Plagiarism Detector v. 1921 - Originality Report 24/08/2021 15:20:06

Analyzed document: 237-1183-1-SM.pdf Licensed to: JPKMI

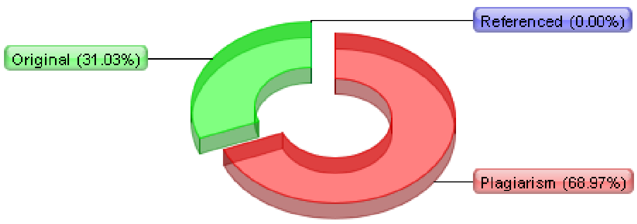
Comparison Preset: Rewrite Detected language: En

Check type: Internet Check

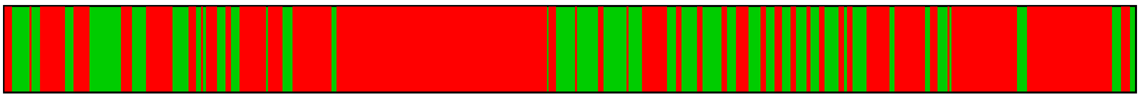
Disclaimer: this report must be correclty interpreted and analyzed by a qualified person who bears the evaluation responsibility!
Any information provided in this report is not final and is a subject for manual review and analisys!

Detailed document body analysis:

Relation chart:



Distribution graph:



Top sources of plagiarism: 75

96%	3720	1.	http://icsejournal.com/index.php/JCSE/article/download/172/88
90%	3353	2.	http://icsejournal.com/index.php/JCSE/article/download/26/28
37%	1315	3.	https://www.researchgate.net/publication/348604301_Loan_default_prediction_using_decision_trees_and_random_forest_A_compa

Processed resources details: 268 - Ok / 27 - Failed

Important notes:

Wikipedia:	Google Books:	Ghostwriting services:	Anti-cheating:
Wiki Detected!	[not detected]	[not detected]	[not detected]

Active References (Urls Extracted from the Document):

No URLs detected

Excluded Urls:

No URLs detected

Included Urls:

No URLs detected

Detailed document analysis:

Plagiarism detected: **0,79%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **1**

Available online at http://icsejournal.com/index.php/JCSE
Journal of Computer Science an Engineering (JCSE)

e-ISSN 2721-0251
Vol. X, No. X, February 2020, pp. xx-xx

1
http://dx.doi.org/10.
35671/jcse

Prediction of Loan Behaviour with Machine Learning Models
for Secure Banking

Mayank Anand1*, Arun Velu2, Dr. Pawan Whig3

1Research Scientist, The Research World, New Delhi, India
2Director, Equifax, Atlanta, USA
1mayankan@gmail.com*; 2gctarun@gmail.com ; 3 pawanwhig@gmail.com
* corresponding author

Plagiarism detected: **0,43%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **2**

A R T I C L E I N F O

ABSTRACT

Article History:

Received
August 23, 2021

Revised XXX
Accepted XXX

Available online XXX

Given loan default prediction has such a large impact on earnings, it

Plagiarism detected: **2,16%** https://www.researchgate.net/publication/33908... + 3 resources! id: **3**

is one of the

most essential and significant concerns that banks and other financial

organisations confront. Although several traditional methods for mining

information about a loan application exist, most of these methods appear to be

underperforming, as the number of problematic loans has increased. For loan

default prediction, a variety of techniques such as Logistic Regression, Decision

Trees, Random Forests, Support Vector Machines, and others are presented in this

research work. The prediction is based on
loan data from multiple internet sources

such as Kaggle, as well as data sets from the applicant's loan application.

Significant evaluation measures including

Plagiarism detected: **1,62%** https://www.researchgate.net/publication/33908... id: **4**

Accuracy, Recall, Precision, F1-Score,

and the ROC analysis area has been done and shown in the results section. It is

found that Extra Trees Classifier and Random Forest has Accuracy of Using

predictive modelling, this study presents an effective basis for loan credit approval

in order to detect vulnerable consumers from a large number of loan applications.

