Noname manuscript No.
(will be inserted by the editor)

WOM-C : Novel Method for Determining Class Membership for Data Points of One Class Data, and using it to Detect Severity of being COVID-19 Positive

Aadhar Gupta · Rashmi Maheshwari · Samkit Jain · Kusum Kumari Bharti

Received: date / Accepted: date

Abstract In the current scenario of explosive spread of COVID-19 and limited testing resources it becomes very crucial to properly manage the use of available COVID-19 testing kits . Capability to identify the people having higher chances of being COVID-19 infected, will enable prioritization of people for testing and containment. This will allow most efficient utilization of testing resources and will also inform the administration about severity of the current situation. The present situation has the demand for the analysis of individuals based upon the clinical features observed. However, classifying individuals simply into COVID-19 positive or negative class fails to give any priority information within the positive class and negative class. As a person without any sign or symptom can be COVID-19 positive as well, hence classification of a person as COVID-19 negative can be misleading. A better solution would be to determine a person's likelihood (severity) of being COVID-19 positive based on exhibited symptoms. This problem is one example of a class of problems where the available data corresponds to a single class and the belongingness of data points to that class is to be determined. In this paper we present **WOM-C**, a novel **W**eightage based **O**ne class **M**embership measurement technique for **C**ategorical features. The weightage of each categorical feature is computed, which is then used to evaluate how strongly a data point belongs to the class (COVID-19 positive class in our case).

A. Gupta
PDPM IIITDM Jabalpur, M.P., India
E-mail: aadhar.innovate@gmail.com ; 1911001@iiitdmj.ac.in
ORCID iD: 0000-0002-6312-8947

R. Maheshwari
PDPM IIITDM Jabalpur, M.P., India
E-mail: 1921601@iiitdmj.ac.in

S. Jain
PDPM IIITDM Jabalpur, M.P., India
E-mail: 1911603@iiitdmj.ac.in

K. K. Bharti
PDPM IIITDM Jabalpur, M.P., India
E-mail: kusum@iiitdmj.ac.in
ORCID iD: 0000-0003-2039-7589

Keywords Class membership · class belongingness · COVID-19 testing · COVID-19 diagnosis · coronavirus · one class data

1 Declarations

1.1 Funding

Not applicable

1.2 Conflicts of interest/Competing interests

Not applicable

1.3 Availability of data and material

The dataset analysed during the current study has been extracted from the file **latestdata.csv** of the dataset- **COVID-19 Open Access Data** available at www.kaggle.com, [https://www.kaggle.com/mkohlegger/COVID19-open-access-data].

The data generated during the current study i.e the preprocessed form of the extracted data and the inter-feature correlation value map, along with compressed latestdata.csv file are available in our Github repository, https://github.com/AadharGupta/WOM-C-feature-weightage-datapoint-membership-one-class-data.git

1.4 Code availability

The source code is available as an .ipynb file in our Github repository , https://github.com/AadharGupta/WOM-C-feature-weightage-datapoint-membership-one-class-data.git

2 Introduction

The coronavirus disease 2019 (COVID-19) has been declared as a pandemic by the World Health Organisation (WHO). Corona viruses are a family of viruses that cause illness such as respiratory diseases or gastrointestinal diseases. The infected people may show up with variety of symptoms or no symptoms at all. The symptoms may be mild, moderate or extreme depending upon the stages of infection and the immune system of the body. The symptoms commonly visible as stated by various doctors are – respiratory problems , flu, cough and cold, difficulty in breathing, fatigue and sore throat. From the COVID-19 patients' dataset (see Sect. 1.3) we have identified 45 distinct symptoms that appeared in the patients.

The rate at which coronavirus is spreading, the detection of infected people needs to be done as soon as possible. Many common symptoms of COVID-19 like fever and cough are very common ailments in general so it won't be accurate to classify someone as COVID-19 positive due to presence of such symptoms. Asymptomatic individuals can also be COVID-19 positive so no individual can be classified as COVID-19 negative. Hence, instead of classifying patients into COVID-19 positive or negative, it is more practical and useful to identify those people who are most likely to be coronavirus (COVID-19) positive. The testing of COVID-19 is a challenging issue everywhere due to the unavailability of diagnosis systems. The current situation has the demand for the detection of COVID-19 based upon the clinical features as observed in people to take necessary action as soon as possible and prevent the spread of disease.

