

Dynamic Nearest Neighbor: An Improved Machine Learning Classifier and Its Application in Finances

Oscar Camacho-Urriolagoitia ¹, Itzamá López-Yáñez ^{1,*}, Yenny Villuendas-Rey ^{1,*}, Oscar Camacho-Nieto ^{1,*} and Cornelio Yáñez-Márquez ^{2,*}

¹ Centro de Innovación y Desarrollo Tecnológico en Cómputo del Instituto Politécnico Nacional, Juan de Dios Bátiz s/n, GAM, Mexico City 07700, Mexico; ocamachou1000@alumno.ipn.mx

² Centro de Investigación en Computación del Instituto Politécnico Nacional, Juan de Dios Bátiz s/n, GAM, Mexico City 07700, Mexico

* Correspondence: ilopez@ipn.mx (I.L.-Y.); yvilluendas@ipn.mx (Y.V.-R.); ocamacho@ipn.mx (O.C.-N.); coryanez@gmail.com (C.Y.-M.)

Abstract: The presence of machine learning, data mining and related disciplines is increasingly evident in everyday environments. The support for the applications of learning techniques in topics related to economic risk assessment, among other financial topics of interest, is relevant for us as human beings. The content of this paper consists of a proposal of a new supervised learning algorithm and its application in real world datasets related to finance, called D1-NN (Dynamic 1-Nearest Neighbor). The D1-NN performance is competitive against the main state of the art algorithms in solving finance-related problems. The effectiveness of the new D1-NN classifier was compared against five supervised classifiers of the most important approaches (Bayes, nearest neighbors, support vector machines, classifier ensembles, and neural networks), with superior results overall.

Keywords: finance risk prediction; machine learning; supervised classification

Citation: Camacho-Urriolagoitia, O.; López-Yáñez, I.; Villuendas-Rey, Y.; Camacho-Nieto, O.; Yáñez-Márquez, C. Dynamic Nearest Neighbor: An Improved Machine Learning Classifier and Its Application in Finances. *Appl. Sci.* **2021**, *11*, 8884. <https://doi.org/10.3390/app11198884>

Academic Editor: Chintan Amrit

Received: 01 September 2021

Accepted: 16 September 2021

Published: 24 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Finance involves checking and savings accounts, credit cards and consumer loans, investments in the stock market, retirement plans, social security benefits, insurance policies and administration of taxes, among others [1]. The key role financial businesses have in society today is indisputable. In this context, the support represented by the applications of machine learning techniques in topics related to financial risk assessment, among other topics of financial interest, is relevant and is an active area of research [2,3].

Machine learning, artificial intelligence, data analytics and related disciplines have more and more presence in everyday environments [4,5]. It is not difficult to infer that the presence of intelligent algorithms will be relevant in the daily life of human beings of the future. The same will happen with the scientific and technological disciplines that are cultivated in the academic and industrial fields. A tangible example of this is that more and more machine learning algorithms are applied to patterns generated in different application areas [6].

The conceptual bases of machine learning are varied: Bayes theorem, distances and comparisons, imitation of the behavior of nature, metaheuristics, mathematical and logical models of the neurons of the human brain, mathematical functions and high-dimensional spaces are some examples, among others [7].

Several applications of machine learning algorithms have been made for financial business [8–10], and for the automatic risk assessment in the financial environment [11,12], with good results. However, despite the efforts made by researchers, there is no one best machine learning algorithm for all classification problems, as stated in the “no

free lunch” theorems [13]. This is why we aim at introducing a novel algorithm, based on the well-known nearest neighbor classifier [14].

The proposed algorithm, named Dynamic Nearest Neighbor (D1-NN) was compared against five state of the art machine learning classifiers with different conceptual foundations. The analysis of the results allows us to declare that D1-NN has an overall better performance for finance-related datasets.

2. Related Works on Machine Learning for Financial Business

Finance is a branch of economics, which is the science in charge of the study of money and capital markets, the institutions and participants that intervene in them, the policies for attracting resources and distributing the results of funds, economic agents, the study of the time value of money, the theory of interest and the cost of capital. Therefore, finance refers to the conditions and the opportunity with which capital is obtained, the uses of it, the inherent risks and the profit that an investor obtains [1].

If the general concepts of finance are restricted to the activities of individuals, the branch of personal finance arises. Regarding personal finance, some issues of significant importance for the human being are related to savings accounts, credit cards and consumer loans, among others. Regarding personal finances, it is a fact that people must make financial decisions, despite the uncertainty inherent in the management of their assets, whose management could have consequences for their economic situation.

