IST-Africa 2020 Conference Proceedings
Miriam Cunningham and Paul Cunningham (Eds)
IST-Africa Institute and IIMC, 2020
ISBN: 978-1-905824-65-6

Towards Small-Scale Farmers Fair Credit Scoring Technique

Benjamin OTIENO[1], Franklin WABWOBA[1] George MUSUMBA[2]
[1]Kibabii University, P.O. Box 1699, Bungoma, 50200, Kenya
Tel: +254 733365384, Email: otienobenjamin@gmail.com, fwabwoba@kibu.ac.ke
[2]Dedan Kimathi University of Technology, Private Bag-10143, Nyeri, Kenya
Email: musumbaggw@gmail.com

Abstract: In consumer loans, where lenders deal with masses, use of algorithms to classify borrowers is fast catching up. Classification based on predictive models tend, to adversely affect borrowers. In this paper, we study the extent to which various algorithms disenfranchise borrowers lying on the boundaries of decision making. In the study, the data used for loan appraisal, and decisions made by the lenders are subjected to a set of select algorithms. The bias suffered by borrowers in each case is determined using mean absolute error (MAE) and relative absolute error (RAE). The results show that FURIA has the least bias with the MAE of 0.2662 and 0.1501 and RAE of 64.19% and 30.31% for the German and Australian data sets respectively. Consequently, FURIA is modified to remove the hard boundaries which results in even lower MAE of 0.2535 and 0.1264 and RAE of 64.14% and 27.73% for the German and Australian data sets respectively.

Keywords: small-scale farmers, risk scoring, fuzzy logic, FURIA

1. Introduction

Small-scale farmers produce over 70% of world food supply [1]. Feeding nine billion people in the world is no mean feat. Some of the challenges facing small scale farmers include lack of farm inputs, lack of labour, lack of farm machinery, lack of finance, use of uncertified seed, poor storage facilities, low adoption of technology, lack of fertilizer, late farm operations, subsistence mentality by farmers, lack of awareness of improved agricultural practices and lack of technical know-how [2].

The lack of financing is one key challenge that limits farming activities of small-scale farmers to a large extent. Access to finance for many business enterprises is facilitated by financial institutions. Retail banks play a big role in providing this financing. To provide the finance however, lenders will be interested in having the borrower guarantee his ability to repay. This is where risk appraisal or risk assessment becomes important.

The need for banks to appraise credit risk for large loans like mortgages or construction loans is just as much as is the need to appraise small and micro loans. For large loans, banks tend to use personalized data and the borrower ordinarily provides adequate information to be used in risk appraisal [3]. However, this becomes untenable for thousands of clients who borrow small amounts. It would be expensive to appraise each client on an individual basis. This is the category under which small-scale farmers fall. They are many in number and the information they provide for risk appraisal is either incomplete or not entirely accurate [4]. This forms the basis for the need of automated methods to perform risk scoring on this set of clients. Both banks and fin-techs have resorted to computer-based credit risk appraisals to speed up service delivery and reduce the cost of lending.

When small-scale farmers go to the financial institutions – retail banks or fin-techs, they are primarily perceived as high-risk borrowers. Most banks don't have specific products for this class of borrowers. Those that are willing to lend to them do so at a premium. The decision to lend or not to lend is arrived at after weighing the risk of the borrower.

The key aspect in decision-making is uncertainty [5]. Decision-making where there is uncertainty is akin to probabilistic inference [6]. In cases where there are a number of criteria to be considered in making a decision, each with its own uncertainty, the complexity of the decision making process increases [7]. In any instance where a decision is made based on probability, there are possibilities of making an erratic decision. There are two types of errors: an error committed when the null hypothesis is rejected when it is actually true, otherwise called false positive, or not rejecting the null hypothesis when it is actually false, otherwise called false negative. The false positive is also called type I error whereas the false negative is called type II error. In credit risk scoring, the risk of classifying a good client as bad hence reject their loan application is the type I error while type II error classifies bad customer as good, making banks end up in a loss position [8].

