Introduction to Credit Risk Modelling

Credit Scoring & Credit Control XVI
Pre-conference workshop

Dr Galina Andreeva, Galina.Andreeva@ed.ac.uk
The workshop is an introduction for complete novices, newcomers to the field of credit scoring or people with limited experience. It covers the basic principles and standard methodology of scorecard development.

1. Concepts, principles and objectives
2. What is needed
3. Stages in credit model (scorecard) development
   - Characteristics transformation
   - Modelling algorithms
     - Logistic regression
   - Measuring predictive accuracy.
4. Where to find more information.
What is credit scoring?

Decision support systems used in consumer credit
Aims at risk assessment of potential borrowers (application scoring) and existing borrowers (behavioural scoring)
Risk/creditworthiness is usually measured by default probability
Default probability is predicted from observed borrower’s characteristics on the basis of the analysis of known performance of previous customers.
40-50 years ago – relationship lending

• Judgmental assessment based on 3(5) Cs (Character, Capacity, Collateral + Capital, Conditions)
• Very slow, subjective, often inconsistent

Now – transactional lending

• Automated impersonal decisions
• Quick and consistent
• Huge volumes of unsecured lending
The Basic Idea

Need the data on characteristics and performance of previous borrowers
Assign points to each category of borrowers’ characteristic so that they reflect the level of risk of this category
For a new borrower, sum all the points over the relevant categories of all characteristics \(\rightarrow\) get an overall score – a summary of this borrower’s creditworthiness.
Example of a scoring table (part)

<table>
<thead>
<tr>
<th>Time at current address</th>
<th>Less than 6 months</th>
<th>6m – 2 years</th>
<th>2 – 6 years</th>
<th>6 - 10 years</th>
<th>10 + years</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>25</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residential Status</th>
<th>Owner</th>
<th>Tenant</th>
<th>With parents</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Banking</th>
<th>Current account</th>
<th>Saving account</th>
<th>Current and saving</th>
<th>No account</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Retired</th>
<th>Full-time</th>
<th>Part-time</th>
<th>Self-employed</th>
<th>Student</th>
<th>Other</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>18-25</th>
<th>26-31</th>
<th>32-40</th>
<th>41-54</th>
<th>55+</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Lewis M. (1992) *An Introduction to Credit Scoring*, FICO: San Rafael, CA
• “Future is going to be like the past”

• Scoring models are based on associations / correlations, they are not causal models
  • It is predictive not explanatory.
  • Anything that helps predict can be used (but there is still a desire to understand risk drivers).

• It has to interact easily with organisations information system.
  • So variables that can be used must be easily and cheaply obtained and automatically updated.

Birds flying low, expect rain and a blow
### Credit risk approaches

<table>
<thead>
<tr>
<th>Lending to individuals</th>
<th>Lending to businesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Relatively small amounts of money lent to a large number of customers</td>
<td>- Large amounts of money lent to a relatively small number of businesses</td>
</tr>
<tr>
<td>- focus more on prediction, less on causality</td>
<td>- focus more on causality, less on prediction</td>
</tr>
<tr>
<td>- Management Science and Data Mining</td>
<td>- Finance and Accounting</td>
</tr>
<tr>
<td>- limited research and a few textbooks</td>
<td>- extensive research, loads of textbooks</td>
</tr>
</tbody>
</table>
What is scoring used for

New customers (application scoring):
• who is likely to default?
• who is likely to respond?
• who is likely to use the product?
• who is likely to switch to a competitor?

Existing customers (behavioural):
• shall we increase a credit limit?
• can we cross-sell other products?

Problem customers (collections):
• shall we sell the debt off?
• if not, what method of contacting to be used?

Newer Uses:
• Risk-based pricing
• Securitisation
• Profit Scoring
• Customer Scoring
• Capital Adequacy (Basel New Accord)
• IFRS9
Process of scorecard development

Data preparation

Sample selection

Characteristics analysis

Fitting Good/Bad model

Cut-off selection

Evaluation of model performance

Reject inference
Classification methods

Statistical
- Discriminant analysis (DA or LDA),
- Linear regression (LR),
- Logistic regression (LogR),
- Nearest-neighbours approach (k-NN),
- Classification trees.

Machine-Learning
- Support Vector Machines (SVM),
- Neural networks (NN),
- Genetic algorithms,
- Random Forests,
the list is not exhaustive…
Measures of predictive accuracy

- Kolmogorov-Smirnov (K-S) statistic
- Area under the ROC curve (AUROC)
  - Gini coefficient
- Error rate
- Bad rate.
Books on credit scoring


Selected References