The Z-Score is one of the first attempts to systematize financial data. It is a method that allows analyzing the financial strengths of a company, to make objective predictions about a possible bankruptcy. Z-Score was created in 1960 by Edward Altman, a professor at New York University. In developing the Z-score model, Altman performed a multiple discriminant analysis of data from a sample of companies. Half of those companies under the study had gone bankrupt two decades ago, while the rest were still in operation at the date of the creation of the Z-score model [15].

The possible bankruptcy of a company is a recurring theme in investigations about financial risks. This occurs because financial institutions (and investors) need to reduce the possible risk of not obtaining the expected dividends or even of not recovering their capital [16]. The problem of determining the possible bankruptcy of a company can be approached as a pattern classification problem, as described in [17]. In such research work, a bankruptcy prediction model is proposed by modifying the nearest neighbor classifier [14]. Several authors had address the problem of automatic bankruptcy prediction using machine learning algorithms [18], with an increased interest in using deep learning [19], metaheuristic algorithms [20] and classifier ensembles [21].

In addition to bankruptcy, credit risk has been the subject of various scientific publications [22]. In this regard, a typical research work is the prediction of good clients (those who will fulfill their obligations), and bad clients (those who are likely not to fulfill their obligations). Among machine learning methods, researchers have used deep neural networks with metaheuristics [23], and transfer learning [24] for this task. Ampountolas et al. compare several simple and ensemble classifiers for credit scoring, resulting in Random Forest being the most accurate one [5]; while Dastile et al. focused on comparing statistical based models [22]. Other studies are focused on using fuzzy sets and decisions trees [25]. Another interesting finance-related topic is bank campaigns [26]. In [4], machine learning algorithms are used to predict the selling of long-term deposits, and a dataset containing information from a telemarketing campaign carried out by a banking institution in Portugal from 2008 to 2013 was created. However, as stated before, there is no best machine learning algorithm for all classification problem.

3. Proposed Method

In this paper, we propose a novel model for classification, with an embedded feature selection approach. The algorithm is based on the well-known nearest neighbor classifier and is named as Dynamic 1-NN (D1-NN) because it reflects the idea behind the simple

but effective modification of the classical algorithm: we have stopped considering the set of attributes as a static set, to make it dynamic.

The new idea consists in the dynamic consideration of subsets extracted from the set of attributes $A = \{A_1, \dots, A_m\}$ throughout the dataset under study. Therefore, the training or learning phase remains exactly the same as in the classic model. The novelties are reflected in the classification phase and in a better performance, so that the new D1-NN algorithm is competitive in the state of the art of pattern classifier algorithms.

The initial assumptions for the proposed algorithm are exactly the same as those specified for the original algorithm. That is, for the creation and operation of the new D1-NN algorithm it is assumed that:

- In a specific pattern classification problem, there are m, n , and p which are fixed positive integers greater than 1.
- There is a training or learning set of cardinality n , $T = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, which is made up of patterns.
- Each of the n patterns in the set T is made up of m attributes $A = \{A_1, \dots, A_m\}$.
- The j -th component of the i -th training or learning pattern \mathbf{x}_i is denoted by x_{ij} .
- The patterns do not contain categorical, mixed, or missing values. There are only numeric values.
- Each of the patterns in the training or learning set belongs to a single class in the set of p classes $C = \{C_1, \dots, C_p\}$.

The D1-NN classifier training phase is the same as the one of the classical 1-NN algorithm. Therefore, the D1-NN classifier training or learning phase consists of storing in memory the training or learning set T .