Many risk scoring algorithms used by banks and fin-techs aim at reducing mostly type II error, but in the process end up with a lot of type I error. In the process of doing that, deserving small-scale farmers are denied the much needed credit to finance their farming activities. This research explores an algorithm that has hitherto not been used in risk scoring, but has a potential of not only reducing errors, but can accommodate more borrowers, though not at the expense of the lender. The key weakness in the existing risk scoring algorithms being addressed in this paper is the crisp thinking. Lenders make binary decision – to lend or not to lend. This way, many who fall on the boundary on the rejection side fall there unjustifiably. And even some who fall on the rejection side justifiably could still be extended credit albeit at a premium rather than being denied the credit facility in totality.

The rest of the paper is organized with section two detailing the objective, methodology in section three and developments on the field under study in section four. Section five has the results with conclusion being discussed in section six.

2. Objective

Small-scale farmers particularly suffer from this anti-selection bias as a result of the kind of information they are able to avail to lenders for the purpose of risk scoring. Most small-scale farmers don't 1) have titles to their farms, they don't 2) keep books of accounts, 3) have access to collateral, 4) have guaranteed markets, and so the lenders may to some extent be justified in classifying them as high risk since the actual risk may be unknown in some cases. But even with the information they are capable of providing, they are still pushed to further marginal positions by the lenders. The objective of this paper is to determine the extent of bias introduced by the various algorithms used in risk scoring and to propose an algorithm that can reduce this bias. It is the researchers' hypothesis that if applied by retail banks and fin-techs, the proposed algorithm has the potential of increasing financial access to small-scale farmers, while not disadvantaging the lenders.

3. Methodology

Using quantitative data alone for credit risk scoring can be disadvantageous to low-income population, for example, where borrowers with fewer assets are deemed to be more at risk of default [1]. It has been shown that for micro-lending, qualitative analysis is of great importance in arriving at the credit score of a borrower [2]. Some banks use pure judgmental methods, like Teba Bank in South Africa, Unibanka in Latvia, and United Bulgarian Bank in Bulgaria whereas others use a combination of statistical methods and judgmental methods such as CAC Leasing in Slovakia and Credit Indemnity in South Africa [3]. For the micro borrowers, the more complicated the risk scoring method, the more costly the credit appraisal gets [1]. Automation of the risk scoring process using various algorithms can greatly reduce the cost of financial services.

This research explores the efficacy of various automated risk scoring algorithms that have the potential of reducing the cost of credit scoring. The data used depicts the kind of data that would ordinarily be sourced from small-scale farmers – has both quantitative and qualitative data and has several missing values.

Two independent data sets are used in the experiment. The experiment is set up to determine the effect of the various algorithm on the borrowers lying on the boundary of the decision making. One set of data is german credit data[1] whereas the other is Australian credit data[2]. This is a correlational research, a form of quantitative research where the correlation between the classification decision and the algorithm used is determined.

The first data set (German credit) has one thousand (1000) entities (applicants) whereas the second data set (Australian credit) has six hundred ninety (690) applicants. Both data sets have a combination of numerical, real and nominal data, typical of credit appraisal data that would be provided by small-scale farmers. Also, each data set has instances of missing data, another typical phenomenon with small-scale farmers' data availed for credit appraisal.

Each set of data is analysed independently against eight different algorithms. Seven of the algorithms are those commonly used in credit scoring, whereas the eighth is an experimental one. For both data sets, applied to the eight algorithms, the data was divided into 66% of the data used as the training set and the remaining 34% as the test set. This gave a uniform chance to all the algorithms. When comparing algorithms, such factors as accuracy, complexity and computational expense may be considered [4]. The key focus of analysis in this research was accuracy, as this is what constitutes the major risk to lenders. The more the accurate the prediction, the lower the probability of default.

4. Developments

Credit scoring is a typical classification problem which turns out to have a number of uncertainties. Incorrect classification, especially a classification that results in credit denial causes unwanted bias. Such biases can be detrimental to people with lower bargaining power, such as small-scale farmers. It is therefore important to have unbiased credit scoring classifiers. Credit scoring classifiers can be classified as individual classifiers, homogeneous ensemble classifiers or heterogeneous ensembles [5]. According to Yu, Wang and Lai [9], the algorithms can be classified as statistical (such as Linear Discriminant Analysis, Logistic Regression, Probit Regression, K-Nearest Neighbors (KNN), Decision Trees), mathematical programming (such as Linear Programming, Quadratic Programming, Integer Programming), artificial intelligence (such as Artificial Neural Networks, Support Vector Machines, Genetic Algorithm, Genetic Programming, Rough Set), hybrid approaches (such as ANN and Fuzzy Systems, Rough Set and ANN, Fuzzy System and support vector machines) or ensemble approaches (such as ANN Ensemble, support vector machines Ensemble, Hybrid Ensemble).

