Dr. Ingo Mierswa
Founder & President

Gartnerは、2018年には世界中の大企業の半数以上が競争力維持のために先進的なデータ分析を用いることになり、産業界は変革を迫られると予測。

Gartner predicts that by 2018, more than half of large organizations globally will compete using advanced analytics […], causing the **disruption of entire industries**.

Source: Gartner, Predicts 2016: Advanced Analytics Are at the Beating Heart of Algorithmic Business, 06 November 2015 , G00292891


473 ms

もし、あなただけが 知ってたら

If you only knew


クッション1個$10
$10 per cushion


雨が降ると知っていたら:
あなたは傘を
持ちますか?
If you know it is going to rain:
Should you bring
your umbrella?

データ分析は正しいビジネスアクションをどのように実現するのか？

How can analytics help you to find the right business **action**?

BI(ビジネスインテリジェンス)から処方的分析へ

徒歩 → 傘を持つ
車 → 長めの通勤時間を考慮
Walking → Bring umbrella
Car → Expect longer commute

処方的分析
PRESCRIPTIVE ANALYTICS

Smith氏に電話し、
3%の更新割引を提案
Call Mr. Smith to offer
3% renewal discount

明日の降水確率は95%
Probability for rain tomorrow is 95%

予報的分析
PREDICTIVE ANALYTICS

Jon Smith氏は92%の確率で
我々の顧客ではなくなる。
We will lose Jon Smith tomorrow (92%)

来年、少なくとも200日間
は雨が降る。
It will rain at least 200 days next year

BI的な “予測”
BI-STYLE “PREDICTIONS”

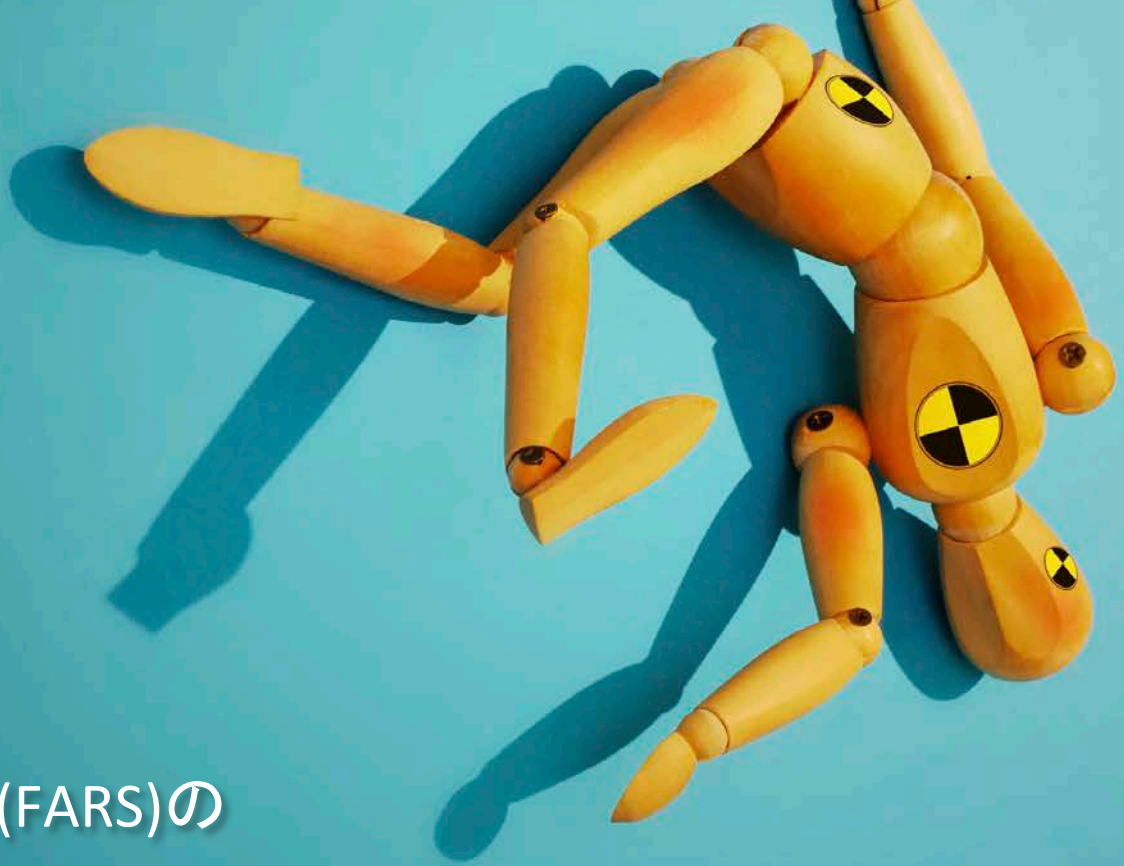
来年、約20%の顧客を
失うことになる。
We will lose about 20% of
our customers next year