Keywords:

Loan Default
Data Science

Data Mining
Machine Learning

Correspondence:

Telephon: +91 9560291169

E-mail: mayankan@gmail.com

1. Introduction (Heading 1) (bold, 11 pt)

The banking industry's credit lending sector has seen tremendous growth and fierce competition from

a slew of new credit start-ups. At the same time, the rise in loan applications and consumption has

resulted in an increase in bad credit losses. Credit loans are loans issued by banks or financial

institutions to people or customers that are repayable at a certain period with or without

Plagiarism detected: **0,94%** https://www.researchgate.net/publication/33908... + 4 resources! id: **5**

interest.

Credit loans are commonly used for a variety of purposes, including personal use, educational

purposes, medical purposes, travel, and business. Financial institutions must design effective models

that can capture the information in

the existing data and produce robust prediction models that can

assist minimise the likelihood of bad credit as a result of the increase in loan applications and the

rapidly increasing competition. Financial institutions can gain insights on applicants' habits,

Plagiarism detected: **2,07%** https://www.researchgate.net/publication/33908... + 3 resources! id: **6**

consumption patterns, default predictors, and attributes using a variety of sophisticated predictive

modelling techniques. Numerous studies have been performed in order to determine the key

characteristics that influence loan repayment; these studies are significant since they assist banks

optimise profit. According to Manjeet et al (2018), there are seven sorts of characteristics that may

influence consumer loan default: the customer's annual income, debt-income ratio, occupation, home

ownership, work duration, and whether or not

the customer has a savings/checking account. Chang,

Cow, and Liu (2002) also suggested that an applicant's personal attributes, such as age and attitude,

may influence borrowers' risk behaviour. Borrower's age, location, resident/work duration, phone owner, monthly income, loan term, whether or not applicant works in the

Plagiarism detected: **0,65%** https://mail.irjet.net/archives/V5/i4/IRJET-V5I4... id: **7**

public sector, house

ownership, and loan numbers are among the important characteristics that may impact loan default,

according to Steenackers and Goovaerts (1989

). Ali Bangherpour (2015) found that loan age was the

Plagiarism detected: **0,17%** https://www.wsj.com/articles/SB100014240527... id: **8**

most important factor in predicting loan

default when using a large dataset from 2001 to 2006, while

market loan-

Plagiarism detected: **1,08%** https://www.researchgate.net/publication/33908... id: **9**

to-value was the most effective factor for mortgage loan applications. There is a need to

develop strong and effective machine learning models that can help capture key trends in credit data

in addition to identifying factors that may

influence loan default. The model chosen is critical since

it determines the accuracy, precision, and efficiency of a prediction system. For loan default

Plagiarism detected: **0,6%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **10**

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

2
http://dx.doi.org/10.
35671/jcse

prediction, a variety of models have been employed, and while there is no one best model, certain

models clearly outperform others.

Plagiarism detected: **2,33%** https://www.researchgate.net/publication/33908... + 10 resources! id: **11**

In this paper, we use all the classification machine learning

algorithms present today to study and analyze bank loan dataset and suggest some of the important

factors/variables that may influence loan repayment. We also present evaluation statistics (Accuracy,

Precision, Recall, Confusion Matrix, F1-Score ad ROC area) of our model. The remainder of the

paper is organized as follows: section 2 provides a background and overview of earlier research on

bank loan prediction. Section 3 explains the methodology of the data. In Section

4 we discuss our

Analysis and Results and

Plagiarism detected: **1,28%** https://www.researchgate.net/publication/34860... + 3 resources! id: **12**

section 5 concludes with future scope.

2. Literature Review

This section summarises some of the previous work on developing ML and DL models using various

algorithms to improve the loan prediction process and assist banking authorities and financial firms

in selecting an eligible candidate with

a low credit risk.

The subject of loan prediction is hotly debated in the banking and financial industries. In today's

competitive financial environment,

Plagiarism detected: **3,6%** https://www.researchgate.net/publication/34860... + 4 resources! id: **13**

credit scoring has become a critical tool for the same.

