Credit appraisal for individuals and institutions generally measures the financial condition and the ability to pay back a loan i.e. to establish the credit worthiness of the individual. Factors that are taken into consideration are

* Age
* Income
* Number of dependents
* Continuity of employment
* Repayment capacity
* Previous loans
* Recurring payments such as Gas, Telephone, Electricity & water

While appraising the credit worthiness of an individual the 3 C’s of Credit must be kept in mind

* Character
* Capacity
* Collateral.

If any of these are missing, then the application should be scrutinized.

# Identity Verification

There are multiple options to verify identity with variable degrees of difficulty towards integration and implementation

1. **Malaysian Identity Card (MyKad):** Malaysians Identity Card number contains some information about the card holder. From the number submitted, we can get:

* Date of birth
* Place of birth
* Gender

The information can be extracted using “Malaysia Identity Card Detail Extractor” API found at <https://github.com/syahzul/mykad-detail-extractor>

1. **MyKad Command Set:** The MyKad Command Set is a low-level programming instruction to allow MyKad to read or write information into the MyKad. The information that can be accessed includes name, address, identity card number, photograph, fingerprint minutiae, driving license and passport. The applicant must be either a private limited company, sole proprietorship, partnership, government or private agency incorporated in Malaysia.
2. [**MyKad Reader**](http://www.salesapp.biz/mobile-mykad-reader-ocr/): EMAS™ supports various methods of photo ID reading and verification. The Mobile ID Reader and Optical Character Recognition (OCR) approach is used to help enterprises minimize and prevent identity fraud – which otherwise would be a costly damage to any business operations.
   1. The reader hardware provides the following information

* Photo
* Full name
* New MyKad (IC) no.
* Old IC no.
* Date of birth
* Place of birth
* Gender
* Race
* Religion
* Citizenship
* Address , Postcode, City ,State
* Date of issue
* East Malaysia origin
  1. The OCR Reader provides the following detail
* Full name
* New MyKad (IC) no.
* Address
* Gender
* Whole photo image of the MyKad front page

# Background Check and Identity verification

**Identity verification (AML and KYC):** Malaysia recently enacted new tougher anti-money laundering (AML) and know your customer (KYC) regulations. This will likely increase the likelihood of it gaining full membership into the Financial Action Task Force (FATF) and provide official recognition of Malaysia as a safe and attractive environment for offshore investors. Malaysian credit header data is now available as a data source for identity verification via GlobalGateway, [Trulioo’s AML/KYC](https://www.trulioo.com/blog/2015/08/05/malaysia-credit-data-now-available-in-globalgateway/) compliant electronic identity verification (eIDV) service. Trulioo’s AML/KYC compliant electronic identity verification (eIDV) service will help businesses adhere to Malaysia’s strict due diligence rules. GlobalGateway is used by e-commerce, finance, insurance, gaming and social media clients worldwide for all their compliance, risk mitigation and age verification needs. Trulio has flexible [integration](https://www.trulioo.com/developers/identity-verification/) options through their “Normalized API” or “XML Direct” methods.

# Credit Worthiness

Malaysia has a central database that contains credit related information. This information is held in a computerized database system known as the *Central Credit Reference Information System* (CCRIS). At present, the database system contains credit information on about 9 million borrowers in Malaysia. CCRIS automatically processes the credit data received from the financial institutions and synthesizes the information into credit reports, which can be made available to the financial institutions upon request. Access of the database is strictly governed and imposes severe penalty for unauthorized access. The information is only disseminated to financial institutions to evaluate a credit facility application.

**NOTE:** No free databases or free access to databases (via API) are available WRT to CTOS (FICO equivalent) or CCRIS.

The following credit bureaus are currently operating in Malaysia.

1. **Credit bureau Malaysia (CBM)**: The Bureau accesses and maintains both positive and negative credit information in order to provide a more comprehensive view of an SME’s or Consumer’s credit standing. Like all other credit bureau in the world, the Credit Bureau essentially collects credit information on borrowers from lending institutions and furnishes the credit information collected back to the institutions in the form of credit report via an on-line system known as Central Credit Reference Information System (CCRIS).