Classifiers can also be classified as parametric and non-parametric [6] or supervised and unsupervised [7]. Supervised classifiers include Bayesian network (B-Net) [8], multivariate Gaussian [10], Naïve Bayes [11], Decision trees [12], support vector machines [13] and classification trees [14]. B-Net and Naïve Bayes use belief network built on probabilistic graphical model with naïve Bayes applying strong (naïve) independence. Simple decision

[1] ftp.ics.uci.edu/pub/machine-learning-databases/statlog/

[2] http://archive.ics.uci.edu/ml/datasets/statlog+(australian+credit+approval)

trees can also be used in decision making. Statistical distributions such as normal distribution, multivariate Gaussian distribution and linear regression can also be used in decision problems.

Supervised classifiers include minimal distance classifiers [15], KNN [16], Locally weighted learning (LWL) [17], Additive logistic regression boosting (ALRB) [18], Kernel Density estimation [19], Logistic regression (LR) [20], multi-layer perceptron [21], Artificial Neural Network (ANN) [22], Euclidian Distance [23] and Neural Network [24]. In KNN, the k closest members are used to train the model and the outcome used to generalize the classification of the test data. LWL is an intuitive learning model where a decision is arrived at by relating similar occurrences from a database. On the other hand, unsupervised classifiers include Fuzzy Logic [25], Bootstrapping local [26], K-Means [27], Fuzzy Unordered Rule Induction Algorithm (FURIA) [4] and Genetic algorithms.

Algorithms that are commonly used in credit scoring include Bayesian Networks [28, 29], Linear Discriminate analysis [30, 31], Logistic Regression [31], Artificial Neural Networks [32, 33], K-Nearest Neighbor [34, 35], Deep Learning [36, 37], Decision Trees [38] and support vector machines [9, 39]. All these are classification algorithms that end up classifying an applicant into one of the two categories – grant loan or deny loan.

The regression analysis considers the relationship between borrowers' behavior and default variable. The simple equation is represented by equation 1.

$$yi = \beta' \cdot x + ui \qquad (1)$$

yi denotes whether the borrower has defaulted or not (taking binary 0 or 1). Further improvement of the regression model can take the form of LR. This is a deformation of the linear regression to take into account qualitative indicators alongside the quantitative indicators already taken care of in linear regression [40]. The logistic regression equation takes the form:

$$ln\left\{\frac{p_i}{1-p_i}\right\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots\dots\dots + \beta_n x_n \qquad (2)$$

Equation 2 makes the assumption that the variables are linearly related. This assumption is not usually true and in order to remove the non-linearity, in this research, the equation 2 is modified to equation 3.

$$P_j = \frac{e^{XiBj}}{\sum_{j=1}^{k-1} e^{XiBj} + 1} \qquad (3)$$

The LR ordinarily does not deal with instance weights but in this experiment the algorithm is modified to equation 3 to handle instance weights. In the equation, pi represents the probability defining the bounds of risk classification. However, this model assumes linear relationship between the variables. On the other hand, Naïve Bayes classifier assumes conditional independence of the independent variables (IV) [5] with the predictor taking the form in equation 4.

$$p(x|y) = \prod_{j=1}^{m} p(x_j|y) \qquad (4)$$

B-Net are a modification of Naïve Bayes, with the conditional independence of the IV relaxed. Instead, a correlation between the IV is determined. B-Net is commonly employed in credit scoring [41]. This method is ideal where there is limited data as it provides similar results for large data sets as well as small data sets [42]. Equation 5 represents B-Net algorithm.

$$p(\theta|y) \propto p(y|\theta)*p(\theta) \qquad (5)$$

More precisely, equation 5 can be represented as in equation 6.

$$p(\theta|y) = \frac{p(y|\theta)*p(\theta)}{p(y)} \qquad (6)$$

where θ, is the parameter of interest, p(θ|y) denotes the conditional density of θ for given y [43]. B-Net is ideal where there is need to use diverse sources of data on a single model. For example, use of customer bank statement items can be used to model credit risk for the customer [44]. Loffler, Posch, and Schone (2005) conclude from their research that due to the accuracy of the Bayesian estimators, the Bayesian statistic is deemed to be more accurate as compared to LR and Linear Discriminant Analysis (LDA) in risk estimation.

