Comparative Analysis of Data mining Methods to Analyze Personal Loans using Decision Tree and Naïve Bayes Classifier

Menuka Maharjan
Assistant Professor, Computer Engineering Department, Nepal Engineering College, Bhaktapur, Kathmandu, Nepal
Menuka[at]nec.edu.np

Abstract: The data mining classification techniques and analysis can enable banks to move precisely classify consumers into various credit risk group. Knowing what risk group a consumer falls into would allows a bank to fine tune its lending policies by recognizing high risk groups of consumers to whom loans should not be issued, and identifying safer loans that should be issued on terms commensurate with the risk of default. So research en for classification and prediction of loan grants. The attributes are determined that have greatest effect in the loan grants. For this purpose C4.5, CART and Naïve Bayes are compared and analyzed in this research. This concludes that a bank should not only target the rich customers for granting loan but it should assess the other attributes of a customer as well which play a very important part in credit granting decisions and predicting the loan defaulters.

Keywords: C4.5, CART, Naïve Bayes, Type II error

1. Introduction

The decision-making of accepting or rejecting a client’s credit by banks is commonly executed via judgmental techniques and credit scoring models. Most banks and financial institutions use the judgmental approach which is based on the 3C’s, 4C’s or 5C’s which are character, capital, collateral, capacity and condition. However, to improve assessment of credit applicants, banks can use credit scoring or predictive models to classify the applicants[1]. A bank loans officer needs analysis of his/her data in order to learn which loan applicants are safe" and which are risky” for the bank. To understand that information, classification is a form of data analysis that can be used to extract models describing important data classes or to predict future data trends. Several classification techniques have been proposed over the years e.g., neural networks, genetic algorithms, Naïve Bayesian approach, decision trees, nearest-neighbour method etc [2].

The classification is dependent on characteristics of the borrower (such as age, education level, occupation, marital status and income), the repayment performance on previous loans and the type of loan. In this study, my attention is restricted to C4.5, CART and Naïve Bayes classification considering its advantages like efficiency with respect to time accuracy data, etc and analyze different parameters (age, income, credit rating job etc.) those influence the loan grants.

2. Methodology

Classification is learning a function that maps an item into one of a set of predefined classes. It is the type of data analysis that can be used to extract models to describe important data classes or to predict future data trends. The classification process consists of two phases; the first phase is learning process, the training data will be analyzed by the classification algorithm. The learned model or classifier is represented in the form of classification rules. Next, the second phase is classification where the test data are used to estimate the accuracy of the Classification model or classifier. If the accuracy is considered acceptable, the rules can be applied to the classification of new data [3]. This section is about the framework for comparing the performance of the classification algorithms of decision trees: CART, C4.5 and Naïve Bayes classification with the role play of the attributes in them to predict loan grants data is taken from data sets[9]. It consists of 1000 data, among which 60% are used for training and remaining 40% are utilized for testing purpose that are work

Loan Prediction using C4.5
C4.5 is an algorithm used to generate a decision tree developed by Ross Quinlan. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier [7].

C4.5 algorithm:
For the classification the total number of good and bad in loan grants is found out from the data set. Information gain is calculated for the whole dataset i.e. Info (D) and then for each attribute the normalized information gain is calculated individually i.e. Info(D). Gain (A) is calculated subtracting...
the information gain and information gain of individual attribute for that particular attribute.

\[ IG(A) = H(S) - \sum_{t \in T} P(t)H(t) \]

Where,

\[ H(S) - \text{Entropy of set } S, \text{ and } H(S) = -\sum_{x \in X} p(x) \log_2 p(x) \]

T- The subsets created from splitting set \( S \) by attribute \( A \) such that \( P(T) \)- The proportion of the number of elements in \( T \) to the number of elements in set \( S \)

\[ H(t) = -\sum_{i} p_i \log_2 p_i \]

The process is repeated for all the attributes and selected the highest normalized information gain for a decision node. The features of the attribute may be nominal or categorical like if age is attribute with its category like

- age 0-18
- age 19-30
- age 31-40
- age 41-50
- age 51 above

The process is repeated for each of the resulting attributes and made the table for only each features in both sides. Recursion is done until leaf node is not found.

