UNIVERSITI TEKNOLOGI MARA

CLASSIFICATION OF CREDIT CARD HOLDER BEHAVIOR USING K NEAREST NEIGHBOR ALGORITHM

AHMAD FARIS BIN RAHIMI

Thesis submitted in fulfilment of the requirement s for
Bachelor of Computer Sciences (Hons.) Faculty of
Computer and Mathematical Sciences

JANUARY 2017

ACKNOWLEDGEMENT

ABSTRACT

Credit card issuer is the authority that have to responsible of their customer behaviors. Therefore, a lot of customer behavior that be faced everyday. Percentage of credit card users are not able to repay the debts was increased year by year. Then, the debt be considered as the bad debt. There are a number of credit card issuers who are unable to bear the loss that lead them to bankrupt. Therefore, classification of the credit card holder behavior using K Nearest Neighbor be proposed. There are three objectives in the development of this proposed project. The first objective is to explore of k Nearest Neighbors technique for solving the classification of credit card holder behavior problem. The second one is to develop prototype for classification of credit cardholder behavior based on k Nearest Neighbors Algorithm. The third one is to evaluate the accuracy of the k Nearest Neighbors algorithm in the classification credit card holder behavior. The significance of the project is to help the credit card issuer in classifying the customer behavior in the payment pattern of credit card. The customer's behavior is to be classified into credible or non-credible user. This classification could detect customers with default payments earlier, so that actions could be taken by the credit card issuers. This could also avoid or reduce the loss of credit card issuer by bad debt of the user of credit card. This project consists of five phases which are preliminary study, data collection, system design, implementation, result and analysis. In the implementation phase, Bubble Sort, Euclidean Distance, 10-Fold Cross validation, and K Nearest Neighbor algorithm are developed. This application is using the data from a Taiwan bank which is obtained from the UCI data repository website. Testing of the application are using 10-Fold Cross Validation technique. 90% of the data be used as a training and another 10% be used for the testing part. The highest accuracy is 78.33% of the 10-fold Cross Validation method are found in data is 80% training and 20% testing when k is equal to 5. This indicates that the performance of KNN is acceptable and promising in this classification problem. Since KNN is the simplest form of artificial intelligence, future work could combine this algorithm with other classification algorithm. One of the suitable way to make this project better by extend it to predict the new applicant of credit card based on their behavior. The more the algorithm use in a project the better performance will be in result.

TABLE OF CONTENTS

CHAPTER THREE: METHODOLOGY