Classification is a machine learning technique to identify the membership of data points in the group [21]. There are different types of classifications based on the number of classes, namely- one class classification(OCC), binary class classification and multi-class classification. The OCC models have become the active areas of research in the field of Machine Learning [15] [17]. The one class classification deals with single target class. It comes in to play when there exist two classes in reality but data is available for only one class or it is available for the other class in comparatively much scarce amount. In such cases, we can't train a model to learn the difference between the data distribution of the two classes. Instead, a model is trained to learn the distribution of the single available class (or the class with more available data). Then this model detects whether a given test data point is from the learnt distribution or is an outlier. If an outlier then it is classified into the other class.

The one class classification problem can be identified as a special case of binary class classification. It helps to solve the problems in which either the data of second class is too costlier or its data is unavailable or the data is present in very small amount. This can help to solve complex problems in medical diagnosis, in which the data from only one class is available(healthy/unhealthy), for example - breast cancer detection through mammograms [5][24] , one-class classification of cognitive brain functions [4], lung tissue categorization [7]. Our problem statement is similar to one class problem in the sense that we also have data available for just one class, COVID-19 positive confirmed class ([26],section 4.1) and model will learn, using data from one class only. However, our problem is different from one class classification in the sense that rather than classifying test data point into positive or negative class, we categorise it into degree of belongingness to the positive class. Our problem aims at determining how strongly a given feature combination belongs to the positive class, on the basis of all the feature combinations corresponding to the positive class in the available data.

Corona virus infected patients have presently been detected using machine learning and deep learning techniques on the chest X-ray images [1] and CT images [20] [8] [2] [22]. The proposed method tries to check the severity of being COVID-19 positive depending upon the patient's symptoms by assigning weightage to the symptoms.

The rest of the section is organised as follows: The section 3, Related Work briefly covers the related work, section 4, Proposed Methodology discusses the novel approach presented, along with the relevant terms and concepts and lastly the section 6, Conclusions And Future Work concludes this work and gives some insights for further development of our approach and possible improvements in this area.

3 Related Work

The One Class Classification (OCC) [13] issues are not quite the same as conventional binary class or multi-class classification problems. In one class classification problems either the negative class is missing or not appropriately inspected. The issue emerges in classifying positive case or target class without negative class or inappropriately portrayed negative class. Some other terms such as Outlier Detection [18], Novelty Detection [3] or Concept Learning [11] are used by different researchers to present similar concepts. These terms originate as a result of different applications to which OCC has been applied. In OCC the learning algorithm is being used to differentiate between data that appears normal and abnormal with respect to the distribution of the training data that's why it has also called outlier (or novelty) detection [9].

OCC Verses Multi-class Problems [13] - The problems that we are facing in the conventional classification, such as the estimation of error rates, measuring the complexity of a solution, the curse of dimensionality, the generalization of the method, and so on, also appear in OCC, and sometimes become even more prominent.

Decision Making in OCC - OCC defines a classification boundary around the positive (or target) class, such that it accepts as many objects as possible from the positive class, while it minimizes the chance of accepting non-positive (or outlier) objects. Since just one side of the boundary can be determined, in OCC, it is difficult to choose, based on only one class how firmly the boundary should fit in each of the directions around the data. It is also harder to decide which attributes should be used to find the best separation of the positive and non-positive class objects. In particular, when the boundary of the data is long and non-convex, the required number of training objects might be very high [25].

Applications of OCC -

(i)Text Classification - The ability to build classifiers without negative training data is useful in a scenario when one needs to extract positive documents from many text collections or source. Liu et al. [12] propose a method (called Spy EM) to solve this problem in the text domain.It is based on Na ıve Bayesian classification(NB)and the Expectation Maximization (EM) algorithm [6].

(ii)Bio-Medical Classification – OCC has been applied for the classification of tumor for the cancer detection by various authors using Support Vector Machine, Gaussian model, a mixture of Gaussians models, or the Parzen density estimators [3][14][10].