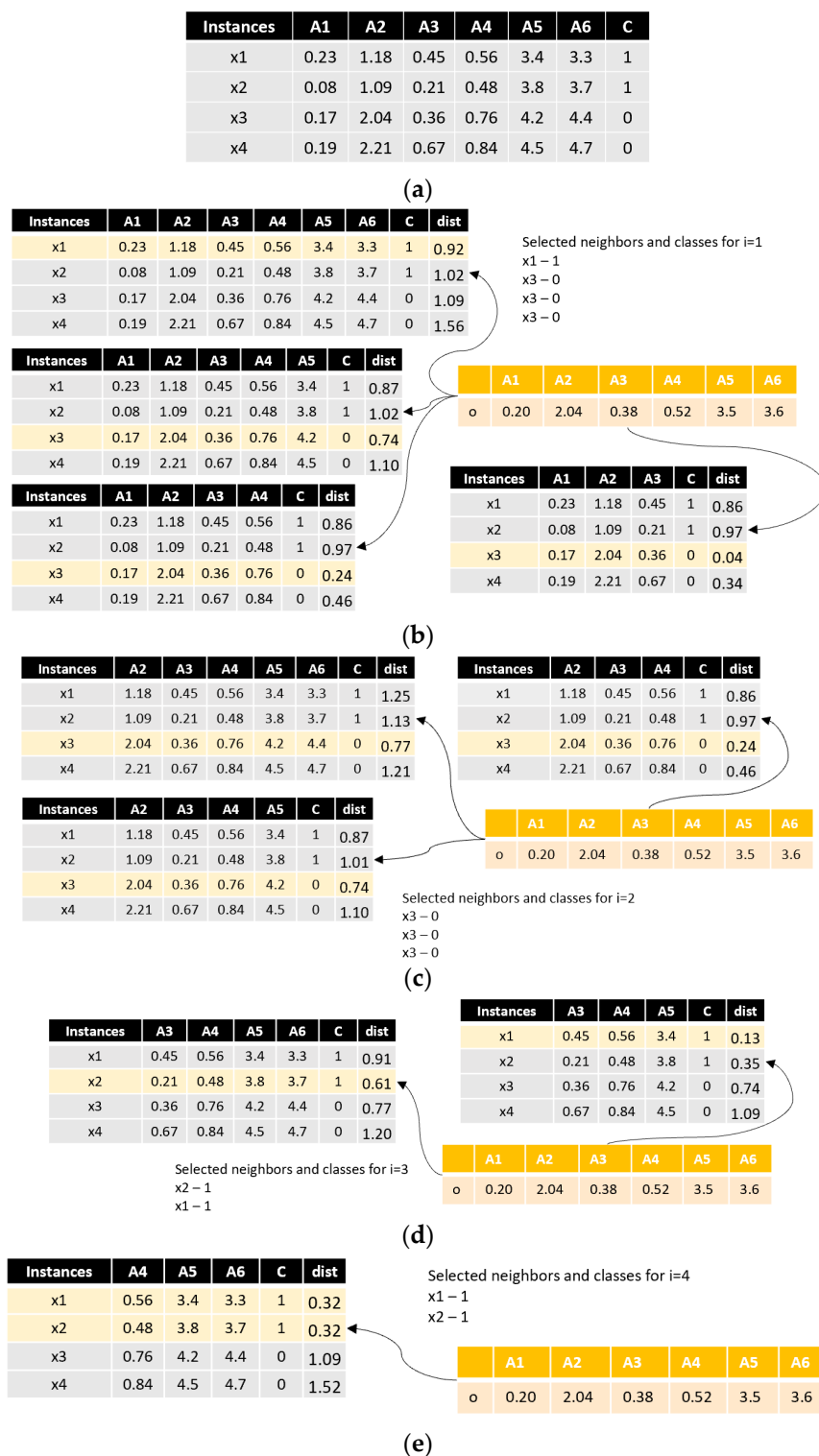
When the training or learning phase concludes, what proceeds is to test the proposed algorithm with patterns whose class the system does not know. To conduct the D1-NN classification phase, it is assumed that there is a new pattern to be classified, which is formed by the same m attributes of the training or learning patterns. This is the test pattern, which is denoted by \mathbf{o} , and whose class is unknown. In this phase, the D1-NN will attempt to assign the correct class to the test pattern \mathbf{o} .

The introduction of a parameter L to the proposed algorithm is the main contribution of our research to the original algorithm 1-NN. This parameter L gives the possibility to choose attributes in subsets of cardinality L . By moving a window of size L through the set of attributes A , the proposal becomes a dynamic algorithm, the D1-NN.

The classification phase of D1-NN has three steps:

1. A parameter L is set, $L \in \mathbb{N}, 0 \leq L \leq m$ (typically $0 \leq L \leq 3$).
2. For each index i such that $1 \leq i \leq m - L$, and for each index j such that $i + L \leq j \leq m$ do:
 - 2.1 Create a new learning set $T_{[ij]}$ with the patterns of T restricted to attributes A_k such that $i \leq k \leq j$
 - 2.2 Apply the 1-NN classifier to the pattern to be classified \mathbf{o} , with the set $T_{[ij]}$
 - 2.3 Store the class that delivers the 1-NN in step 2.2, as $C_{[ij]} \in C$
3. Assign to the pattern \mathbf{o} the most frequent class in the set of all values $C_{[ij]}$

In the following, we exemplify the functioning of the proposal with an example dataset (Figure 1a), and the classification phase of D1-NN is showed in Figure 1b to 1f. In the example of Figure 1 the number of attributes is $m = 6$ and we set the parameter $L = 2$. Let us consider an object to classify as $\mathbf{o} = [0.20, 2.04, 0.38, 0.52, 3.5, 3.6]$.



LAST STEP: Votes for each class: Class 1–5 votes; Class 0–6 votes. Decision: Assign class 0

Figure 1. Example of the classification phase of the proposed D1-NN. Note that the original training set is decomposed into several training sets, and the instance to classify (in orange) is presented to each of them. The distance (dist column) is computed, and the nearest instance in each of the resulting training set is highlighted in light yellow. The classes corresponding to the nearest instances are stored. (a) Training set; (b) Procedure for $i = 1$; (c) Procedure for $i = 2$; (d) Procedure for $i = 3$; (e) Procedure for $i = 4$. Finally, the LAST STEP: Decision.

