WEBINAR@LUNCHTIME

„STATE OF THE ART IN CREDIT RISK MODELING”

BY PROF. BART BAESENS
PROF. DR. BART BAESSENS

• Studied at KU Leuven (Belgium)
  • Business Engineer in MIS, 1998
  • PhD. in Applied Economic Sciences, 2003
• PhD. Title: Developing Intelligent Systems for Credit Scoring Using Machine Learning Techniques
• Professor at KU Leuven, Belgium
• Lecturer at the University of Southampton, UK
• Research: analytics, credit risk, fraud, marketing, …
• YouTube/Facebook/Twitter: DataMiningApps
• www.dataminningapps.com
OVERVIEW

• Introduction
• Data Quality
• Model requirements
• Model discrimination versus calibration
• Model validation
More than ever before, analytical models steer strategic decisions of financial institutions!

Minimum equity (buffer capital) and provisions a financial institution holds are directly determined, a.o., by:
- credit risk models
- market risk models
- operational risk models
- fraud risk models
- insurance risk models
- model risk metamodels (?)
- ...

Analytics typically used to build all these models!

Often subject to regulation (e.g. Basel II/Basel III, Solvency II, …)!

Model errors directly affect profitability, solvency, shareholder value, macro-economy, …, society as a whole!
CREDIT RISK COMPONENTS

- **Probability of default (PD) (decimal):** probability of default of a counterparty over a one year period (Art. 4, EU)

- **Exposure at default (EAD) (currency):** amount outstanding

- **Loss given default (LGD) (decimal):** ratio of the loss on an exposure due to default of a counterparty to the amount outstanding (Art. 4, EU)

- **Expected loss** = PD × LGD × EAD

- **Unexpected loss** = f(PD, LGD, EAD)
CREDIT RISK MODEL ARCHITECTURE

0. Prepare the data
   - internal data
   - external data
   - expert judgment

1. Create the model
   - application scorecard
   - behavioral scorecard

2. Define ratings and calibrate the model
   - PD calibration
   - risk ratings definition

PD Modeling

LGD Modeling

EAD Modeling
### TRADITIONAL ANALYTICS: PERFORMANCE BENCHMARKS

<table>
<thead>
<tr>
<th>Context</th>
<th>Number of Characteristics</th>
<th>AUC ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Credit Scoring</td>
<td>10-15</td>
<td>70%-85%</td>
</tr>
<tr>
<td>Behavioural Credit Scoring</td>
<td>10-15</td>
<td>80%-90%</td>
</tr>
<tr>
<td>Fraud detection (insurance)</td>
<td>10-15</td>
<td>70%-90%</td>
</tr>
<tr>
<td>Churn detection (Telco)</td>
<td>6-10</td>
<td>60%-80%</td>
</tr>
</tbody>
</table>


IMPROVING TRADITIONAL ANALYTICS: 2 STRATEGIES

• **Strategy 1**: Use complex modeling techniques
  • E.g. neural networks, support vector machines, random forests, …
  • **Pro**: powerful models (e.g. universal approximation)
  • **Con**: loss of interpretability, marginal performance gains

• **Strategy 2**: Enrich your data
  • External data (FICO score, bureau data, …)
  • **Data quality!**
  • **Pro**: model still interpretable
  • **Con**: additional resources needed (ICT)
DATA QUALITY

• GIGO principle
  • Garbage in, Garbage out; messy data gives messy models

• In many cases, simple analytical models perform well, so biggest performance increase comes from the data!

• “The best way to improve the performance of an analytical model is not to look for fancy tools or techniques, but to improve DATA QUALITY first”
EXAMPLE DATA QUALITY CRITERIA

• **Data accuracy**
  • E.g., outliers
  • Age is 300 years versus Income is 1.000.000 Euro (not the same!)

• **Data completeness**
  • Are missing values important?

• **Data bias and sampling**
  • Try to minimise, but can never totally get rid of

• **Data definition**
  • Variables: what is the meaning of 0?
  • Target: fraud, churn, default, customer lifetime value (CLV), ....

