Application of machine learning in motor insurance pricing

James Rakow – Partner, Deloitte (email: jrakow@deloitte.co.uk)
Sulabh Soral – Director, Deloitte (email: ssoral@deloitte.co.uk)
Agenda

• Machine Learning – Why now?

• Some examples

• How does it impact motor insurance pricing?

• How does it impact our day jobs?
Machine Learning – Why now?
Information rich ecosystem

2.5x10^18 bytes per day
6.4bn IoT devices

75% is unstructured data
3.4bn connected people

7 June 2016
More information helps!

Opportunities

- Geographical opportunities
- Emerging risks
- New products
- Avoid the “bad” risks
- Understanding clients
- Individual solutions

Who owns the clients wins?
Does more information mean ready insights?

Ever increasing massive data lakes need to be harvested

Many different sources and formats to transform

High noise so need to look deeper

Many more combinations possible
Introducing machine learning

Source – SAS Institute
Introducing machine learning

Machine Learning vs. Statistical Modelling

• Emphasis on learning in machine learning

• Emphasis on inference in statistical models

• Process to build models is different because –
  – Come from different backgrounds (Subjects and history)
  – Many more a priori assumptions/expectations in statistical modelling
  – Focus on generalisation vs. fitting to distribution
  – Big data vs. limited sets
  – Machine learning vs. Hypothesis/human effort
  – Techniques used
    – In statistical modelling we start with an understanding of correlations and distribution and then optimise a predictor but in machine learning we start with no such understanding
Introducing machine learning

Types of ML

• Supervised - Is used with labelled data (All inputs are paired with outputs in a training set)

• Unsupervised - Is used with un-labelled data (Inputs not paired with any preconceived output)

• Reinforced - Learning the process to reach a desired outcome (In a way a mix of supervised and unsupervised)

• Others
Unsupervised Learning - Clustering

Unlabelled data

Clustered into “look-a-like” segments
A Game of Noughts (and crosses)
An example of reinforced learning

![Graph showing number of wins per 10 games, first win in under 5 minutes, and getting better at beating Alex.](Graph.png)
Some Examples
Machine learning toolkit

Core Technologies

Artificial Intelligence
Deep Learning
Machine Learning
Predictive APIs
Image Recognition
Speech Recognition

Innovative Industries

Finance
Bloomberg
alphasense
Education
coursera
declara
Automotive
BMW
Mobileye
Retail
Bay Sensors
PRISM SKYLABS
Oil and Gas
kaggle
AYASDI
Adtech
ADBRAIN
Yieldmo

Supporting Technologies

Hardware
NVIDIA
Qualcomm
Data Collection
IBM Bluemix
TRIFACTA
Software
Java
Python
Scala
Operating System
MySQL
mongDB
Hadoop
Database
Cloud
elastic
amazon web services
Microsoft Azure
Some examples of data based innovations supported by machine learning

Social data
inVenture

Sensors and images
Sentrian
Remote Patient Intelligence

Auditory
AUGURY
Machines talk, we listen.

Open
OpenCorporates
Impact on motor insurance pricing
Motor insurers will have to embrace machine learning

Adoption of machine learning for motor insurance pricing

- Changes in automotive technology
- Hybrid segments - due to multiple factors

Competition – Data Enrichment, Customer Experience
Notes from our ML experiences in motor pricing

Data Enrichment

• New sources of data – Machine learning used to transform data and find new patterns (Open, third party, others)
• Better lift curves

Better point of quote experience

• Reduction in number of policy holder inputs (“traditional data”) required – Machine learning used to find patterns in alternative data sources
• Without reduction in predictive accuracy

Better model building

• Relaxation in model assumptions as compared to traditional models
• More predictive accuracy
• Easy to integrate alternative data sources with Machine Learning in pricing models – Traditional tools are limited
• Interactions machine led not hypothesis led
• Non restrictive processes – can use multiple different models at pace
Notes from our ML experiences in motor pricing (continued)

Quicker model building and testing

- Faster turnaround time – Minimum human intervention required as compared to traditional tools: choosing variables, transforming/smoothing, setting interactions etc.
- Faster in testing too

Newer insights

- Cutting the data in a new way led to newer insights – needs segments that can develop new products

Other experiences from machine learning (apart from pricing)

- Commercial and operational excellence – Claims and Fraud
- Better customer relationships – Recommender systems
- New products and markets – Pay per mile
Things to consider

Prediction accuracy vs. inference
• Most machine learning algorithms are complex and blackbox
• Hard to explain – Note regulatory requirements

Technology
• High powered computing - costs
• However a lot of machine learning infrastructure is open

Talent
• Expertise in data science
• Programming

Pricing in motor insurance will undergo a process enhancement and as a first step machine learning can be used for data enrichment and open machine learning platforms can speed up building/testing GLMs - The way forward is a mix between the 2 approaches
How does it impact our day jobs?
Training

What do actuaries learn about Analytics?

Training

• How much actuarial work is – or should be – here?

Image borrowed from Drew Conway’s blog
Training

- As opposed to here?

Image borrowed from Drew Conway’s blog
“Insurance is one of the most data intensive industries. Historically actuaries and underwriters have always dealt with large data sets and have used these to make decisions. But in spite of this, insurance analytics rarely features on the business school curriculum. It is time that business schools recognise that there needs to be specialised courses on insurance analytics. A lack of skilled analysts in this sector is preventing the insurance industry from advancing.”

Insurance analytics – The missing link: Financial Times, October 2014
Questions?