Fuzzy Unordered Rules Induction Algorithm (FURIA) is a rule based classification algorithm derived from the well-known Repeated Incremental Pruning to Produce Error Reduction (RIPPER) algorithm [4]. The addition made to RIPPER is the fuzzy rules, a move from the crisp rules used in RIPPER. FURIA trapezoidal function is represented by four parameters, $I^F = \emptyset^{s,L}, \emptyset^{c,L}, \emptyset^{c,U}, \emptyset^{s,U}$. The trapezoidal function is of the form in equation 7.

$$I^F(v) \overset{\text{df}}{=} \begin{cases} 1 & \phi^{c,L} \leq v \leq \phi^{c,U} \\ \frac{v-\phi^{s,L}}{\phi^{c,L}-\phi^{s,L}} & \phi^{s,L} < v < \phi^{c,L} \\ \frac{\phi^{s,U}-v}{\phi^{s,U}-\phi^{c,U}} & \phi^{c,U} < v < \phi^{s,U} \\ 0 & \text{else} \end{cases} \qquad (7)$$

In equation 7, $\emptyset^{s,L}, \emptyset^{c,L}$ represent the lower bound and upper bound of the elements with membership 1 whereas $\emptyset^{c,U}, \emptyset^{s,U}$ represent lower and upper bound of elements with membership >0. FURIA is easy to apply due to its closeness to human reasoning [45]. The fact that FURIA can process linguistic data make it close enough to human reasoning [46]. The trapezoidal function of I^F takes the shape in Figure 1.

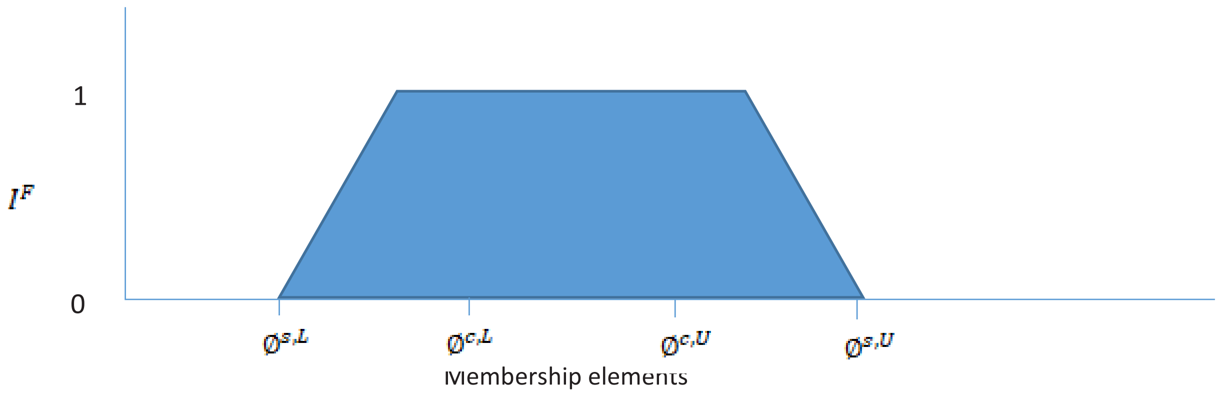

Figure 1: FURIA interval

There are a number of metrics that can be used to evaluate the accuracy of a machine learning algorithm. They include, MAE [47], Classification Accuracy, Logarithmic Loss, Confusion Matrix, Area under Curve, F1 Score, Mean Squared Error [48], Kappa values [49], and RAE [50].

