

ON THE IMPACT OF WEIGHTING INFORMATIVE FEATURES IN THE DECISION-MAKING PROCESS

Tatur M. , Chen J. ¹, Kruykov A. 

Abstract The article serves as a methodological and educational note, viewing decision-making as a multi-stage process, encompassing phases of data preparation and normalization, weighing informative features, executing machine learning algorithms, and analyzing results. The work aims to show that assigning weights to informative features is an informal procedure that can significantly influence decision-making outcomes. Experiments were conducted on a test dataset using formal classification, clustering, and ranking algorithms with variations in the weight coefficients of informative features. It was assumed that changes in class labels, clusters, and ranks indicate changes in decision-making. It is visually demonstrated that decision-making based on data classification using discriminant function methods does not depend on the selection of weights for informative features. Conversely, decisions made by classification, clustering, and ranking algorithms that rely on calculating distances between objects are affected by the choice of weights and the data distribution in the feature space. Therefore, selecting machine learning algorithms and weights for informative features must be conducted thoughtfully, considering the context of solving a specific applied problem.

Key words: Decision-making, Informative features, Weight coefficients, Clustering, Classification, Ranking

AMS Mathematics Subject Classification: 62H30, 68T05.

DOI: 10.32523/2306-6172-2026-14-3-141-150

1 Introduction

In its most general form, decision-making consists of two stages: the first involves building a decision model F , based on the applied problem, and the second focuses on making decisions Y using the input data X . In machine learning theory, various "unified" mathematical methods are employed to construct the model F , ranging from classical data mining techniques to artificial neural networks (ANN). With the proper selection of a mathematical method and by tuning the parameters W of the model F , an acceptable quality level of decision-making $Y=F(X, W)$ can be achieved (Fig. 1).

Traditionally, parameters W are associated with neural network weights. However, in the context of data analysis, they often serve as interpretable feature weights that determine their contribution to the final decision. As research shows [1],[2], feature weighting critically affects model robustness, especially in multi-criteria tasks. For example, works [3]-[5] demonstrate that adaptive weight assignment based on entropy or analytic hierarchy improves prediction accuracy, whereas incorrect weight distribution [6]-[8] distorts results even with minor natural data anomalies.

Despite the abundance of works in this field, including deep scientific research [9], simple recommendations for applying machine learning algorithms, assigning weights to informative features, and explaining model dependencies within the context of applied decision-making

¹Corresponding Author.

problems are rare in the literature. Therefore, this methodological note does not aim to introduce novel machine learning algorithms or evaluate the statistical efficiency (e.g., Accuracy) of classifiers on large datasets. Instead, the primary focus is to allow novice researchers to visually trace the trajectory of decision boundary shifts and observe how the class, cluster, or rank of specific objects changes under varying weights. We hope this explicit demonstration of the decision-making process will be particularly interesting for novice researchers in machine learning, both in scientific and methodological aspects.

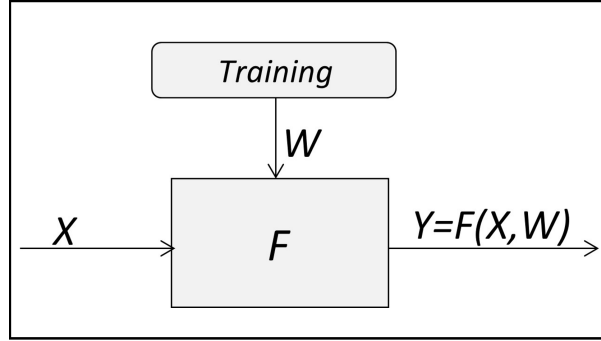


Figure 1: Generalized decision-making model

2 Formulation of the Research Problem

The problem statement can be presented in its original form (i.e., in terms of the applied domain) or in a formalized form. Methodologically, it is crucial to transition correctly from the applied to the formal problem statement.