去年、雨が降った
のは231日
It rained 231 days last year

ビジネスインテリジェンス(BI)
BUSINESS INTELLIGENCE

昨年、500万人の
顧客を失った
We lost 5 million
customers last year

VALUE
付加価値
行動
Action
知見
Insights



Demo:
死亡率分析システム (FARS)のデータの分析

Demo:
Analyzing the Fatality Analysis Reporting System (FARS) Data

Crash Report / Overview

Info

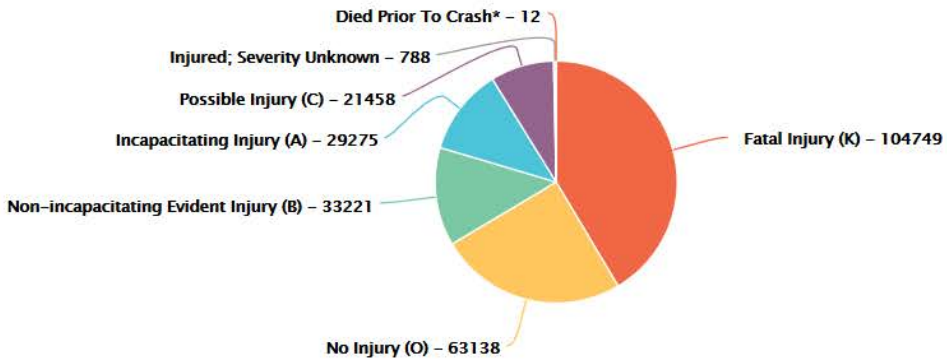
The **Overview** tab display the Fatality Analysis Reporting System (FARS) data from the National Highway Traffic Safety Administration (NHTSA).

We analyzed data for more than **250,000 persons** who experienced an accident in the **years 2009 - 2013**.

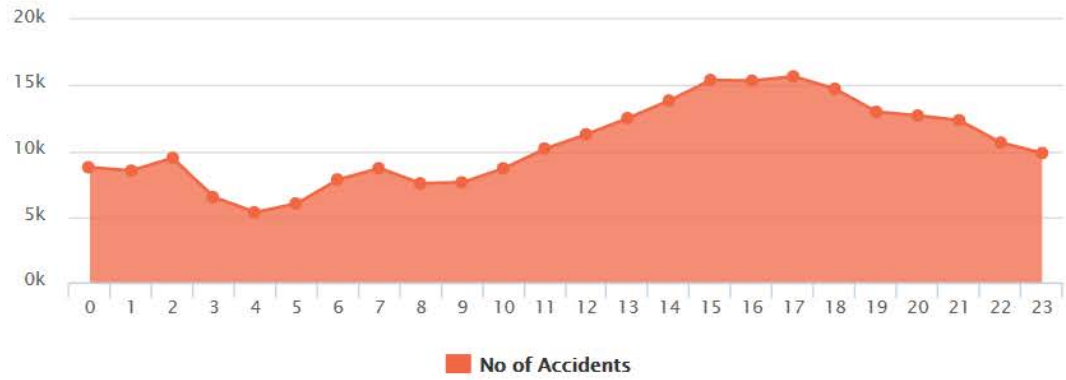
This tab provides a general overview of the data. For example, it illustrates accident time patterns (higher frequency in the afternoon, on weekends, and during the summer) and gender distribution of accidents.

This app also analyzes **geographic distributions** and discovers which **influence factors** likely cause fatalities. We will inspect the predictive models that can be used to **predict the outcome of an accident** with respect to fatality.