Kappa value is the estimate that two algorithms agree [49]. This would be ideal where there is one algorithm that has been determined to be efficient and another is being benchmarked to it. MAE is a general measure of the difference between two continuous variables. On the other hand, RAE is an extension of the mean relative error divided by the exact value, expressed as a percentage. Logarithmic loss measures the misclassification by penalizing the false classifications [48]. Confusion matrix is a simple Yes/No matrix with the actual compared to the predicted. In confusion matrix, the measure on the efficacy of the algorithm is based on the combination of the actual and predicted outcomes. The possible outcomes are comprised of a combination of true positives, true negatives, false positives or false negatives.

It is very seldom that decisions are made with full knowledge of all the alternatives. In practical worlds, there is always some level of uncertainty in the decision making process. There are instances where the uncertainty can be precise (such as in a game of chance) and in some instances, that uncertainty is not knowable (such as the case of credit appraisal) [51]. In all these algorithms, what clearly comes out is that given the uncertainty, there is some learning done using the known data and then a decision is made on where to classify the unclassified entities. The classification is made based on how closely related the attributes of the entity being classified is to the entities from which the learning has been done. In all the algorithms, there are chances of incorrect classification of the entity. This is the main problem in this research – the misclassification of entities as a result of the calculations being done by any given algorithm. The research compares the cases of misclassifications committed by various algorithms, with a goal of determining which algorithm has the least misclassification error. This research uses mean absolute error and relative absolute error to measure the accuracy of the machine learning algorithms.

5. Results

In the German data set (Table 1), FURIA records the lowest MAE of 0.2662 and also the lowest RAE of 61.19%. However, Logistic regression and Naïve Bayes have the lowest incorrectly classified entries at 22.65%. Modification of the FURIA to remove the crisp decision for the entities at the boundary results in a reduction of incorrect classification for the FURIA from 87 to 79. 14 clients who would have wrongly been classified on the rejection or acceptance side however in the modified FURIA are given a chance to access loans, at a premium or a discount.

Table 1: German Credit results

		Type I error	Type II error	Total incorrectly classified	% of incorrectly classified	MAE	RAE
1	Logistic regression	35	42	77	22.65%	0.2986	72.02%
2	Naïve Bayes	36	41	77	22.65%	0.2811	67.77%
3	K-NN	57	45	102	30.00%	0.3006	72.49%
4	B-Net	43	**42**	85	25.00%	0.3021	72.85%
5	Additive logistic regression boosting	29	57	86	25.29%	0.3228	77.83%
6	Adaptive Boosting	28	58	86	25.29%	0.3274	78.94%
7	Locally weighted learning	**10**	88	98	28.82%	0.3598	86.76%
8	FURIA	25	62	**87**	**22.59%**	**0.2662**	**64.19%**

Table 2 shows the Australian data set, where FURIA turns out to have the lowest MAE of 0.1501 and also the lowest RAE of 30.31%. However, B-Net still has lower total incorrectly classified entities making up 14.04%. Some modification to the strict binary decision process, achieved by modifying the FURIA to treat the borrowers on the borderline differently results in 18 borrowers who would have either incorrectly been on the acceptance side or the rejection side, being given a variable pricing as a result of their new risk rating. The modified FURIA turns out to have the lowest total incorrectly classified entries at 26, MAE of 0.1264 and RAE of 27.73%. This ends up with more applicants having access to credit, and lower risk exposure to the lenders.

Table 2: Australian credit results

		Type I error	Type II error	Total incorrectly classified	% of incorrectly classified	MAE	RAE
1	Logistic regression	24	15	39	16.60%	0.1994	40.26%
2	Naïve Bayes	48	10	58	24.68%	0.2560	51.68%
3	KNN	31	20	51	21.70%	0.2183	44.06%
4	B-Net	25	**8**	**33**	**14.04%**	0.3540	34.90%
5	Additive logistic regression boosting	17	17	34	14.47%	0.2106	42.52%
6	Adaptive Boosting	23	13	36	15.32%	0.3367	43.52%
7	Locally weighted learning	**10**	24	34	14.47%	0.2259	45.61%
8	FURIA	20	15	35	14.9%	**0.1501**	**30.31%**