Furthermore, this area has gained greater attention and academic interest as a result of recent

advances in data science and some important discoveries in the field of artificial intelligence. It has

drawn more attention in recent years to research on loan prediction and credit risk assessment. Thanks

to the sudden high demand for loans, there is a huge increase in need for additional improvements in

credit scoring and loan prediction models. A variety of methodologies have been used to give credit

scores to individuals, and much research has been done on the subject throughout the years. Unlike

in the past, when experts were hired and models relied on professional judgments to assess an

individual's creditworthiness, the focus

now is on an automated means of performing the same thing.

Researchers and banking regulators have been focusing on using

Plagiarism detected: **19,1%** http://icsejournal.com/index.php/JCSE/article/d... + 15 resources! id: **14**

machine learning algorithms and

neural networks for credit scoring and risk assessment in recent years. Many notable results have

been reached in this area, which might be used as stepping stones for further study and investigations.

The Random Forest Classification Algorithm was adopted by Lin Zhu et al. in paper [4] and Nazeeh

Ghatasheh in paper [5] to construct a model for loan default prediction. Paper [4] concluded that

random forest has much better accuracy (98%) than other algorithms like logistic regression (73%),

decision trees (95%), and support vector machines (75%). The results of the paper [5] concluded that

the random forest algorithm is one of the best options for credit risk prediction. Paper [5] also talked

about the advantages of the algorithms, which are the competitive classification accuracy and

simplicity.

Paper [6] reviewed many methods available like logistic regression, k- nearest neighbours, random forest, neural networks, support vector machines, stochastic gradient boosting, Naive Bayes, etc. and

concluded that it is nearly impossible to declare one best method of all.

Nikhil Madane and Siddharth Nanda in paper [7] reviewed credit scoring of mortgage loans and

made the following conclusions:

Credit applications that do not pass certain requirements are often not accepted because the

probability of them not paying back is high.

Low-income applicants are more likely to get approval, and they are more likely to pay back their

loans in time.

Pidikiti Supriya et al. [8] used Decision Trees as a machine learning tool to implement their model.

They started their analysis with data cleaning pre-processing, missing value imputation, then

exploratory data analysis, and finally model building and evaluation. The authors on a public test set

managed to achieve the best accuracy of 81%.

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

3
http://dx.doi.org/10.35671/jcse

The conducted tests using the C4.5 algorithm in decision trees in paper [9] showed that the maximum

precision value achieved was 78.08% with data partition of 90:10 and the biggest recall value was

96.4% with data partition of 80:20. Therefore partition of 80:20 was concluded to be the best due to

its highest accuracy and high recall value.

The authors in paper [12] selected 4 different models:

M1: Logistic Regression model

M2: Random Forest model

M3: Gradient Boosting model

D1-D4: Multilayer Neural Network models (deep learning)

And using these models they showed that data quality check is important, i.e., data analysis and

cleaning before modelling to omit redundant variables. The paper also concluded that the choice of

features and the algorithm are two major aspects when deciding whether to give an individual a loan

or not.

In paper [13], Aboobyda Jafar Hamid, and Tarig Mohammed Ahmed used Data Mining to build a model for classifying loan risk. They used three algorithms for it:

J48

Bayes Net

Naive Bayes

The paper concluded that J48 was the best algorithm for the purpose because of its high accuracy

(78.3784%) and low mean absolute error (0.3448).

Aditi Kacheria et al. [14] used the Naive Bayesian algorithm for their model. And to improve the

classification accuracy, they used the k-NN and binning algorithms to improve the quality of the

data. K- NN was used to deal with the missing values and the binning algorithm was used to remove

the anomalies from the data set.

Martin Vojtek and Evzen Kocenda concluded that most local banks in the Czech Republic and

Slovakia are using logit method-based models in paper [15]. Other methods like CART or neural

networks are primarily used as support tools in the variable selection process or the process of model

quality evaluation. The authors also concluded that the k-NN method is not used at all or is very

rarely used.

