

Metrics for Identifying Bias in Datasets

Alessandro Simonetta¹, Andrea Trenta², Maria Cristina Paoletti¹ and Antonio Vetrò³

¹Department of Enterprise Engineering, University of Rome "Tor Vergata", Via del Politecnico n.1, 00133, Rome, Italy

²UNI CT533 Commission member

³Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi n. 24, Turin, Italy

Abstract

Nowadays automated decision-making systems are pervasively used and more often, they are used for taking important decisions in sensitive areas such as the granting of a bank overdraft, the susceptibility of an individual to a virus infection, or even the likelihood of repeating a crime. The widespread use of these systems raises a growing ethical concern about the risk of a potential discriminatory impact. In particular, machine-learning systems trained on unbalanced data could rise to systematic discriminations in the real world. One of the most important challenges is to determine metrics capable of detecting when an unbalanced training dataset may lead to discriminatory behaviour of the model built on it. In this paper, we propose an approach based on the notion of data completeness using two different metrics: one based on the combinations of the values of the dataset, which will be our benchmark, and the second using frame theory, widely used among others for quality measures of control systems. It is important to remark that the use of metrics cannot be a substitute for a broader design that must take into account the columns that could lead to the presence of bias in the data. The line of research does not end with these activities but aims to continue the path towards a standardised register of measures.

Keywords

data quality, iso 25000, bias, Artificial Intelligence, AI, machine learning, ML, standard, square

1. Introduction

Recently, machine learning has become more and more popular in different sectors [1][2][3][4][5], thanks to both the possibility of having large amounts of data available [6][7] and the possibility of increasingly efficient implementations [8][9][10][11].

The world around us is gradually being enriched with functionalities that seem "normal" to us, but which are actually the result of sophisticated learning algorithms belonging to the field of AI [12]. This may seem rational and neutral, but sometimes, it can lead to harmful situations such as discrimination [13] or to worse predictions than what could be achieved with balanced data [14]. Equity and discrimination risks arise mainly due to disproportionate data sets: learning algorithms build the model from training data, so such a disproportion can lead to conclusions that are out of line with reality [15][16][17]. On the other hand, in some situations it is difficult to have homogeneous and proportional data because if the population is unbalanced, the sample used for learning will also be unbalanced too. In literature there are different techniques that allow to fill the dataset with miss-

ing data [14][18][19]. The idea of the present study is to verify the balance of the data even before the onset of discriminatory problems by the prediction algorithm through an index that measures the degree of balance of the data. At the same time the proposed metric can be considered within the framework of the measures of the ISO 25000 series of standards [20][21]. In the following, reference will be made to the data quality "characteristics" (e.g. completeness, accuracy) defined in the ISO/IEC 25024 standard [22]. This article will also use the terminology of relational algebra: the term column identifies the field or attribute of the dataset (intensional level), while the term tuple or row identifies an instance of the dataset (extensional level).

2. The proposed methodology

Let us consider a dataset containing distinct instances of an entity, in *one-to-many* relation with other entities. The classic example is the entity *Person* that can belong to different categories or classes describing gender, ethnicity, income bracket and religion (Fig. 1).

A simple, but sometimes inefficient, way of realising this concept is through the introduction of as many columns as there are concepts linked by the set relation since it is a one-to-many relation. In a relational database, these relationships are realised through foreign keys or columns with discrete domain controlled through row constraints. For the problem at hand, some of these columns are special, i.e. they may contain an identifier (primary key), a protected attribute (e.g., gender or reli-

ICYRIME 2021 @ International Conference of Yearly Reports on Informatics Mathematics and Engineering, online, July 9, 2021

✉ alessandro.simonetta@gmail.com (A. Simonetta);

andrea.trenta@dataqualitylab.it (A. Trenta);

mariacristina.paoletti@gmail.com (M. C. Paoletti);

antonio.vetro@polito.it (A. Vetrò)

ORCID 0000-0003-2002-9815 (A. Simonetta); 0000-0001-6850-1184

(M. C. Paoletti); 0000-0003-2027-3308 (A. Vetrò)

© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)