In the example due to $m = 6$ and $L = 2$, the index i takes values that meet this inequality $1 \leq i \leq 6 - 2$; that is, $1 \leq i \leq 4$.

Also, for each of the four values of the index i , the index j takes the values determined by this inequality: $i + 2 \leq j \leq 6$. For each subset, we create a new training set, we find the nearest neighbor of the instance \mathbf{o} and store the corresponding classes.

- (i). For $i = 1$ the index j takes the values 3, 4, 5, and 6. That is, the following four subsets of attributes are formed:

$$\{A_1, A_2, A_3\}, \{A_1, A_2, A_3, A_4\}, \{A_1, A_2, A_3, A_4, A_5\}, \{A_1, A_2, A_3, A_4, A_5, A_6\}$$

- (ii). For $i = 2$ the index j takes the values 4, 5, and 6. That is, the following three subsets of attributes are formed:

$$\{A_2, A_3, A_4\}, \{A_2, A_3, A_4, A_5\}, \{A_2, A_3, A_4, A_5, A_6\}$$

- (iii). For $i = 3$ the index j takes the values 5, and 6. That is, the following two subsets of attributes are formed:

$$\{A_3, A_4, A_5\}, \{A_3, A_4, A_5, A_6\}$$

- (iv). Finally, for $i = 4$ the index j takes only the value 6. That is, the following subset of attributes are formed:

$$\{A_4, A_5, A_6\}$$

After conducting a large number of experiments with real-world datasets, we have found empirically that it is convenient to choose several windows in the indexes of attribute set A . Specifically, it is convenient to work with all the windows that are obtained by taking the following values of the parameter L : $0 \leq L \leq 3$.

4. Results

In order to achieve impact applications of the new classifier, eight finance-related datasets have been selected, and described in Section 4.1. Section 4.2 contains the description of five machine learning classifiers of the most important approaches: instance based classifier, decision trees, neural networks, support vector machines and Bayesian theory. The results obtained by these state of the art algorithms are presented in Section 4.3, which also includes discussions and comparative analysis.

4.1. Datasets

The eight datasets described in this subsection contain patterns taken from real-world activities, whose attributes are closely related to personal finance. All datasets, without exception, have two classes and all but Iranian credit are available in the machine learning repository of the University of California at Irvine [27]. The Iranian credit dataset was shared by the authors of the corresponding research [28].

Due to the restriction of the compared classifiers, the datasets were preprocessed, transforming all data to numerical data and using missing values imputation, in order to be able to apply state-of-the-art algorithms and compare those results with those obtained by the D1-NN in the same datasets.

In addition, an algorithm was applied to eliminate the imbalance of the classes in all the datasets and thus be able to use accuracy as a performance measure. The imbalance ratio (IR) is an index that measures the degree of imbalance. It is defined as [29]:

$$IR = \frac{|majority_class|}{|minority_class|} \quad (1)$$

where $|majority_class|$ represents the cardinality of the majority class in the dataset, while $|minority_class|$ represents the cardinality of the minority class. A dataset is considered balanced, if its IR value is close to or less than 1.5 (note that the IR value is always greater than 1).

To illustrate the concept of the IR index, the Qualitative Bankruptcy Dataset [30] will be taken as an example, which has 250 patterns included into two classes. The majority class contains 143 patterns while in the minority class there are 107 patterns.

$$IR = \frac{143}{107} = 1.33 \quad (2)$$

Thus, the Qualitative Bankruptcy Dataset is a balanced dataset.

In order to address the categorical and missing data, a transformation filter provided by the well-known WEKA (Waikato Environment for Knowledge Analysis) platform was applied [31]. After applying these filters, there are datasets with only numeric components and no missing values.

In the case of datasets with $IR > 1.5$, it was decided to apply the SMOTE (Synthetic Minority Over-sampling Technique) to compensate for the class imbalance [32]. By reducing class imbalance, it is now possible to apply accuracy as a performance measure.

After applying a pattern classifier to a dataset, it is necessary to calculate the performance, in order to know the benefits of this classifier with respect to the other algorithms of the state of the art. The accuracy performance measure is useful to be applied in balanced datasets, and is defined as [33]:

$$Accuracy = \frac{\text{Number of hits}}{\text{Total tested patterns}} \quad (3)$$

The following briefly describes the datasets in alphabetical order:

Australian credit approval

This dataset [34] contains data for people applying for credit cards. The two classes refer to the acceptance or rejection of each of the 690 applications. There are 14 attributes that each pattern contains. There are attributes of type continuous, nominal and missing values. Although the IR value indicates that there is no imbalance (1.24), the complexity of data at the attribute level that this dataset represents is a good challenge for classifiers.

