### How can a risk score model of traffic accidents reduce the amount of accidents? POC on city-level.

Business Understanding

In 2008, accident related costs in Germany were 30 billion € [[0](https://web.archive.org/web/20120330144056/http://www.bast.de/DE/Publikationen/Forschung-kompakt/2011-2010/2010-17.html)]. A major percentage of these costs are covered by car insurances. It is evident that these insurances have an interest in reducing the quantity of accidents to save costs. A recent development in the insurance market are ‘telemetry contracts’, where driving behaviour according to traffic regulations is incentivised. Cars of contractors are equipped with small hardware which monitors driving behaviour. If e.g. the speeding limit is not exceeded, the telemetry contract gets 10% cheaper as risk of accident for the insurance company is reduced.

In the same manner we propose a telemetry contract based on the statistical risk of accidents at the current place of a car and environmental conditions. The driver gets warned that he is approaching a risk zone prompting a higher attention level and thus a reduction in accidents.

French version :

En 2015, les coûts liés aux accidents aux USA se sont élevés à 242 milliards d'euros [0]. Un pourcentage important de ces coûts est couvert par les assurances automobiles. Il est évident que ces assurances ont un intérêt à réduire le nombre d'accidents pour économiser des coûts. Une évolution récente sur le marché de l'assurance est celle des "contrats de télémétrie", qui incitent à adopter un comportement de conduite conforme au code de la route. Les voitures des entrepreneurs sont équipées d'un petit matériel qui surveille le comportement de conduite. Si, par exemple, la limite de vitesse n'est pas dépassée, le contrat de télémétrie est 10 % moins cher car le risque d'accident pour la compagnie d'assurance est réduit.

De la même manière, nous proposons un contrat de télémétrie basé sur le risque statistique d'accident à l'endroit où se trouve la voiture et sur les conditions environnementales. Le conducteur est averti qu'il s'approche d'une zone à risque, ce qui lui permet d'être plus attentif et donc de réduire le nombre d'accidents.

**Mathieu**

We show some numbers to demonstrate the attractiveness of this idea to insurance companies.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Type** | **QTY**  **(abs.)** | **QTY**  **(%)** | **Cost per Acc. (€)** | **Insurer Cost per Acc. (0.5) (€)** | **Insurer Cost per Acc. Type (Mio. €)** |
| **fatal** | 22 | 0.05 | 1,226,624 | 612,312 | 13.4 |
| **severe injury** | 585 | 1.29 | 133,501 | 66,750 | 39.1 |
| **slight injury** | 4683 | 10.30 | 11,433 | 5,716 | 26.8 |
| **without injury** | 40164 | 88.36 | 6,479 | 3,239 | 130.1 |
| **Total** | **45454** |  |  |  | **209.4** |

Cost of accidents for Munich in the year 2017.

Source : Qty from [[1](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj7urb3nLfsAhWqx4UKHUGUDBcQFjABegQIAxAC&url=https%3A%2F%2Fwww.muenchen.de%2Frathaus%2Fdam%2Fjcr%3A71bb021f-4dec-4eeb-a3da-ed55499b612f%2FMue.Sta_3Q18_2_Verkehrsunf%25C3%25A4lle.pdf&usg=AOvVaw2fXFYPXRcp3ZYVVm70OMKI)], cost calculation from [[2](https://www.safetycube-project.eu/wp-content/uploads/SafetyCube-D3.2-Crash-costs-estimates-for-European-countries.pdf)]. The cost from [[2](https://www.safetycube-project.eu/wp-content/uploads/SafetyCube-D3.2-Crash-costs-estimates-for-European-countries.pdf)] include socio-economic cost (human capital loss) which are not all covered by car insurances. We apply the factor 0.5 based on the ‘without injury’ category as they are the main block in the distribution of accidents. [[3](https://www.fischer2.com/nachrichten/112-versicherung/kfz/680-die-schadensh%C3%B6he-bei-verkehrsunf%C3%A4llen-steigt.html#:~:text=Durchschnittlich%20kostete%20jeder%20Kfz-Haftpflichtschaden,Prozent%20mehr%20als%20in%202014.)], [[4](https://de.statista.com/statistik/daten/studie/38507/umfrage/durchschnittliche-kosten-je-autounfall-seit-1995/)] report that the average cost for basic insurance in Germany was around 3300€. However, costs for partial and full insurance as well as technical experts and lawyer fees are not taken into account. With a factor of 0.5 we get an overall weighted average cost per accident of ca. 4600€ which is reasonably close to the reported 3300€ (which does exclude certain costs.)