• **Data recency/latency**
  • Refresh frequency
## DATA QUALITY CRITERIA (MOGES, LEMAHIEU, BAESSENS, 2011)

<table>
<thead>
<tr>
<th>Cat.</th>
<th>DQ dimensions</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic</td>
<td>Accuracy (AC)</td>
<td>The extent to which data are certified, error-free, correct, flawless and reliable</td>
</tr>
<tr>
<td></td>
<td>Objectivity (OBJ)</td>
<td>The extent to which data are unbiased, unprejudiced, based on facts and impartial</td>
</tr>
<tr>
<td></td>
<td>Reputation (REP)</td>
<td>The extent to which data are highly regarded in terms of its sources or content</td>
</tr>
<tr>
<td>Contextual</td>
<td>Completeness (COM)</td>
<td>The extent to which data are not missing and covers the needs of the tasks and is of sufficient breadth and depth of the task at hand</td>
</tr>
<tr>
<td></td>
<td>Appropriate-amount (APM)</td>
<td>The extent to which the volume of the information is appropriate for the task at hand</td>
</tr>
<tr>
<td></td>
<td>Value-added (VAD)</td>
<td>The extent to which data are beneficial and provides advantages from its use</td>
</tr>
<tr>
<td></td>
<td>Relevance (REL)</td>
<td>The extent to which data are applicable and helpful for the task at hand</td>
</tr>
<tr>
<td></td>
<td>Timeliness (TIM)</td>
<td>The extent to which data are sufficiently up-to-date for the task at hand</td>
</tr>
<tr>
<td></td>
<td>Actionable (ACT)</td>
<td>The extent to which data is ready for use</td>
</tr>
<tr>
<td>Representation</td>
<td>Interpretable (INT)</td>
<td>The extent to which data are in appropriate languages, symbols, and the definitions are clear</td>
</tr>
<tr>
<td></td>
<td>Easily-understandable (EU)</td>
<td>The extent to which data are easily comprehended</td>
</tr>
<tr>
<td></td>
<td>Representational-consistent (RC)</td>
<td>The extent to which data are continuously presented in same format</td>
</tr>
<tr>
<td></td>
<td>Concisely-represented (CR)</td>
<td>The extent to which data compactly represented, well-presented, well-organized, and well-formatted</td>
</tr>
<tr>
<td></td>
<td>Alignment (AL)</td>
<td>The extent to which data is reconcilable</td>
</tr>
<tr>
<td>Access</td>
<td>Accessibility (ACC)</td>
<td>The extent to which data is available, or easily and swiftly retrievable</td>
</tr>
<tr>
<td></td>
<td>Security (SEC)</td>
<td>The extent to which data access to data is restricted appropriately to maintain its security</td>
</tr>
<tr>
<td></td>
<td>Traceability (TRA)</td>
<td>The extent to which data is traceable to the source</td>
</tr>
</tbody>
</table>
SURVEY: DATA QUALITY FOR CREDIT RISK ANALYTICS

• 50+ banks participating world-wide
• Focus on credit risk analytics
• Initial findings:
  • Most banks indicated that between 10-20 percent of their data suffer from data quality problems
  • Manual data entry one of the key problems
  • Diversity of data sources and consistent corporate wide data representation main challenge for data quality
  • Regulatory compliance key motive to improve data quality
DATA QUALITY: SHORT TERM VERSUS LONG TERM IMPACT

• No short term solution
  • Deal with in a statistical way using e.g. data transformations
    • Outlier truncation, missing value imputation, data enhancement
  • Buy external data (data poolers!)

• Structural solutions in the long term
  • Re-design data entry processes
  • Master data management
ANALYTIC MODEL REQUIREMENTS

- **Statistical performance**
  - Lift curve, ROC curve, Gini coefficient, …
  - R-squared, MSE, …

- **Interpretability + Justifiability**
  - Very subjective, but CRUCIAL!
  - Often need to be balanced against statistical performance

- **Operational efficiency**
  - How much effort is needed to evaluate/monitor/re-train the model(s)?

- **Economical cost**
  - What is the cost to gather the model inputs and evaluate the model?
  - Is it worthwhile buying external data and/or models (e.g. BKR score)?

- **Regulatory compliance**
  - In accordance with regulation and legislation
  - E.g., Basel II\Basel III, Solvency II
MODEL DISCRIMINATION VERSUS MODEL CALIBRATION

• Model discrimination
  • Rank order (score) entities with respect to likelihood of event occurring
  • Examples
    • Rank order customers in terms of likelihood to default on their obligation
    • Bart is more risky to default than Victor!
  • However, despite traditional focus in data mining, this is no longer sufficient!
  • We need to know the **EXACT** probability of the event occurring!

