SOCIETY OF ACTUARIES®

2019 ANNUAL MEETING & EXHIBIT

October 27-30
Toronto, Canada

Session 089: Unleashing the Power of Claims Data Through Machine Learning

SOA Antitrust Compliance Guidelines
SOA Presentation Disclaimer

Unleashing the power of claims data through Machine Learning

Dave Keirstead
VP, Head of Advanced Analytics
Manulife Canada

Enormous opportunity

Opportunities

Claims

Data Capture
Apply structure & meaning to claims data

Workflow
Triage, decisioning to support straight through processing / case management

Behaviour Prediction
Leverage claims patterns to understand and predict outcomes

Data Capture

- Apply structure & meaning to claims data (or supporting documents)
- Common challenge across Markets

  - Claims forms
  - Receipts
  - Health Records
  - Attending Physician Statements
  - Other documents

- Approaches - Computer Vision, Natural Language Processing
- Classification and standardization

Workflow

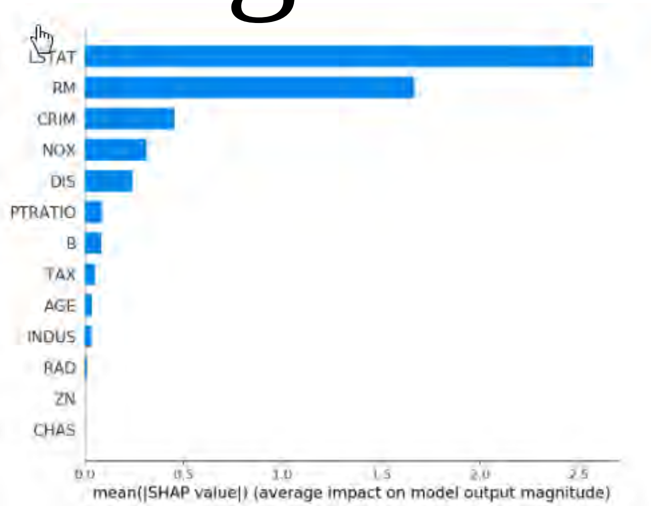
- Claims processing, triage and adjudication
  - Integrate machine learning based predictions with robotic processing automation
  - Claims classification/segmentation
    - Preliminary processing
    - triage for complexity
    - Map to appropriate action
  - Recommendation engine (Next best action)

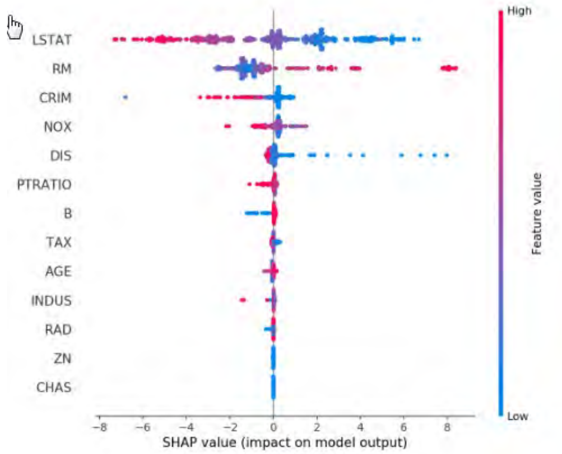
Pattern Detection & Prediction

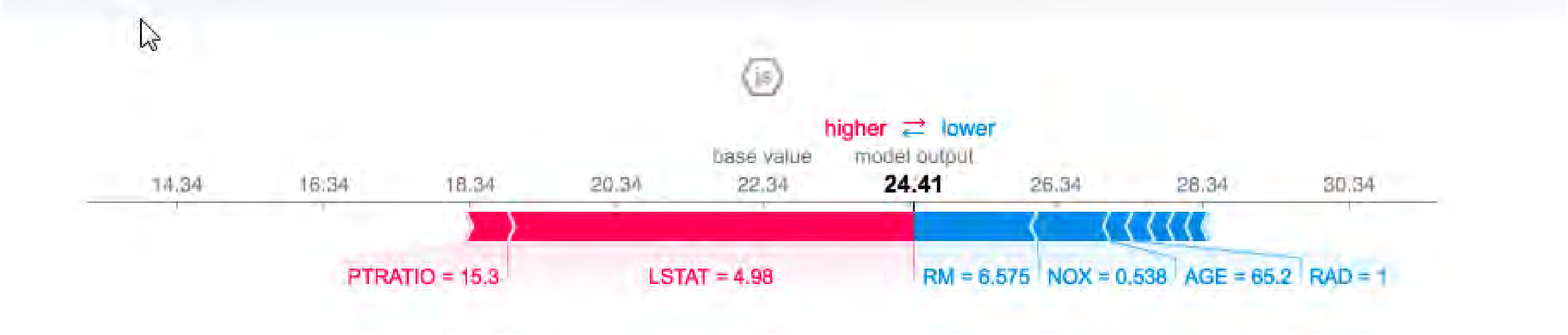
- Predict claims (incidence/severity/drivers)
  - Understand
    - Pricing, experience studies
    - Product/service development
    - Plan Design modifications