In the following table we present how the insurer’s cost changes with an assumed reduction of accidents of 1%. The difference between total cost shows a 2.1 Mio. € opportunity in cost reduction.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Type** | **QTY**  **(abs.)** | **QTY**  **(%)** | **Cost per Acc. (€)** | **Insurer Cost per Acc. (0.5) (€)** | **Insurer Cost per Acc. Type (Mio. €)** |
| **fatal** | 22 | 0.05 | 1,226,624 | 612,312 | 13.3 |
| **severe injury** | 579 | 1.29 | 133,501 | 66,750 | 38.7 |
| **slight injury** | 4,636 | 10.30 | 11,433 | 5,716 | 26.5 |
| **without injury** | 39,762 | 88.36 | 6,479 | 3,239 | 128.8 |
| **Total** | **44999** |  |  |  | **209.4** |

Cost of accidents for Munich in the year 2017 with an assumed 1% reduction of accidents. Greyed out columns are the same as in the previous table.

**Venkastesh**

A major question is on what data the risk score will be based on and even more important the cost of data-acquisition. The model will be based on plain police reports which poses several advantages:

* data already exists and due to legislation and continuous stream of new data is guaranteed
* due to the responsibility of police reports in legislative processes the data quality is sufficient
* Cities can sell their police report data to insurances, creating an additional income stream. Besides the financial aspect, cities have an intrinsic motivation to limit the number of accidents and their negative consequences to keep the city attractive for living and industry. Insurers will use the report data to do exactly that
* Lastly, not much effort is required from the city. No additional infrastructure to manage and police reports need little modification for anonymization to comply with privacy regulations

This triangle between insurance companies, cities and citizens creates a Win-Win-Win situation:

* The insurer reduces cost and stays competitive with a new and cheaper contract offer for customers
* The city generates additional income with little effort and gets indirect help in managing accidents
* The customer profits from cheaper contracts and additional safety while driving.

To summarise, just for the inner city of Munich, with a conservative assumption about the impact of a risk score based warning signal the cost reduction opportunity is 2.1 Mio €. The cost of data acquisition must be subtracted but will be only a small fraction of the potential cost reduction.

Of course there exists not only one insurance company in Munich which could yield the cost reduction. The key is that the risk score can easily be scaled to a higher level than a city e.g. to state- or even country-level as police reports exist everywhere.

**Louis**

**Data Understanding**

There are several interesting datasets of police data on accidents. We opted for a version on accidents in Chicago from 2015-present. With close to 450,000 there is no risk of scarcity. Also, it includes rich properties as presented in the following table.

|  |  |
| --- | --- |
| Crash\_Date | MM/DD/YYYY HH:MM:SS AM/PM |
| Weather\_condition | clear, rain, cloudy, snow, sleat, hail |
| Lighting\_Condition | Daylight, Dawn, Darkness, Lighted Road |
| Roadway\_Surface\_Cond | Wet, Dry, Snow/Slush |
| Latitude. Longitude |  |

**Risks**

2 major risks:

I) Performance of model. We identify 2 categories of influences on accidents a priori. Human behaviour (attention level, tiredness, intoxication, mood) and environmental factors (road condition, time, place, weather condition). Data on environmental factors is available in contrast to human behaviour which is difficult to model today. We will have to check in detail how much accidents depend on human behaviour and adjust the performance expectation of the model accordingly.

II) Acceptance of the model. The model needs to be well tuned, finding the balance between warning before entering the dangerous zones and not warning all the time as this would lower user acceptance.

If human behaviour influence too large/correlation of environment too small, system will warn continuously and not get taking serious by human.

Methods:

One-group classification/SVM, EDA

<https://stats.stackexchange.com/questions/237538/what-if-i-train-a-classifier-with-only-positive-example>

introduction/Business model -> Julian (Mathieu)

presentation of business case -> Venkatesh

Jeu de données -> Louis

What we want to do and risks -> Mathieu(Julian)

???  
Model of car correlates with accident?

* Cannot consider all factors. Is also not necessary as risk can be sufficiently modelled from other (available) factors.

cela ne peut fonctionner que si l'info est transmise en temps réel ; est-ce faisable ?

* the model is updated in certain intervals. The interval is based on the statistical implications from frequency of accidents.
* For now we focus on creating the model. If time permits in the mid-term future we will also work on details of the deployment of the model.
* Also without details on deployment the model result could already be used by vehicle/road authorities to implement counter-measures such as additional street signs.