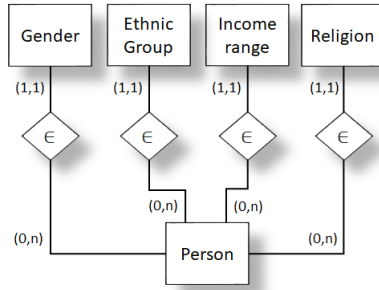


Figure 1: The Entity-Relationship diagram representing the person's membership of several categories

gion, or any other personal characteristic that should not be object of discrimination [23]), a *target* variable (usually denoted by Y) and an *output* of a prediction (usually denoted by R) if it has already been computed. However, the approach proposed is agnostic with respect to this classification because for our purposes we are interested in evaluating those columns that assume values in finite and discrete intervals, which we will call categorical with respect to the row data. This characteristic will allow us to consider the set of their values as the digits constituting a number in a variable base numbering system. The idea of the present study is based on the principle that a learning system provides predictions consistent with the data with which it has been trained. Therefore, if it is fed with non-homogeneous data it will provide unbalanced and discriminatory predictions with respect to reality. For this reason, the methodology we propose starts with the analysis phase of the reality of interest and of the dataset, an activity that must be carried out even before starting the pre-training phase in line with previous studies where bias measures in automated decision making systems were proposed [24][25][26][22]. In particular, during this phase, it is necessary to identify all the independent columns that define whether the instance belongs to a class or category. Suppose we have a structured dataset as follows:

$$DS = \{C_0, C_1, \dots, C_{n-1}\} \quad (1)$$

Indicating with the set S the positions of the columns categorising the instances, functionally independent of the other columns in the dataset:

$$S \subseteq \{0, 1, \dots, n-1\}, \dim(S) = m, m \leq n \quad (2)$$

we can analyze the new dataset consisting of the columns $C_{S(j)}$ with $j \in [0, m-1]$. Having said that, we can decide to use two different notions of completeness: maximum or minimum. In the first case the presence in the dataset of a greater number of distinct instances that

belong to the same categorising classes constitutes a constraint for all the other instances of the dataset. That is, one must ensure that one has the same number of replicas of distinct class combinations for distinct instances. Instead, in the second case it is sufficient to have at least one combination of distinct classes among those possible for each instance. For simplicity, but without loss of generality of the procedure, we will explore the minimum completeness of the dataset, then we will reduce the dataset to just the columns $C_{S(j)}$ by removing duplicate rows. We will use the Python language to explicate the calculation formulas and make the mathematical logic implied less abstract. The Python language has the *pandas* library, which makes it possible to carry out analysis and data manipulation in a fast, powerful, flexible and easy-to-use manner. Through the *DataFrame* class it is possible to load data frames from a simple csv file:

```
import pandas as pd
df=pd.read_csv(<filename>,
               delimiter=" ; ")
```

We will use two different metrics to measure the degree of completeness: one is based on combinatorial calculus and the other on frame theory.

2.1. Combinatorial metric

The ideal value of minimum completeness for the combinatorial metric is when in the dataset there is at least one instance that belongs to each distinct combination of categories. The absence of some combination could create the lack of information that we do not want to exist. To calculate the total number of distinct combinations we need to calculate the product of the distinct replicas per single category ($C_{S(j)}$).

```
k=len(df['CS0']).unique() *
len(df['CS1']).unique() * ... *
len(df['CSm-1']).unique()
```

On the other hand, in the dataset we only have the characterising columns so we can derive the true number of distinct instances in order to determine how far the data in our possession deviates from the ideal case.

```
len(df.drop_duplicates())/
(len(df['CS0']).unique() *
 len(df['CS1']).unique() * ... *
 len(df['CSm-1']).unique())
```

The value for maximum completeness is calculated from the maximum number of duplicates of the same combinations of characterising columns. For this reason it is necessary to maintain in the dataset in addition to the columns $C_{S(j)}$ a discriminating identification field of the rows with the same values in these columns. To determine the potential total, once the maximum number of

duplications (M) has been determined, it is necessary to extend this multiplication factor to all other classes.

```
M= df.groupby([ 'CS0' ,... , 'CSm-1' ] ).
      size().reset_index(name='counts').
      counts.max()
```

```
len(df)/(M*k)
```

2.2. Metric based on frame theory

What we have been described till now is just a combinatorial calculations exercise, however we would like to have a more sophisticated calculation model. For this reason, we have found in Frame theory, which finds its natural application in the field of signals and control systems theory, a promising calculation system [27][28]. In a H_m finite-dimensional Hilbert space, of m dimension and with an inner product $\langle \cdot, \cdot \rangle_{H_m}$, a frame of H_m is defined as a finite collection of vectors $(v_i)_{i \in I} \subset H_m$ if there exist two constants θ and Θ , with $0 < \theta \leq \Theta$, such that:

$$\theta \|v\|_{H_m} \leq \sum_{i \in I} |\langle v_i, v \rangle_{H_m}|^2 \leq \Theta \|v\|_{H_m}^2 \quad (3)$$