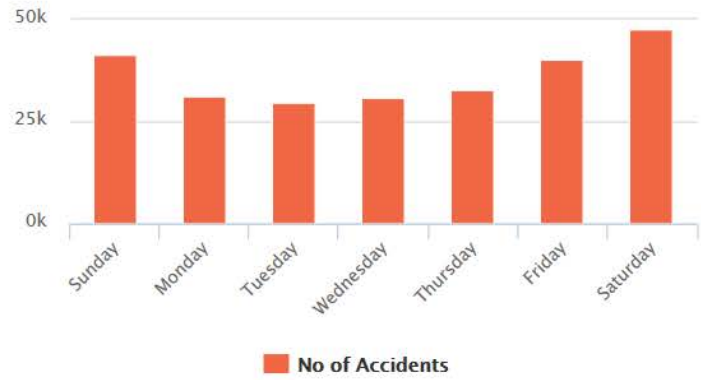
Injury Severity



Accidents by Hour



Accidents by Day of Week

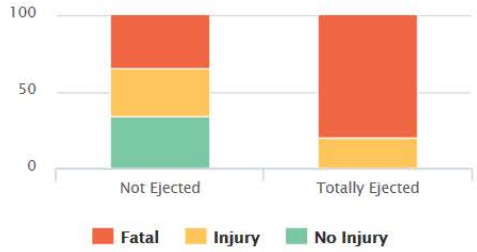


Info

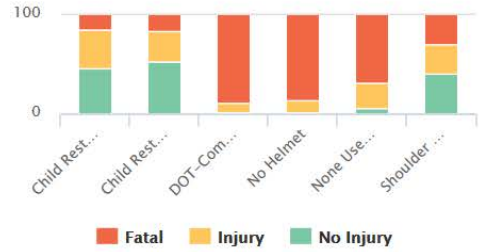
The **Influence Factor**s tab illustrates the most influential factors for fatality. The data clearly shows, for example, that the Ejection, Restraint System, and Age factors have a large impact on survival rate.

Most Important Influence Factors

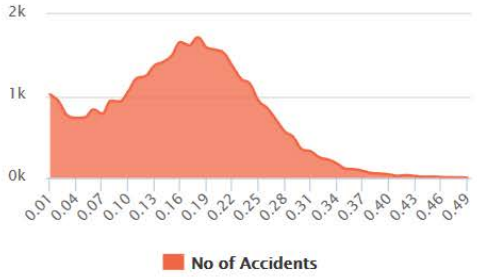
Ejection Influence on Fatality



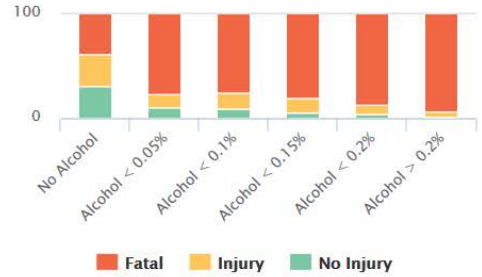
Restraint System Influence on Fatality



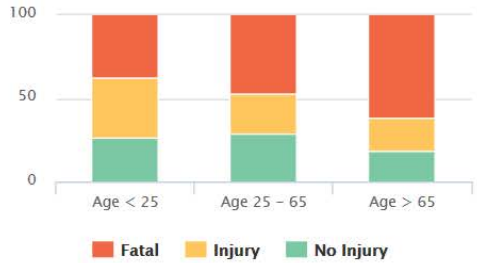
Alcohol Test Result



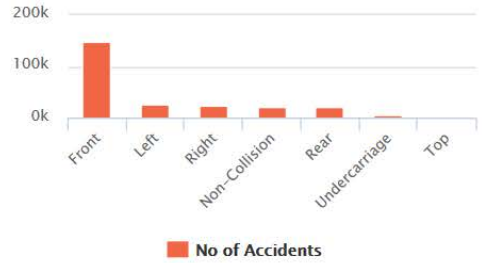
Alcohol Influence on Fatality



Age Influence on Fatality



Area of First Impact



Crash Report / Advanced Analytics

Info

The **Advanced Analytics** tab illustrates the best models for predicting whether an accident will result in fatalities, injuries, or no injuries.

The decision tree correctly predicts three classes with a roughly 76% accuracy. The best predictive model – **Boosted Naïve Bayes** – is correct in **more than 84% of all cases**.

Decision Tree Performance

Accuracy of Bayesian Boosti
Naive Bayes

84.4 %

Accuracy of Random Forest

83.7 %

Accuracy of Naive Bayes plus Feature Selection

83.2 %

250,000件の事故の死亡率を予測度

加えて、それぞれの事故は死亡率の確からしさを算出“この事故に似た事例の21%”
Accuracy for predicting fatality correctly on all 250,000 cases

In addition, each case gets a likelihood for fatality: “21% for an accident similar to this.”