Yu Li in paper [2] did a comprehensive study comparing the XGBoost algorithm's performance with

the performance of logistic regression. The paper concluded that the model discrimination and model

stability of the XGBoost model was substantially higher than that of the logistic regression model.

Plagiarism detected: **0,57%** https://shsfeapi1.pdc-gate2.com/get_doc.php?... + 6 resources! id: **15**

Finally, complete content and organizational editing before formatting. Please take note of the

following items when proofreading spelling and grammar:

3. Methodology

The dataset represents default payment data of a bank. The dataset has 850 individual records. Many

preprocessing techniques were applied such as cleaning, data integration, data formatting, data

normalization etc. on the said dataset to bring out a well formed dataset. Then, we have applied

different classification

Plagiarism detected: **0,2%** https://www.irjet.net/archives/V8/i6/IRJET-V8I6... id: **16**

algorithms such as Logistic Regression, Decision Tree,

K-Nearest Neighbors,

SVM, Random Forest, and different types of Ensemble Boosting techniques to find out the accuracy

of our predictive model. A dichotomous dependent variable default payment (Yes = 1, No = 0) has been employed in this study. Then we have compared our classification results with the destination

value of the dataset. This research was implemented using Python on the Jupyter Kernel on a local

Plagiarism detected: **0,6%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **17**

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

4
http://dx.doi.org/10.
35671/jcse

machine. This research includes 8 important explanatory variables. The details of these variables are

as follows:

4. X1: Age

5. X2: Educational Background Category

6. X3 : Employment Status(or Years of Experience)

7. X4 : Address – Demographic Area converted to Numeric Equivalent

8. X5 : Income

9. X6 : Debt Income

10. X7 : Credit to Debt Ratio

11. X8 : Other Debt

12. Results and Discussion

Based

Plagiarism detected: **0,23%** https://www.researchgate.net/publication/33908... id: **18**

on the pre-processed data sets, we use the

15 Classification algorithms which included various

Classifiers, Boosting, Logistic Regression, SVM and Naive Bayes implemented in python programming

language. Out of which, we will show the results of top 5 algorithms which showed optimistic results in the

scenario. We trained our classifier

Plagiarism detected: **2,3%** https://www.researchgate.net/publication/33908... + 3 resources! id: **19**

on the clean data set using the feature default as target. For evaluation, we

use five metrics; Accuracy, F1-Score, Recall, Precision and ROC area.

12.1 Confusion Matrix

The confusion matrix is an important 2-dimensional matrix that contains information about the actual classes

and the predicted classes of a classifier. In the loan application data used in this paper, the number 0 represents

the loan default category, and 1 represents the normal category. Table 1-5. shows the calculated confusion matrix of

top 5 algorithms with the data.

Predicted Class

Actual Class 0 – Not Default 1 - Default

0 – Not Default 81 21

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

5
http://dx.doi.org/10.
35671/jcse

1 - Default 5 98

Table 1. Confusion Matrix for Extra Trees Classifier

Predicted Class

Actual Class 0 – Not Default 1 - Default

0 – Not

Default

81 21

1 - Default 8 95

Table 2. Confusion Matrix for Random Forest Classifier

Predicted Class

Actual Class 0 – Not Default 1 - Default

0 – Not

Default

77 25

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

6
http://dx.doi.org/10.
35671/jcse

1 - Default 9 94

Table 3. Confusion Matrix for CatBoost Classifier

Predicted Class

Actual Class 0 – Not Default 1 - Default

0 – Not

Default

79 23

1 - Default 10 93

Table 4. Confusion Matrix for Extreme Gradient Boosting

Predicted Class

Actual Class 0 – Not Default 1 - Default

0 – Not

Default

77 25

Plagiarism detected: **0,6%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **22**

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

7
http://dx.doi.org/10.
35671/jcse

1 - Default 13 90

Table 5. Confusion Matrix for Light Gradient Boosting Machine Classifier

12.2. Accuracy

Accuracy

Plagiarism detected: **1,36%** https://www.researchgate.net/publication/33908... + 21 resources! id: **23**

measures the proportion of correctly classified predictions; it is defined by the formula:

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives

and FN is the number of false negatives.