### Loan Prediction Using Cart

Classification and regression trees (CART) is a non-parametric decision tree learning technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numerical, respectively.

#### CART algorithm:

It will search for all possible variables and all possible values in order to find the best split – the question that splits the data into two parts with maximum homogeneity. The process is then repeated for each of the resulting data fragments which use impurity functions like Gini splitting index and Towing splitting index [6]. Here Gini splitting rule (or Gini index) is used for the loan prediction. It uses the following impurity function:

### Splitting Criteria:

Gini index is measured to find the impurity of \( D \), a data partition or set of training tuples, as

\[ \text{Gini}(D) = 1 - \sum_{i=1}^{m} p_i^2 \]

where \( p_i \) is the probability that a tuple in \( D \) belongs to class \( Ci \). The sum is computed over \( m \) classes.

Here, splitting is compulsory binary so, data \( D \) is splitted into \( D1 \) and \( D2 \). The partitioning is done as follows

\[ \text{Gini}_X(D) = \frac{\text{Gini}(D1)}{|D1|} + \frac{|D2|}{|D|} \text{Gini}(D2) \]

The reduction in impurity that would be incurred by a binary split on a discrete or continuous-valued attribute \( A \) is

\[ \Delta \text{Gini}(A) = \text{Gini}(D) \cdot \text{Gini}_X(D) \]

The process is repeated for each attributes and decision for the rootnode is made for the lowest valued Gini \( D \) [6]. Again if the attribute purpose is chosen as the root node then its features like personal loan and business loan is splitting binary and made the table for only each features in both sides. Recursion is done until leaf node is found.

### Loan Prediction using Naive Bayes

A Naïve Bay’s classifier estimates the class-conditional probability by assuming that the attributes are conditionally independent, given the class label \( y \). The conditional independence assumption can be formally stated as follows:

\[ P(X|y) = \prod_{i=1}^{d} P(X_i|y) \]

Where each attribute set \( X = \{X_1, X_2, \ldots, X_d\} \) consists of \( d \) attributes. [8 ]’

#### Algorithm

1) From data set \( D \) associated class label \( n \) dimensional attribute vector \( X = (x_1, x_2, x_3, \ldots, x_n) \), depiction \( n \) measurement made on the tuple from \( n \) attributes. \( A1, A2, A3 \ldots An \)

2) Suppose we have \( m \) classes \( c1, c2, \ldots, km \) giving tuple \( X \), classifier will predict \( X \) belongs to the highest posterior probability, condition on \( X \).

\[ X \in Ci \text{ if } P(Ci|X) > P(Cj|X) \text{ for } 1 \leq j \leq m, j | Ci \text{, for which } P(Ci) \text{ is maximized is called maximum posterior hypothesis; } \]

\[ P(Ci|X) = \frac{P(X|Ci)P(Ci)}{P(X)} \]

3) \( P(X) \) is constant for all classes maximize \( P(X|Ci)P(Ci) \).

\[ P(Ci) = \frac{|C|}{|D|} \]

4) Calculate \( P(X|Ci) \) is extremely expensive Naïve assumes class conditional independence is made.

\[ P(X|Ci) = \prod_{k=1}^{d} P(X_k|Ci) \]

\[ = P(X_1|Ci).P(X_2|Ci) \ldots P(X_d|Ci) \]

Where \( X_k \) is the value of attribute, \( A_k \) for \( X \).

If \( A \) is category

\[ P(X|A = c_i) = \frac{\# \text{of tuple of class}_C \in D \text{ that have value} X_k}{|C|D|} \]

#### 3. Results and Discussion

The German loan dataset consist of 1000 dataset 60% of data is used for the train set and 40% is used for the test set. Experiments for CART and C4.5 using German data set are summarized below:

| Table 1: C4.5 train |
|-------------------|
| Attributes | Confusion Matrix | Precision | Recall | F_Score | Accuracy | CCI | Time | No. of Leaf | Size of Tree |
| category1       | 398.25 67 110     | 85.5914 94.08983 | 89.63964 | 398.18333 508 | 0.2 | 52 | 75 |
| category2       | 404.19 54 123     | 88.20961 95.50827 | 91.71396 | 404.205 527 | 0.1 | 7 | 13 |
| category3       | 398.25 67 110     | 85.5914 94.08983 | 89.63964 | 398.18333 508 | 0.2 | 52 | 75 |
| category4       | 394.29 110 67     | 78.1746 93.14421 | 85.5914 394.11176 461 | 0 | 16 | 23 |
| category5       | 359.64 70 107     | 83.68298 84.86998 | 85.5914 359.17833 466 | 0.5 | 18 | 27 |
| category6       | 423 0 177 0       | 70.5 100 | 82.69795 | 423 | 0 | 1 | 1 |
Here out of 600 data are used for training in both the C4.5 and CART method. Category 1, 2,3,10 is better for correctly classified instances out of 600 data during train phase. The categories 4,6,7,8 shows that the false positive rate is large compared to other categories. Categories 6,8,11 shows all data are true positive so, there will be loss if the banks take true negative data as good one . The precision ,accuracy is higher for category 1,2,5,10 compared to other categories.

### Table 2: CART train

| Attributes | Confusion Matrix | Precision | Recall | F_score | Accuracy | CCI | Time | No of leaf | Size of tree |
|------------|------------------|-----------|--------|---------|----------|-----|------|------------|-------------|
| category1  | 389 34 91 86     | 91.962175 | 91.962175 | 389.14333 | 475 | 2.78 | 6 | 11 |
| category2  | 399 24 97 80     | 94.326241 | 94.326241 | 399.13333 | 479 | 1.61 | 7 | 13 |
| category3  | 389 34 91 86     | 91.962175 | 91.962175 | 389.14333 | 475 | 3.22 | 6 | 11 |
| category4  | 417 6 149 28     | 98.58156  | 98.58156  | 417.04667 | 445 | 0.42 | 3 | 5 |
| category5  | 411 12 111 66    | 97.16312  | 97.16312  | 411.11111 | 477 | 1.09 | 12 | 23 |
| category6  | 423 0 177 0      | 100      | 100      | 423      | 423 | 0.69 | 1 | 1 |
| category7  | 423 0 177 0      | 100      | 100      | 423      | 423 | 1.05 | 1 | 1 |
| category8  | 423 0 177 0      | 100      | 100      | 423      | 423 | 0.66 | 1 | 1 |
| category9  | 398 25 90 87     | 94.089835 | 94.089835 | 398.145  | 485 | 1.39 | 9 | 17 |
| category10 | 399 24 97 80     | 94.326241 | 94.326241 | 399.13333 | 479 | 1.59 | 7 | 13 |
| category11 | 423 0 177 0      | 100      | 100      | 423      | 423 | 0.44 | 1 | 1 |

From the above table correctly classified instance out of 600 instances is higher in categories 9,10,2 in the case of CART.in confusion matrix category 6, 7, 8, 11 shows the worst case as false positives are 177 and true positive values are 423 for all these categories.

### Table 3: Naive Bayes train

| Attributes | Confusion Matrix | Precision | Recall | F_score | Accuracy | CCI | Time |
|------------|------------------|-----------|--------|---------|----------|-----|------|
| category1  | 375 48 82 95     | 82.05689 | 88.65248 | 85.22727 | 78.33333 | 470 | 0.02 |
| category2  | 383 40 94 83     | 80.2935  | 90.54374 | 85.11111 | 77.66667 | 466 | 0    |
| category3  | 375 48 82 95     | 82.05689 | 88.65248 | 85.22727 | 78.33333 | 470 | 0.02 |
| category4  | 393 30 117 60    | 77.05882 | 92.9078  | 84.24437 | 75.5    | 453 | 0.02 |
| category5  | 368 55 117 60    | 75.87629 | 86.99764 | 81.05727 | 71.33333 | 420 | 0.03 |
| category6  | 414 9 164 13     | 71.6263  | 97.87234 | 82.71728 | 71.16667 | 427 | 0.03 |
| category7  | 388 35 128 49    | 75.1938  | 91.72577 | 82.64111 | 72.83333 | 437 | 0.02 |
| category8  | 400 23 157 20    | 71.81329 | 94.56265 | 81.63265 | 70   | 420 | 0.02 |
| category9  | 379 44 88 88     | 80.98291 | 89.59811 | 85.07295 | 77.83333 | 467 | 0.02 |
| category10 | 379 44 94 83     | 80.12685 | 89.59811 | 84.59821 | 77   | 467 | 0.02 |
| category11 | 423 0 177 0      | 70.5    | 100      | 82.69795 | 70.5 | 423 | 0.02 |