Bank

A Portuguese banking institution ordered direct marketing campaigns which were based on phone calls [35]. Through a questionnaire structured in 16 items applied to 4521 people, it is intended to predict if a potential client will subscribe to a term deposit. The two classes are Yes/No. The dataset patterns contain mixed traits, that is, both numeric and categorical attributes, with no missing values. The IR index is 7.67, which indicates that the dataset is severely unbalanced.

Bank additional

This is an additional dataset to the Bank dataset. It also emerged from a marketing campaign which was also based on phone calls. The marketing campaign was carried out by a Portuguese banking institution, through a questionnaire structured in 20 items applied to 4119 people. As in the Bank dataset, the purpose is to predict if a potential client will subscribe to a term deposit [35]. Also, the two classes are Yes/No, and the dataset patterns contain mixed attributes, that is, both numerical and categorical attributes, with no missing values. The IR index is 8.13, which indicates that the dataset is severely unbalanced.

Banknote authentication

The creation of this dataset is intended to detect fraudulent banknotes [34]. To do this, data were extracted from images that were taken from genuine and forged banknote-like specimens. A total of 1327 patterns were generated, where each pattern consists of four attributes: wavelet variance, wavelet skewness, wavelet kurtosis and image entropy, which were extracted from the images. The IR index is 1.24, which indicates that the dataset is balanced.

Credit approval

This dataset [34] classifies people described by a set of attributes representing cases of people who were granted credit (383 patterns) and who were not granted credit (307

patterns) in a bank. In total the dataset contains 690 patterns that contain 15 mixed attributes (numeric and categorical). It has no missing values and the *IR* index is 1.24, which indicates that the dataset is balanced.

German credit data

This dataset classifies people described by a set of attributes as good or bad credit risks [34]. There are 1000 patterns containing 24 mixed attributes (numeric and categorical). Class 1 contains 700 patterns, while class 2 contains 300 patterns. It has no missing values and the *IR* index is 2.33, which indicates that the dataset is imbalanced.

Iranian credit

This dataset classifies people described by a set of attributes that represent cases of people who are good customers (950 patterns) and bad customers in terms of paying their credit (633 patterns) in an Iranian private bank. In total the dataset contains 1583 patterns containing 28 mixed attributes (numeric and categorical) [28]. It has no missing values and the *IR* index is 1.50, which indicates that the dataset is balanced.

Qualitative bankruptcy

With this dataset it is possible to predict a future bankruptcy based on qualitative attributes [30]. There are two classes, bankruptcy cases (107 patterns) and non-bankruptcy cases (143 patterns). In total the dataset contains 250 patterns that contain six categorical attributes). It has no missing values and the *IR* index is 1.33, which indicates that the dataset is balanced.

Table 1 contains a summary (in alphabetical order) of the specifications for the eight datasets.

Table 1. Specifications of the datasets (in alphabetical order).

Dataset	Number of Patterns	Number of Features	IR	Number of Classes
Australian credit approval	690	14	1.24	2
Bank	4521	16	7.67	2
Bank additional	4119	20	8.13	2
Banknote authentication	1372	4	1.24	2
Credit Approval	690	15	1.24	2
German credit data	1000	24	2.33	2
Iranian credit	1583	28	1.50	2
Qualitative bankruptcy	250	6	1.33	2

4.2. State of the Art Classifiers for Comparison

This subsection contains brief descriptions of five of the most important classifiers of the state of the art, which were implemented on the WEKA platform [31]. These are the classification algorithms against which our proposal, the D1-NN classifier, will be compared.

Naïve Bayes

The Naïve Bayes classifier [36] is a probabilistic algorithm with a superstructure based on Bayes' theorem. The classifier assumes that the features are independent of each other, and that is why it contains the word "naïve" in its name.

Nearest Neighbor (kNN)

This is one of the first supervised classifiers [14], and despite its simplicity, has very good performance. Nearest neighbor is an instance-based classifier, which uses dissimilarity or distance functions to select the closest instance to the pattern to classify.

Logistic Regression (Logistic)

Logistic regression is a statistical learning technique [31], used for both regression and classification problems. It has been widely used, and has a small computational complexity.

Multi-Layer Perceptron (MLP)

The multi-layer perceptron with backpropagation classifier [37] is an artificial neural network consisting of multiple layers, which allows solving non-linear problems. Although neural networks have many advantages, they also have their limitations. If the model is not trained correctly, it can give inaccurate results, in addition to the fact that the functions only look for local minima, which causes the training to stop even without having reached the percentage of allowed error.

