Machine Learning in Life Insurance

Searching for Patterns in Cash Flow Models

Pierre Joos

Hannover, 16. May 2019
Issue: Life insurance valuation is extremely complex

- **External data**
  - Capital market scenarios

- **Internal data**
  - Balance sheet data
  - Company strategy

---

1. **Analysis:**
   What drives the results?

2. **Prediction:**
   Can we predict the results?
Analysis:
Long-term low interest rates are worst case for insurer
Prediction:
GBMs can predict the PVFP for a given market scenario quite good
Life insurance policy

(Permanent) Life insurance:
Financial contract between policy holder and insurer where premiums are accumulated and paid out later

Cash-flows:
- Policy holder pays premiums until retirement
- Returns from investments are accumulated
- Insurer pays annuity after retirement

Uncertainties:
- Investment returns
- Longevity of policy holder

→ What is the total profit / loss of this policy?
Life insurance balance sheet in Solvency II

Policy view

Cash-flow profile of life insurance

Portfolio view

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital Gains</td>
<td>Own Funds</td>
</tr>
<tr>
<td>Market Value of investments</td>
<td>Market Value of future payments to PH</td>
</tr>
<tr>
<td>Other Assets</td>
<td>Other Liabs</td>
</tr>
</tbody>
</table>

1. What is the value of all assets?
2. Who will get this value?
Market changes impact balance sheet

1. Increase of capital gains on bonds
2. Stronger increase of market value of future payments to PH due to guaranteed payments with long duration

Low interest rate

Assets
- Market Value of investments
- Capital Gains
- Other Assets

Liabilities
- Market Value of future payments to PH
- Own Funds
- Other Liabs

Own Funds
- Capital Gains
- Market Value of future payments to PH

Interest rate on German government bonds with a duration of 10 years

© Allianz SE 2017
Cash flow models (CFMs) are central to valuation

Cash Flow Model

Monte-Carlo Simulation

- Assets $A_1, \ldots, A_N$
- Reserves $R_1, \ldots, R_N$
- Cash flows $C_1, \ldots, C_N$
- Profits $P_1, \ldots, P_N$

$\rightarrow$ Present Value of Future Profits (PVFP)

External data
- Capital market scenarios

Internal data
- Balance sheet data
- Company strategy

Detailed representation of investments and insurance products

Future management behaviour included (e.g. profit sharing, future asset allocation)

Options & guarantees are explicitly valued using financial math concepts

Distribution of PVFPs

Worst case

Risk capital

Own Funds

PVFP

Frequency
CFMs project complex annual decision processes…

- Shareholder
  - Available capital and buffers (e.g. free RfB), Solvency
  - Dividend
  - Change of buffers
  - book and market value of investments

- Investments
  - Buying and selling of assets, accounting effects
  - Adjustment of reserves

- Crediting Strategy
  - Guaranteed interest rates, technical results, reserves
  - Premium refunds, usage of technical buffers
  - Force realisation of gains/losses
  - Availability of capital gains/losses, investment result

- Liabilities
  - Projection horizon typically over 50 years
... and are applied to a variety of different market scenarios

- 1000 – 5000 market scenarios
- Typically 10+ market factors

- Cash flow model of Willis Towers Watson
- Average German insurer

- 1000 – 5000 PVFPs as output
- Stochastic nature of results
Machine learning is now able to tackle complex problems…

- Classification
- Regression
- Time series analysis
- Clustering
- Dimension reduction
- Image analysis
- Other
  - Text & speech analysis
...depending on the quantity and quality of the available data

Classical (linear) methods have advantages
- Better interpretability
- Faster calibration

Data quantity and quality is critical for the model choice
- Representative period of time
- Covering relevant special cases and not only the “average”
- Target variable as objectively as possible

Risk of overfitting with too complex ML models

⇒ Simpler models are always preferable
Agenda

1 Cash Flow Models
2 Analysis
3 Prediction
Which scenarios are responsible for extreme losses?
Traditional approaches provide limited insights

