My idea is:**(Statistical) Credit Model**

Credit risk: Credit risk refers to *the risk that a borrower will default on any type of debt by failing to make required payments.* (from Wikipedia)

There are many ways to assess credit risk. The typical approaches are,

* Qualitative Analysis (Traditional) by credit analysts
* Quantitative Model

Why we use quantitative model?

→There are so many companies in the market (Let’s say 1,000). Since qualitative analysis is time consuming, we cannot apply this approach to every company. Therefore we need to shrink the universe (a set of companies we consider to do investment or not). If we have a statistical model that automatically choose the companies whose default probability is less than 1%, we save time and money.

Please see <http://en.wikipedia.org/wiki/Credit_analysis>.

In my understanding, quantitative models are categorized into

* Statistical model
  + Time series data (ex. How long does company A survive? How CDS spread changes in future?)
    - Hazard rate model
    - Survival analysis (Cox proportional hazard rate model, …)
    - …
  + **Two period data (ex. Estimation of 1 year default probability)**
    - **Altman Z-score (Discrimination Analysis)**
    - **Logistic Regression**
    - **Decision Tree**
    - **…**
  + Credit Rating (Transition)
    - Ordered Logit
    - Neural Network
    - Genetic Algorithm
    - …
  + …
* Structural model (Modeling default mechanism by using option theory)
  + Merton model (Moody’s KMV)
  + …
* Reduced-form model (Modeling credit risk itself by using some stochastic process)
  + LS model
  + …

We are interested in statistical model.

We assume that default probability can be expressed by a function.

Prob(Default at T|t)=f(Liability/Capital, Liquidity, Company Size, Sector, …)

Therefore, we need to choose appropriate function (parametric, semi-parametric, non-parametric) and explanatory variables.

* Data
  + Chinese Companies
    - Not Default
    - Default
      * When did they default?
  + … and their corporate data (Capital, Liability, …)
* Data Source
  + Bloomberg

Reference

The following paper uses logistic regression model to predict default probability.

<http://www.stat.fsu.edu/~jfrade/HOMEWORKS/STA5168/FRADE_STA5168_paper.pdf>