RULING THE WORLD OF CAR INSURANCE

Strata Data Conference, New York, 2017

Riccardo Gianpaolo, Corbella | Big Data Engineer
Beniamino, Del Pizzo | Big Data Engineer
MEETING AGENDA

- Introduction
  - Traditional car insurance
  - Telematics

- The Scenario
  - Exploiting the data

- Our Solution
  - Architecture Overview
  - Crash Validation
  - Take advantage of trip informations

- Conclusions
YOUR SPEAKERS

Riccardo Gianpaolo, Corbella
Big Data Engineer @ Data Reply IT
M.Sc. In Computer Science

Beniamino, Del Pizzo
Big Data Engineer @ Data Reply IT
M.Sc. In Computer Engineering
- 100+ Big Data Engineers & Data Scientists
- Strong partnership with Cloudera, Hortonworks and Microsoft
- Strongly investing on Apache Spark (sponsor of the first EU Spark summit)
- Huge investments in training & recruiting of young talents (strong partnership with the major European Universities)
INTRODUCTION
CAR INSURANCE offers protection against physical damage or bodily injury resulting from traffic collision and against liability that could also arise from it.

Optionally protects against theft and damage to the vehicle sustained from events other than traffic collisions.
THE PREMIUM

**INSURANCE PREMIUM** is the amount of money that an individual or business must pay for an insurance policy.

Factors that influence the premium calculation are:

- Driver profile (age, history, …)
- Car characteristics (engine size, horse power, …)
- Selected coverage
- Geographic area
WHAT’S WRONG WITH IT?

- High possibility of fraudulent intents by a driver
- Little knowledges about the customers
- Few data and only standard informations used to compute the premium
- Standardized policies not tailored to the customers
- Claims evaluated manually (long times for resolution)
- High risk of loss for companies
The combination of telecommunications and informatics to facilitate an efficient transfer of information over vast networks gave birth to **TELEMATICS**.

This is **DISRUPTING THE TRADITIONAL CAR INSURANCE** business model and companies are learning from the experience of the first movers in order to develop their own telematics experience.
BLACK BOX IS THE KEY FACTOR
IT’S ALL ABOUT DATA

BLACK BOX

Enables the collection of raw data which is transformed into actionable knowledge.

Affects all components in the insurance value chain and bring an unprecedented level of innovation to motor insurance.

KNOWLEDGE MEANS INNOVATION

TURN ON THE WIN WIN WIN MODEL
WIN WIN MODEL

Insurance company point of view

- Fraud reduction and/or inflated claims
- Risk-based selection / pricing
- More efficient claim handling process
- Customer loyalty

Customer perspective

- Driving assistance (e.g. proactive assistance in case of crash)
- Faster claims resolution
- Anti-theft service
- Alerts (i.e. weather, speed limits, road dangerousness, …)
- Drive gamification
IN THE END IS ALL ABOUT MONEY...
MORE THAN 5.4 MILLION OF BLACK BOXES IN ITALY

Highest penetration rate of telematics clients in the world (16%).

Number of telematic clients is quickly growing up: in the years 2015-2020 +15% of new customers per annum with an expected penetration rate of 28% in 2020. [ANIA]*

*Associazione Italiana Imprese Assicuratrici
THE RULER OF ITALIAN MARKET

Installed Black Boxes in Italy

- Our customer: 63%
- Others: 37%

OUR CUSTOMER IS THE MAIN PLAYER

3,4 million of black boxes
Has been estimated that will able to reach more than 4 million of clients within the end of 2018

[ANIA]
THE SCENARIO
DATA ON HAND

- Trip
- Theft
- Failure
- Crash
- Alert
- SOS Call
EXPLOITING THE DATA
FOR THE COMPANY

- Knowledge creation
- Customer behaviour and habits
- Capture the client needs
- Roads dangerousness
- Benefits of the insurance bottom line
- Development of a new risk model
- Better claims description
- Claims cost reduction
EXPLOITING THE DATA
FOR THE CUSTOMER

- Reward mechanism and Gamification
- Pay as you drive
- Pay how you drive
- Value added services
- App on smartphone
- Car maintenance monitoring system
- Real time coaching
<table>
<thead>
<tr>
<th>Message type</th>
<th># messages per day</th>
<th>Size per day (Byte)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crash GPS positions</td>
<td>560 K</td>
<td>~ 3 GB</td>
</tr>
<tr>
<td>Crash gyroscopes</td>
<td>250 M</td>
<td>~ 50 GB</td>
</tr>
<tr>
<td>Crash accelerations</td>
<td>300 M</td>
<td>~ 60 GB</td>
</tr>
<tr>
<td>Trip GPS positions</td>
<td>4 M</td>
<td>~ 30 GB</td>
</tr>
</tbody>
</table>
Double challenge