Traditional approaches:
- Sensitivity analysis
- Plot worst scenarios
- Multidimensional plots
- (Linear) regression

Issues:
- Only linear or 1-dim dependencies are identified
- Analysis often requires additional runs of the CFM
- Time-consuming manual analysis

Potential for Machine Learning?
The evolution of data analytics

Statistics

Analyzing quantitative information
- Descriptive analysis and hypotheses testing
- Scientific sub-topic of mathematics
- “Data generating process”

Data Mining

Explaining patterns in the data
- Information extraction from large data sets
- Visualization and structuring
  - “Patterns”

Machine Learning

Prediction of future based on experience
- Flexible models for complex data sets
- Model learn from data / experience
  - “Prediction”
What is a pattern?

- The algorithm searches for patterns which are in the scenarios of certain target (e.g. lowest OFs)

- A pattern was found for interest rates (Item IR)

- Colored horizontal bar indicates time range, gauge and value range:
  - Horizontal bar = average over that time within a certain range
  - Width: time range between year 0 and 15
  - Height: value range between ~3.5% and 0%

- Quality of this pattern is shown visually (color of bar) and as text in heading:
  - Coverage (cov=74%): fraction of targeted scenarios covered by this pattern
  - Purity (pur=65%): fraction of scenarios in this pattern belonging to the target
How to read the detailed output

- All scenarios matching the pattern are shown as lines
  - Blue lines = scenario in target
  - Grey lines = scenario not in target
- Total range of all scenario is shown as light grey background
- 65% of all scenarios matching this pattern are in our target (purity)
  - Visually represented by color of the bar
- 74% of all targeted scenarios are covered by this pattern (coverage)

Interpretation:
Falling interest rates are driver for bad OFs
Second pattern describes bad OFs

- Line indicates that interest rates change (Δ) over a the first 15 years
- Scenarios match this pattern if interest rates fall by 0.4% per year
  - Interest rates are down by approx. 6% after 15 y
  - Bandwidth is at ±5%
  - In this case this is equal to “IR at year 15 between −10% and 0%” (due to fixed rate at t=0)
- A scenario meeting this requirement has a probability of 63% to be in the target (purity)
- This pattern covers 52% of the targeted scenarios (coverage)
- Pattern #4 seems to describe same scenarios as pattern #3 before
Additional scenarios covered by new pattern

- Pattern #4 covers partly different scenarios than pattern #3
- E.g. IR scenarios, that
  - Rise first
  - Then fall abruptly
- Pattern #4 covers approx. 11% of scenarios in target that are not covered by pattern #3
- Marginal coverage can be only seen when sorting the found patterns
Overview plot shows all found patterns

- Projection horizon was segmented into three buckets (1–15, 16–30 and 31–50 years) to reduce computational workload
- Search tries to find a combination of pattern with best coverage and purity
  - Each pattern has an identifier (#)
  - Different patterns are alternative explanations (OR)
- Quality plot on the right hand shows coverage and purity of all patterns together
  - Size and color of individual points show coverage and purity of individual pattern
  - Pattern #4 increases total coverage but decreases total purity
- Interpretation: Worst OFs are driven by declining interest rates

→ Can we improve coverage further?
Residual plot shows scenarios that did not match a pattern

- Two groups of interest rate developments can be identified visually
  - Scenarios in the middle
  - Rising interest rates
- Search for further patterns should include combinations of variables and has to improve total coverage without diluting total purity too much
Pattern #5 has two conditions:
- Interest rates have to be between 8% and 12% in year 15 (vertical line)
- Equity index has to be less than 50% after 15 years (logarithmic scale!)