In this work, as an example of an applied problem, we will evaluate players of a basketball team and make corresponding decisions. Suppose a qualifying game has taken place, and performance indicators $X(x^1, x^2)$ have been collected for all $n=11$ team athletes, where x^1 is the number of points scored and x^2 is the number of rebounds. The coach subjectively classified players with numbers 1,5,7,10,11 as "attacking" (A-class) if $x^1 \geq 20$, and all others as B-class.

The initial dataset is presented in Table 1.

Table 1: Distribution of initial monitoring data

n	1	2	3	4	5	6	7	8	9	10	11
x^1	22	16	17	18	20	18	24	19	19	24	23
x^2	7	8	12	8	7	11	12	9	13	6	9
Class	A	B	B	B	A	B	A	B	B	A	A

From the applied domain perspective, the following decision-making tasks can be articulated:

1. How can players be evaluated in the context of the qualifying game data?
2. What goal player groups can be formed based on the game results?

To address these (and similar) questions, it is essential to move from the specific domain to a formal problem statement.

It is necessary to design the decision making model in which each feature x^j is assigned a weight w^j based on the goal priority of "points" or "rebounds." We aim to evaluate how

varying the weight coefficients $W(w^1, w^2)$ of features $X(x^1, x^2)$ affects the final decisions. To achieve this, a series of experiments involving classification, clustering, and ranking tasks with different weights W are conducted.

Changes in cluster labels, classes, and object ranks will be interpreted as influence of weights $W(w^1, w^2)$ on decision making.

The computational experiments were implemented using Python programming language, specifically utilizing the pandas library for data manipulation and scikit-learn for machine learning algorithms.

The preliminary data processing stage involves normalization. We selected the standard z-score normalization approach (1)[10]:

$$z_i^j = \frac{(x_i^j - u^j)}{s} \quad (1)$$

where

$i = 1, \dots, 11; j = 1, 2$. z_i^j is the standardized feature value,

x_i^j is the original player performance indicators,

u^j is the mean value of the features, $u^j = 1/n \times \sum_{i=1}^n x_i^j$,

s is the standard deviation of the features, $s = \sqrt{\left\{1/n \times \sum_{i=1}^n [(x_i^j - u^j)]^2\right\}}$.

Normalized data are presented in Table 2, and their distribution is shown in Fig. 2. The figure illustrates that objects are uniformly distributed in the 2D feature space without distinct groups when feature weights are equal ($w^1 = w^2$).

Table 2: Normalization results

n	1	2	3	4	5	6	7	8	9	10	11
z^1	0.74	-1.48	-1.11	-0.74	0.00	-0.74	1.48	-0.37	-0.37	1.48	1.11
z^2	-1.01	-0.56	1.21	-0.56	-1.01	0.76	1.21	-0.12	1.65	-1.45	-0.12
Class	A	B	B	B	A	B	A	B	B	A	A

3 Conducting Experiments and Interpretation of Results

Assume that the coach (or expert) sets weight values $W(w^1, w^2)$ for parameters x^1 ("points") and x^2 ("rebounds") as follows:

Case 1: $w^1 = 1, w^2 = 5$;

Case 2: $w^1 = 5, w^2 = 1$.

We will solve the clustering, classification, and ranking tasks using these weights. Weighted Euclidean distance [11] is used for calculating distances between objects and centroids as follows:

$$d(c, z_i) = \sqrt{w^1(z_i^1 - z_c^1)^2 + w^2(z_i^2 - z_c^2)^2} \quad (2)$$

where z_i is the i -th object with coordinates (z_i^1, z_i^2) ; w^1, w^2 are weights; (z_c^1, z_c^2) are the current centroid coordinates.

