for text classification. *Pattern Recogn.* **35**(12), 2705–2710 (2002)
10. Lane, S., MacDonald, N.E., Marti, M., Dumolard, L.: Vaccine hesitancy around the globe: Analysis of three years of WHO/UNICEF joint reporting form data-2015–2017. *Vaccine* **36**(26), 3861–3867 (2018)
11. Lange, C., Lange, J.: Applying machine learning and AI explanations to analyze vaccine hesitancy. arXiv preprint [arXiv:2201.05070](https://arxiv.org/abs/2201.05070) (2022)
12. Loomba, S., de Figueiredo, A., Piatek, S.J., de Graaf, K., Larson, H.J.: Measuring the impact of Covid-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat. Hum. Behav.* **5**(3), 337–348 (2021)
13. Lyu, H., et al.: Social media study of public opinions on potential Covid-19 vaccines: informing dissent, disparities, and dissemination. *Intell. Med.* **2**(1), 1–12 (2022). <https://doi.org/10.1016/j.imed.2021.08.001>. <https://www.sciencedirect.com/science/article/pii/S266710262100036X>

14. Osman, S.M.I., Sabit, A.: Predictors of Covid-19 vaccination rate in USA: a machine learning approach. *Mach. Learn. Appl.* **10**, 100408 (2022). <https://doi.org/10.1016/j.mlwa.2022.100408>. <https://www.sciencedirect.com/science/article/pii/S2666827022000834>
15. Riad, A., et al.: IADS-SCORE: Universal predictors of dental students' attitudes towards Covid-19 vaccination: machine learning-based approach. *Vaccines* **9**(10), 1158 (2021)
16. Ruiz, J., Featherstone, J.D., Barnett, G.A.: Identifying vaccine hesitant communities on twitter and their geolocations: a network approach. In: *Proceedings of the 54th Hawaii International Conference on System Sciences*, p. 3964 (2021)
17. Sallam, M.: Covid-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. *Vaccines* **9**(2), 160 (2021)