Usage-Based Insurance
The International View

Marcus Looft (Milliman, German office in Munich, Bavaria)
Chris Cooksey (EagleEye Analytics, US office in Columbia, SC)

Casualty Actuarial Society:
Ratemaking and Product Management Seminar

Concurrent Session 6, 9:30am-10:45am
11 March 2015
Dallas

Milliman

High level structuring

1. UBI business in Europe
2. One technical challenge: (Big) Data
3. Another technical challenge: Statistical analysis of UBI portfolios
4. Case study: Differentiating UBI client profiles using Machine Learning on top of classical GLM models
5. Conclusions / Questions

Milliman

UBI business in Europe – A cultural thing

UK: 35 Mio cars, 1% UBI
Young driver mileage accounts, Mileage bonus for safe driving

CH: 4.3 Mio cars, >1% UBI
Crash recorder (Young driver)
Theft protection (Luxury cars)
Emergency call (Safety)

France: 32 Mio cars, <1% UBI
Attractive offerings for safe drivers
Roadside assistance

Spain: 22 Mio cars, <1% UBI
Fleets
Track&Trace

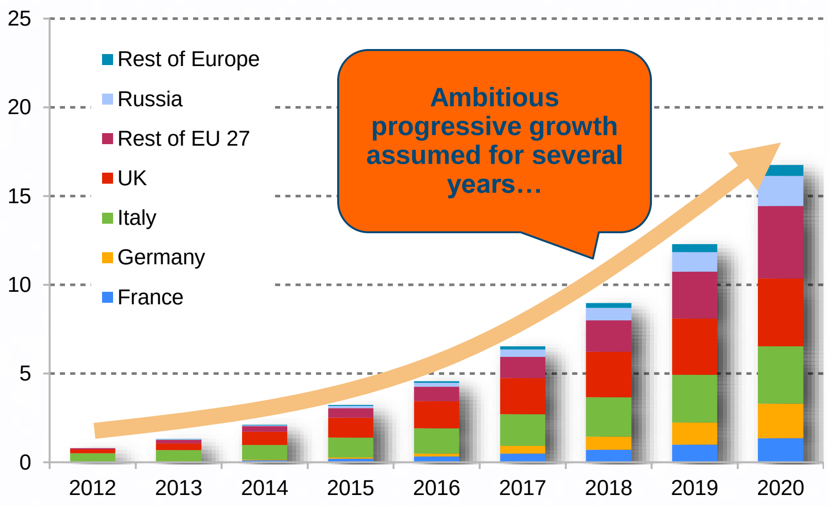
Germany: 43 Mio cars, 0.005%
Stubborn&conservative market, investment costs deemed high, waiting for strong car industry

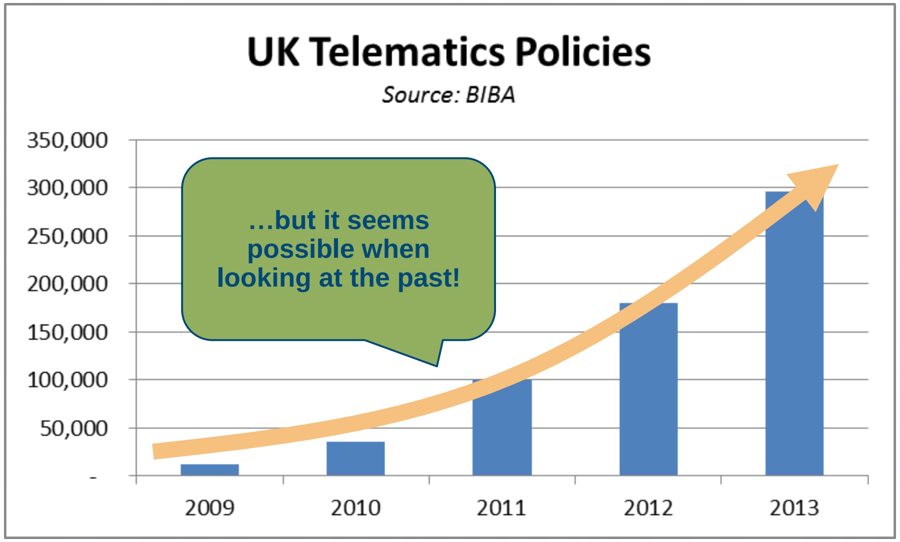
Austria: 4.5 Mio cars, >1% UBI
Technology-friendly country, Safety first

Italy: 37 Mio cars, 4% UBI
Early adopters, Attractive low mileage discounts (South of Italy)

Sources: 1) PRNewswire, Insight Report: Technology in Action - A Roadmap for Insurance Telematics
2) United Nations Economic Commission for Europe

Expected UBI growth in Europe





The Insurance Telematics (or UBI) will represent more than **35 million policies in 2020** or around **15% of the European personal lines market.**

Source: 1) Ptolemus USAGE-BASED INSURANCE Global Study 2013
2) British Insurance Brokers' Association (BIBA)