Here, out of 600 data sets the higher correctly classified instances is high in categories 1,3,i.e 470 and lower in category 5,6,8 i.e. 420, 414 and 400 respectively. The accuracy is high in categories, 1,3 i.e. it is 78.3333% and lower in category 8 . FP rate is rate is high in categories 4,5,6,7,8,11 and lower in 1,2,3,9,10.

### Table 4: C4.5 test

| Attributes | Confusion Matrix | Precision | Recall | F_score | Accuracy | CCI | Time |
|------------|------------------|-----------|--------|---------|----------|-----|------|
| category1  | 266 11 45 78     | 85.5305   | 96.029 | 90.47619 | 86    | 344 | 1.89 |
| category2  | 264 13 56 67     | 82.5      | 95.307 | 88.44221 | 82.75 | 331 | 0.75 |
| category3  | 265 12 101 22    | 72.4044   | 95.668 | 82.42613 | 71.75 | 287 | 0.25 |
| category4  | 236 41 65 58     | 78.4053   | 85.199 | 81.6609  | 73.9  | 294 | 0.22 |
| category5  | 269 8 96 27      | 73.6986   | 97.112 | 83.8062  | 74   | 296 | 0.28 |
| category6  | 277 0 123 0      | 69.25     | 100    | 81.83161 | 69.25 | 277 | 0.45 |
| category7  | 277 0 123 0      | 69.25     | 100    | 81.83161 | 69.25 | 277 | 0.45 |
| category8  | 277 0 123 0      | 69.25     | 100    | 81.83161 | 69.25 | 277 | 0.45 |
| category9  | 257 20 62 61     | 80.5643   | 92.78  | 86.24161 | 79.5  | 318 | 0.5 |
| category10 | 267 10 62 61     | 81.155    | 96.39  | 88.11881 | 82   | 328 | 0.86 |
| category11 | 277 0 123 0      | 69.25     | 100    | 81.83161 | 69.25 | 277 | 0.11 |

Volume 10 Issue 12, December 2021

www.ijsr.net
Licensed Under Creative Commons Attribution CC BY

Paper ID: SR211207105339
DOI: 10.21275/SR211207105339
423
Here out of 400 data are used for testing in both the C4.5 and CART method. The categories 6, 7, 8, 11 are not good for attributes for classification as there precision 69% only. The category 8 shows the best as its precision and accuracy is 95%.

### Table 5: CART test

| Attributes | Confusion Matrix | Precision | Recall | F_score | Accuracy | CCI | Time |
|------------|------------------|-----------|--------|---------|----------|-----|------|
|            | TP | FN | FP | TN |               |         |      |      |
| category 1 | 244 | 33 | 48 | 75 | 83.5616 | 88.08664 | 85.7645 | 79.75 | 319 | 1.89 |
| category 2 | 259 | 18 | 62 | 61 | 80.6854 | 93.50181 | 86.62207 | 80 | 320 | 0.75 |
| category 3 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0.25 |
| category 4 | 238 | 39 | 56 | 67 | 80.9524 | 85.92058 | 83.36252 | 76.25 | 305 | 0.22 |
| category 5 | 260 | 17 | 80 | 43 | 76.4706 | 93.86282 | 84.27877 | 75.75 | 303 | 0.28 |
| category 6 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0.45 |
| category 7 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0.45 |
| category 8 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0.22 |
| category 9 | 246 | 31 | 55 | 68 | 81.7276 | 88.80866 | 85.12111 | 78.5 | 314 | 0.5 |
| category 10 | 259 | 18 | 62 | 61 | 80.6854 | 93.50181 | 86.62207 | 80 | 320 | 0.86 |
| category 11 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0.11 |

In the above table category 3, 6, 7, 8, 11 shows higher false positive values so these are the worst attributes while category 2, 5, 10 are the best categories. Category 5 consists of only 4 attributes.