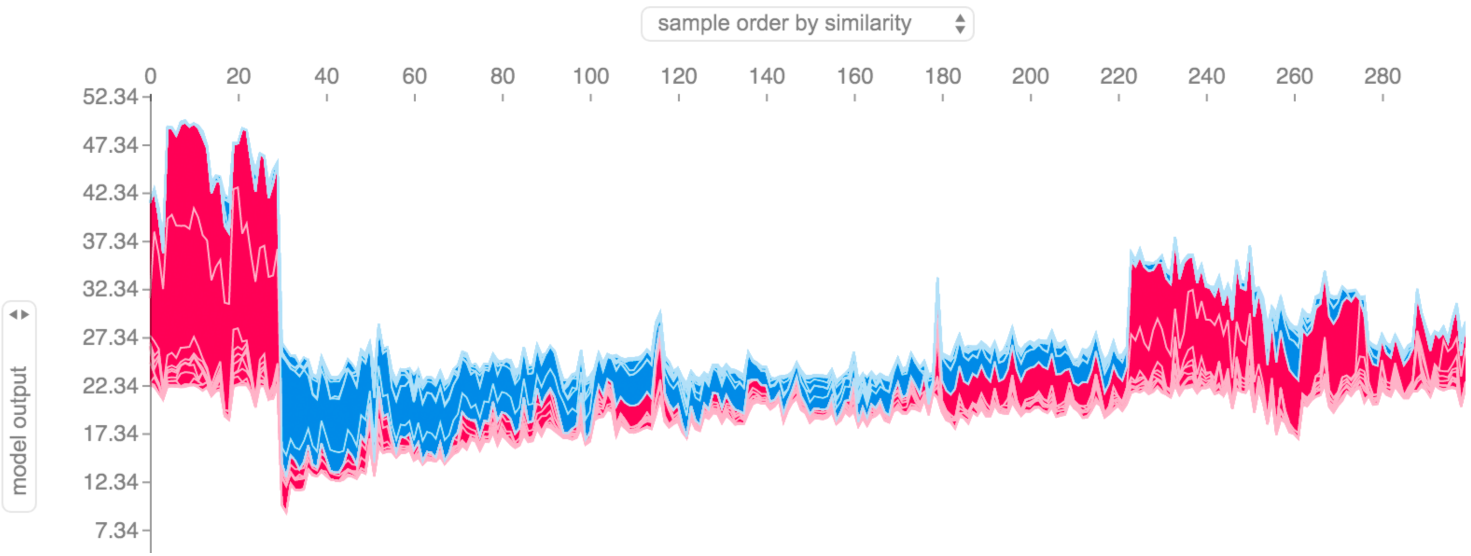
  - Act:
    - Case Management
    - Interventions/Drug Adherence