FURIA learns fuzzy rules instead of the conventional crisp rules and the rules are unordered, rather than appearing as rule lists [4]. FURIA has an advantage of having soft boundaries, whereas other classification algorithms have hard boundaries. However, for decision making, the soft boundaries are again converted into crisp boundaries. To obtain fuzzy rules, fuzzification of the crisp rules is done using training set DT ⊆ D to obtain the best fuzzy extention of each rule. The pseudocode for the FURIA in the experiment is as follows:

```
let A be the set of antecedents of r
while A<> null, do
        a_min ← null { a_min denotes the antecedent with the lowest risk}
        risk_min ←0 { risk_min is the lowest risk value so far}
        for i = 1 to size(A) do
                compute the best fuzzification of A(i) in terms of risk
                risk_A(i) ← be the risk value of this best fuzzification
                if Risk_A(i) > risk_min then
                        risk_min ← risk_A(i)
                        a_min ← A(i)
                end if
        end for
        A ← A\ a_min
        update r with a_min
end while
```

To achieve the crisp decisions from FURIA, rule stretching is usually done for the entities not covered by any rule. As much as rule stretching works better than default classification (predicting the most frequent class) [4], it still introduces unwanted bias. So instead of disenfranchising borrowers using the rule stretching, a modofication of FURIA is done to assign risk score assuming the given applicants are of unknown risk rather than assigning them untrue risk score. For these, a fixed premium is loaded to the pricing rather than total classification on the rejection.

Table 3 shows the results of the modified FURIA where there are 14 applicants in the German data set and 18 applicants in the Australian data set that could not be classified by FURIA now classified. The loan applicants at the boundary who in each case get the opportunity to access loans. These applicants cannot be necessarily classified as high risk, rather they have failed to be classified by any of the rules. Loading a premium to the pricing has the effect of smoothing the possibility of there being a number in the lot who are of high risk.

Table 3: Modified FURIA

		Type I error	Type II error	Total incorrectly classified	% of incorrectly classified	MAE	RAE
	Modified FURIA – German data set	19	60	79	23.23%	**0.2535**	**64.14%**
	Modified FURIA – Australian data set	16	10	26	11.06%	**0.1264**	**27.73%**

The quality of a lender's loan book is determined by the probability of default, sometime measured as value at risk (VAR) [52]. The more the accuracy with which the probability of default is measured, the lower the VAR. Use of the modified FURIA promises increased accuracy in prediction of client default, hence lower probability of default.

The application of the modified FURIA would lead to small-scale farmers ceasing to be seen as high-risk clients as the algorithm would measure their credit risk to near-accurately measure. Applied to loan applications – by retail banks or fin-techs – the modified FURIA has a potential of improving access to loans by small-scale farmers, without compromising the credit risk exposure by the lenders. This would go a long way in enhancing the productivity of the small-scale farmers, a major boost to food security.

6. Conclusions

Small scale farmers have suffered discrimination at the hands of lenders, denying them the much needed credit facilities. The lenders on the other hand have found it difficult to correctly appraise the risk of the small scale farmers. Existing risk scoring algorithms have so far worked well for the system with lenders limiting their credit risk exposures. But the algorithms have had serious biases against borrowers falling at the boundary of acceptance and rejection. These tend to be a substantial number of farmers, who end up being denied credit facilities. From the experiments in this paper, FURIA promises better results in terms of reducing the credit risk exposure to the lenders and at the same time including in the lending net more small-scale farmers. The modified FURIA promises even better results with the removal of sharp boundaries and giving risk based pricing to the borrowers who fall at the boundary.

The achievement of an optimal credit scoring for small-scale farmers, especially in developing nations has a potential of increasing the food production hence improving food security. In the experiments, the data used was from Germany and Australia. Further experimentation using data from developing nations should be done to confirm the consistency of the outcomes. Since credit scoring has reached higher maturity levels in Europe, further works involving collaboration between Europe and Africa on the application of the established credit scoring methods and the unique needs in Africa, applying new form of data in the scoring should be explored. Application of such technologies as smart ledgers, internet of things and big data can be explored in future works. This could result in retail banks and fin-techs opening up more room for small-scale farmers to access credit facilities. This way, the level of food production should improve considerably.