### Table 6: Comparison of accuracy for C4.5, CART and Naïve Bayes

| Attributes | Confusion Matrix | Precision | Recall | F_score | Accuracy | CCI | Time |
|------------|------------------|-----------|--------|---------|----------|-----|------|
|            | TP | FN | FP | TN |               |         |      |      |
| category 1 | 237 | 40 | 49 | 74 | 82.8613 | 85.55957 | 84.19183 | 77.75 | 311 | 0.06 |
| category 2 | 240 | 37 | 61 | 62 | 79.73422 | 86.6426 | 83.04498 | 75.5 | 302 | 0.03 |
| category 3 | 248 | 29 | 97 | 26 | 71.88406 | 89.53069 | 79.74277 | 68.5 | 274 | 0 |
| category 4 | 246 | 31 | 81 | 42 | 75.22936 | 88.80866 | 81.45695 | 72 | 288 | 0 |
| category 5 | 254 | 23 | 88 | 35 | 74.26901 | 91.69675 | 82.06785 | 72.25 | 289 | 0 |
| category 6 | 266 | 11 | 111 | 12 | 70.55703 | 96.02888 | 81.34557 | 69.5 | 278 | 0 |
| category 7 | 254 | 23 | 94 | 29 | 72.98851 | 91.69675 | 81.28 | 70.75 | 283 | 0 |
| category 8 | 264 | 13 | 111 | 12 | 70.4 | 95.30686 | 80.9816 | 69 | 276 | 0 |
| category 9 | 240 | 37 | 63 | 60 | 79.20792 | 86.6426 | 82.75862 | 75 | 300 | 0 |
| category 10 | 238 | 39 | 59 | 64 | 80.13468 | 85.92058 | 82.92683 | 75.5 | 302 | 0.02 |
| category 11 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0 |

Here, from the above figure we can see the accuracy is higher in C4.5 for the category in comparison to CART and Naïve Bayes. The category 4 performed good because it contains only four attributes and its accuracy is higher. The category 6, 7, 11 are the worst ones and accuracy is same in C4.5, CART and Naïve Bayes.

### Table 7: Comparison of correctly classified instances for C4.5, CART and Naïve Bayes

| Attributes | Confusion Matrix | Precision | Recall | F_score | Accuracy | CCI | Time |
|------------|------------------|-----------|--------|---------|----------|-----|------|
|            | TP | FN | FP | TN |               |         |      |      |
| category 1 | 237 | 40 | 49 | 74 | 82.8613 | 85.55957 | 84.19183 | 77.75 | 311 | 0.06 |
| category 2 | 240 | 37 | 61 | 62 | 79.73422 | 86.6426 | 83.04498 | 75.5 | 302 | 0.03 |
| category 3 | 248 | 29 | 97 | 26 | 71.88406 | 89.53069 | 79.74277 | 68.5 | 274 | 0 |
| category 4 | 246 | 31 | 81 | 42 | 75.22936 | 88.80866 | 81.45695 | 72 | 288 | 0 |
| category 5 | 254 | 23 | 88 | 35 | 74.26901 | 91.69675 | 82.06785 | 72.25 | 289 | 0 |
| category 6 | 266 | 11 | 111 | 12 | 70.55703 | 96.02888 | 81.34557 | 69.5 | 278 | 0 |
| category 7 | 254 | 23 | 94 | 29 | 72.98851 | 91.69675 | 81.28 | 70.75 | 283 | 0 |
| category 8 | 264 | 13 | 111 | 12 | 70.4 | 95.30686 | 80.9816 | 69 | 276 | 0 |
| category 9 | 240 | 37 | 63 | 60 | 79.20792 | 86.6426 | 82.75862 | 75 | 300 | 0 |
| category 10 | 238 | 39 | 59 | 64 | 80.13468 | 85.92058 | 82.92683 | 75.5 | 302 | 0.02 |
| category 11 | 277 | 0 | 123 | 0 | 69.25 | 100 | 81.83161 | 69.25 | 277 | 0 |

We can see the CCI using Naïve Bayes is remarkably higher compared to C4.5 and CART. The categories 1, 4, 9, 10 are good ones while category 5, 11 are the worst ones.