Milliman

UBI Business stories & interest

Story	Insurers Interest

Some challenges will remain

Some battle fields

Milliman

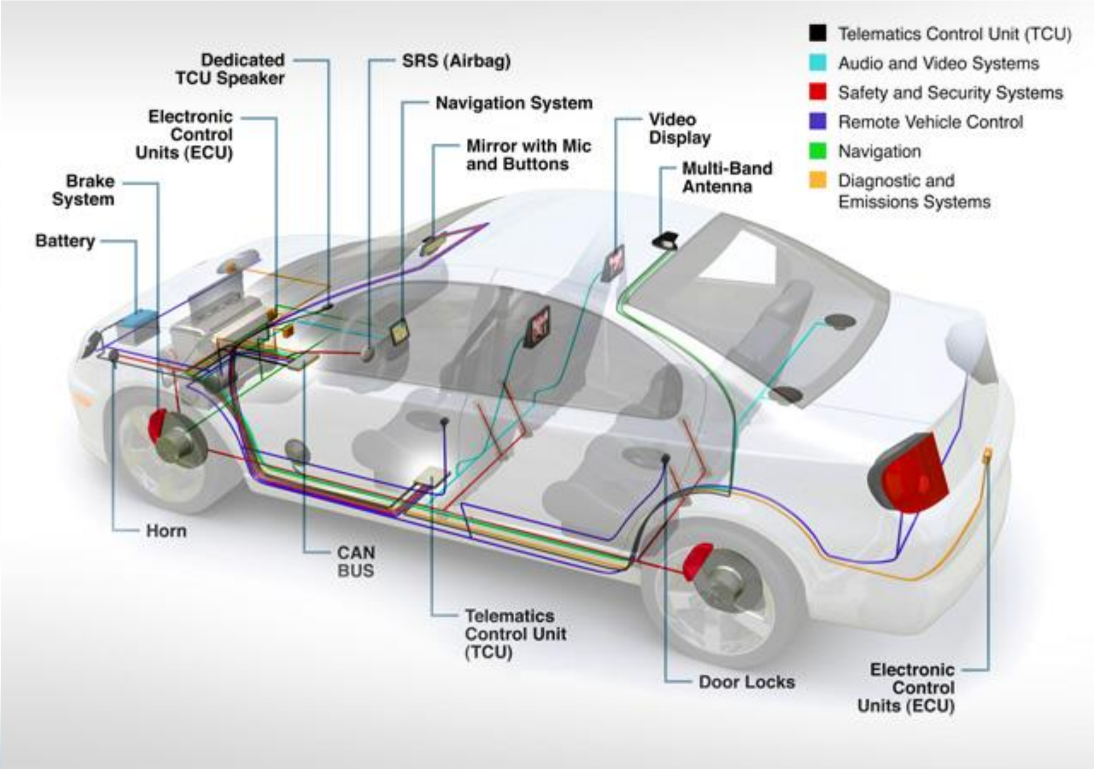
One technical challenge: (Big) Data

What companies want to achieve with UBI/Telematics data?

1. Actuaries: better assess the covered risks
2. Business: Offer additional services to enhance customer relationship

This is basically done in three steps:

1. Data is measured in the car
2. Data is transmitted from the car to a server
3. Data is processed & analyzed



Source: *http://servicesangle.com/blog/2012/07/09/carmakers-anxious-to-use-big-data-tech-from-big-biz-to-personal-perks/*

Data Preparation

Telematics & Big Data

Since the amount of data in telematics applications is enormous usually much of the measured data is discarded and condensed. This happens at multiple stages.

Devices that record acceleration values for accident investigation usually record these values at very high frequency (say 100 times per second) but only save the information if an acceleration threshold is exceeded (i.e. an accident happened).

Measuring a quantity every second, you get 300*3600=1’080’000 data points for one client driving 300 hours a year (average in UK)

Achieving Pay HOW You Drive

Sample Trip Summary Data – One Day

State Date	Start Time	Motorway Yards	Urban Yards	Other Yards	Motorway Seconds	Urban Seconds	Other Seconds	Total Speeding Yards	Total Speeding Seconds
3/3/2012	12:12:00	-	31						
3/3/2012	14:17:11	-	3,355						
3/3/2012	14:34:03	39,566	39,010						
3/3/2012	15:47:59	-	11,346						
3/3/2012	17:21:11	31,426	43,634						
3/3/2012	19:36:07	-	4,501						
3/3/2012	21:57:27	-	14,255						
3/3/2012	22:24:43	-	-						

A lot of room for modern pattern recognition. However this is a non-trivial task and certainly requires some use of Machine Learning algorithms!