A frame is said to be *tight* if $\theta = \Theta$. So we can say that, a frame is tight when its vectors are as spread out in space as possible. A notable result that we will exploit in the following is that whenever a sequence $(v_i)_{i \in I} \subset H_m$ of vectors constitute a frame, the constants θ and Θ are the smallest and the largest eigenvalues respectively of the corresponding frame operator. Because the frame operator is self adjoint, a frame is tight if and only if its frame operator is an appropriate multiple of the identity operator. If we consider now as row vectors the component tuples of the original dataset, reduced to the m columns $C_{S(j)}$, we represent the new dataset with the matrix $V(k \times m)$:

$$V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} \quad (4)$$

Before calculating the Gramian matrix, the values must be centred with respect to the mean row, in the following way:

$$W = V - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \cdot \frac{1}{k} \sum_{i=1}^k v_i \quad (5)$$

The Gramian matrix is obtained by performing the matrix product:

$$G = W^T \cdot W \quad (6)$$

At this point, we can calculate the eigenvalues, which represent for each column a measure of variance of the values within that column.

$$\Lambda = (\lambda_1, \dots, \lambda_m) \quad (7)$$

With reference to the different categories of frames [29], we recall that for some of them it is possible to expect results on the values and multiplicity of the eigenvalues [30][31]. In the following we derive considerations about the tightness of the frame by analysing the distribution of the eigenvalues. The ideal case is when the frame is tight and the eigenvectors are uniformly distributed with respect to the vector space. In this case all the eigenvalues assume the same value, and an easy way to check this is when the difference between the maximum and minimum eigenvalues is zero:

$$\Delta = \max(\Lambda) - \min(\Lambda) = 0 \quad (8)$$

However, this index cannot be adopted since it does not have a predefined range of value and it would not allow comparisons between datasets. In this regard several measures of homogeneity and balance of a distribution have been proposed in literature (e.g., Shannon, Gini-Simpson, Theil, etc.). In particular we will use the Gini-Simpson index for categorical data which we will normalise to obtain values in the range $[0, 1]$:

$$f_i = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i} \quad (9)$$

$$Gini_{index} = \frac{m}{m-1} \cdot \left(1 - \sum_{i=1}^m f_i^2 \right) \quad (10)$$

What we have seen till now is not directly applicable to any dataset because the columns representing the categories may not be defined on a numerical domain. For example, the category *sex* is usually defined on the domain *char* or on the presence of null values inside columns. Moreover, since the method is based on the position of the eigenvectors in the vector space, it is sensitive with respect to their Euclidean distance. Distance depending on dataset values. Therefore, in order to apply the methodology correctly, the following algorithm should be followed:

1. analyse the reality of interest and the dataset by selecting the columns that classify the rows and that are independent of each other ($C_{S(j)}$);
2. mapping the domain of column values that classify the rows into an ordinal set of natural numbers \mathbb{N} by means of a transformation function that matches each value in the domain (numeric, alphanumeric or null) to a natural progressive number while preserving the ordering of the values;

3. construct the matrix W from V by subtracting the mean vector;
4. calculate the matrix G and its eigenvalues;
5. calculate the heterogeneity index of the eigenvalues;
6. analyse the results obtained and return to the initial point if any error situations arise (dependency between columns, incorrect mapping,...).

The Fig. 2 shows the two different algorithms needed to calculate the metrics described in this article.