Algorithm Accuracy

Extra Trees Classifier 86.17

Random Forest Classifier 85.55

CatBoost Classifier 84.92

Light Gradient Boosting 84.49

Extreme Gradient Boosting 83.87

Table 6. Accuracy of top 5 algorithms in descending order

Journal

Plagiarism detected: **0,57%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **24**

of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

8
http://dx.doi.org/10.
35671/jcse

Fig.1. Comparison of Accuracy of all the classification models performed on the data

12.3 Recall

Recall

Plagiarism detected: **0,71%** https://www.researchgate.net/publication/33908... id: **25**

measures the ability of the classifier to predict correctly instances of a certain class; it is

also called the TPR (true positive rate):

Recall

= TP/(TP+FN)

Algorithm Recall

Extra Trees Classifier 88.20

CatBoost Classifier 87.35

Random Forest Classifier 85.59

Light Gradient Boosting 84.68

Extreme Gradient Boosting 82.91

Plagiarism detected: **0,6%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **26**

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

9
http://dx.doi.org/10.
35671/jcse

Table 7. Recall of top 5 algorithms in descending order

Fig.2. Comparison of Recall of all the classification models performed on the data

12.4 Precision

Precision

Plagiarism detected: **0,43%** https://www.researchgate.net/publication/33908... id: **27**

measures the proportion of predictions made by the classifier as positive that are

actually positive:

Precision = TP/(TP+FP)

Algorithm Precision

Random Forest Classifier 85.06

Extra Trees Classifier 84.27

Extreme Gradient Boosting 83.93

Plagiarism detected: **0,6%** http://icsejournal.com/index.php/JCSE/article/d... + 2 resources! id: **28**

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

10
http://dx.doi.org/10.
35671/jcse

Light Gradient Boosting 83.46

CatBoost Classifier 82.57

Table 8. Precision of top 5 algorithms in descending order

Fig.3. Comparison of Precision of all the classification models performed on the data

12.

5 F1-Score

F1-score is the harmonic average of precision and recall:

F1-Score = 2*Precision*Recall/ (Precision+Recall
)

Algorithm F1-Score

Extra Trees Classifier 85.97

Random Forest Classifier 85.03

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

11
http://dx.doi.org/10.
35671/jcse

CatBoost Classifier 84.57

Light Gradient Boosting 83.83

Extreme Gradient Boosting 83.07

Table 9. F1-Score of top 5 algorithms in descending order

Fig.4. Comparison of F1-Score of all the classification models performed on the data

12.6 ROC

The

ROC (receiver operating characteristics) curve is a visualization technique for showing a

classifier’s performance. It represents the sensitivity and specificity of the classifier. The ROC

curve is a two-dimensional curve with the FPR (false positive rate) as the X-axis and the TPR (true

positive rate) as the Y-axis. The ranges of the ROC curve runs from (0, 0) to (1, 1). To compare

models, we calculate the area under the ROC curve (AUC). The larger the AUC, the better the

model.

Figure 5-9 shows ROC-AUC curve for top 5 algorithms.

Plagiarism detected: **2,78%** http://icsejournal.com/index.php/JCSE/article/d... + 7 resources! id: **32**

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

12
http://dx.doi.org/10.35671/jcse

Fig.5. ROC Curve for Extra Trees Classifier

Fig.6. ROC Curve for Random Forest Classifier

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

13
http://dx.doi.org/10.35671/jcse

Fig.7. ROC Curve for Cat Boost Classifier

Fig.8. ROC Curve for Extreme Gradient Boosting Classifier

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

14
http://dx.doi.org/10.
35671/jcse

Fig.9. ROC Curve for Light Gradient Boosting Machine Classifier

12.7 Feature Importance

One

Plagiarism detected: **0,77%** https://www.researchgate.net/publication/33908... id: **33**

of the important results of this study is to determine the important features that help the

classifier to correctly predict loan default. This helps in business intelligence

and decision making.