Daylight ride

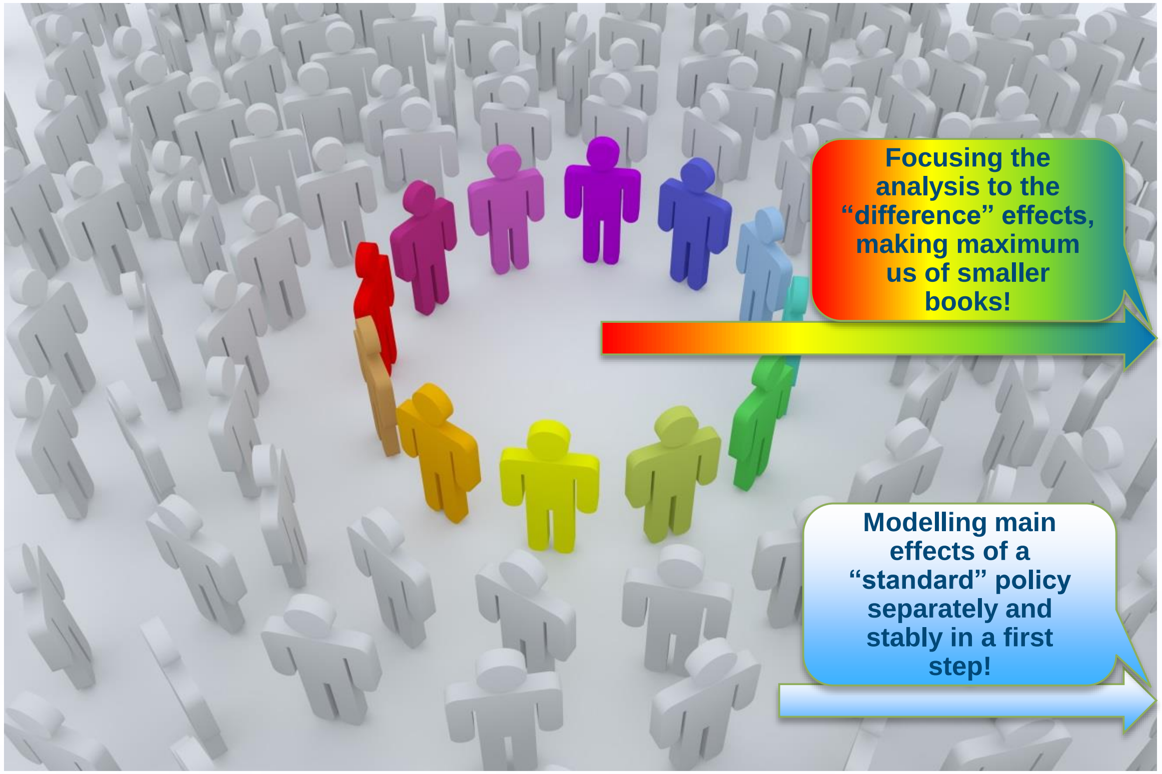
Urban yards

Aggregated statistics per year and risk record

Data source: *Beginner’s Roadmap to Working with Driving Behavior Data*; Jim Weiss, Jared Smollik

Another technical challenge:
Statistical analysis of UBI portfolios

- Rapid Pricing “Difference“ Diagnostics using Machine Learning:



Combine regular policy pricing with Telematics data analysis*:

This needs new modeling technology!

Price Difference = Loss Ratio = Telematics Claims / TP **(Telematics)**

Technical Price = TP **(Standard Policy)**

***Such analysis cannot be done with classical methods like GLMs because**
a) Cost effecting, complex interactions within the Telematics data can only be detected automatically (through Machine Learning)
b) The price difference cannot be fitted by a GLM-Distribution
c) Correlation between Telematics and Non-Telematics effects will disturb clarity of results in a single GLM. Furthermore distribution over different frequency and severity models confuse the difference analysis of Telematics policies.

Side bar: Why Machine Learning?

- Telematics data is new to the industry
- Automated approaches can be useful for field selection
- Data mirrors real life, and real life is about interactions
  - Drivers with a speeding violation are worse risks (on average).
  - This pool of drivers with a speeding violation are not homogenous…
    - Some speed on highways; some on rural road
    - Some speed constantly and got caught once; some just had a bad day
    - Some speed during the day; some speed at night
  - In other words, the importance of this indicator (having a speeding violation) will be different for different drivers, and the dependencies become ever more important with additional data.
- The importance of local interactions

Side bar: Why Trees?

Machine learning has many, many approaches. Trees are useful because:

- Trees are all about local interactions.
- Single trees can be simple and transparent. Relationships are there to see.
- Boosted trees can be smooth and powerful, the results stable.
- Even boosted trees are transparent, even if they are complex.

Remember that all automated routines run an extra risk of overfitting the data. You *must* validate these models.

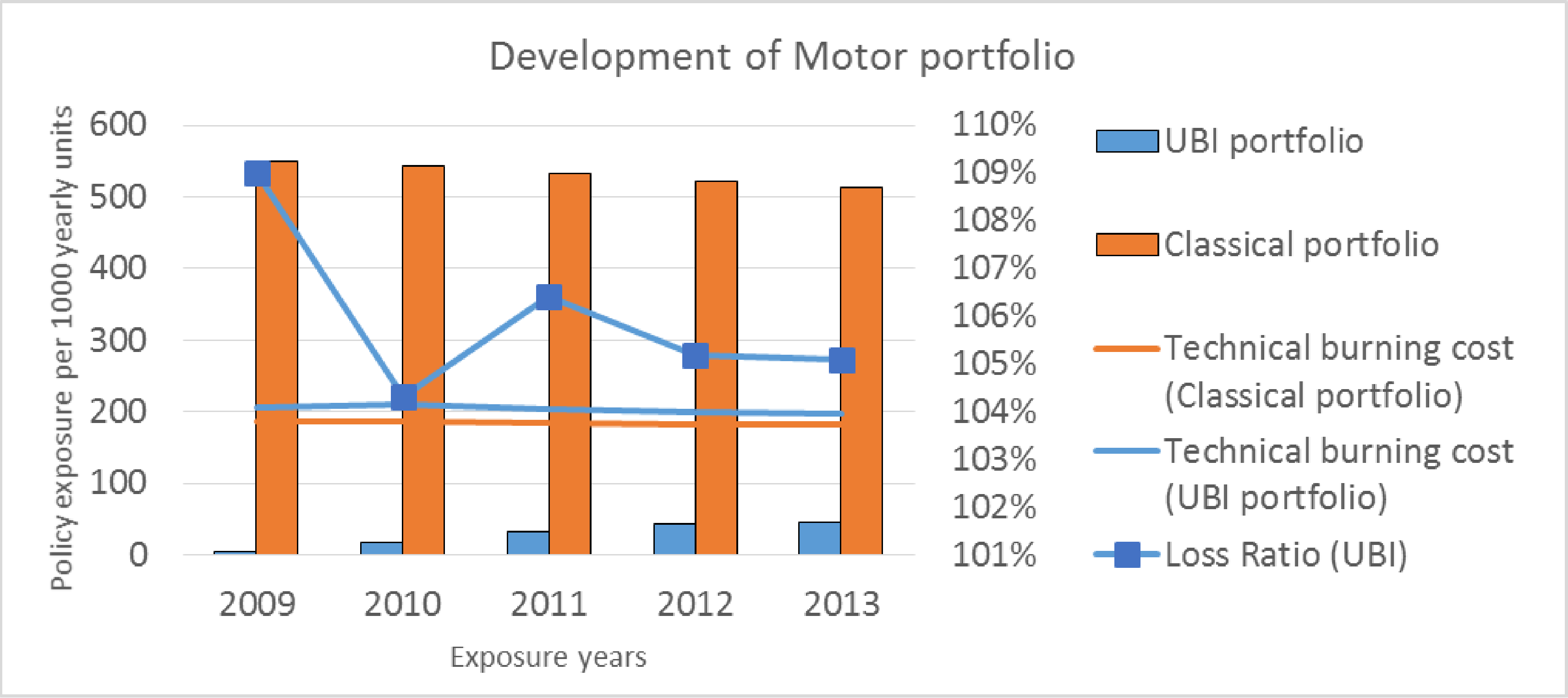
Milliman

Case study: Differentiating UBI client profiles using Machine Learning on top of classical GLM models