• Model calibration
  • Provide well-calibrated and accurate projected probabilities based on
    • Historical data
    • Expectations with respect to the future (e.g. GDP contraction versus expansion)
  • Losses only make sense in an **ABSOLUTE** way!
  • Example
    • $P(\text{Bart defaults})=0.90$; $P(\text{Victor defaults})=0.75$

*BRING THE MACRO-ECONOMY INTO THE MODEL!*
MODEL DISCRIMINATION VERSUS MODEL CALIBRATION

Model Discrimination

<table>
<thead>
<tr>
<th>Characteristic Name</th>
<th>Attribute</th>
<th>Scorecard Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE 1</td>
<td>Up to 26</td>
<td>100</td>
</tr>
<tr>
<td>AGE 2</td>
<td>26 - 35</td>
<td>120</td>
</tr>
<tr>
<td>AGE 3</td>
<td>35 - 37</td>
<td>185</td>
</tr>
<tr>
<td>AGE 4</td>
<td>37+</td>
<td>225</td>
</tr>
<tr>
<td>GENDER 1</td>
<td>Male</td>
<td>90</td>
</tr>
<tr>
<td>GENDER 2</td>
<td>Female</td>
<td>180</td>
</tr>
<tr>
<td>SALARY 1</td>
<td>Up to 500</td>
<td>120</td>
</tr>
<tr>
<td>SALARY 2</td>
<td>501-1000</td>
<td>140</td>
</tr>
<tr>
<td>SALARY 3</td>
<td>1001-1500</td>
<td>160</td>
</tr>
<tr>
<td>SALARY 4</td>
<td>1501-2000</td>
<td>200</td>
</tr>
<tr>
<td>SALARY 5</td>
<td>2001+</td>
<td>240</td>
</tr>
</tbody>
</table>

Example application scorecard

Model Calibration

Historical probability of default (PD) calibration for customer segment B!
MODEL CALIBRATION: EXAMPLE APPROACH

- Analytical models typically built using a snapshot at a given period in time!
- Cluster data mining model outputs (e.g. scores) into pools
  - Scores are too fine granular anyway!
  - Essentially, a semi-supervised learning exercise
  - Score 200-300: pool A; score 301-500: pool B, score 501-650: pool C, …
- For each pool, calibrate event probability using
  - Time series analysis techniques (ARIMA, VAR, …)
  - Dynamic models/Markov Chains
  - Simulations
  - Projected macro-economic scenarios
- Model transitions between pools
  - Gives an idea about customer volatility/model stability
  - Do I have a point-in-time (PIT) or through the cycle (TTC) analytical model?
SUMMARISING: MODEL ARCHITECTURE

Calibration

Dynamic macro-economic models

Discrimination

Data Mining Model Scorecard

Data

Internal/External Data Expert Input

<table>
<thead>
<tr>
<th>Characteristic Name</th>
<th>Attribute</th>
<th>Scorecard Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE 1</td>
<td>Up to 26</td>
<td>100</td>
</tr>
<tr>
<td>AGE 2</td>
<td>26 - 35</td>
<td>120</td>
</tr>
<tr>
<td>AGE 3</td>
<td>35 - 37</td>
<td>185</td>
</tr>
<tr>
<td>AGE 4</td>
<td>37+</td>
<td>225</td>
</tr>
<tr>
<td>GENDER 1</td>
<td>Male</td>
<td>90</td>
</tr>
<tr>
<td>GENDER 2</td>
<td>Female</td>
<td>180</td>
</tr>
<tr>
<td>SALARY 1</td>
<td>Up to 500</td>
<td>120</td>
</tr>
<tr>
<td>SALARY 2</td>
<td>501-1000</td>
<td>140</td>
</tr>
<tr>
<td>SALARY 3</td>
<td>1001-1500</td>
<td>160</td>
</tr>
<tr>
<td>SALARY 4</td>
<td>1501-2000</td>
<td>200</td>
</tr>
<tr>
<td>SALARY 5</td>
<td>2001+</td>
<td>240</td>
</tr>
</tbody>
</table>
SIDE BENEFIT: STRESS TESTING

- By introducing the macro economy into the model, one can do stress testing
  - “evaluate the potential impact on a firm of specific adverse events and/or movements in a set of financial variables” (BIS, 2005)
- Sensitivity analysis
  - Single variable versus multiple variables
  - E.g. assume all credit scores decrease by 5%
- Scenario analysis
  - Historical or hypothetical
  - E.g. 3 successive years of GDP contraction, house prices drop by 5%, …
  - Could be a 1/25 years event (e.g. in the United Kingdom)
- Common challenges/problems:
  - Lack of historical data
  - Correlations break down during stress (need to have data on downturn periods)
  - Integrate risks
  - What is stress??
  - What to do with the results? Strategic impact?
MODEL RISK

• “Essentially, all models are wrong, but some are useful“ (George E. P. Box, 1987)

• Models are not perfect, some are actually VERY bad, but what’s the alternative???
  • Default risk/fraud prediction: good performance (Gini coefficients around 50 to 80%)!
  • Loss/LGD prediction: awful performance (R² of 0.30 already great!)