References

[1] Women's World Banking , "Credit scoring in microfinance," Women's World Banking , 2003.

[2] A. Merger and M. Barrera, "Credit scoring for microenterprise lenders," The Aspen Institute, 2007.

[3] USAID, "A handbook for developing credit scoring system in microfinance context," USAID, 2006.

[4] J. Hühn and E. Hüllermeier, "FURIA: An Algorithm For Unordered Fuzzy Rule Induction,"

Data Mining and Knowledge Discovery, 2007.

[5] S. Lessmann, B. Baesensb, H.-V. Seow and L. C. Thomas, “Benchmarking state-of-the-art classification algorithms for credit scoring,” Journal of the Operational Research Society , 2017.

[6] Y. Kumar and G. Sahoo, “Analysis of Parametric & Non Parametric Classifiers for Classification Technique using WEKA,” I.J. Information Technology and Computer Science, pp. 43-49, 2012.

[7] B. C. Love, “Comparing supervised and unsupervised category learning,” Psychonomic Bulletin & Review, 2002.

[8] N. Friedman, D. Geiger and M. Goldszmidt, “Bayesian network classifiers,” Machine Learning, 1997.

[9] L. Yu, S. Wang and K. K. Lai, “Bio-inspired credit risk analysis: Computational intelligence with support vector machines,” Springer, 2008.

[10] K. Mardia, “Multi-dimensional multivariate Gaussian Markov random fields with application to image processing,” Journal of Multivariate Analysis, 1988.

[11] T. Patil and S. Shereka, “Performance analysis of Naive Bayes and J48 classification algorithm for data classification,” IEEE Computer Society Conference, 2012.

[12] M. Friedl and C. Brodley, “Decision tree classification of land cover from remotely sensed data,” Remote sensing of environment, 1997.

[13] S. Gunn, “Support vector machines for classification and regression,” ISIS technical report, 1998.

[14] W. Buntine, “Learning classification trees,” Statistics and computing, 1992.

[15] W. Duch, N. Grudzi and G. Diercksen, “Neural minimal distance methods,” 3-rd Conf. on Neural Networks and Their Applications, pp. 183-188, 1997.

[16] K. Weinberger, J. B. Q. and L. K. Saul, “Distance metric learning for large margin nearest neighbor classification,” Advances in neural information processing systems, pp. 1473-1480, 2006.

[17] C. G. Atkeson, A. W. Moore and S. Schaal, “Locally weighted learning for control,” Lazy learning, pp. 75-113, 1997.

[18] G. Ridgeway, D. Madigan and T. Richardson., “Boosting methodology for regression problems,” AISTATS, 1999.

[19] Z. I. Botev, J. F. Grotowski and D. P. Kroese, “Kernel density estimation via diffusion,” The annals of Statistics, pp. 2916-2957, 2010.

[20] XuejunLiao, Y. Xue and L. Carin, “Logistic regression with an auxiliary data source,” in Proceedings of the 22nd international conference on Machine learning. ACM, 2005.

[21] S. Mitra and S. K. Pal, “Fuzzy multi-layer perceptron, inferencing and rule generation,” IEEE Transactions on Neural Networks , pp. 51-63, 1995.

[22] J. M. Zurada, Introduction to artificial neural systems, St. Paul: West publishing company, 1992.

[23] M. Bakonyi and C. R. Johnson, “The Euclidian distance matrix completion problem,” SIAM Journal on Matrix Analysis and Applications, pp. 646-654, 1995.

[24] L. K. Hansen and P. Salamon, “Neural network ensembles,” IEEE Transactions on Pattern Analysis & Machine Intelligence, pp. 993-1001, 1990.

[25] L. A. Zadeh, “Fuzzy logic,” Computer, 21(4), pp. 83-93, 1988.

[26] G. Claeskens and M. Aerts, “Bootstrapping local polynomial estimators in likelihood-based models,” Journal of statistical planning and inference, 86(1), pp. 63-80, 2000.

[27] J. A.Hartigan and M. A. Wong, “Algorithm AS 136: A k-means clustering algorithm,” ournal of the Royal Statistical Society. Series C (Applied Statistics), 28(1), pp. 100-108, 1979.