Classification of COVID-19 – The classification of COVID-19 has been performed by various authors using various domains like chest CT images [20][8][2], chest X-Ray images [1] performed using deep neural networks and machine learning methods. Sun et al. [22] have proposed the feature selection guided deep forest construction using the CT images. Randhawa et al. [16] have proposed a novel way of machine learning using intrinsic genomic signatures can provide rapid alignment-free taxonomic classification of novel pathogens. The method delivers accurate classifications of the COVID-19 virus without a priori biological knowledge, by a simultaneous processing of the geometric space of all relevant viral genomes.
Sethy et al.[19] proposed the deep feature plus support vector machine (SVM) based methodology is suggested for detection of coronavirus infected patient using X-ray images. For classification, SVM is used instead of deep learning based classifier, as the later one need a large dataset for training and validation. The SVM classifies Xray images of COVID-19 patient, pneumonia patient and healthy people with the use of deep features extracted from fully connected layer of the pre-trained network.

4 Proposed Methodology

As clinical data for only COVID-19 positive class is available therefore it is not possible to determine whether a given data point belongs to the negative class. All that can be done on the basis of available data(one class data) is to predict how strongly a given sample belongs to the known COVID-19 positive class, on the basis of trends observed in the available data for this class.

So we introduce a new method to determine how decisive each feature is for confirming this class for the data point. We use weight assignment strategy. In the dataset used, the class is "a person being COVID-19 positive" and "symptoms" are the categorical features. Each categorical feature can have either of the two values- present or absent (whether that symptom is present in a patient or not). Here it is necessary to introduce two concepts- first, a new term "Decisiveness" w.r.t. a feature and second, the Theil's U correlation.

Decisiveness – It is the value denoting the contribution of a feature towards determining the target class (let be 'C') for the data point (individual). In other words how capable the feature is in solely determining the class C for the data point. If a feature "f " always occurs in the presence of many other features then it alone probably isn't much responsible (decisive) for the class to be C i.e. it needs support from other features to decide that class for the data point. On the other hand, if a feature often occurs with all other features absent or with only a few other features present, then it plays a decisive role in determining the class C for the data point and can be considered a good indicator for that class.

Correlation between any feature "k" and all other features plays an important role towards Decisiveness of feature "k". Suppose there is a feature "S" which always has the same value as feature "k" , then we need not consider "S" among other features that occurred in a data point when "k" occurred. As the presence of "S" implies the presence of "k", the presence of "S" is not adding any new informa-

tion. For a feature "k" present in the data point, we don't represent the 'additional features present' in that data point by the count of total features present other than "k". Instead, we use correlation to find the total new information provided by each of these features when the value of "k" is known. This new information will be considered as 'additional features present' value. In this paper, we use Theil's U correlation [23] which is both, suited for categorical variables and free from curse of symmetry. Theil's U(f,k) yields a value of 1 if the value of "f " is fully determined by knowledge of k, and 0 if knowledge of value of "k" doesn't provide any help at all in determining the value of feature "f " .

We begin by first assigning weightage to each feature using the decisiveness property and then we compute the total weight of a data point using each individual feature's weight. On the basis of the total weight evaluated for a data point, we determine how strongly it belongs to the class or in our case how severely the patient is COVID 19 positive.

4.1 About Dataset

The dataset used [26] has been taken from **www.kaggle.com** (see sect. 1.3). We have extracted the required data from the file **latestdata.csv** which is a part of the dataset- "COVID-19 Open Access Data". The file latestdata.csv contains information of total 4,37,635 confirmed COVID-19 patients. Among these, 431 patients have a non- null entry against symptoms. The symptom information of these 431 patients has been used for this research.

Data Preprocessing: Using the dataset mentioned above we extract the symptoms of COVID-19 patients. We identified a total of 45 features (symptoms) that appeared in the 431 patients. For each individual we create a linear vector of size 1*F where F represents set of all the features identified(symptoms). Each feature f ϵ F are arranged in columns. The one-hot vector is generated for the features which exist in a patient. The value 0 is assigned for feature that is missing and 1 is assigned to feature that is present.