Only scenarios meeting both conditions (AND) are included in this pattern

This mechanism allows to identify interactions between variables

Individual purities of both conditions (ca. 20%) are weaker than joint purity (54%)

This pattern covers only a small part of our targets (coverage 7%)
Other combination found: RE decrease relevant when IR flat

Pattern # 6: IR,Ø 0-15 ∈ [-0.021...0.010] ^ RE, @ 15-30 ∈ [0.492...0.833] with Quality (n=94): cov=48%, pur=51% and score=-0.10
Overview of results

- Analysis of worst OFs
  - Data: Towers Watson, internal industry model
  - Closing: 2015 Q4
  - Scenarios: 1000 iterations (with neg. interest rates)
  - Target: 10% lowest quantile of OFs
  - Patterns with max 2 conditions
  - Patterns with min 50% purity

- Plot shows all found patterns
  - Falling interest rate as main risk driver visible
  - Negative real estate relevant if interest rates are flat
  - High interest rates only relevant if there is a significant equity shock at the same time (crafted scenario)
Only 2 unexplained scenarios remain

Found: 4 patterns with total coverage 98% and purity 52%
**How does it work?**

1. **Simplify data**
   - Segment time data into buckets (here 0–15, 15–30 and 30–50)
   - Define general patterns of scenarios within a bucket, e.g.
     - IR drops from 1% to around –5% within first bucket
     - IR in first bucket on average around –3%

2. **Find optimal parameters of patterns**
   - Optimal w.r.t. coverage and purity to a given target (worst PVFPs)
   - Using classical optimisation algorithm

3. **Find optimal combination of patterns**
   - Test all combinations of patterns in order to find combination effects, e.g. low interest rates together with losses on real estate
   - Criteria needed which combinations are preferred, e.g. as few conditions as possible (simple is better)
1. **Static analysis**
   - All funds are analysed based on closing runs
     - Model runs automatically in batch mode
     - Saves all graphics onto hard-drive
   - Allows quick first analysis of sensitivities

2. **Analysis of changes**
   - Analyse the changes from one quarter to the next
     - Run analysis on all funds for previous and current quarter
     - Analyse the changes in PVFP by scenario (if scenarios have same seed)
   - Can also be applied to what-if-calculations

3. **Model validation**
   - Using the analysis tool to validate cash flow model after model change
     - Run analysis on all funds and on multiple targets (low PVFP, high PVFP, …)
   - Results of analysis gives hints for further validation steps
Example 1: Protection product with IR up sensitivity

Found: 2 patterns with total coverage 93% and purity 72%
Example 2: Classic product with high guarantee

Found: 2 patterns with total coverage 92% and purity 73%
Example 3: Classic product with guarantee and fixed surrender value

Found: 2 patterns with total coverage 60% and purity 79%
Example 4: Fund with significant equity investments

Found: 1 patterns with total coverage 62% and purity 81%

Item EQ_EUR_TR

Quality

Level

Purity

Time

Coverage
Agenda

1. Cash Flow Models
2. Analysis
3. Prediction
### The evolution of data analytics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Data Mining</th>
<th>Machine Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyzing quantitative information</td>
<td>Explaining patterns in the data</td>
<td>Prediction of future based on experience</td>
</tr>
<tr>
<td>- Descriptive analysis and hypotheses testing</td>
<td>- Information extraction from large data sets</td>
<td>- Flexible models for complex data sets</td>
</tr>
<tr>
<td>- Scientific sub-topic of mathematics</td>
<td>- Visualization and structuring</td>
<td>- Model learn from data / experience</td>
</tr>
<tr>
<td>- “Data generating process”</td>
<td>- “Patterns”</td>
<td>- “Prediction”</td>
</tr>
</tbody>
</table>

\[
s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}
\]
What do we want to achieve

- Run time of cash flow models is very high
  - Projection of huge number of internal fields
  - Depends on granularity of assets and liabilities (model points)
- Store results from all calculations
- Apply Machine Learning algorithms to predict PVFP for a given capital markets scenario
- Use these models for quick calculations of PVFPs within the same quarter

Huge speed enhancement
Linear regression

Model assumes linear dependency between PVFPs and the capital market factors

\[ \hat{Y} = bX + a \]
Decision trees

A decision tree defines a hierarchical sequence of rules (decisions) on the capital market factors which branches out to a predicted value.