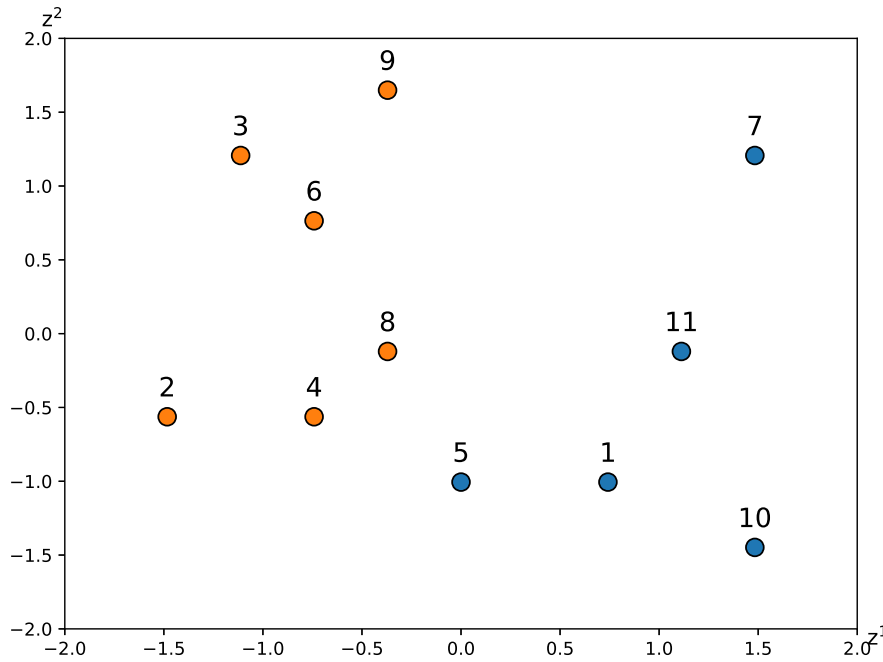


Figure 2: Distribution of initial data in normalized form

3.1 Classification by LDA method and centroid comparison method

Suppose that a new (12th) player joins the team with parameters $x_{12}^1 = 21$, $x_{12}^2 = 10$. It is needed to determine its class using the monitoring precedents.

First, normalize the data using formula(1). Normalized data are in Table 3.

Table 3: Normalization results with the new object

n	1	2	3	4	5	6	7	8	9	10	11	12
z^1	0.74	-1.57	-1.19	-0.80	-0.03	-0.80	1.51	-0.42	-0.42	1.51	1.12	0.35
z^2	-1.07	-0.61	1.23	-0.61	-1.07	0.77	1.23	-0.15	1.69	-1.53	-0.15	0.31
Class	A	B	B	B	A	B	A	B	B	A	A	???

Next, we construct two models of decision making as formal classification problems. In the first model, we apply the linear discriminant analysis (LDA) algorithm, and in the second model, we apply the centroid comparison algorithm. In each model, we will perform calculations with different weights w^1 , w^2 of 1:5 and 5:1, respectively. The results of the data classification experiments are shown in Figs. 3, 4. As follows from the classification results by the linear discriminant analysis algorithm, the twelfth object belongs to class A, in both cases of weights, i.e., changing the weights w^1 , w^2 from 1:5 to 5:1 did not affect the result of the decision. This is easily explained by the fact that the discriminant function (3), which is also the decision function, depends on the same weights w^1 , w^2 . It means that at different scaling along the abscissa and ordinate axes, the position on the plane of the graph (line) of the discriminant function changes with the change in the distribution of patterns.

$$y = w^1 z^1 + w^2 z^2 + C \quad (3)$$

where C is a constant.

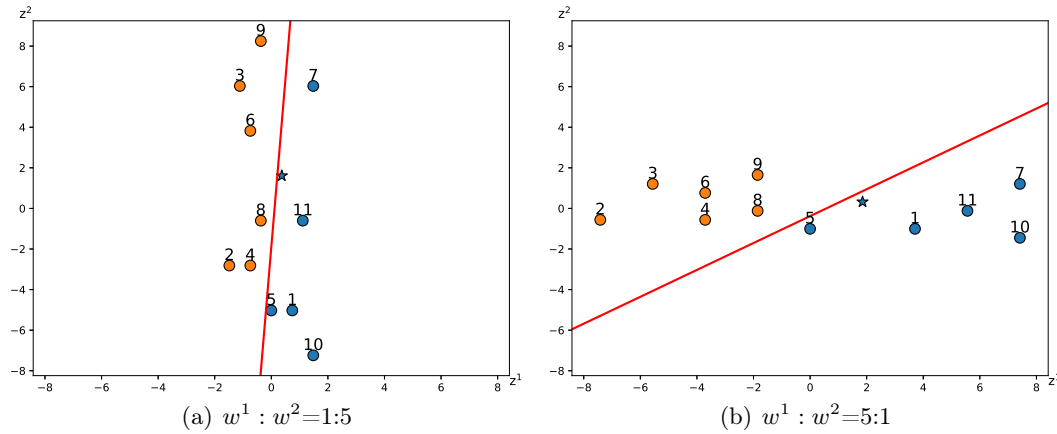


Figure 3: LDA classification results with different weights. The 12th object is marked as a star