- European client with private UBI Motor business. Next to classical risk data the UBI had some aggregated telematics data
  - E.g. yearly mileage, %day rides, Preferred road type, Number of trips per year, etc.)
- Overall the UBI business for this client is more expensive than the standard business (currently a show-stopper for further sales boosting)
- The goal was to find better risk differentiators using Machine Learning algorithms than the ones that had already been found with the GLM techniques

Case study – UBI portfolio vs classical portfolio

Development of technical burning costs (based on claims paid and IBNR)



The UBI portfolio has 5% higher technical burning costs compared to the normal business
UBI investment not included in this calculation!
Incentive discounts on UBI price to grow this segment also not included!

Case study: Europe

For the client we have built two Loss Ratio (LR) models, namely

1. Boosted tree*: LR = UBI claims / GLM (GLM with Telematics data on whole portfolio)

   Hereby testing the general strength of Machine Learning on top of the Telematics effects already in a GLM

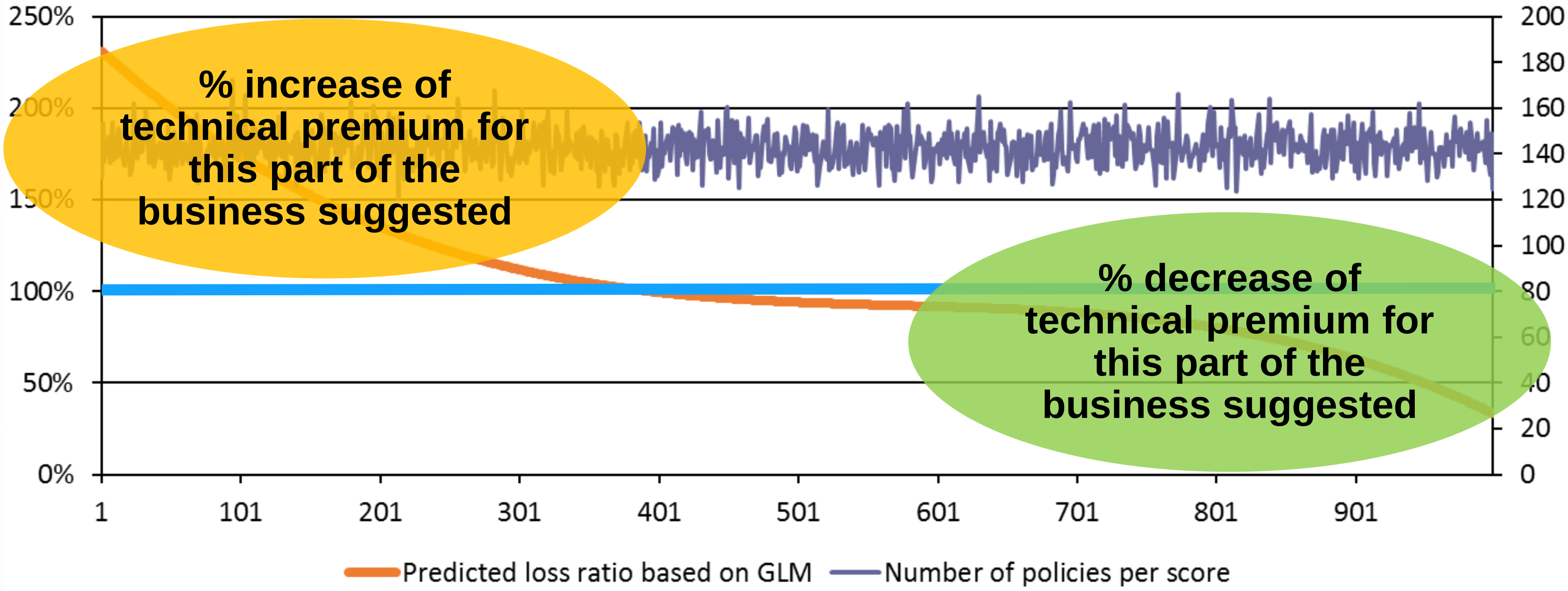
2. Regression tree*: LR = UBI claims / boosted tree (boosted tree without Telematics data only on standard policies)

   Hereby indicating the very profitable and unprofitable UBI client segments relative to the technical pricing for a standard policy

*Created with EagleEyeAnalytics Talon Pricing Software on the Telematics policy data

1. Model: How much additional "signal" was found with Machine Learning on the UBI book?

The score on the x-axis represents the ordered Machine Learning segmentation* ranging from much more expensive UBI policies to much less expensive UBI policies relative to the technical burning cost coming from the GLM (Telematics data included as main effects)