Method: rpart, $R^2$ (in sample) = 77.16%
Decision trees are quite flexible but weak

Properties

- Decision trees are generated very fast
- Typically not too complex (binary splits and few leaves)
- Are readable for a human
Random Forest

Random = Trees based on random subsets of features and data

Forest = Many trees (ensemble)

Prediction of a Random Forest is average of predictions of the individual trees
Neuronal network

Structured data processing

- Input layer has nodes (neurons) for each feature at each time step
- Output layers represents prediction
- Hidden layers react to patterns in the input
  - Exact pattern cannot be prescribed
  - Number of hidden layers determines predictive power → deep neural networks
Boosting + Bagging

Boosting (forward stagewise modelling)

Boosted tree

Method: blackboost, R² (in sample) = 91.3%

Bagging (Subsampling)

Bagged neuronal net

Method: brnn.bagged, R² (in sample) = 86.99%
Hyper parameter tuning

Without tuning

Method: regr.xgboost, $R^2$ (in sample) = 35.26%

With tuning

Method: regr.xgboost, $R^2$ (in sample) = 99.95%
Machine learning is not complicated

#1. Define data
mydata = CF-Model-Output-Data combined with Scenario Information

#2. Define tasks
tasks = list(makeRegrTask(data=mydata, target="PVFP"),…)

#3. Define methods
learners = list(makeLearner("regr.rpart"), …)

#4. Make the analysis (combine all tasks with all methods, do cross validation, compare e.g. MSE, MAPE, R^2,…) (bmr = benchmark(learners, tasks, cv10, measure= list(mse, mape, rsq)))

<table>
<thead>
<tr>
<th>task.id</th>
<th>learner.id</th>
<th>mse.test.mean</th>
<th>mape.test.mean</th>
<th>rsq.test.mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>mydata</td>
<td>1.183934e+16</td>
<td>0.9909817</td>
<td>0.2960641</td>
</tr>
<tr>
<td>2</td>
<td>mydata</td>
<td>6.879358e+15</td>
<td>1.1763703</td>
<td>0.6037440</td>
</tr>
<tr>
<td>3</td>
<td>mydata</td>
<td>4.693410e+15</td>
<td>0.8590164</td>
<td>0.7381586</td>
</tr>
<tr>
<td>4</td>
<td>mydata</td>
<td>3.775355e+15</td>
<td>0.8320113</td>
<td>0.7650239</td>
</tr>
</tbody>
</table>
Comparison of approaches

- Modern ML models are often easier to use and yield as good or better results than classical methods
- Overfitting is an issue which has to be addressed (e.g. using cross-validation, bagging, randomisation)
- ML models are often black boxes
Fund 1: protection product

**Statistics**

<table>
<thead>
<tr>
<th></th>
<th>NN</th>
<th>GBM_segment</th>
<th>GBM_full</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>0.996</td>
<td>0.994</td>
<td>0.986</td>
</tr>
<tr>
<td>AVG (%)</td>
<td>2827</td>
<td>3211</td>
<td>4168</td>
</tr>
<tr>
<td>MSE (%)</td>
<td>0.0073</td>
<td>0.0083</td>
<td>0.0108</td>
</tr>
<tr>
<td>STD</td>
<td>3697</td>
<td>4250</td>
<td>6986</td>
</tr>
<tr>
<td>NEG%</td>
<td>53.09</td>
<td>52.0</td>
<td>52.0</td>
</tr>
</tbody>
</table>
Fund 1: next quarter prediction

Statistics

<table>
<thead>
<tr>
<th></th>
<th>NN</th>
<th>GBM_segment</th>
<th>GBM_full</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{R2}</td>
<td>0.918</td>
<td>0.917</td>
<td>0.914</td>
</tr>
<tr>
<td>\textit{AVG} (%)</td>
<td>16835</td>
<td>16679</td>
<td>16725</td>
</tr>
<tr>
<td>\textit{MSE} (%)</td>
<td>17422</td>
<td>17494</td>
<td>17896</td>
</tr>
<tr>
<td>\textit{STD}</td>
<td>4558</td>
<td>5275</td>
<td>6442</td>
</tr>
<tr>
<td>\textit{NEG%}</td>
<td>99.82</td>
<td>100.0</td>
<td>99.27</td>
</tr>
</tbody>
</table>