- Acquire skills to manage a huge amount of data
- Introduce a new approach for product development
OUR SOLUTION
REAL-TIME MONITORING

**KAFKA** is the **HEART** of our solution since it provides scalability and fault tolerance.

Real-time analysis and metrics about the received messages with Spark:
- Conduct of crash validator module
- TSP performances and meet of SLA requirements

Kafka lag – how our streaming apps are keeping the pace with incoming data:
- Allows to deal with failures in real-time
- Enable the supply of critical services without delays (time constraints preserved)
CRASH VALIDATION
DESIGN

ENRICHMENT MODULE

Enriched Crash

Confirmed Crash

Call Center

kafka

CRASH VALIDATOR

Spark

HBASE
A black box samples GPS coordinates and accelerations every 10/20 ms BUT if nothing in particular occurs it just send a subset of the informations collected (~ 1 position every 2 Km).

**IF A CRASH EVENT IS DETECTED:**

- All the sampled positions and accelerations that refer to the 3s before and after the climax are considered
- Collected informations are grouped in chunks and sent as crash messages
- Crash messages pass trough the tsp
### CRASH VALIDATION

#### MESSAGE FORMAT

<table>
<thead>
<tr>
<th>HEADER</th>
<th>Customer ID</th>
<th>Crash ID</th>
<th>Chunk Number</th>
<th>TSP Timestamp</th>
<th>TSP Crash Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAYLOAD</td>
<td>GPS Position</td>
<td>Acceleration</td>
<td>Gyroscope</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Real-time validation of crashes: false positive detection.

Validation requires the computation of the followings KPIs:

- Max speed
- Max acceleration
- Acceleration variance
- Impact time

The accuracy of the algorithm is improved by:

- Tracing out damaged black boxes to discard distorted informations
- The whole signal is reconstructed (messages can be received with delay or out of order)
For each true positive we send an alert to the call center. Every assistance sent for a false positive event means losses for the company. The reached accuracy is 84.6% compared to 75% of the TSP, furthermore we are able to fire 20% less alerts related to false positive.
NOT EVERY TRIP ENDS IN A CRASH
PAY HOW YOU DRIVE

Stop to classify customers using only static data such as age, location, health!
Street Dangerousness is identified by computing and keeping updated the followings KPIs:

- Number of crashes
- Travel speed
- Common weather conditions
- Local driver’s habits

A mobile app provides statistics (Pay as You Drive) and reports about street and trip to the customers.

HBase is used as central hub for KPIs management and to feed the app.
TO SUMMARIZE
While telematics is all about data, the actual data registered by the connected car’s sensors is only raw material.

The car insurance sector has established the most successful use case for telematics data so far by transforming the collected data into actionable knowledge affecting all components in the insurance value chain and bringing an unprecedented level of innovation to motor insurance.
BLACK BOX - THE EFFECTS

- Faster and more accurate claims management process
- Behavioural changes of drivers
- Reduction of:
  - Frequency of claims
  - Risk for companies
  - Fraud and/or inflated claims
  - Overall premium cost
The Italian market has showed that black boxes in vehicles reduce the claims frequency by an average of about 20% compared to the previous era, all other factors being equal. [ANIA]

The claims frequency of vehicles with black boxes and drivers aged between 18 and 25 living in Naples is 6.08%. For vehicles without black boxes and drivers of the same age living in the same province, the claims frequency is 8.31%. These values are equivalent to a total raw difference of 26.8%. [ANIA]
Italy is at the forefront in this digital transformation due to the innovative usage of telematics data to offer new services and improve the existing ones.
OUR EXPERIENCE

The architecture has proved to be fault tolerant and effective to manage the huge amount of data and to offer cutting edge services:

- Conduct of crash validator module
  - Allows saving money for the companies by cutting the costs of unnecessary rescues
  - Provide instantaneous assistance in case of accident

- Driver ability score:
  - Pay How You Drive

- Statistics
  - Pay As You Drive
Consumers will continue to gain a better understanding of telematics and companies will find innovative ways to leverage the data collected by the corresponding devices far beyond the motor landscape.
A LOOK AT THE FUTURE

Telematics can be seen as a door opener to many new opportunities for both consumers and customers of different insurance fields.
THANK YOU

Strata Data Conference, New York, 2017
www.reply.com