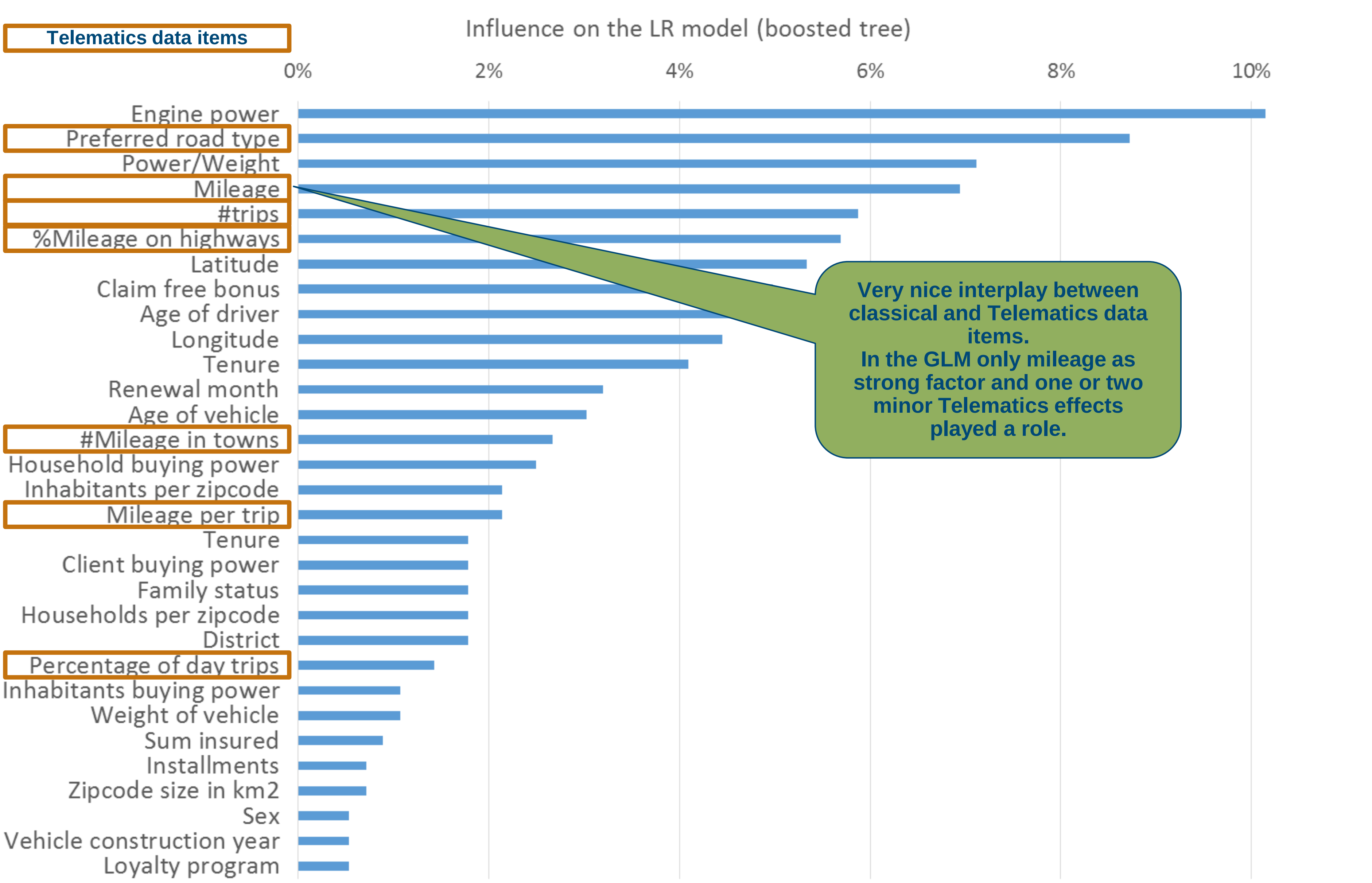


Using Machine Learning found local interactions among the risk factors even though mileage was already included in the GLM as the strongest Telematics factor!
We can learn about interesting new segments in the UBI book usually not visible through GLMs

***Created with EagleEyeAnalytics Talon Pricing Software using a boosted tree (Ensemble method)**

Milliman

Significance of classical and UBI factors



2. Model: Identify interesting UBI segments

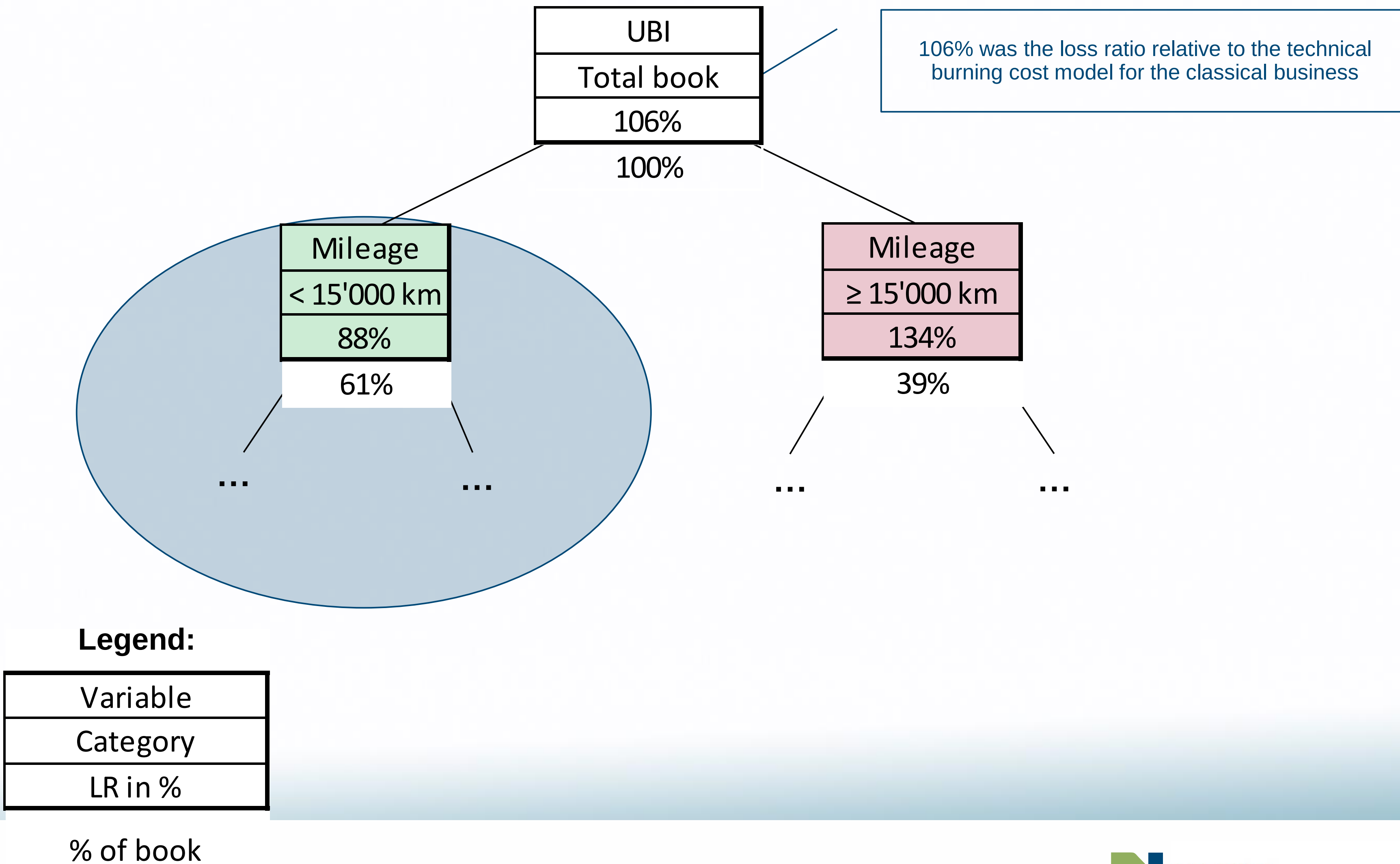
- The UBI portfolio had a 5% worse technical loss ratio than the classical portfolio*
- But there are interesting segments that had a very good and a very poor loss ratio. Some are listed in the following:

LR*	Volume	Description

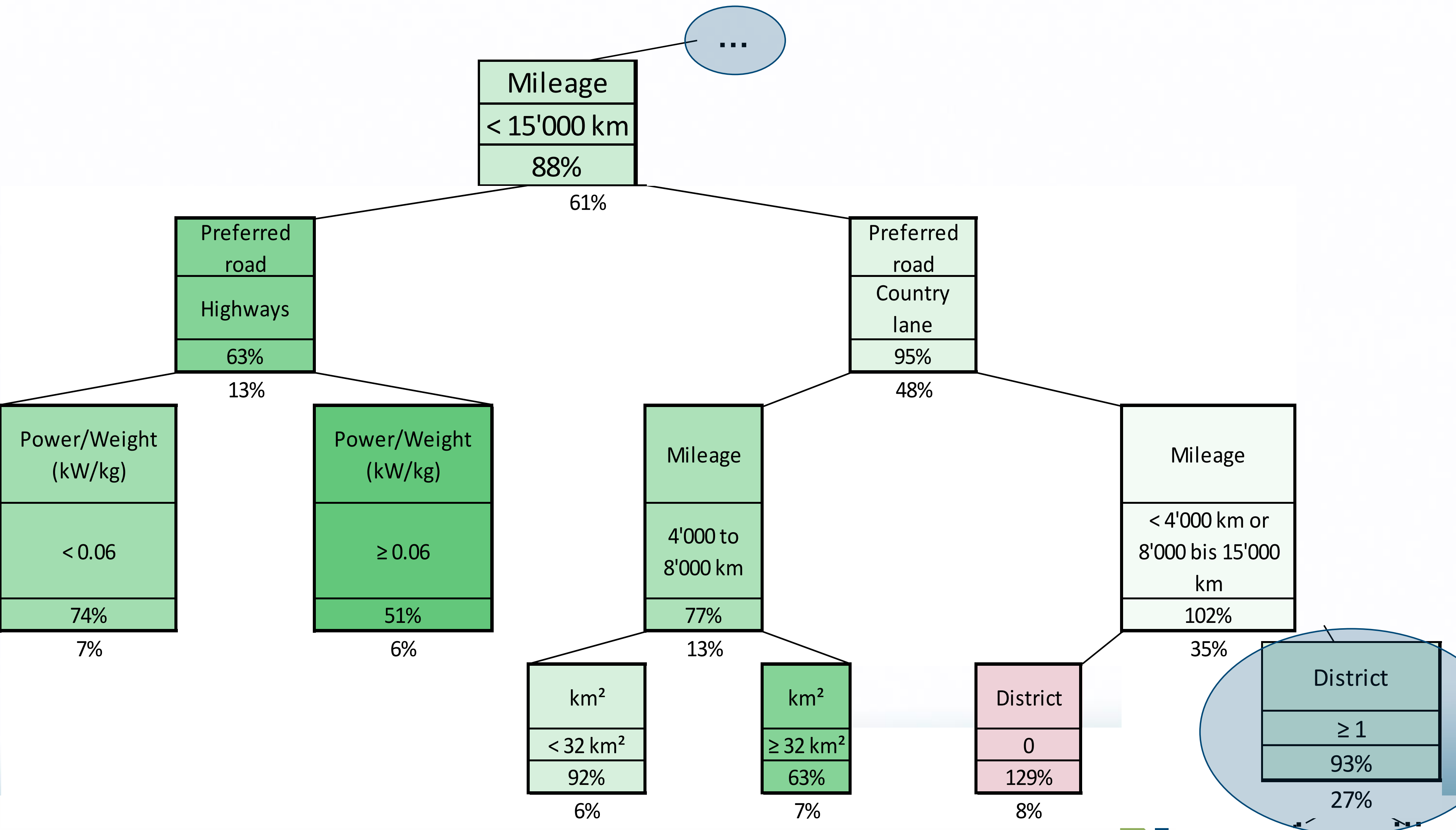
*Compared to the technical burning cost of the classical portfolio