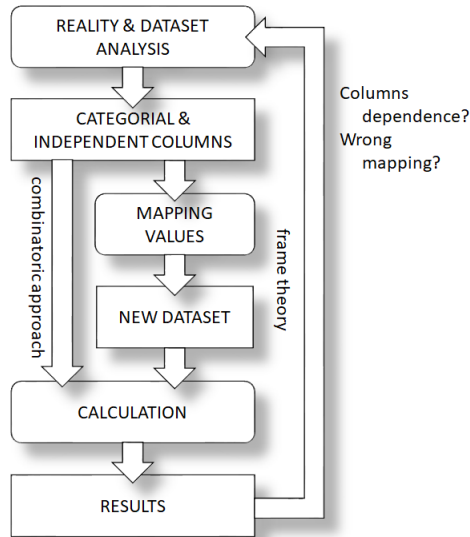


Figure 2: The two algorithms compared in terms of the steps to follow

3. Case Study

3.1. Designing

Due to the high sensitivity of the frame-based method to domain values, we preferred to test its effectiveness using a dataset with synthetic data. This allowed us to start from an ideal model. Knowing in advance the presence of bias in the dataset will allow us to verify the goodness of performance of the two metrics, without having the influence of any side effects. We will use a dataset with m categorising columns and assume that we have the same number of values in the categories that is to say X_B . Computing all X_B^m possible arrangements with repetitions is a simple task [32]. In fact, it is sufficient to compose all the numbers ranging from 0 to $X_B^m - 1$ expressing them in terms of m digits of the base X_B . In our study, we used a spreadsheet to automatically and easily create the various datasets, data sources that were then

given to the Python interpreter. This heuristic approach allow us to understand the limitations of the adopted metrics and their strengths. To make the case more realistic we used a dataset with 5 categorising columns on a domain based on 5 discrete values (0..4), so $m = 5$ and $X_B = 5$. The m categorising columns in our simulation are A, B, C, D and E (Table 1).

Table 1

The complete dataset in the base 5-value domain

other instance columns	A	B	C	D	E
...	0	0	0	0	0
...	0	0	0	0	1
...
...	0	0	0	0	4
...	0	0	0	1	0
...
...	4	4	4	4	3
...	4	4	4	4	4

With this assumption the distinct rows of the dataset are $X_B^m = 5^5 = 3,125$. The methodology can be replicated with any value of m and X_B . The method remains valid also in the more general case in which the m columns belong to categories with different range. In fact, in the latter case it will be sufficient to adopt a variable base numbering system (X_{B_i}). In the case study we will consider the case of minimum completeness, since maximum completeness can be easily extended by the latter.

The matrix V corresponds to the rows in Table 1:

$$V = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 4 & 4 & 4 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 \end{bmatrix} \quad (11)$$

The idea is to start from the complete schema by introducing, step by step, distortion effects to verify the behaviour of the two metrics. The combinatorial metric can be applied directly to the set of columns reduced from the original dataset without any transformation (Fig.5 2). Indeed, the advantage of this method is that it acts directly on the values regardless of their nature (datatype). The first index can be calculated using the following instructions:

```

len(df.drop_duplicates())/
(len(df['A'].unique()) *
 len(df['B'].unique()) *

```

```
len(df['C']).unique() *
len(df['D']).unique() *
len(df['E']).unique())
```

An index closer to unity means a better (minimum) completeness of the data set, while more it tends towards zero, greater is the incompleteness of the data. An index closer to unity means a better (minimum) completeness of the data set, while more it tends towards zero, greater is the incompleteness of the data. The second metric requires the Python library *numpy.linalg* to calculate the eigenvalues of the Gramian matrix. Once the eigenvalues have been calculated, the Gini-Simpson index can be determined:

```
from numpy import linalg as LA
df=df.drop_duplicates()
df=df.subtract(df.mean())
G=np.dot(df.T,df)
w, v = LA.eig(G)

# Gini-Simpson Index
(w.size/(w.size-1))*
(1-((w/w.sum())**2).sum())
```

3.2. Results

3.2.1. Case 1: no bias

With reference to Table 1, the presence of all possible digits in the 5 positions guarantees that there is no combination of the categories that is not represented, so the data set contains exhaustively all the allowed combinations [33]. It is interesting to verify that the value of the two metrics does not change with respect to the position of the rows and columns, so that a random shuffle of them does not alter the final value of the two metrics. This demonstrates the invariance of the two metrics with respect to the position of the rows and columns. Indeed, we cannot make any assumptions in advance about the sorting (horizontal or vertical) of the data.

3.2.2. Case 2: introduction of bias

First, we are interested to know what will be the behaviour of the two metrics if a subinterval of values is missing in the data set. This circumstance could lead to discrimination in a category. We will consider three different situations in which the rows of Table 1 are reduced:

1. deleting of a value in a column: e.g. all rows that have the value zero in column *A*;
2. deleting a minimum portion of values in a column: for example, the first 10 rows out of 625 that have in column *A* the value zero;

3. deleting a large portion of the values in a column: e.g. the first 615 rows out of 625 that have in column *A* the value zero.