4.2 Working

1. Computing feature decisiveness :
   The decisiveness of every feature is computed separately. To compute the decisiveness of a feature(symptom) "k", only those rows(individual persons) are considered which have the feature "k" as present. For Decisiveness of "k" in a row, the 'new information' provided by a feature "f ", given the value of "k" is known(f also being present in the row), is evaluated by subtracting Theil's U(f,k) from 1. We sum the 'new information' obtained from every feature present in the row(other than k) and then add 1 to it. This final sum is placed as denomination with numerator as 1. The aim of adding 1 in the denominator term is to prevent the overall term from becoming indeterminate in case sum of the correlation values is zero (all features except "k" were absent in the row or no new information was obtained from other features). Then we raise this

term to a power "p". This is to increases the difference between Decisiveness when the feature appears alone, and when it appears along with many other features. It increase the impact of the decisiveness when feature appears alone or with very few features. In our problem of COVID-19 we have demonstrated the impact of using power p=2 and p=3 on the final feature weights (Table 1). We found p=2 to be more suitable for this problem (more description under Sect. 5), hence used it to perform our computations.

$val(k, row_i) = \frac{1}{1+\sum(1-corr(f_j,k))}$

$\forall\ f_j \epsilon F$ such that $f_j = 1$ in row_i and $f_j \neq$k
F stands for set of all features
"corr" function used above represents Theil's U correlation function[23]

D(k,row_i) = $val(k, row_i)^p$

D(k,row_i) represents Decisiveness of k for i^{th} row
We use power p= 2 for our problem.

2. Computing the weight of feature "k":
We compute the average of decisiveness values (obtained after raising to power "p") for all rows "k" is present in. This averaged Decisiveness becomes the final weight assigned to the feature "k".

Sum(D(k)) = $\sum_i D(k, row_i)$

Where i iterates over all rows where k is present

Avg(D(k)) = $\frac{Sum(D(k))}{N}$

N in the denominator is the total number of rows where k is present.
Final weight of feature k : W(k)= Avg(D(k))

3. The above steps are repeated for each feature to obtain the corresponding weight.

4. Total weight of each data point (each individual's set of symptoms):
Finally, we compute the weight corresponding to each data point by adding the weights corresponding to each feature (symptom) present. Assuming the available data sample to represent the actual distribution of COVID-19 patients, we measure the highest total weight (1.1845) and lowest total weight (zero ; asymptomatic patient) among all data points. Then we divide this interval (0 - 1.1845) into 4 smaller intervals (Table 2). The last (fourth) interval is an open interval because even if the total weight of an unseen data point exceeds the highest total weight from seen data, we still place it in the last interval. Depending on which interval the total weight of a data point falls in, the data point is assigned a category representing the degree of belongingness to the class.

Unseen Data: For any unseen data instance T_i (the data of new patient), compute the following:
Data Point weight W(T_i)= $\sum$ W(k_j)
$\forall$ k_j ϵ F such that $k_j = 1$ in T_i
F stands for the set of all features

The interval in which the weight W(T_i) falls, indicates its degree of belongingness to the class. (Table 2)

The way of assigning categories is optional. We can have as many intervals we feel suitable for the problem, we can also express the computed total weight of a data point as a percentage of the highest total weight observed but all the data point total weights exceeding the highest total weight will have to be assigned 100 percent and there can be big margin by which the unseen data point's total weight exceeds our seen highest total weight. So this approach is not recommended.

The Theil's U correlation [23] used previously as "corr(f,k)" may be defined as:-
Entropy for single distribution, H(X) = $-\sum_x P_X(x) log P_X(x)$

While conditional entropy, $H(X|Y) = -\sum_{X,Y} P_{X,Y}(x,y) log P_{X|Y}(x|y)$

Uncertainity Coefficient, $U(X|Y) = \frac{H(X)-H(X|Y)}{H(X)}$
$= \frac{I(X;Y)}{H(X)}$ where I(X;Y) is the mutual information

Instead of using correlation in the beginning to eliminate the redundant features, using correlation later on to reduce the impact of redundancy- Reason: If suppose a feature A has same value as feature B, 9 out of 10 times and we eliminate one of these, say B, due to high redundancy. Now suppose a patient has that one exceptional case where A is not equal to B, and the feature we have considered, A, is absent from patient while the feature we don't consider, B, is present. Now suppose no feature other than B is present in the patient to give any information and B was the only chance but we missed it. The patient may appear healthy - COVID-19 negative but was actually COVID-19 positive. In medical cases FALSE NEGATIVE is very DANGEROUS. Therefore we don't eliminate any feature in the beginning so that we do not miss any information. We use correlation to reduce the impact of redundant features during the procedure.