Milliman

2. Model: As a tree (Root)

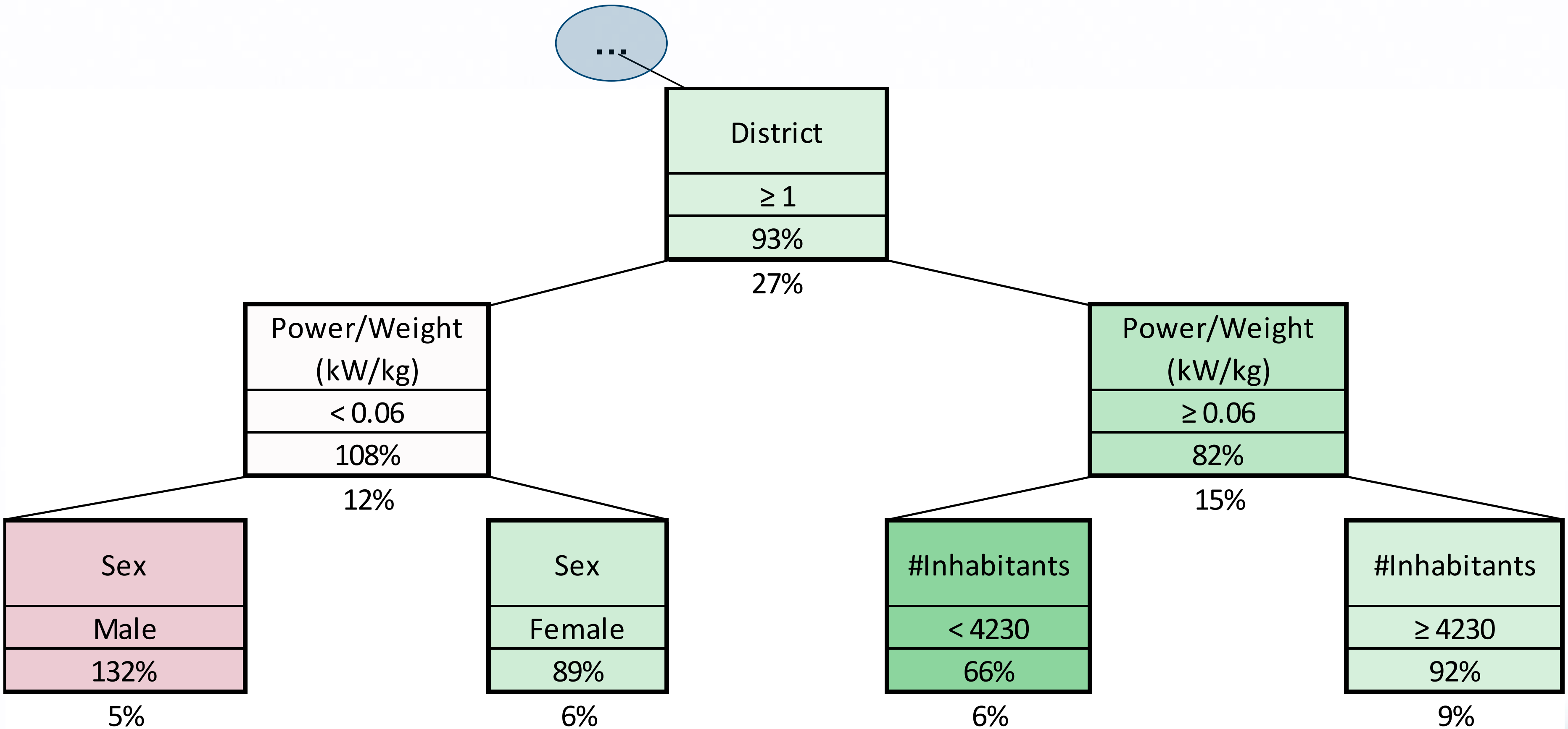


2. Model: The good guys (left branch)

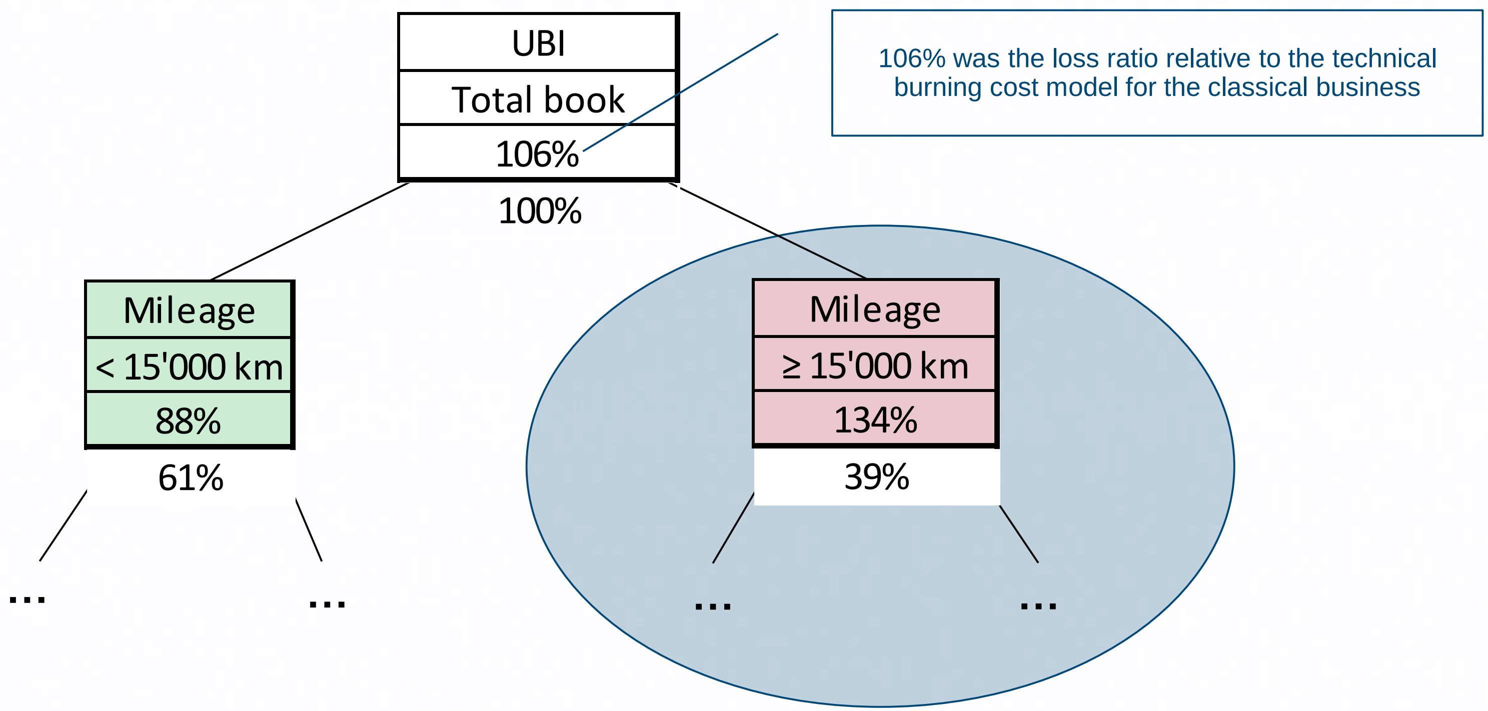


Milliman

2. Model: Still some good guys (left branch)



2. Model: The poor guys (right branch)



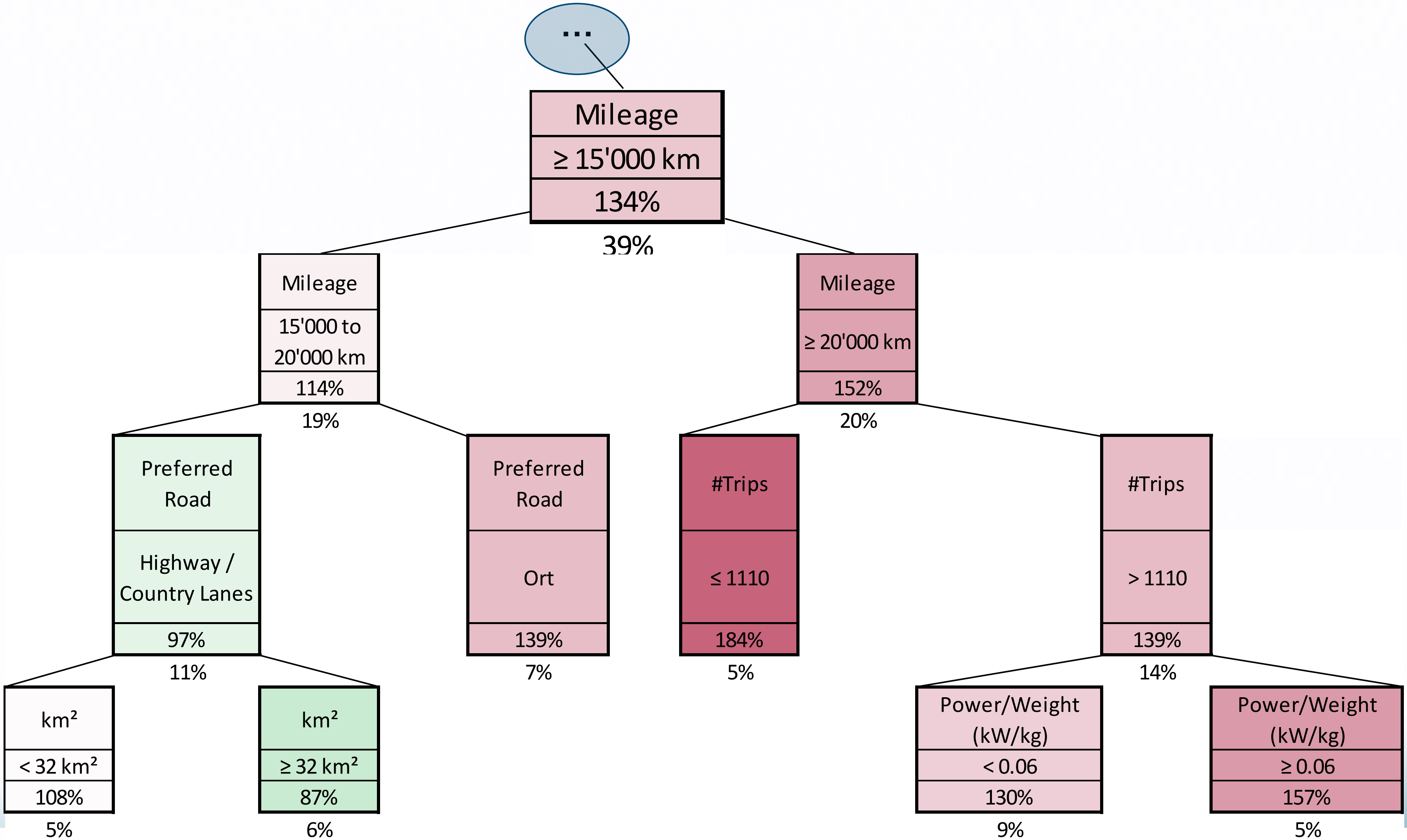
Legend:

Variable
Category
LR in %

% of book

Milliman

2. Model: The poor guys (right branch)



Conclusions

- UBI business is still a child (not fully grown up at least in Europe)
- UBI comes with big data and this mine is barely tapped
- Using Machine Learning can bring new insights that are truly natural and intuitive (and not necessarily artificial)
- Comments / Questions ?

Thank you!

marcus.looft@milliman.com