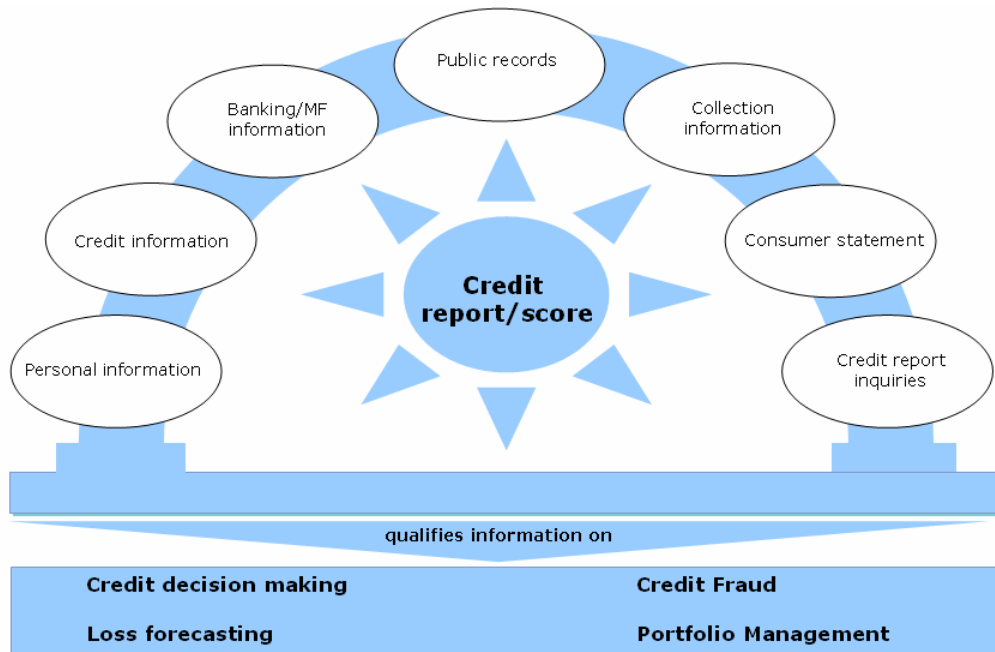


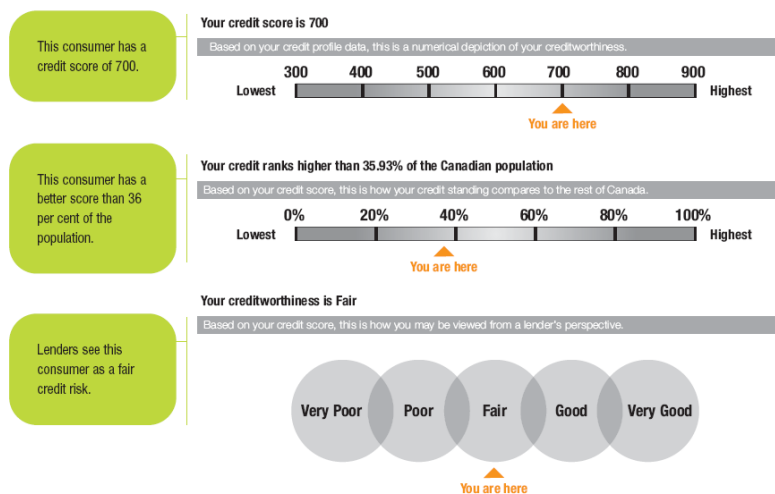
Credit scoring

Credit scoring is based on consumer specific information which aids lenders and other organizations to make qualitative decisions.



Source: Aaum research

The benefits of credit scoring helps the credit institutes to make qualified decisions on the customer's present/future credit requirements. This is illustrated with TransUnion's Credit Score report below



Source: Understanding your credit report and credit score, Financial Consumer Agency of Canada

This document contains proprietary information and data of AAUM. Users are permitted to print or download extracts of this document for non-commercial or personal use. No part of this document or any material appearing on it may be reproduced on, stored in or transmitted to any form without prior written permission of Aaum Research and Analytics Private Limited. Copyright © 2009 Aaum Research and Analytics Private Limited. All Rights Reserved.

The score eventually qualifies the customer into high risk, average and low risk. There are many statistical techniques available to model the scoring.

An analysis by Aaum’s researchers revealed that model formulation depends on data availability, data characteristics and the organization's comfort. The below diagram illustrates the advantages and disadvantages of various scoring methodologies.

Suitability of a model depends on data availability, data characteristics and the organization's comfort

LDA	<ul style="list-style-type: none"> Simple, very easily estimated. Often used by banks for credit-scoring purposes. 	<ul style="list-style-type: none"> LDA requires normally distributed data but the credit data are often non-normal (and categorized).
Logistic	<ul style="list-style-type: none"> Most successful predictive model. Works well with categorized data 	<ul style="list-style-type: none"> Very sensitive to missing values. Like LDA, a Parametric model
CART	<ul style="list-style-type: none"> Non parametric model and employs binary trees for classification Very intuitive, easy to explain to management, and it is able to deal with missing observations. 	<ul style="list-style-type: none"> Computational burden in case of large datasets (at each node every characteristic has to be examined) Optimizes only locally on a single variable at a time Often the trees are not stable
K-nn	<ul style="list-style-type: none"> Non parametric model and out performs other non parametric approaches. 	<ul style="list-style-type: none"> Lack of formal framework for choosing k Can make only discrete predictions
Neural	<ul style="list-style-type: none"> Non linear optimization and can achieve a high prediction accuracy rate Suitable for classification 	<ul style="list-style-type: none"> Lack of explanation capability

The key insights coming out from the Aaum research are

- Logistic regression and LDA are most widely used scoring methodologies.
 - Much simpler and convenient.
- CART or neural networks are used mainly as support tools, either in the process of selecting variables or in the process of the model-quality evaluation.
- k-NN method is not used at all or is used very rarely.
- The facts are very surprising!



- The alternative (nonparametric) methods have excellent potential in pattern recognition and they are very competitive with logit regression.
- It seems that this potential is unrecognized by the current financial institutes!
- Logit method is the most favored method in practice, mainly due to (almost) no assumptions imposed on variables, with the exception of missing values and multicollinearity among variables.
- Contrary to this, non-parametric methods can deal with missing values and multicollinearity (or correlations) among variables, but often are computationally demanding.
 - Rules that are constructed on the basis of some of these methods can be hard to explain to a manager as well as to a client.

Please write to info@aaumanalytics.com if you are looking for expertise in credit scoring.