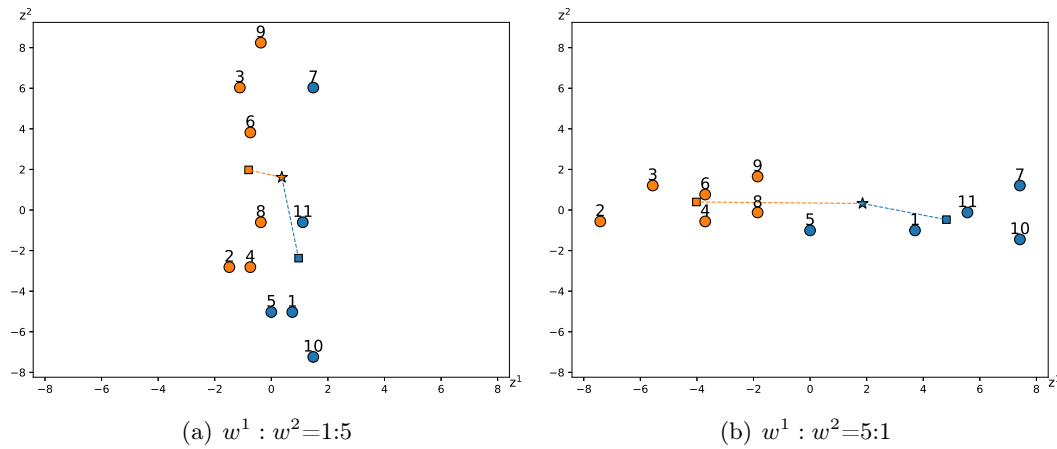


Figure 4: Centroid comparison classification results with different weights. The 12th object is marked as a star

However, in results of classification by the centroid comparison algorithm with different weighting ratios, the Euclidean distance from the unknown pattern to the centroids of the classes changes with the change in the distribution of patterns. Therefore, in the case of 1:5 weight ratio, the unknown pattern will be assigned a class B label and in the case of 5:1, it will be assigned a class A label. Thus, we conclude that the weights of informative features influence decision making.

3.2 Clustering by k-means method

Let's continue to consider the influence of weight coefficients on decision making when forming target (semantically defined) groups. For this purpose, let us conduct an experiment on data clustering using the k-means algorithm with different weight coefficients w^1 , w^2 for features X . The results of the experiment are presented in Figure 5.

In comparison with Fig. 2, the clustering results more clearly formed the groups of defensive players - patterns 3,6,7,9 in Fig. 5(a) and attacking players - patterns 1,7,10,11 in Fig. 5(b). It is useful to compare these results with the initial data from Table 1 and Fig. 2, as well as with the initial proposal of the "coach" to classify the players.