5 Results and Discussion

We observe in Table 1 that the Decisiveness of some features comes out to be 1. This happens if all the other features were absent every time that feature appeared. So even if that feature's presence is very rare, whenever it occurs, the feature was itself capable of deciding the class of the data point without any contribution from other features. Sometimes this may be true and sometimes misleading. If it is possible for the class to occur without presence of any feature (asymptomatic

COVID-19 positive patients) then it can happen that a feature, not at all related to the class was present by chance when no other feature was. In such a case the feature appears to be single-handedly deciding the class which is not true. Due to this reason, the actual Decisiveness for a few symptoms may be somewhat lesser than computed. This depends on the dataset fed to the algorithm and the problem being dealt. If the class has capability of becoming independent of all features and irrelevant features get captured into the dataset then some misleading observations might occur.

In Table 1 we also observe that computing Decisiveness with p=3 makes some features very insignificant in comparison to other features. The difference between Decisiveness of higher decisive features and lower decisive features became so large that even multiple of the lesser decisive features together won't be able to match the impact of a single high decisive features. This may lead to biasness. Hence we find power p=2 more suitable for our problem. The choice of value for parameter "power p" in denominator of decisiveness term depends on the user, there is no rule for as to which value will be suitable , but it is subject to the nature of problem and data at hand , therefore it requires some experimentation to find which value would be the most appropriate.

It is worth mentioning that this approach is slightly different from our original approach, which we had to abandon for two reasons. In the original approach we considered two properties of a feature -decisiveness and frequency. Frequency being the number of times the feature occurred in the dataset (i.e. the number of data points the feature was present in). For each feature we computed the decisiveness values for all rows it was present in, and then added all row decisiveness values instead of taking average. And then some power "q" was applied (like 1/3 i.e. cube root) on this sum and this would yield the final feature weight. But the shortcoming of this approach was that rare features even if extreme indicators of the class, were prone to getting overshadowd by the commonly occurring less important ones. For example cough is a very common symptom in COVID 19 positive patients ,but there can be twice more COVID-19 negative patients around, having cough due to allergies, seasonal change etc. because it is a very common ailment. The approach did not account for this fact. There may be a feature very commonly occurring with a class but that doesn't mean that it is a trustable indicator of the class. Second reason being that even if multiple great indicators occurred in a data point simultaneously, still the overall indication of being COVID positive may be less than fever alone, because of the frequency of occurrence of fever.To reduce the great impact of frequency we tried reducing power "q" applied after summing, to lower values like 1/4 , 1/10 and increased "p" for decisiveness but didn't get satisfactory results. Both these reasons, especially the first one eventually lead us to average each row's decisiveness to get the final feature weight , thus eliminate the frequency factor completely.

Discussion: Our algorithm can incorporate modifications. If new symptoms need to be considered then the developer will need to add a new column corresponding to each new symptom, and fill the the appropriate values corresponding to patients. To add the data of more patients we just need to add new row to the patient symptom dataset for each new patient and define values corresponding to

the predefined columns (symptoms). After this patient symptom dataset is prepared. We just have to run the algorithm to compute the weight of each symptom, which can then be used to evaluate the weight of data points. It Should be kept in mind that addition of any new row (data point) or column (feature) will change the weights of all the symptoms hence all the symptoms' weights need to be calculated again.

The potential of this method is not limited to just the detection of chances of having a disease. We came across a problem of determining how greatly a data point belongs to a class, given the data belonging to only that class i.e. to determine if a data point is an outlier or an inlier and by what degree . This methodology will be applicable for any similar problem with "one class data" .

We have used this methodology using binary values for features i.e. a feature is present(1) or absent (0). If suppose this method is to be used for a data where the features have multiple categorical values or values other than absent/present, then we will have to split that feature into multiple columns, one for each possible categorical value that feature can have, and then place present or absent (the column corresponding to the actual categorical value will be assigned present (1) and remaining columns will be assigned absent (0)). Then the data will be ready for this algorithm.