[28] C. K. Leong, “Credit Risk Scoring with Bayesian Network Models,” Computational Economics, 2016.

[29] M. W. Triki and Y. Boujelbene, “Bank Credit Risk: Evidence from Tunisia using Bayesian

Networks," Journal of accounting, finance and auditing studies, 2017.
[30] G. Mircea, M. Pirtea, M. Neamłu and S. Băzăvan, "Discriminant analysis in a credit scoring model," Recent Advances in Applied & Biomedical Informatics and Computational Engineering in Systems Applications, 2011.
[31] Z. Khemais, D. Nesrine and M. Mohamed, "Credit Scoring and Default Risk Prediction: A Comparative Study between Discriminant Analysis & Logistic Regression," International Journal of Economics and Finance, 2016.
[32] H. A. Bekhet and S. F. K. Eletter, "Credit risk assessment model for Jordanian commercial banks: Neural scoring approach," Review of Development Finance, 2014.
[33] V. Pacelli and M. Azzollini, "An Artificial Neural Network Approach for Credit Risk Management," Journal of Intelligent Learning Systems and Applications, 2011.
[34] R. Babu and A. Satish, "Improved of K-Nearest Neighbor Techniques in Credit Scoring," International Journal For Development of Computer Science & Technology, 2013.
[35] A. K. Abdelmoula, "Bank credit risk analysis with k-nearest neighbor classifier: Case of Tunisian banks," Accounting and Management Information Systems, 2015.
[36] P. M. Addo, D. Guegan and B. Hassani, "Credit Risk Analysis Using Machine and Deep Learning Models," Risks, 2018.
[37] J. A. Sirignano, A. Sadhwani and K. Giesecke, "Deep Learning for Mortgage Risk," arXiv, 2018.
[38] M. Sudhakar and C. V. K. Reddy, "Two step credit risk assesment model for retail bank loan Applications using decision tree data mining technique," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 2016.
[39] L. Hongjiu, H. Yanrong and Wuchong, "An Application of Support Vector Machine for Evaluating Credit Risk of Bank," Proceedings of the 7th International Conference on Innovation & Management, 2017.
[40] X.-L. Li and Y. Zhong, "An overview of personal credit scoring: techniques and future work," International journal of intelligence science, pp. 181-189, 2012.
[41] W. Abramowicz, M. Nowak and J. Sztykiel, "Bayesian Networks as a Decision Support Tool in Credit Scoring Domain," The Poznan University of Economics, 2003.
[42] G. Loffler, P. N. Posch and C. Schone, "Bayesian Methods for Improving Credit Scoring Models," Department of Finance, University of Ulm, Helmholtzstrasse 18, 89069 Ulm, Germany., 2005.
[43] A. Ziemba, "Bayesian updating of generic scoring models," StatConsulting Ltd., Warsaw, Poland, 2005.
[44] R. P. Bunker, M. A. Naeem and W. Zhang, "Improving a Credit Scoring Model by Incorporating Bank Statement Derived Features," Auckland University of Technology, 55 Wellesley St E, Auckland 1010, New Zealand, 2013.
[45] M. Gasparovica and L. Aleksejeva, "Using Fuzzy Unordered Rule Induction Algorithm For Cancer Data Classification," Mendel, 2011.
[46] D. Hand and W. Henley, "Statistical classification methods in consumer credit scoring: a review," J.R. Statist. Soc. A, 1997.
[47] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," Climate research, 2005.
[48] A. Mishra, "Metrics to evaluate your machine learning algorithm," 12 March 2019. [Online]. Available: https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234.
[49] R. S. DeFries and J. C.-W. Chan, "Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data," Remote Sensing of Environment 74.3, pp. 503-515, 2000.
[50] K. Chen, S. Guo, Y. Lin and Z. Ying, "Least absolute relative error estimation," Journal of the American Statistical Association, 105(491), pp. 1104-1112, 2010.

[51] Decision Innovation, “Making effective decisions in high uncertainty,” 13 March 2019. [Online]. Available: https://www.decision-making-solutions.com/decision-making-in-uncertainty.html.
[52] C. Zazzara, “Credit risk in the traditional banking book: a VaR approach under correlated default,” Research in Banking and Finance 2, 2000.