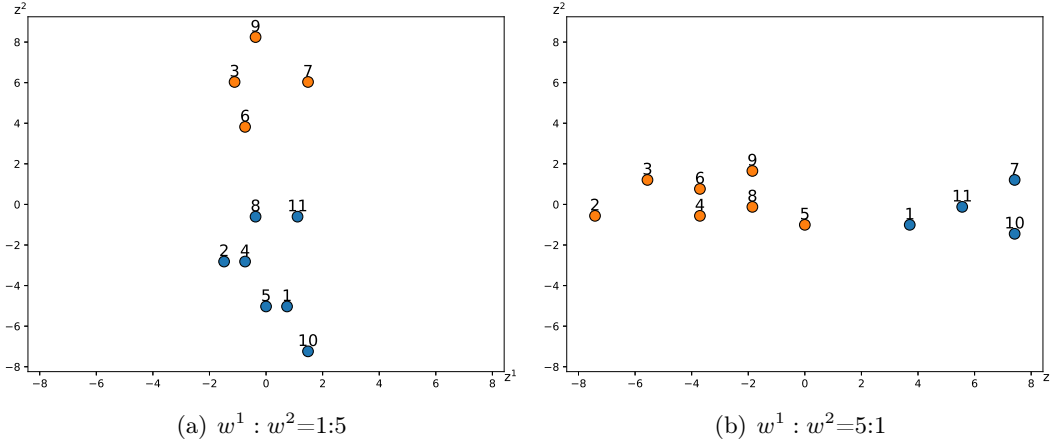


Figure 5: k-means clustering results with different weights.

Comparison of clustering results shows that the labels of object clusters (2, 4, 5, 7, 8) have changed. When the feature weights change, the composition of clusters and the distribution of objects in space change. This allows researchers to visually observe the trajectory of cluster boundary shifts, confirming that such transformations directly affect the possible decisions made at the individual object level.

3.3 Ranking

Ranking is one of the most commonly used data mining tasks. It allows determining the relative position of each object on one axis using a ranking function, which simplifies their interpretation and, ultimately, decision-making. The choice of the ranking function is determined by the specific application problem, and in most cases is an informal procedure. In this study, we applied the simplest and intuitive version of ranking by the sum of points and rebounds with weighting coefficients. This approach aligns with standard multiple attribute decision making methods [12], where we assume that this function (4) expresses the overall utility of players in both offensive and defensive aspects:

$$R(z_i) = w^1 z_i^1 + w^2 z_i^2 \quad (4)$$

The results of ranking the total utility of players, with different weights for scoring and rebounds are shown in Table 4.

Table 4: Results of object ranking with different weighting coefficients

RANK	1	2	3	4	5	6	7	8	9	10	11
w^1, w^2 (1:5)	8	9	3	7	10	4	2	6	1	11	5
w^1, w^2 (5:1)	4	11	10	9	6	8	1	7	5	2	3

As can be seen from the table of ranking results, the feature weights have a strong influence on the ranking results. If the feature weights are changed, the ranking solution may change completely.

4 Validation of the Obtained Conclusions

To validate whether the conclusions drawn from the synthetic data (based on basketball) can be generalized to independent real data, we employ the UCI Wine Quality (red) dataset [13]. Also, we will use other algorithms of decision making.

For the sake of direct observation of object-level changes, we need a small-sample, low-dimensional setting that allows object-level inspection of decision changes. Therefore, we selected only the eight records from the dataset, and only two features are retained, namely fixed acidity (x^1) and residual sugar (x^2). Initially, wine objects 2 and 3 are assigned to class A (high residual sugar, i.e., residual sugar > 2.0), while objects 1, 4, 5, 6, and 7 are assigned to class B. It is necessary to classify the 8th wine sample. This dataset is presented in Table 5.

Table 5: Distribution of initial Wine data

n	1	2	3	4	5	6	7	8
x^1	7.4	7.8	7.8	7.4	7.4	7.9	7.3	7.8
x^2	1.9	2.6	2.3	1.9	1.8	1.6	1.2	2.0
Class	B	A	A	B	B	B	B	???

Table 6: Normalization results

n	1	2	3	4	5	6	7	8
z^1	-0.87	0.87	0.87	-0.87	-0.87	1.31	-1.31	0.87
z^2	-0.03	1.74	0.98	-0.03	-0.28	-0.79	-1.80	0.22
Class	B	A	A	B	B	B	B	???

In this section, we employ Support Vector Machine(SVM) and K-Nearest Neighbor(KNN,k=3) classification methods to classify the 8th object under different feature weight settings(Case 1: $w^1 : w^2 = 1 : 2$, Case 2: $w^1 : w^2 = 2 : 1$) and analyze the impact of weight variation on decision outcomes. The features have been standardized using formula (1), as applied earlier in the paper, to obtain z^1, z^2 . The standardized data are shown in Table 6. The classification results are shown in Fig. 6 and Fig. 7.