6 Conclusions and Future Work

In this paper we have presented a methodology to find the measure of membership of a given data point in a class. We divide the total range of weight a data point can have, into various intervals, each representing a different degree of belongingness to the class. The category of belongingness to the class is allotted to a data point on the basis of the interval it falls in. In our problem of COVID-19, a new patient is assigned the chances of being COVID-19 positive i.e. low, moderate, high or very high, on the basis of its symptoms. This is done on the basis of the metric 'total weight' corresponding to that individual's symptom list, which we get by summing the weights of all the symptoms occurring in the list.

Future Work: In this paper we have dealt with only categorical features. We have not considered continuous value features. Although some modifications may allow us to use this algorithm for continuous value features as well. While computing the weightage of purely categorical features, we just needed inter-feature correlation to obtain the Decisiveness value. But for continuous value features, we may also have to consider the distribution of values over the entire continuous range of a feature in the training data to determine what value contributes towards being an outlier or an inlier (more the inlier-ness, higher the belongingness to class). How to compute the Decisiveness for a continuous value feature still needs to be studied.

Here we are pioneering a new method to deal with one class data problems with a different approach and hope there will be much progress in this field in the coming time.

Table 1 Weightage of Symptoms

Features(symptoms)	Feature Weights (p= 2)	Feature Weights (p= 3)
Pharyngeal dryness	1	1
somnolence	1	1
pharyngitis	0.7001	0.6501
abdominal pain	0.4551	0.3883
anorexia	0.4548	0.3881
flu	0.4192	0.3672
dizziness	0.4028	0.3218
respiratory distress	0.3743	0.295
heart attack	0.3721	0.3024
fever	0.3514	0.2747
fatigue	0.3104	0.2398
lesions	0.2611	0.1334
organ failure	0.2542	0.1282
diarrheoa	0.2509	0.1257
digestive discomfort	0.2505	0.1254
eye irritation	0.2505	0.1254
pneumonia	0.2428	0.1425
chest disomfort	0.2394	0.1442
cough	0.216	0.125
shortness of breath/dyspnea	0.2155	0.1428
rhinorrhea	0.2151	0.1316
dry mouth	0.2047	0.0963
chills	0.2013	0.1128
weakness	0.1873	0.1141
acute respiratory viral infection	0.1858	0.0847
muscle pain	0.182	0.1082
heart failure	0.1788	0.0814
headache	0.1778	0.1037
dry cough	0.1591	0.072
body malaise	0.1574	0.0679
inappetence	0.1572	0.0708
phlegm	0.1509	0.0638
septic shock	0.1481	0.0609
esophageal reflux	0.1376	0.051
sepsis	0.1375	0.0575
cold	0.1344	0.0532
arrhythmia	0.1246	0.0495
dysphagia	0.1224	0.0428
lethargy	0.1206	0.0419
kidney injury	0.1202	0.0481
cardiopulmonary arrest	0.1161	0.0395
Headache	0.1131	0.038
cardiogenic shock	0.0973	0.0312
sore body	0.068	0.0177
anhelation	0.0648	0.0165

Table 2 Data Point Categorisation (belongingness to the class) on the basis of Weightage Range

Data Point Weightage Range	Category (Belongingness to class)
0- 0.2961	Low
0.2961 - 0.5922	moderate
0.5922 - 0.8883	high
0.8883 - above	very high