Fig.10-14 shows the most important features as a plot for top 5 algorithms. The employment

experience and debt income

Plagiarism detected: **0,43%** https://www.researchgate.net/publication/33908... id: **34**

of the customer are the two most important features from the result

of this plot.

Fig.10.

Plagiarism detected: **2,41%** http://icsejournal.com/index.php/JCSE/article/d... + 8 resources! id: **35**

Feature Importance Plot for Extra Trees Classifier

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

15
http://dx.doi.org/10.35671/jcse

Fig.11. Feature Importance Plot for Random Forest Classifier

Fig.12. Feature Importance Plot for Cat Boost Classifier

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

16
http://dx.doi.org/10.35671/jcse

Fig.13. Feature Importance Plot for Extreme Gradient Boosting Classifier

Fig.14. Feature Importance Plot for

Plagiarism detected: **3,46%** https://www.researchgate.net/publication/33908... + 3 resources! id: **36**

Light Gradient Boosting Machine Classifier

13. Conclusion

In this paper, we have successfully used the various classification algorithms for bank loan default

prediction. The task was to predict if a loan applicant will default on loan payment or not. The

analysis was implemented in the python programming language, and performance metrics like

accuracy, recall, precision, f1-score were calculated. From the analysis, we found out that the most

important features used by our model for predicting if a customer would default in payment or not

depends heavily on the employment or job experience in years and debt income of the customer. This

paper provides an effective basis for loan credit approval in order to identify risky customers from a

large number of loan

applicants using predictive modeling.

Acknowledgment

I am grateful to The Research World(ThRews) for contributing in my research over this paper.

Plagiarism detected: **8,2%** http://icsejournal.com/index.php/JCSE/article/d... + 15 resources! id: **37**

References

[1] Aslam U, Aziz H I T, Sohail A and Batcha N K 2019 An empirical study on loan default prediction models

Journal of Computational and Theoretical Nanoscience 16 pp 3483–8

Journal of Computer Science an Engineering (JCSE)
Vol. X, No. X, February 2020, pp. xx-xx

e-ISSN 2721-0251

17
http://dx.doi.org/10.35671/jcse

[2] Li Y 2019 Credit risk prediction based on machine learning methods The 14th Int. Conf. on Computer

Science & Education (ICCSE) pp 1011–3

[3] Ahmed M S I and Rajaleximi P R 2019 An empirical study on credit scoring and credit

scorecard for

financial institutions Int. Journal of Advanced Research in Computer Engineering & Technol. (IJARCET)

8 275–9

[4] Zhu L, Qiu D, Ergu D, Ying C and Liu K 2019 A study on predicting loan default based on the random

forest algorithm The 7th Int. Conf. on Information Technol. and Quantitative Management (ITQM) 162

pp 503–13

[5] Ghatasheh N 2014 Business analytics using random forest trees for credit risk prediction: a comparison

study Int. Journal of Advanced Science and Technol. 72 pp 19–30

[6] Breeden J L 2020 A survey of machine learning in credit risk

[7] Madane N and Nanda S 2019 Loan prediction analysis using decision tree Journal of The Gujarat Research

Society 21 p p 214–21

[8] Supriya P, Pavani M, Saisushma N, Kumari N V and Vikas K 2019 Loan prediction by using machine

learning models Int. Journal of Engineering and Techniques 5 pp144–8

[9] Amin R K, Indwiarti and Sibaroni Y 2015 Implementation of decision tree using C4.5 algorithm in decision

making of loan application by debtor (case study: bank pasar of yogyakarta special region) The 3rd Int.

Conf. on Information and Communication Technol. (ICoICT) pp 75–80

1. Introduction (Heading 1) (bold, 11 pt)
2. Literature Review
3. Methodology
12. Results and Discussion
13. Conclusion

Plagiarism detected: **1,05%** https://www.researchgate.net/publication/33908... + 2 resources! id: **38**

In this paper, we have successfully used the various classification algorithms for bank loan default prediction. The task was to predict if a loan applicant will default on loan payment or not. The analysis was implemented in

the python programming la...
Acknowledgment
References