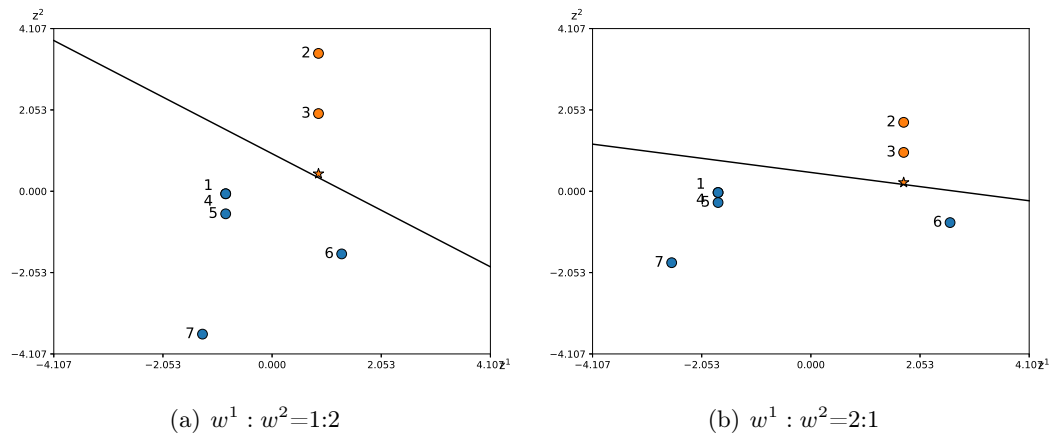


Figure 6: SVM classification results with different weights. The 8th object is marked as a star

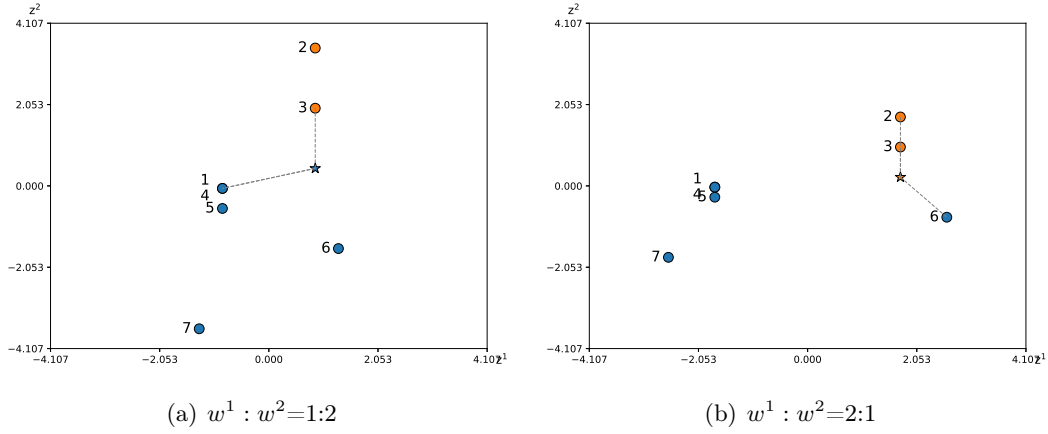


Figure 7: KNN classification results with different weights. The 8th object is marked as a star

According to the SVM classification results, under both weight settings (Case 1 and Case 2), the 8th object is consistently assigned to class A. This indicates that altering the feature weights does not affect the decision outcome. The reason is that rescaling feature axes by weight coefficients corresponds to a uniform transformation of both the object distribution in the feature space and the location of the separating line, so the set of support vectors remains unchanged. Thus, in this case, like the LDA case, the decision is robust to weight variation.