• Model imperfection is typically dealt with by
  • Conservative parameter calibration (aka economic downturn calibration)
    • E.g. assume statistically estimated probability of default is 3%.
    • Use 5% for strategic decisions to capture model risk!
  • Create equity buffer/provisions for model risk
    • Hard to quantify!
MODEL MONITORING

• Why analytical models may degrade in performance?
  • Sample effects (models estimated on limited samples)
  • Macro-economy (downturn versus upturn)
  • Internal effects (e.g. strategy change, population drift, M&A)
  • In reality: a very nice (?) mixture of these!

• Need to constantly monitor outcomes of models

• Crucial since models more and more steer strategic decisions of the firm (cf. supra)
  • E.g. equity calculation in a Basel II/Solvency II environment
  • Risk based pricing

• Quantitative versus Qualitative validation
MODEL VALIDATION

• Quantitative validation
  • Backtesting
  • Benchmarking

• Qualitative validation
  • Data quality
  • Model design
  • Documentation
  • Corporate governance and management oversight
BACKTESTING

• Contrasting ex-post realized numbers with ex-ante predictions
• Using statistical tests and performance measures
• Examples
  • Use binomial test for comparing default/fraud rates
  • Monitor decrease in AUC (Gini) over time
• Challenges
  • Which test statistics to use?
  • Which confidence levels to adopt?
  • How to deal with correlated behavior (portfolio effects)?
  • When to take action and what action?
## Backtesting: Examples

### Score Range

<table>
<thead>
<tr>
<th>Score Range</th>
<th>Expected (training) %</th>
<th>Observed (actual) % at t</th>
<th>Observed (actual) % at t+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-169</td>
<td>6%</td>
<td>7%</td>
<td>6%</td>
</tr>
<tr>
<td>170-179</td>
<td>10%</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>180-189</td>
<td>9%</td>
<td>7%</td>
<td>10%</td>
</tr>
<tr>
<td>190-199</td>
<td>12%</td>
<td>9%</td>
<td>11%</td>
</tr>
<tr>
<td>200-209</td>
<td>12%</td>
<td>11%</td>
<td>10%</td>
</tr>
<tr>
<td>210-219</td>
<td>8%</td>
<td>11%</td>
<td>9%</td>
</tr>
<tr>
<td>220-229</td>
<td>7%</td>
<td>10%</td>
<td>11%</td>
</tr>
<tr>
<td>230-239</td>
<td>8%</td>
<td>12%</td>
<td>11%</td>
</tr>
<tr>
<td>240-249</td>
<td>12%</td>
<td>11%</td>
<td>10%</td>
</tr>
<tr>
<td>250+</td>
<td>16%</td>
<td>14%</td>
<td>15%</td>
</tr>
<tr>
<td>SSI versus Expected</td>
<td>0.0605</td>
<td>0.494</td>
<td></td>
</tr>
<tr>
<td>SSI versus t-1</td>
<td>0.0260</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Number of Observations and Defaulters

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Observations</th>
<th>Number of Defaulters</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR model</td>
<td>5866</td>
<td>105</td>
<td>0.85</td>
</tr>
<tr>
<td>AR 2006</td>
<td>5677</td>
<td>97</td>
<td>0.81</td>
</tr>
<tr>
<td>AR 2005</td>
<td>5462</td>
<td>108</td>
<td>0.80</td>
</tr>
<tr>
<td>AR 2004</td>
<td>5234</td>
<td>111</td>
<td>0.83</td>
</tr>
<tr>
<td>AR 2003</td>
<td>5260</td>
<td>123</td>
<td>0.79</td>
</tr>
<tr>
<td>AR 2002</td>
<td>5365</td>
<td>113</td>
<td>0.79</td>
</tr>
<tr>
<td>AR 2001</td>
<td>5354</td>
<td>120</td>
<td>0.75</td>
</tr>
<tr>
<td>AR 2000</td>
<td>5306</td>
<td>119</td>
<td>0.82</td>
</tr>
<tr>
<td>AR 1999</td>
<td>4970</td>
<td>98</td>
<td>0.78</td>
</tr>
<tr>
<td>AR 1998</td>
<td>4501</td>
<td>62</td>
<td>0.80</td>
</tr>
<tr>
<td>AR 1997</td>
<td>3983</td>
<td>60</td>
<td>0.83</td>
</tr>
<tr>
<td>Average AR</td>
<td>5111.2</td>
<td>101.1</td>
<td>0.80</td>
</tr>
</tbody>
</table>
ACTION PLANS