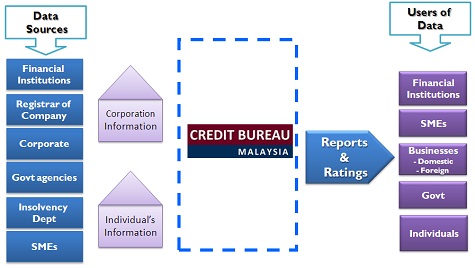
The Bureau obtains its credit information from the following data sources;

a) Central Credit Reference Information System (CCRIS) from Bank Negara Malaysia

b) Dishonored Cheques information (DCHEQS) from Bank Negara Malaysia

c) Non-bank Credit Information which is provided and updated by CBM subscribers on a voluntary basis

Based on the above information and the Bureau’s customized credit score-cards, a credit report with credit rating is produced. Subscribers of Credit Bureau Malaysia which includes Financial Institutions, Multinational Corporations, Government Agencies, Businesses and Credit Grantors will then use such reports to assist in their credit evaluation process.



Singapore in particular has several Credit Bureaus which are

1. **Credit Bureau (Singapore) Pvt. Ltd (CBS)** is Singapore’s most comprehensive consumer credit bureau that has full-industry uploads from all retail banks and major financial institutions. It is a joint venture between The Association of Banks in Singapore (ABS) and Infocredit Holdings Pvt. Ltd. CBS aggregates credit-related information amongst participating members and presents a more complete risk profile of a customer to credit providers. This helps credit providers to determine the likelihood of the customer repaying, thus enhancing their risk assessment capabilities.

In 2002, Infocredit Holdings established a second consumer credit bureau in Singapore known as CreditScan. Based on proven international standards, it is tailored to meet local requirements. The credit bureau will allow grantors of consumer credit in Singapore the access to relevant and accurate credit history information.

CreditScan is available to all industries (excluding members of the Association of Banks) that grant credit to individuals, to ensure that a wide and varied spectrum of information is accessible. The hierarchy of debt clearly shows that debtors default their payments with non-banks first before defaulting to the banks. This means credit grantors outside of the banks already experience increased risks as compared to the banks. Information from industries outside the banks will therefore provide an earlier indication of a consumer’s inability to meet future commitments.

Information will be offered on the basis of reciprocity. This will ensure that credit grantors are in a position to make fully informed decisions and that all current information on the consumer is available. CreditScan information will assist credit grantors in assessing new applications, reviewing existing relationships, collecting overdue accounts and preventing fraudulent applications.

1. **CONSUMER CREDIT BUREAU (DP CREDIT BUREAU)**
   1. **Consumer**: DP CREDIT BUREAU works in partnership with **Experian®,** their stakeholder, on a consumer scoring solution, designed to facilitate the risks and rewards of commercial and financial decisions. DP Info’s consumer scoring solution is currently being used by leading consumer lenders to manage their portfolio risks over time, providing them with proactive indicators on their consumer clients. DPCB helps Singapore’s banks and financial institutions make better lending decisions by providing comprehensive information on an applicant’s payment history and credit standing.
   2. **Commercial**: SME COMMERCIAL CREDIT BUREAU (DP Info Group) operates an SME Commercial Credit Bureau, a neutral entity that collates payment information of companies and businesses, developing a common information-sharing platform. In 2002, DP Info Group established a formal information-sharing platform to improve credit transparency in Singapore. The SME Commercial Credit Bureau was set up to benefit all corporations in Singapore by providing a neutral and formal member-based platform where the payment records of the members’ clients and suppliers are shared. This fosters more responsible payment behavior and enhances a company’s ability to manage and enhance their credit evaluation process over time.
2. **Singapore Commercial Credit Bureau:** In co-operation with the Association of Small and Medium Enterprises (ASME), **D&B Singapore** signed a memorandum of understanding on 16 October 2004 to establish the Republic’s first SME Credit Bureau (Contributor), a web-based platform specifically designed to enhance risk management capabilities. Effective 4 October 2010, SME Credit Bureau was rebranded to Singapore Commercial Credit Bureau (SCCB). An all-in-one credit assessment portal providing access to various tiers of information on commercial entities incorporated in Singapore with access to basic business registry information, litigation, detailed accounts on corporate payment patterns and topping up the information with their New Credit Risk Index (NCRI) as well as their Payment Index (PI).
3. [**FIS DATA REFERENCE**](http://www.fis.com.my/services_member2.htm) **SDN BHD (988222-H):** FDRmaintains databases to help businesses loan profitably and minimize credit risk by allowing their subscribers to access those records through their enquiry services. FDR has collated, accumulated and designed a credit data base system of information which, among others, includes the following items:

* Identification particulars of motor vehicles, goods and equipment which are subjects of hire purchase and leasing agreements;
* Particulars of hirers, lessees, borrowers and their guarantors;
* Problematic vehicles

The source of the database is mainly provided by their Subscribers. FDR is not a rating agency and does not attempt to rate, rank or give opinion on any individual or company. FDR entertains the following inquiries from its members

* Credit Enquiry
* Vehicle double financing enquiry
* Vehicle ownership claim registration
* Vehicle ownership claim discharge

All enquiries are furnished on an explicit (written consent) permission obtained from the subject.