In contrast, the KNN classification results clearly demonstrate sensitivity to feature weighting. Under Case 1 ($w^1 : w^2 = 1 : 2$), the three nearest neighbors of the new object are samples 1 (B), 4 (B), and 3 (A), leading to classification as class B. Under Case 2 ($w^1 : w^2 = 2 : 1$), the three nearest neighbors shift to samples 2 (A), 3 (A), and 6 (B), resulting in classification as class A. This outcome illustrates that methods based on calculating distances between patterns depend on the weight balance of features directly, thereby changing the neighborhood composition and ultimately the decision result, as in the case with the Centroid comparison algorithm. Overall, the experiments on this independent dataset and alternative classification algorithms are consistent with the methodological findings obtained from the synthetic dataset in Section 3. While the sample size is intentionally kept minimal to preserve object-level interpretability, it effectively illustrates that decision boundaries in distance-based methods undergo a measurable shift (e.g., changing the class label of a specific object) when feature weights are altered.

5 Discussion and Future Work

This paper does not claim to present new scientific knowledge in the field of machine learning. However, the authors sought to demonstrate the methodology of the study, which included the following steps:

1. Application problem statement from the subject area;
2. Going from applied to formal problem statement from machine learning;
3. A hypothesis or plan for conducting the research;
4. Conducting the experiments;
5. Interpreting the results of the experiments in accordance with the hypothesis.

In general, the present work can be considered as a kernel for conducting further research in machine learning with increasing depth of immersion, but adhering to the presented structure

and methodology. So, one of the first questions that arises is, how to choose the right feature weights? This question may not be so simple unless one considers it "like training a neural network". Next, it is reasonable to investigate whether the results and conclusions obtained change if other ways of normalizing the data are used: if not, on what considerations should one or another method be chosen? The next stage of immersion could be to investigate the impact of the metrics used to compare and evaluate the proximity of patterns in the feature space: traditionally Euclidean distance is used, but how well do the results adequately reflect the problem in the subject area compared to other metrics? This step will be closely related to the dimensionality of the feature space, so it is useful to investigate the impact of the number of informative features on the accuracy of decision making. Here it is also worth considering ways to reduce dimensionality in the context of the ranking problem already discussed.

Furthermore, in practical applications, relying on expert subjective weighting can introduce bias. Future research within this methodological framework should explore objective weighting mechanisms. Techniques such as the Analytic Hierarchy Process (AHP), Entropy weight methods, and automated feature selection algorithms [3]-[5] provide formal mathematical foundations for weight distribution. Integrating these automated weighting techniques with the distance-based algorithms discussed in this paper represents a logical next step to reduce subjective bias while maintaining high model interpretability [14]-[16].

This list can be continued if desired, but for us it is of interest to explore these issues from a unified methodology and intuitive background, so that a novice researcher can see the basic principles, methods and dependencies of Machine Learning and DataMining in a complex and interconnected way. Then, in general, such a cycle of experiments performed within the framework of a common methodology can become the basis for a training course.

6 Conclusion

The development of a mathematical decision-making model encompasses both formal and informal stages. The selection of weight coefficients for informative features in tasks involving distance calculations between patterns constitutes one such informal stage. This paper demonstrates the mechanism of this influence through examples of classification, clustering, and ranking tasks. Decision-making is influenced by the distribution of patterns, the metric employed, and the dimensionality of the feature space. This relationship can be complex; however, it is evident that decisions are most susceptible to change for patterns located close to each other under the given metric. In classification methods based on constructing a separating or discriminant function, altering the weights of informative features (i.e., scaling the distribution of patterns) correspondingly shifts the position of the discriminant function within the feature space. Consequently, changes in feature weights do not impact classification results (or decision-making outcomes). In ranking tasks, changes to the weights of informative features exert the most pronounced influence on decision-making, as ranking functions typically reflect domain-specific considerations and are not bound by formal metrics for assessing distances between patterns. For educational and methodological purposes, tracking these shifts on a minimal, interpretable set of objects is strictly necessary, as employing massively large datasets would obscure the visual trajectory of individual decision changes.

The present study did not directly investigate ANNs. While the W (learned during ANN training) play a key role in decision-making, with the ANN reproducing the discriminant function (in classification tasks). In ANN, the weights of neurons do not reflect the proportional influence of features on decision making. This implies that the parameter settings W are non-interpretable.