References

1. Abbas, A., Abdelsamea, M.M., Gaber, M.M.: Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. arXiv preprint arXiv:2003.13815 (2020)
2. Barstugan, M., Ozkaya, U., Ozturk, S.: Coronavirus (covid-19) classification using ct images by machine learning methods. arXiv preprint arXiv:2003.09424 (2020)
3. Bishop, C.M.: Novelty detection and neural network validation. IEE Proceedings-Vision, Image and Signal processing **141**(4), 217–222 (1994)
4. Boehm, O., Hardoon, D.R., Manevitz, L.M.: Classifying cognitive states of brain activity via one-class neural networks with feature selection by genetic algorithms. International Journal of Machine Learning and Cybernetics **2**(3), 125 (2011)
5. Costa, M., Moura, L.: Automatic assessment of scintmammographic images using a novelty filter. In: Proceedings of the Annual Symposium on Computer Application in Medical Care, p. 537. American Medical Informatics Association (1995)
6. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society: Series B (Methodological) **39**(1), 1–22 (1977)
7. Depeursinge, A., Iavindrasana, J., Hidki, A., Cohen, G., Geissbuhler, A., Platon, A., Poletti, P.A., Müller, H.: Comparative performance analysis of state-of-the-art classification algorithms applied to lung tissue categorization. Journal of digital imaging **23**(1), 18–30 (2010)
8. Hasan, A.M., AL-Jawad, M.M., Jalab, H.A., Shaiba, H., Ibrahim, R.W., AL-Shamasneh, A.R.: Classification of covid-19 coronavirus, pneumonia and healthy lungs in ct scans using q-deformed entropy and deep learning features. Entropy **22**(5), 517 (2020)
9. Hempstalk, K., Frank, E., Witten, I.H.: One-class classification by combining density and class probability estimation. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 505–519. Springer (2008)
10. Irigoien, I., Sierra, B., Arenas, C.: Towards application of one-class classification methods to medical data. The Scientific World Journal **2014** (2014)
11. Japkowicz, N.: Concept-learning in the absence of counterexamples: an autoassociation-based approach to classification, praca doktorska. New Brunswick Rutgers, The State University of New Jersey (1999)
12. Liu, B., Lee, W.S., Yu, P.S., Li, X.: Partially supervised classification of text documents. In: ICML, vol. 2, pp. 387–394 (2002)
13. Noumir, Z., Honeine, P., Richard, C.: On simple one-class classification methods. In: 2012 IEEE International Symposium on Information Theory Proceedings, pp. 2022–2026. IEEE (2012)
14. PEH, D.: Ro duda, pe hart, and dg stork, pattern classification (2001)
15. Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C., et al.: Estimating the support of a high-dimensional distribution. Technical Report MSR-T R-99–87, Microsoft Research (MSR) (1999)
16. Randhawa, G.S., Soltysiak, M.P., El Roz, H., de Souza, C.P., Hill, K.A., Kari, L.: Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: Covid-19 case study. Plos one **15**(4), e0232391 (2020)
17. Ridder, D.: An experimental comparison of one-class classification methods. In: Proc. 4th Annual Conference of the Advanced School for Computing and Imaging, 1998 (1998)
18. Ritter, G., Gallegos, M.T.: Outliers in statistical pattern recognition and an application to automatic chromosome classification. Pattern Recognition Letters **18**(6), 525–539 (1997)
19. Sethy, P.K., Behera, S.K.: Detection of coronavirus disease (covid-19) based on deep features. Preprints **2020030300**, 2020 (2020)
20. Singh, D., Kumar, V., Kaur, M.: Classification of covid-19 patients from chest ct images using multi-objective differential evolution–based convolutional neural networks. European Journal of Clinical Microbiology & Infectious Diseases pp. 1–11 (2020)
21. Soofi, A.A., Awan, A.: Classification techniques in machine learning: applications and issues. Journal of Basic and Applied Sciences **13**, 459–465 (2017)
22. Sun, L., Mo, Z., Yan, F., Xia, L., Shan, F., Ding, Z., Shao, W., Shi, F., Yuan, H., Jiang, H., et al.: Adaptive feature selection guided deep forest for covid-19 classification with chest ct. arXiv preprint arXiv:2005.03264 (2020)
23. Taha, A., Hadi, A.S.: Pair-wise association measures for categorical and mixed data. Information Sciences **346**, 73–89 (2016)

24. Tarassenko, L., Hayton, P., Cerneaz, N., Brady, M.: Novelty detection for the identification of masses in mammograms (1995)
25. Tax, D.: One-class classification; concept-learning in the absence of counterexamples. ph. d thesis. delft university of technology, asci dissertation series, 2001. 146 p (2001)
26. Xu, B., Kraemer, M.U., Gutierrez, B., Mekaru, S., Sewalk, K., Loskill, A., Wang, L., Cohn, E., Hill, S., Zarebski, A., et al.: Open access epidemiological data from the covid-19 outbreak. The Lancet Infectious Diseases **20**(5), 534 (2020)