Modeling Probability of Default and Credit Limits

ZAŁA HERGA, JAN RUPNIK, PRIMOŽ ŠKRABA, BLAŽ FORTUNA
ARTIFICIAL INTELLIGENCE LABORATORY
JOŽEF STEFAN INSTITUTE

SIKDD, OCTOBER 10TH, 2016
Introduction

Default: clients not meeting their debt obligations

- Challenge: compute Probability of Default (PD)

How to limit default risk?

- Credit Limit
Outline

Data

PD model
- Computation
- Weight of evidence
- Results

Credit limits model
- Computation
- Variation of inputs
Data

Financial data (publicly available in several European countries)

Monthly trading data (private information)
- Sum of trades
- Outstanding debts
- Delayed payments
- Disputed claims
PD model

- Simple and easy to understand
- Logistic regression

\[ F(x) = \frac{1}{1 + e^{-\left( \beta_0 + \beta_1 \cdot woe(x_1) + \ldots + \beta_n \cdot woe(x_n) \right)}} \]
Challenge

How should default be defined?

• What if a client is late for one day?
• What if a client owes 10€?
• What if a client didn’t pay one bill, but paid all bills since?
Weight of Evidence (WOE)
Weight of Evidence (WOE)

Transformation of financial indicators into feature vectors using WOE
Weight of Evidence (WOE)

Transformation of financial indicators into feature vectors using WOE

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Fin. Indicator Value</th>
<th>Bins</th>
<th>Feature Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>log(2/1)</td>
<td>log(2/1)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>log(2/1)</td>
<td>log(2/1)</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>log(2/1)</td>
<td>log(2/1)</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>log(1/2)</td>
<td>log(1/2)</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>log(1/2)</td>
<td>log(1/2)</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
<td>log(1/2)</td>
<td>log(1/2)</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>log(1/3)</td>
<td>log(1/3)</td>
</tr>
<tr>
<td>0</td>
<td>8</td>
<td>log(1/3)</td>
<td>log(1/3)</td>
</tr>
<tr>
<td>0</td>
<td>9</td>
<td>log(1/3)</td>
<td>log(1/3)</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>log(1/3)</td>
<td>log(1/3)</td>
</tr>
</tbody>
</table>
Weight of Evidence (WOE)

Transformation of financial indicators into feature og  \( P(\text{company}=\text{good}) \)
\( P(\text{company}=\text{bad}) \)

og  \( P(\text{company}=\text{good}) \) \( P(\text{company}=\text{bad}) \)
\( P(\text{company}=\text{bad}) \) \( P(\text{company}=\text{bad}) \) og
\( P(\text{company}=\text{good}) \) \( P(\text{company}=\text{bad}) \) tors using WOE

1. Create n bins
2. Assign each company to corresponding bin
3. Count the number of bad and good companies in each bin
4. Assign WOE to companies of a corresponding bin as \( \log P(\text{company} = \text{good}) P(\text{company} = \text{bad}) \)
Weight of Evidence (WOE)

Transformation of financial indicators into feature
\[ \log P(\text{company} = \text{good}) \]
\[ P(\text{company} = \text{bad}) \]

\[ \log P(\text{company} = \text{good}) P(\text{company} = \text{bad}) \]
\[ (\text{company} = \text{good}) P(\text{company} = \text{bad}) \]
\[ P(\text{company} = \text{bad}) P(\text{company} = \text{bad}) \]
\[ \log P(\text{company} = \text{good}) P(\text{company} = \text{bad}) \]

1. Create n bins
2. Assign each company to corresponding bin
3. Count the number of bad and good companies in each bin
4. Assign WOE to companies of a corresponding bin as \[ \log \frac{P(\text{company} = \text{good})}{P(\text{company} = \text{bad})} \]

Using WOE:

1. Create n bins
2. Assign each company to corresponding bin
3. Count the number of bad and good companies in each bin
4. Assign WOE to companies of a corresponding bin as \[ \log \frac{P(\text{company} = \text{good})}{P(\text{company} = \text{bad})} \]
Results

2 models:
- full
- stepwise

Comparison
- Disputed claims (true negatives)
- Amount of missed trading volume (false positives)

Cutoff?
- 1€ disputed claims vs 1€ trading volume (margin)
- In addition to profit: risk aversion
Results

2 models:
  ◦ full
  ◦ stepwise

Comparison
  ◦ Disputed claims (true negatives)
  ◦ Amount of missed trading volume (false positives)

Cutoff?
  ➢ 1€ disputed claims vs 1€ trading volume (margin)
  ➢ In addition to profit: risk aversion
Credit limits model

How to handle identified risky clients?

- Credit limit
Credit limits model (2)

Optimal portfolio based on
- VaR
- Max CVaR
- Margin
- PDs
- Credit limit upper- and lower bounds

Optimization is based on tradeoff between expected profit and risk [1]

Relative amount of approved credit
CVaR decreased by factor 10
Margin increased by factor 10
Lower bound $> 0$
Conclusion and future work

PD model
- More complex methods
- Use of additional features extracted from trading data

Portfolio optimization
- Additional parameters e.g. insurance compensations
- Correlation between clients

Efficient optimal portfolio calculation based on simple PD model and standard financial risk measures