Support Vector Machines (SVM)

The optimization of analytical functions serves as a theoretical basis in the design and operation of SVM models, which attempts to find the maximum margin hyperplane to separate the classes in the attribute space. SVMs [38] continue to occupy the first places in performance in classification problems. The relevance of SVMs has been emphasized as one of the most appreciated classifiers by the international scientific community. For this reason, we have included a representative model of the SVM in the experimental section of this paper.

Adaptive Boosting (AdaBoost)

Ensemble classifiers are methods which aggregate the predictions of a number of diverse base classifiers in order to produce a more accurate predictor, with the idea being that “many heads think better than one”. Ensemble models are valuable tools in pattern classification, routinely achieving excellent results on complex tasks. In this paper, we have selected a boosting ensemble, the AdaBoost algorithm [39].

Table 2 contains a summary of the specifications for the five algorithms.

Table 2. Algorithms against which the D1-NN classifier will be compared.

Algorithm	Conceptual Basis
Naïve Bayes	Bayesian theory
kNN	Instance-based
Logistic	Statistic-based
MLP	Artificial Neural Networks
SVM	Finding a kernel-based hyperplane
AdaBoost	Ensemble of classifiers

4.3. Performance and Comparative Analysis

In this subsection, the advantages of the proposal of this paper, the D1-NN classifier, are experimentally shown. It is necessary to carry out a systematic comparison of the performance of the new classifier with the classifiers of the state of the art. For this purpose, eight datasets related to various aspects inherent to personal finance have been selected. The results obtained by the D1-NN are compared against some of the most important classifiers of the state of the art. The accuracy is computed as in Equation (3). The results are shown in Table 3.

Table 3. Accuracy values (in %) obtained by the compared classifiers (best values in bold).

Datasets	Naïve Bayes	Logistic	kNN	MLP	SVM	AdaBoost	D1-NN
Australian credit approval	77.10	86.52	80.72	84.34	55.50	86.08	82.89
Bank	76.70	80.08	85.58	85.33	60.34	83.98	86.56
Bank additional	76.80	85.76	86.25	88.01	76.63	86.58	91.38
Banknote authentication	75.79	99.12	80.57	84.49	55.36	84.63	83.18
Credit approval	83.89	86.52	99.85	100.00	100.00	94.09	99.92
German credit data	71.78	77.00	73.58	71.61	63.37	75.55	71.69

Iranian credit	50.72	75.93	91.34	88.37	61.59	85.59	95.13
Qualitative bankruptcy	98.00	99.20	99.60	99.20	99.60	99.60	99.60
Average Accuracy	76.35	86.26	87.19	87.67	71.55	87.01	88.79

The results obtained are very promising and are discussed in the next section.

5. Discussion

In conducting the experiment, we use the leave-one-out (LOO) validation method, which has the advantage that it is a deterministic method, without random biases [40]. This means that the result will never change, regardless how many times the algorithm is repeated. This validation method consists of taking a pattern from the test set dataset and the remaining of the patterns as a training or learning set. Once the test pattern has been classified, the extracted pattern is returned to the dataset and the following pattern is taken as the test pattern in the next iteration.

From the analysis of the results in Table 3, it is possible to see really remarkable situations. It is at once obvious that the proposed classifier D1-NN is ranked first in FOUR of the EIGHT datasets (Bank, Bank additional, Iranian credit, and Qualitative bankruptcy). That is, the D1-NN is better than all other classifiers in 50% of the datasets with which the experiments were performed.

Of the four remaining datasets where D1-NN did not rank first, in no case is it ranked last. It is fair to note that the behavior of the classifiers in the dataset Qualitative bankruptcy is very particular. Certainly, in this dataset, the D1-NN classifier comes first. However, it is not the only one because there are three other classifiers that obtain the same accuracy value, and they also rank first. Those classifiers are: kNN, SVM, and AdaBoost. The parameters of all the compared classifiers are detailed in Table 4.

Table 4. Parameters of the algorithms against which the D1-NN classifier is compared.

Algorithm	Parameters
Naïve Bayes	-
kNN	k = 1, Distance: Euclidean
Logistic	Statistic-based
MLP	Hidden layers: $\frac{attributes+classes}{2}$, Learning rate: 0.2, Momentum: 0.2, Training time: 500, Validation threshold: 20
SVM	Kernel: polynomial, optimization: SMO [41]
AdaBoost	Ensemble size: 10, Base classifiers: C4.5 decision trees [42]

D1-NN is the ONLY classifier of the seven compared that ranks first in four of the eight datasets, the next is logistic regression with three first ranks. In this context, it is worth emphasizing the consistency of the results exhibited by the proposed D1-NN classifier. Furthermore, it is important to note that the D1-NN, being a simple classifier, competes successfully with AdaBoost which is not a simple classifier, but an ensemble where many simple classifiers work in a collaborative environment.