1. [**Credit Reporting Agency Malaysia**](https://www.mycreditinfo.com.my/index.php/home/aboutRAMCI) **- RAM Credit Information (RAMCI):** RAMCI is Malaysia’s leading & Most Trusted Credit Bureau (Credit Reporting Agency). RMCI provides the following services

* Access to Payment Profiles & Behaviors
* Financials, ROC, ROB, Bankruptcy & Winding up Records
* Salary Deduction Information of members from 90% of all cooperatives
* Monitors customers credit term performance
* Identifies Credit-worthy potential customers
* Current and past shareholdings trace & listing of multiple address
* Court litigation cases
* Comprehensive access of bankrupts and wound-up companies
* Instant access to Singapore and Malaysia Company records
* Credit scores & ratings, robustly trained with local credit & trade records

# NLP / Artificial Intelligence/ Neural Networks for Analysis

The outlook for the potential use of EDUK AI technology to provide advanced credit risk assessment and profiling of individuals and companies hinges on the application of the technology in disruptive ways such as correlating an individual’s credit information with hobbies, shopping habits, interaction with friends and lifestyle. The idea is to develop a meaningful association to measure risky behavior based on an individual’s social interactions and connections. Current calculations of credit largely ignore populations that do not have a source of income (such as students, low income workers or people who have never borrowed from a bank). By marrying the social context with credit history a wider pool of applicants would possibly qualify for credit while it may eliminate fringe groups that would normally qualify for credit. This will essentially translate to both big data analysis as well as smart data analysis as a means to control risk.

Developing predictive risk control technologies will be the market differentiator that will propel the future of EDUK AI technology. Emerging artificial intelligence (AI) techniques such as ANNs, genetic algorithm (GA) and support vector machine (SVM) have been applied with positive results for both consumer and company credit ratings (Lai, Yu, Wang & Zhou, 2006a).

## Background

Personal credit scores are normally computed from information available in credit reports collected by external credit bureaus and ratings agencies. Credit scores may indicate personal financial history and current situation. However, it does not tell you exactly what constitutes a "good" score from a "bad" score. More specifically, it does not talk about the level of risk for the purposed lending consideration.

In order to arrive at a better model for factoring credit scores, leading research indicates that the subject needs to evaluated/rated based on other aspects such as behavior, demography and biography aside from published credit scores.

A behavioral score predicts the likelihood of an account going 'bad', based on payment history, usage, delinquency and timing characteristics i.e. risk scores are designed to predict the future performance or behavior of an account. The principle of "the future is like the past" is applied. Subjectivity plays no part in the derivation of these scores. A customer’s behavioral score, which is a function of the behavior of the account holder, is systemically irreversible.

Examples of some typical characteristics that may be used in account level behavioral scores are illustrated below:

* Days in excess for the last six months (weighting 12)
* Overdraft Limit Utilization as a percentage of Minimum Balance Last 6 Months(weighting 18)
* Days in Debit Last 6 Months (weighting 11)
* Current Minimum balance (weighting 8)
* Current Days in Excess (weighting 11)
* Current Days in Credit (weighting 6)
* Average Credit Turnover Last 6 Months (weighting 11)
* Months Account Open (weighting 13)
* Payment Reversals Last 6 months (weighting 10)

Other characteristics (sample) related to demography and biography are listed below

* Age
* Sex
* Nationality
* Marital status
* Spouse employment status
* Number of dependents
* Number of siblings
* Preferred language of correspondence
* Home language
* Population group
* Province of residence
* Time at residence
* Occupation, Job title & date of assumption of current post
* Income per annum
* Economic sector
* Educational qualifications
* Previous credit history
* Account balance and surplus
* Number of defaults
* Outstanding debt
* Property Owner
* Motor Vehicle Owner