As we can see from the results obtained (Table 2) it might be surprising that the combinatorial metric in case 1) returns the ideal value one: because the calculation is based on the values present in the dataset thus the result is that the domain is complete. The question is whether the bias measure should take into account all possible values of the categorising domains or not. If we want to extend the data domain to the all the possible (but not present) values we should increase the denominator of the index formula and the expected value should be 0.80. In cases 2) and 3) the combinatorial metric is consistent with the level of bias in the data: low bias corresponds to a value close to the ideal, while higher bias corresponds to a greater deviation from the optimum. The metric using the Gini-Simpson index applied to the eigenvalues of the *W* matrix identifies case 2) as the best situation, which represents an almost complete domain. While case 1) is correctly identified as the one with the highest bias in the data.

Table 2

The value of the two metrics in the three cases of bias

	combinatorial metric	Gini on eigenvalues metric
1)	1.0000	0.9934
2)	0.9968	0.9999
3)	0.8032	0.9936

3.2.3. Case 3: linear dependence

Suppose that during the design phase of the analysis (Fig. 2) a category that depends functionally on another is wrongly inserted as columns into the new dataset. For example, suppose that there is column *B* which depends on *A* ($B \rightarrow A$) according to the relation:

$$B = (A + 1) \bmod 5 \quad (12)$$

In this case the combinatorial index will change from 1 to 0.2 because the contribution of a column is reduced and the interval is reduced by one fifth ($3,125/5 = 625$) too. The metric based on the Gini-Simpson index applied to the eigenvalues is not affected by functional dependencies since the eigenvalues are the same. Therefore, the index is equal to unity. We observe the same behaviour of the two metrics if the functional dependence is between several columns, such as the case where *D* depends on the other columns ($D \rightarrow A, B, C, E$) according to the relation:

$$D = (A + B + C + E) \bmod 5 \quad (13)$$

3.2.4. Case 4: different ordinal scales

This paragraph demonstrates the necessity of the mapping phase (Fig. 2) in order to use correctly frame theory. In fact, since the method is based on the concept of Euclidean distance, it requires integer, homogeneous and comparable values. This is not directly applicable to real datasets since it is not possible to make a priori assumptions about the data domain of the columns. Having said that, we will study the behaviour of the two metrics in three different random transformations of the initial dataset (Table 1):

1. an uniform transformation T_1 applied to all columns that transforms discrete values into other discrete values that are not necessarily adjacent and have a different Euclidean distance from their initial positions:

$$T_1 : \begin{cases} 0 \rightarrow 44 \\ 1 \rightarrow 11 \\ 2 \rightarrow 5 \\ 3 \rightarrow 111 \\ 4 \rightarrow 9999 \end{cases}$$

2. the same transformation T_1 applied to a single column;
3. a different transformation (T_A, T_B, T_C, T_D, T_E) per individual column (Table 3).

Table 3

The different transformations applied to columns

	T_A	T_B	T_C	T_D	T_E
0	10	90	77	11	1,500
1	20	9	55	8	2,000
2	30	99	66	9	500
3	40	900	44	15	2,500
4	50	999	88	14	1,000

The combinatorial metric is invariant with respect to the distance of the values since it counts the distinct shapes present in each column. So, in all the three cases the dataset is complete and has no bias (value one). Conversely, the metric based on frame theory is sensitive to distances so, except in the first case where the transformation acts consistently and uniformly on all values, in the others it detects a presumed bias (Table 4). This demonstrates the need to reposition the values on an equidistant ordinal scale in case the dataset does not have this feature in the category data.

4. Conclusion

The problem of biased outputs of machine learning systems in domains that impact the rights and freedom of

Table 4

The value of the two metrics in the three transformation cases

	combinatorial metric	Gini on eigenvalues metric
1)	1.0000	0.9999
2)	1.0000	0.0028
3)	1.0000	0.4989

people pose important ethical and social challenges. It is important to be able to anticipate the possible emergence of discriminatory behaviour which, in certain situations, could even violate fundamental human rights. For this reason, we propose to integrate the common data curation pipeline in ML with a preliminary analysis of intrinsic properties of the dataset that could anticipate bias problems downwards the chain. In this paper, building on top of previous work in measures of disproportions in datasets, we propose a measurement process composed of two complementary metrics: combinatorial and Gini on eigenvalues. We show a case study and report on their strengths and limitations. These metrics open up new scenarios for possible future developments in the mitigation of bias since the learning phase.