Model calibration

- NOT OK
  - Model discrimination
    - NOT OK
      - Data stability
        - NOT OK
          - Re-estimate model
        - OK
          - Tweak model
    - OK
      - Re-calibrate model

- OK
  - Continue using model
KEY LESSONS LEARNT

• The best way to improve the performance of an analytical model is to improve **data quality** first
• A good model does more than giving good statistical performance (**model requirements**)!
• Discrimination versus calibration: bring the **macro-economy** into the model!
• Introduced the idea of **model risk**
• Discussed the need for **model validation** and **action plans**!
REFERENCES


• See [www.dataminingapps.com](http://www.dataminingapps.com)
COURSES

• Analytics in a Big Data World
  https://support.sas.com/edu/schedules.html?ctry=us&id=1339

• Advanced Analytics in a Big Data World
  https://support.sas.com/edu/schedules.html?ctry=us&id=2169

• Credit Risk Modeling
  https://support.sas.com/edu/schedules.html?ctry=us&id=2455

• Fraud Detection using Supervised, Unsupervised and Social Network Analytics
  https://support.sas.com/edu/schedules.html?id=1912&ctry=US
Self-Paced E-learning course: Credit Risk Modeling

See: [https://support.sas.com/edu/schedules.html?ctry=us&id=2455](https://support.sas.com/edu/schedules.html?ctry=us&id=2455)

The E-learning course covers both the basic as well some more advanced ways of modeling, validating and stress testing Probability of Default (PD), Loss Given Default (LGD) and Exposure At Default (EAD) models. Throughout the course, we extensively refer to our industry and research experience. Various business examples and small case studies in both retail and corporate credit are also included for further clarification. The E-learning course consists of more than 20 hours of movies, each 5 minutes on average. Quizzes are included to facilitate the understanding of the material. Upon registration, you will get an access code which gives you unlimited access to all course material (movies, quizzes, scripts, ...) during 1 year. The course focuses on the concepts and modeling methodologies and not on the SAS software. To access the course material, you only need a laptop, iPad, iPhone with a web browser. No SAS software is needed. See [https://support.sas.com/edu/schedules.html?ctry=us&id=2455](https://support.sas.com/edu/schedules.html?ctry=us&id=2455) for more details.
Credit Risk Modeling Using SAS

e-Course

Important: Do not close this window while you are taking the course. If you do, the course will stop functioning.

Technical Requirements

Desktop
- Windows 7 or Windows 8 with one of the following browsers:
  - Internet Explorer 9 or later
  - Chrome 34 or later
- MacOS with Chrome 36 or later
- High-speed internet connection
- Screen resolution set to 1024 x 768 or higher

Mobile
- iOS?

SAS Software
To complete the practices, you must have access to SAS software on the same machine where you are viewing the training. SAS software is not included with the training.

Test your system
The following requirements are automatically being tested for you. Check marks mean your system meets the requirements.

✔️ JavaScript
JavaScript is enabled.

✔️ Movie Player
You are ready to play movies.

✔️ Progress Tracking
Your progress is being tracked at SAS.
You will be able to use all tracking features within the course.

☐ Pop-Up Windows
Are pop-up windows enabled? Click below to find out.
  - [ ] Check Pop-Ups

Problems? Try pressing the Ctrl key while clicking. You can also visit our pop-up troubleshooting page.
GUT VORBEREITET IST HALB ZERTIFIZIERT –
ANTWORTEN ZUR SAS VISUAL ANALYTICS ZERTIFIZIERUNG
DOWNLOAD SLIDES AND RECORDING
WWW.SAS.DE/LUNCHTIME

HOW DID YOU LIKE OUR WEBINAR?