### Figure 2: Average value of correctly classified instance for C4.5, CART and Naïve Bayes

From the above figure, the accuracy is higher in C4.5 compared to classifier Naïve Bayes and CART.
The average precision is remarkably higher in C4.5 compared to CART and Naïve Bayes. The average precision of C4.5 is 78%, CART is 75.5% and Naïve Bayes is 75.1.

The average recall value is higher it is 96%, CART is 95% and Naïve Bayes is 90.80% Here C4.5 is better in comparison to CART and Naïve Bayes.

The average F_score is higher in C4.5 i.e. 86%, CART is 83.80% and that of Naïve Bayes is 82%. Therefore we can conclude that C4.5 is better.

The average accuracy is higher in comparison to the classifier C4.5 than CART and Naïve Bayes. The average accuracy for C4.5 is 77.5%, CART is 74% and Naïve Bayes is 72.1%.

Naïve Bayes predicted higher compared to CART and C4.5 in Correctly classified instances. The average precision, recall, F_score, accuracy is high in C4.5 compared to CART and Naïve Bayes.

The main focus is on the false positive value as it is the positive count for the bad customers that it the most risk factor for the loan prediction. The categories 1,2,4,8,9,10 contain the lower FP value in which in category 8, C4.5 has the lowest FP. The category 4 is also acceptable as it contains only 4 attributes in which FP is low. The categories 3, 6, 7, 11 are the worst one.
4. Conclusion and Discussion

If a customer with bad credit is misclassified as a customer with good credit then a bank will suffer. In this research three different classifiers, C4.5, CART and Naïve Bayes have been applied to predict loan grants and the attribute selection in them. More, financial institution is seeking better strategies through the help of credit scoring models. Therefore, it is concluded that categories 4, 8 is the best one and categories 3, 6, 11 are the worst as it counts false positive value is greater in all the C4.5, CART and Naïve Bayes testing. Among the classifier C4.5, CART and Naïve Bayes, C4.5 is the best classifier to predict loan.

References

[1] Pratik Gosar, Paras Kapadia, Niharika, Maheswori, Pramila Chawan K. Chopde, "A study of a Classification Based Credit Risk Analysis Algorithm," International Journal of Engineering and Advanced Technology, vol. 1, no. 2249-8958, p. 3, April 2012.
[2] Sarika Chaudary Sanjay Kumar Maliki, "Comparative Study of Decision Tree Algorithms For Data Analysis," International Journal of Research in Computer Engineering and Electronics, p. 8, June-2013.
[3] A.Yilmaz camurcu Serhat Ozekes, "Classification and Prediction in a Data Mining Application," Journal of Marmarafor Pure Applied Science, pp. 169-174, 2002.
[4] Yu.Zhong Xiao-Lin, "An Overview of Personal Credit Scoring:Techniques and Future Work," International Journal of Intelligence Science, pp. 181-189, August 2012.
[5] Daniela Schiopu Irina Ionita, "Usig Principal Component Analysis in Loan Granting," pp. 88-96, 2010.
[6] Bora Aktan Husey Incea, "A Comparison of data Mining Techniques For Credit Scoring In Banking:A Managerial Perspective," Journal of Business Economics and Management, pp. 233-240, march 2009. [1]
[7] Arun K.Pujari, Data Mining Techniques. Hyderabad, India: Universities press private limited.
[8] P.-N. T.-N. T. Pang-Ning Tan. [Online]. Available: https://wwwusers.cs.umn.edu/~kumar001/dmbook/sol.pdf.
[9] P. D. H. Hofmann. [Online]. Available http://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Data%29
[10] P. T. A. T. Kalyani R. Rawate, "Review on prediction system for bank loan credibility," Scientific Journal of Impact Factor (SJIF): 4.72, vol. 4, no. 12, 2017.
[11] P. C. Abhijit A. Sawant, "Comparison of Data Mining Techniques used for Financial," International Journal of Emerging Technology and Advanced Engineering, vol. 3, no. 6, p. 6, 2013.

Author Profile

Menuka Maharjan is Assistant Professor at Department of Computer Science and Engineering, Nepal Engineering College. She holds M.E. in Computer from Nepal College of Information Technology, Pokhara University. She has been in the teaching field since last 9 years. Her research interest includes data science and machine learning.