References

- [1] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. (Springer Series in Statistics). New York, NY, USA: Springer, 2009, doi: 10.1007/978-0-387-84858-7.
- [2] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, 2014, doi: 10.1016/j.compeleceng.2013.11.024.
- [3] B. Liang, B. Qin, W. Sun, S. Wang, and Y. Dan, "The application of intelligent weighting grey target decision model in the assessment of coal-gas outburst," *J. China Coal Soc.*, vol. 38, no. 9, pp. 1611–1615, 2013 (in Chinese).
- [4] Y. Ren, J. Yang, Z. Guo, and H. Cao, "Self-adaptive entropy weighted decision fusion method for ship image classification based on multi-scale convolutional neural network," *J. Electron. Inf. Technol.*, vol. 43, no. 5, pp. 1424–1431, 2021 (in Chinese), doi: 10.11999/JEIT200102.
- [5] Y. Zhang, X. Zhang, and M. Jia, "Adaptive evaluation method based on analytic hierarchy process," *J. Beijing Univ. Aeronaut. Astronaut.*, vol. 42, no. 5, pp. 1065–1070, 2016 (in Chinese), doi: 10.13700/j.bh.1001-5965.2016.0224.
- [6] L. Yu, X. Song, and Z. Wang, "Impacts and corrections of natural weight on nonlinear sci-tech reviews—Case study of TOPSIS method," *Data Anal. Knowl. Discov.*, vol. 2, no. 6, pp. 48–57, 2018 (in Chinese).
- [7] E. Yilmaz, "Analysis of personal health data breaches: Prioritization with BWM approach," *J. Orig. Stud.*, vol. 5, no. 2, pp. 73–84, 2024, doi: 10.47243/jos.2612.
- [8] D. Soriano, E. Ben-Michael, P. J. Bickel, A. Feller, and S. D. Pimentel, "Interpretable sensitivity analysis for balancing weights," *J. Roy. Stat. Soc. A, Stat. Soc.*, vol. 186, no. 4, pp. 707–721, 2023, doi: 10.1093/jrssi/qnad032.
- [9] V. Pihur, S. Datta, and S. Datta, "RankAggreg, an R package for weighted rank aggregation," *BMC Bioinform.*, vol. 10, Art. no. 62, 2009, doi: 10.1186/1471-2105-10-62.
- [10] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann (Elsevier), 2011.
- [11] N. H. Spencer, *Essentials of Multivariate Data Analysis*. Boca Raton, FL, USA: Chapman and Hall/CRC, 2013.
- [12] G.-H. Tzeng and J.-J. Huang, *Multiple Attribute Decision Making: Methods and Applications*. Boca Raton, FL, USA: CRC Press, 2011.
- [13] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, "Wine quality," UCI Machine Learning Repository, 2009, doi: 10.24432/C56S3T.
- [14] D. Wettschereck, D. W. Aha, and T. Mohri, "A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms," *Artif. Intell. Rev.*, vol. 11, pp. 273–314, 1997, doi: 10.1023/A:1006593614256.
- [15] T. Zuo, J. Zheng, Y. Wang, and Y. Li, "An ensemble learning method based on an evidential reasoning rule considering combination weighting," *Comput. Intell. Neurosci.*, vol. 2022, Art. no. 1156748, 2022, doi: 10.1155/2022/1156748.
- [16] X. Zhang, Z. Wang, L. Jiang, W. Gao, P. Wang, and K. Liu, "TFWT: Tabular feature weighting with transformer," in *Proc. 33rd Int. Joint Conf. Artif. Intell. (IJCAI)*, 2024, pp. 2570–2578, doi: 10.24963/ijcai.2024/284.

Mikhail Tatur
 Belarusian State University of Informatics
 and Radioelectronics,
 Minsk, Belarus.
 Email: tatur@bsuir.by

Jike Chen
 Belarusian State University of Informatics
 and Radioelectronics,
 Minsk, Belarus.
 Email: ts.chen@bsuir.by

Alexey Kruckov
 Belarusian State University of Informatics
 and Radioelectronics,
 Minsk, Belarus.
 Email: kruckaua@gmail.com