5. Acknowledgments

We would like to thank Prof. Marco Torchiano for the interesting discussion on the research articles and to the President of the Italian National Body UNI CT 504 Commission, Dr. Domenico Natale, for his foundational contributions in the field of data quality.

References

- [1] S. A. Ebiaredoh-Mienye, E. Esenogho, T. G. Swart, Artificial neural network technique for improving prediction of credit card default: A stacked sparse autoencoder approach, *International Journal of Electrical and Computer Engineering* 11 (2021) 4392–4402. doi:10.11591/ijece.v11i5.pp4392-4402.
- [2] G. Saranya, A. Pravin, A comprehensive study on disease risk predictions in machine learning, *International Journal of Electrical and Computer Engineering* 10 (2020) 4217–4225. doi:10.11591/ijece.v10i4.pp4217-4225.
- [3] A. Ramachandran, A. Ramesh, A. Sukhlecha, A. Pandey, A. Karuppiah, Machine learning algorithms for fall detection using kinematic and heart rate parameters-a comprehensive analysis, *IAES International Journal of Artificial Intelligence* 9 (2020) 772–780. doi:10.11591/ijai.v9.i4.pp772-780.

- [4] G. Capizzi, G. Lo Sciuto, C. Napoli, E. Tramontana, A multithread nested neural network architecture to model surface plasmon polaritons propagation, *Micromachines* 7 (2016). doi:10.3390/mi7070110.
- [5] G. Lo Sciuto, G. Capizzi, S. Coco, R. Shikler, Geometric shape optimization of organic solar cells for efficiency enhancement by neural networks, *Lecture Notes in Mechanical Engineering* (2017) 789-796. doi:10.1007/978-3-319-45781-9_79.
- [6] A. Boukhalfa, N. Hmina, H. Chaoui, Parallel processing using big data and machine learning techniques for intrusion detection, *IAES International Journal of Artificial Intelligence* 9 (2020) 553-560. doi:10.11591/ijai.v9.i3.pp553-560.
- [7] N. Seman, N. Razmi, Machine learning-based technique for big data sentiments extraction, *IAES International Journal of Artificial Intelligence* 9 (2020) 473-479. doi:10.11591/ijai.v9.i3.pp473-479.
- [8] G. C. Cardarilli, L. D. Nunzio, R. Fazzolari, M. Panella, M. Re, A. Rosato, S. Span, A parallel hardware implementation for 2-d hierarchical clustering based on fuzzy logic, *IEEE Transactions on Circuits and Systems II: Express Briefs* 68 (2021) 1428-1432. doi:10.1109/TCSII.2020.3032660.
- [9] G. De Magistris, S. Russo, P. Roma, J. Starczewski, C. Napoli, An explainable fake news detector based on named entity recognition and stance classification applied to covid-19, *Information (Switzerland)* 13 (2022). doi:10.3390/info13030137.
- [10] A. Simonetta, M. C. Paoletti, M. Muratore, A new approach for designing of computer architectures using multi-value logic, *International Journal on Advanced Science, Engineering and Information Technology* 11 (2021) 1440-1446. URL: http://ijaseit.insightsociety.org/index.php?option=com_content&view=article&id=9&Itemid=1&article_id=15778. doi:10.18517/ijaseit.11.4.15778.
- [11] R. Brociek, G. De Magistris, F. Cardia, F. Coppa, S. Russo, Contagion prevention of covid-19 by means of touch detection for retail stores, volume 3092, 2021, p. 89 - 94. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126138443&partnerID=40&md5=043ee2df651270a6ac0434a39cef92f5>, cited by: 0.
- [12] S. Spanò, G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Matta, A. Nannarelli, M. Re, An efficient hardware implementation of reinforcement learning: The q-learning algorithm, *IEEE Access* 7 (2019) 186340-186351. doi:10.1109/ACCESS.2019.2961174.
- [13] F. Z. Borgesius, et al., Discrimination, artificial intelligence, and algorithmic decision-making, Council of Europe, Directorate General of Democracy (2018). URL: <https://pure.uva.nl/ws/files/42473478/32226549.pdf>.
- [14] A. Watson, Reducing ai bias with synthetic data, 2020. URL: <https://towardsdatascience.com/reducing-ai-bias-with-synthetic-data-7bdc39f290d>.
- [15] C. O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (reprint edition ed.), Broadway Books, New York, 2017.
- [16] S. Barocas, A. D. Selbst, Big data's disparate impact, 2016. URL: <http://dx.doi.org/10.2139/ssrn.2477899>.
- [17] J. Angwin, J. Larson, S. Mattu, L. Kirchner, ProPublica, Machine bias - there's software used across the country to predict future criminals. and it's biased against blacks., 2016. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- [18] G. Menardi, N. Torelli, Training and assessing classification rules with imbalanced data, *Data Mining and Knowledge Discovery* 28 (2012) 92-122. URL: <https://doi.org/10.1007/s10618-012-0295-5>. doi:10.1007/s10618-012-0295-5.
- [19] G. Capizzi, C. Napoli, L. Paternò, An innovative hybrid neuro-wavelet method for reconstruction of missing data in astronomical photometric surveys, in: L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. A. Zadeh, J. M. Zurada (Eds.), *Artificial Intelligence and Soft Computing*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 21-29.
- [20] A. Trenta, ISO/IEC 25000 Quality Measures for A.I.: A Geometrical Approach, in: IWESQ@APSEC, 2020. URL: <http://ceur-ws.org/Vol-2800/paper-05.pdf>.
- [21] A. Trenta, D. Natale, Examples of practical use of ISO/IEC 25000, in: IWESQ@APSEC, 2019. URL: <http://ceur-ws.org/Vol-2545/paper-02.pdf>.
- [22] ISO/IEC 25024:2015, Systems and Software engineering - Systems and software Quality Requirements and Evaluation (SQuARE) - Measurement of data quality, 2015. URL: <https://www.iso.org/standard/35749.html>.
- [23] Charter of fundamental rights of the european union, 2000. URL: https://www.europarl.europa.eu/charter/pdf/text_en.pdf.
- [24] M. Mecati, F. E. Cannavò, A. Vetrò, M. Torchiano, Identifying risks in datasets for automated decision-making, in: *International Conference on Electronic Government*, Springer, 2020, pp. 332-344. doi:https://doi.org/10.1007/978-3-030-57599-1_25.
- [25] E. Beretta, A. Vetrò, B. Lepri, J. D. Martin, Detecting discriminatory risk through data annotation based on bayesian inferences, in: *ACM Con-*