Pareto Front for Best Factors for Unboosted Naive Bayes

Crash Report / Prescriptive & Operationalization

Info

The **Prescriptive & Operationalization** tab allows you to define typical operating conditions. Using those values with the created predictive models, the system will determine the vehicle make and body type with the **lowest probability for fatal accidents**.

RapidMiner's **prescriptive analytics** iterates through tens of thousands of combinations to prescribe the best course of action (i.e., best vehicle selection). With more than 500 connectors to business applications, RapidMiner can **automatically trigger the best course of action** as part of your business processes.

Your Personal Driving Situation

Age	Age 25 - 65
Sex	Male
Drinks Alcohol	☐
Typical Speed	50
Typical Weather	Clear
Typical Road Condition	Dry
Typical Light Condition	Daylight

Submit

Your Prescribed Car & Fatality Probability

Fatality	Vehicle Make	Body Type
19%	Dodge	Standard Pickup

Order new Dodge

予測的

死亡する確率はどの程度か？

Predictive

What is my probability for fatality?

処方的

事故にあって死ぬ確率を下げるためには
どの車種を選べばよいのか？

Prescriptive

What car should I buy for a minimal fatality in case of an accident?

Crash Report / Prescriptive & Operationalization

Info

The **Prescriptive & Operationalization** tab allows you to define typical operating conditions. Using those values with the created predictive models, the system will determine the vehicle make and body type with the **lowest probability for fatal accidents**.

RapidMiner's **prescriptive analytics** iterates through tens of thousands of combinations to prescribe the best course of action (i.e., best vehicle selection). With more than 500 connectors to business applications, RapidMiner can **automatically trigger the best course of action** as part of your business processes.

Your Personal Driving Situation

Age	Age < 25
Sex	Female
Drinks Alcohol	☑
Typical Speed	90
Typical Weather	Rain
Typical Road Condition	Wet
Typical Light Condition	Dark - Lighted

Submit

Your Prescribed Car & Fatality Probability

Fatality	Vehicle Make	Body Type
9%	Datsun/Nissan	2-door Sedan,hardtop,coupe

Order new: Datsun/Nissan


rapidminer
500万人
の顧客が毎日サービスから離れる可能性がある。
50 Million
Utility customers who might churn every day
本当のファン
解約するかも


8400万
のデータポイントが製造プロセスから毎日生み出されている。
84 Million
Data points / day from production processes
rapidminer
この機械を分解せずに
効率を上げることができのだろうか？
Can you increase throughput without
breaking the machine?


9億人
の搭乗客が毎年利用
900 Million
Passengers per year
9 名はアップグレードする可能性がある
9 passengers would upgrade
rapidminer

意思決定を支えるデータサイエンス

Data Science Behind Every Decision

#1 Open Source Data Science Platform

機械学習のセルフサービス

高速プラットフォーム

エクステンションのためのマーケットプレイス

Self-service machine learning

Lightning-fast unified platform

Marketplace for extensions

#1 Marketplace for Data Science Experts

オンデマンド・コンサルタント (“Uber”)

あらゆる業界のエキスパート

世界中に展開

On-demand consultants (“Uber”)

Domain expertise in every industry

Global presence

統一された高速プラットフォーム

Lightning-Fast, Unified Platform

Incorporate all types of data

データ前処理

データ探索の高速化・最適化とタスクのブレンディング・クレンジング

Data Prep

Speed & optimize ALL data exploration, blending & cleansing tasks

モデル化・検証

機械学習を活用したラピッド・プロトタイプと予測用モデルの検証

Model & Validate

Apply machine learning to rapidly prototype & confidently validate predictive models

実装

容易に活用・メンテナンスできるモデルの実装と分析結果の導入

Operationalize

Easily deploy & maintain models and embed analytic results

Embed results in all types of business apps & data visualization tools

ACCELERATES TIME TO VALUE

是非、RapidMinerの
ブースでデモをご覧ください。

Please visit the RapidMiner
booth for a demonstration

Hosted by:

KSK ANALYTICS
Data to Knowledge

KSK Analytics
www.ksk-anl.com

Dr. Ingo Mierswa
imierswa@rapidminer.com

@ingomierswa
@rapidminer