In contrast to the consistency exhibited by the D1-NN classifier, there is one classifier among the six compared, whose behavior is worth mentioning. It is about SVM, which is based on statistical learning theory and optimization of analytical functions, and which is one of the most appreciated models in machine learning, artificial intelligence, data mining and related areas. In the experimental results of Table 3, the SVM model has erratic behavior. On the one hand, it exhibits 100% performance in the credit approval dataset and therefore comes first in this dataset. However, it comes last in five of the eight datasets.

6. Conclusions

The creation and successful testing of the new D1-NN algorithm clearly indicates a situation in scientific research. From the results, it is concluded that it is always possible to improve long-lived and classic algorithms such as kNN. This situation represents good news for those who use machine learning to support decision-making in different activities of the human being, where personal finances excel for obvious reasons.

The analysis and reflections that were included in the text in relation to Table 3, seemed to indicate a broad superiority of the proposed D1-NN algorithm over the other six algorithms selected from the state of the art, which constitutes a contribution to academic research. The main limitation of our study is the number of datasets used, and we want to increase this in the future. As future short-term work, it is proposed to extend the D1-NN algorithm to a family of Dk-NN algorithms, where not only one neighbor is considered to make the decision, but k-nearest neighbors.

Author Contributions: Conceptualization, O.C.-U. and C.Y.-M.; Formal analysis, I.L.-Y. and Y.V.-R.; Investigation, O.C.-N. and C.Y.-M.; Methodology, O.C.-N.; Supervision, I.L.-Y.; Visualization, Y.V.-R. and C.Y.-M.; Writing—original draft, O.C.-U. and Y.V.-R.; Writing—review & editing, Y.V.-R. and O.C.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the used datasets (but Iranian credit) are available at [27]. The Iranian credit dataset was shared by the authors of [28].

Acknowledgments: The authors would like to thank the Instituto Politécnico Nacional (Secretaría Académica, Comisión de Operación y Fomento de Actividades Académicas, Secretaría de Investigación y Posgrado, Centro de Investigación en Computación, and Centro de Innovación y Desarrollo Tecnológico en Cómputo), the Consejo Nacional de Ciencia y Tecnología, and Sistema Nacional de Investigadores for their economic support to develop this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bodie, Z.; Robert, C. Merton and the Science of Finance. *Annu. Rev. Financ. Econ.* **2020**, *12*, 19–38.
2. Alessi, L.; Savona, R. Machine Learning for Financial Stability. In *Data Science for Economics and Finance*; Springer: Cham, Switzerland, 2021; pp. 65–87.
3. Levantesi, S.; Zacchia, G. Machine learning and financial literacy: An exploration of factors influencing financial knowledge in Italy. *J. Risk Financ. Manag.* **2021**, *14*, 120.
4. Moro, S.; Cortez, P.; Rita, P. Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns. *Neural Comput. Appl.* **2015**, *26*, 131–139.
5. Ampountolas, A.; Nyarko Nde, T.; Constantinescu, C. A Machine Learning Approach for Micro-Credit Scoring. *Risks* **2021**, *9*, 50.
6. Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; Galstyan, A. A survey on bias and fairness in machine learning. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–35.
7. Hart, P.E.; Stork, D.G.; Duda, R.O. *Pattern Classification*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2001.
8. Wei, W.; Zhang, Q. Evaluation of rural financial ecological environment based on machine learning and improved neural network. *Neural Comput. Appl.* **2021**, 1–18, <https://doi.org/10.1007/s00521-021-06312-z>.
9. Chen, T.-H.; Chang, R.-C. Using machine learning to evaluate the influence of FinTech patents: The case of Taiwan's financial industry. *J. Comput. Appl. Math.* **2021**, *390*, 113215.
10. Canhoto, A.I. Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective. *J. Bus. Res.* **2021**, *131*, 441–452.
11. Wu, Z. Using Machine Learning Approach to Evaluate the Excessive Financialization Risks of Trading Enterprises. *Comput. Econ.* **2021**, 1–19, <https://doi.org/10.1007/s10614-020-10090-6>.
12. Błaszczczyński, J.; de Almeida Filho, A.T.; Matuszyk, A.; Szlag, M.; Słowiński, R. Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. *Expert Syst. Appl.* **2021**, *163*, 113740.
13. Wolpert, D.H. The supervised learning no-free-lunch theorems. In *Soft Computing and Industry*; Springer: London, UK, 2002; pp. 25–42.

14. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27.
15. Altman, E.I. A fifty-year retrospective on credit risk models, the Altman Z-score family of models and their applications to financial markets and managerial strategies. *J. Credit. Risk* **2018**, *14*, 4. <https://doi.org/10.21314/JCR.2018.243>.
16. Boughaci, D.; Alkhalaf, A.A. Appropriate machine learning techniques for credit scoring and bankruptcy prediction in banking and finance: A comparative study. *Risk Decis. Anal.* **2020**, *8*, 15–24.
17. Chen, H.-L.; Yang, B.; Wang, G.; Liu, J.; Xu, X.; Wang, S.-J.; Liu, D.-Y. A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. *Knowl.-Based Syst.* **2011**, *24*, 1348–1359.
18. Clement, C. Machine Learning in Bankruptcy Prediction—A Review. *J. Public Adm. Financ. Law* **2020**, 178–196.
19. Smiti, S.; Soui, M. Bankruptcy prediction using deep learning approach based on borderline SMOTE. *Inf. Syst. Front.* **2020**, *22*, 1067–1083.
20. Ansari, A.; Ahmad, I.S.; Bakar, A.A.; Yaakub, M.R. A hybrid metaheuristic method in training artificial neural network for bankruptcy prediction. *IEEE Access* **2020**, *8*, 176640–176650.
21. Chen, Z.; Chen, W.; Shi, Y. Ensemble learning with label proportions for bankruptcy prediction. *Expert Syst. Appl.* **2020**, *146*, 113155.
22. Dastile, X.; Celik, T.; Potsane, M. Statistical and machine learning models in credit scoring: A systematic literature survey. *Appl. Soft Comput.* **2020**, *91*, 106263.
23. Pławiak, P.; Abdar, M.; Pławiak, J.; Makarenkov, V.; Acharya, U.R. DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring. *Inf. Sci.* **2020**, *516*, 401–418.
24. Shen, F.; Zhao, X.; Kou, G. Three-stage reject inference learning framework for credit scoring using unsupervised transfer learning and three-way decision theory. *Decis. Support Syst.* **2020**, *137*, 113366.
25. Teles, G.; Rodrigues, J.J.; Saleem, K.; Kozlov, S.; Rabêlo, R.A. Machine learning and decision support system on credit scoring. *Neural Comput. Appl.* **2020**, *32*, 9809–9826.
26. Ghatasheh, N.; Faris, H.; AlTaharwa, I.; Harb, Y.; Harb, A. Business analytics in telemarketing: Cost-sensitive analysis of bank campaigns using artificial neural networks. *Appl. Sci.* **2020**, *10*, 2581.
27. Dua, D.; Taniskidou, E.K. UCI Machine Learning Repository. Available online: <https://archive.ics.uci.edu/ml/index.php> (accessed on 25 August 2021).
28. Sadatrasoul, S.; Gholamian, M.; Shahanaghi, K. Combination of Feature Selection and Optimized Fuzzy Apriori Rules: The Case of Credit Scoring. *Int. Arab. J. Inf. Technol. (IAJIT)* **2015**, *12*, 138–145.
29. López, V.; Fernández, A.; García, S.; Palade, V.; Herrera, F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Inf. Sci.* **2013**, *250*, 113–141.
30. Kim, M.-J.; Han, I. The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms. *Expert Syst. Appl.* **2003**, *25*, 637–646.
31. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA data mining software: An update. *ACM SIGKDD Explor. Newsl.* **2009**, *11*, 10–18.
32. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357.
33. Ballabio, D.; Grisoni, F.; Todeschini, R. Multivariate comparison of classification performance measures. *Chemom. Intell. Lab. Syst.* **2018**, *174*, 33–44.
34. Available online: [http://archive.ics.uci.edu/ml/datasets/statlog+\(australian+credit+approval\)](http://archive.ics.uci.edu/ml/datasets/statlog+(australian+credit+approval)) (accessed on 20 August 2021).
35. Moro, S.; Cortez, P.; Rita, P. A data-driven approach to predict the success of bank telemarketing. *Decis. Support Syst.* **2014**, *62*, 22–31.
36. John, G.H.; Langley, P. Estimating continuous distributions in Bayesian classifiers. *arXiv Prepr.* **2013**, arXiv:1302.4964.
37. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536.
38. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297.
39. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139.
40. Fukunaga, K.; Hummels, D.M. Leave-one-out procedures for nonparametric error estimates. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 421–423.
41. Platt, J. Sequential minimal optimization: A fast algorithm for training support vector machines. In *Advances in Kernel Methods—Support Vector Learning*; Schoelkopf, B., Burges, C., Smola, A., Eds.; Microsoft Research: New York City, NY, USA, 1998.
42. Quinlan, J.R. Bagging, boosting, and C4. 5. In *Proceedings of the Aaai/iaai, Portland, OR, USA, 4–8 August 1996; Volume 1*, pp. 725–730.