- ference on Fairness, Accountability, and Transparency (pp. 794-804), 2021, pp. 794-804. URL: <https://doi.org/10.1145/3442188.3445940>.
- [26] E. Beretta, A. Vetrò, B. Lepri, J. D. Martin, Ethical and socially-aware data labels. in annual international symposium on information management and big data, in: Annual International Symposium on Information Management and Big Data (pp. 320-327). Springer, Cham., 2018. URL: https://link.springer.com/chapter/10.1007%2F978-3-030-11680-4_30.
- [27] K.MishalAssifP., M.R.Sheriff, D. Chatterjee, Measure of quality of finite-dimensional linear systems: A frame-theoretic view, *Syst. Control. Lett.* 151 (2021) 104911.
- [28] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser/Springer, 2016.
- [29] J. Kovacevic, A. Chebira, An introduction to frames, *Foundations and Trends in Signal Processing* 2 (2008) 1-94. doi:10.1561/2000000006.
- [30] M. Fickus, D. G. Mixon, Tables of the existence of equiangular tight frames, 2016. [arXiv:1504.00253](https://arxiv.org/abs/1504.00253).
- [31] E. Tsiligianni, L. Kondi, A. Katsaggelos, Approximate equiangular tight frames for compressed sensing and cdma applications, *Eurasip Journal on Advances in Signal Processing* 2017 (2017). doi:10.1186/s13634-017-0501-0, publisher Copyright: © 2017, The Author(s). Copyright: Copyright 2017 Elsevier B.V., All rights reserved.
- [32] A. Simonetta, M. C. Paoletti, Designing digital circuits in multi-valued logic, *International Journal on Advanced Science, Engineering and Information Technology* 8 (2018) 1166-1172. URL: http://ijaseit.insightsociety.org/index.php?option=com_content&view=article&id=9&Itemid=1&article_id=5966. doi:10.18517/ijaseit.8.4.5966.
- [33] UNI/TS 11725:2018, System and software engineering - Guidelines for the measurement of data quality, 2018. URL: <http://store.uni.com/catalogo/uni-ts-11725-2018>.