Segmented Variable Importance

Time step importance of features
Fund 2: high guarantee product

Statistics

<table>
<thead>
<tr>
<th></th>
<th>NN</th>
<th>GBM_segment</th>
<th>GBM_full</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>0.987</td>
<td>0.986</td>
<td>0.96</td>
</tr>
<tr>
<td>AVG (%)</td>
<td>0.0465</td>
<td>0.0463</td>
<td>0.0474</td>
</tr>
<tr>
<td>MSE (%)</td>
<td>0.0614</td>
<td>0.0651</td>
<td>0.1181</td>
</tr>
<tr>
<td>STD</td>
<td>2725</td>
<td>2895</td>
<td>5245</td>
</tr>
<tr>
<td>NEG%</td>
<td>45.82</td>
<td>45.64</td>
<td>50.73</td>
</tr>
</tbody>
</table>
Fund 3: guarantee and fixed surrender value
Fund 3: next quarter prediction

Statistics

<table>
<thead>
<tr>
<th></th>
<th>NN</th>
<th>GBM_segment</th>
<th>GBM_full</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>0.44</td>
<td>0.459</td>
<td>0.511</td>
</tr>
<tr>
<td>AVG</td>
<td>4028</td>
<td>3299</td>
<td>3213</td>
</tr>
<tr>
<td>(%)</td>
<td>0.0369</td>
<td>0.0302</td>
<td>0.0294</td>
</tr>
<tr>
<td>MSE</td>
<td>4851</td>
<td>4768</td>
<td>4681</td>
</tr>
<tr>
<td>(%)</td>
<td>0.0444</td>
<td>0.0437</td>
<td>0.0429</td>
</tr>
<tr>
<td>STD</td>
<td>4162</td>
<td>4218</td>
<td>4218</td>
</tr>
<tr>
<td>NEG%</td>
<td>73.09</td>
<td>63.64</td>
<td>62.18</td>
</tr>
</tbody>
</table>

Segmented Variable Importance

Time step importance of features
Fund 4: equity focused product

Statistics

<table>
<thead>
<tr>
<th></th>
<th>NN</th>
<th>GBM_segment</th>
<th>GBM_full</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>0.815</td>
<td>0.827</td>
<td>0.857</td>
</tr>
<tr>
<td>AVG (%)</td>
<td>5868</td>
<td>5482</td>
<td>4723</td>
</tr>
<tr>
<td>MSE (%)</td>
<td>0.2586</td>
<td>0.3351</td>
<td>0.2967</td>
</tr>
<tr>
<td>STD</td>
<td>7859</td>
<td>7608</td>
<td>6570</td>
</tr>
<tr>
<td>NNEG (%)</td>
<td>46.0</td>
<td>43.45</td>
<td>49.09</td>
</tr>
</tbody>
</table>

Segmented Variable Importance

Time step importance of features
Fund 4: next quarter prediction

Statistics

<table>
<thead>
<tr>
<th></th>
<th>NN</th>
<th>GBM_segment</th>
<th>GBM_full</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>0.684</td>
<td>0.745</td>
<td>0.775</td>
</tr>
<tr>
<td>AVG (%)</td>
<td>10464</td>
<td>9640</td>
<td>8664</td>
</tr>
<tr>
<td>MSE (%)</td>
<td>13046</td>
<td>11716</td>
<td>10467</td>
</tr>
<tr>
<td>STD</td>
<td>12283</td>
<td>11553</td>
<td>10451</td>
</tr>
<tr>
<td>NEG%</td>
<td>63.45</td>
<td>57.45</td>
<td>54.55</td>
</tr>
</tbody>
</table>

Segmented Variable Importance

Time step importance of features