Weighted averages are assigned to each metric based on some predefined models. The issue with this system is that most of the metrics defined (E.G. Economic Sector) may not have been populated because the bank may not hold such information. Credit Risk Analysis using statistical approaches is usually limited due to competition, confidentiality and privacy, making the statistical approaches difficult to obtain a consistently good result for credit scoring. In order to improve the performance and overcome data shortage, it is therefore imperative to implement an approach that may be able to cope with these challenges. Pattern recognition related decision problems suit Artificial Neural Networks (ANNs) more than to statistical techniques primarily because of the complex nature of problems, which are not well understood mathematically and involve subjectivity. In addition, such problems have qualitative and noise data. Even if the values of every input features are not known, a trained neural network will produce a response.

## Data Cleaning and preparation

The objective of data cleaning is to ‘identify omissions, ambiguities, and errors in the responses’.

For example a key characteristic input required for training of the ANN to recognize credit risk classes is the customer’s annual income. An inspection reveals that some data in this field is inconsistent: it seems that where data is missing in this field, it has been replaced by the value zero. This causes ambiguity since missing data from the annual income field cannot be differentiated from genuine cases where the customer’s income is presently zero.

ANNs could be ideally suited to the turbulent financial risk management field due to their adaptive nature; different combination of factors may be weighed differently for different consumers. Requirements may also vary over time. The system could be periodically updated by retraining with more recent data to reflect changes in consumer behavior and the macro-environment. The ultimate aim would be online training, where the training algorithm would be capable of processing the input piece-by-piece, without having the entire input available to train the network from the start.

The consumer data (behavioral (e.g. past credit history), biographic and demographic (e.g. gender and marital status) are pre-processed and fed as inputs to the **L**earning **V**ector **Q**uantization (LVQ) Classifier.

LVQ Neural Network Classifier

Consumer Credit Score classification (output)

Preprocessed characteristic input vector (per consumer)

Consumer Biographic & demographic Data (Input)

Consumer Behavior (historic Input)

**Schematic of artificial neural network model for credit risk prediction**

Individual characteristics, e.g. gender and marital status, are fed as elements for each consumer i.e. consumer 1 will have a corresponding input which is comprised of variables such as marital status, age and income per annum respectively. The output of the system is the risk category of the consumer. The output credit risk category will be specified to the LVQ system, for each consumer input, during the training phase. It is during this phase that the ANN continually adjusts the weights in order to establish relationships between consumer input attributes and their credit scores. The total number of epochs (iterations) is user-determined. The ANN is ‘learning’ through experience. Once training has been completed the ANN stores the final adjusted weight and bias values in a weight matrix.

LVQ is a method for training competitive networks in a supervised manner. In supervised learning, the correct results (target values, desired outputs) are known and are given to the ANN during training so that the ANN can adjust its weights to try matching its outputs to the target values. The data presented to the network in this stage is called the training data. After training, the ANN is tested by giving it only input values, not target values, and seeing how close it comes to outputting the correct target values. Supervised neural networks are required to learn an input output mapping from existing data. LVQ neural networks will be used to perform the pattern recognition task of the consumer credit classification system. LVQs are closely related to **S**elf-**O**rganizing **M**aps (SOM). It is an algorithm that effectively maps similar patterns (pattern vectors close to each other in the input signal space) onto locations in the output space (Kohonen, 1997). **L**earning **V**ector **Q**uantization is a supervised version of SOM particularly suitable for statistical pattern recognition.

Several software packages exist for ANN analysis. The MATLAB neural network toolbox and simulation environment, by Optimum Solutions, can be used for analysis of the data in this study. Another example of commonly available software for neural network simulation is the Neuralyst function available on Microsoft Excel.

## Conclusion

Credit scores are often obtained by financial institutions from 3rd party credit bureaus or consultants. The actual information used to calculate the credit score is normally not available with the current account performance. This presents a problem with Credit Risk Analysis using Artificial Intelligence, comparing the ANN performance with that employed by the financial institution. ANNs can be used as an additional support system, in conjunction with the existing system, to improve the ability to predict credit default. The ANN system is based on training with the use of past data on the premise that the past predicts the future. Changes in attitudes, economic conditions, etc. over time will affect the accuracy of the system.

# References

Much of the information in this article has been condensed from the report “Credit Risk Analysis using Artificial Intelligence: Evidence from a leading south African Banking institution”, Viresh Moonasar, Nov 2007