Model Interpretability is a game changer

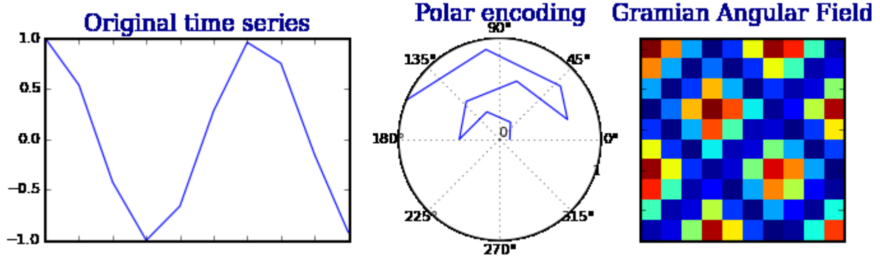


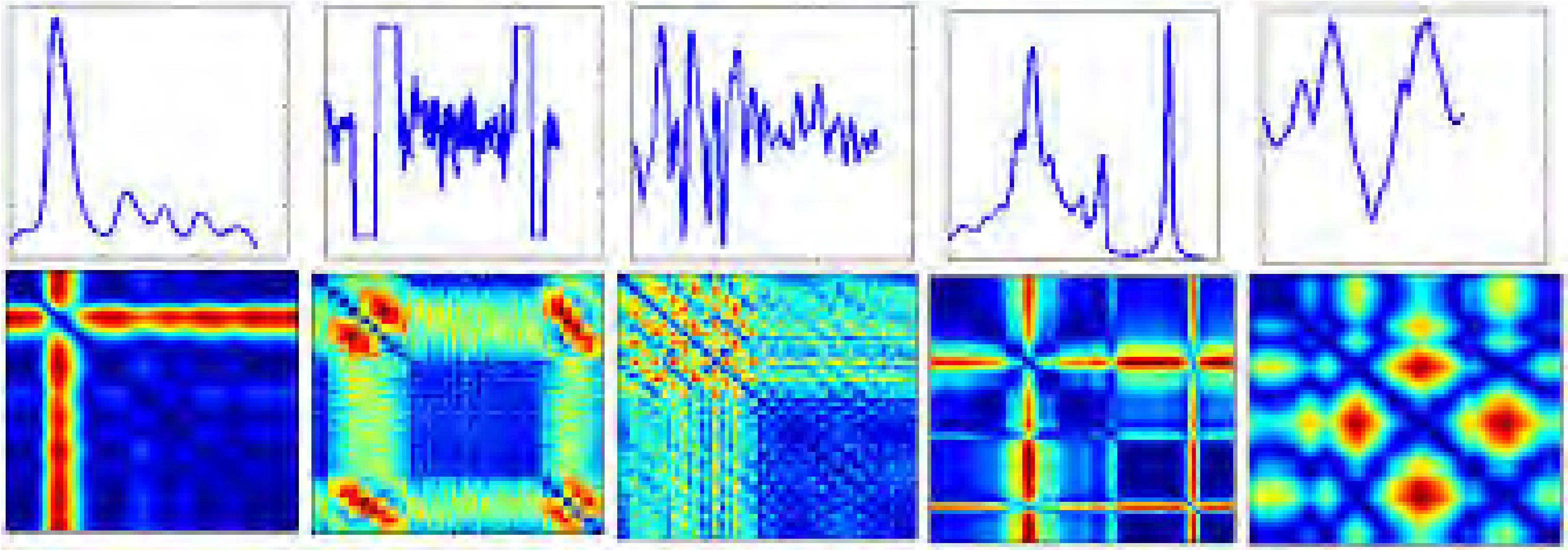






New Methodologies Present opportunities





Thanks!

Munich RE

Claims analytics

Use cases:

- Foreign death fraud models
- Claims risk assessment
  - Critical illness approval/denial and misrepresentation models

FOREIGN DEATH FRAUD

Foreign death fraud - overview

"Warning: increasingly common and sophisticated scams

…In many countries, con artists operate without consequences because local authorities often do not have the physical or financial resources needed to combat Internet crimes…Organized fraud networks are developing more and more innovative and sophisticated approaches to deceive…The criminals conduct extensive searches to create credible documents: complete profiles of fictitious businesses, medical reports, falsified export certificates, etc. The names and logos of reputable organizations, governments and government agencies are often used fraudulently. Websites that appear very authentic are also falsified..."

Government of Canada Advisory, 30 May 2013
http:// Travel.gc.ca/travelling/health -safety/overseas-fraud

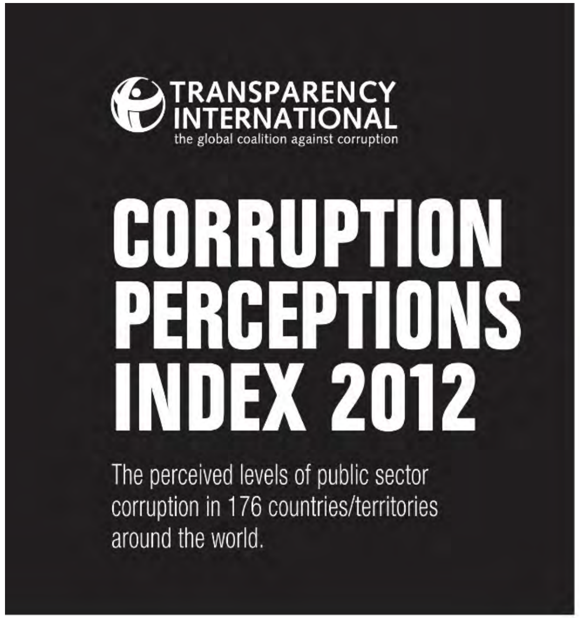
- Misrepresentation of death occurred abroad can be as high as 2% of all individual life claims
- High fraud rates are observed in Middle East, Africa and China with misadventure being the common cause..
- Machine learning methods can leverage multiple data sources to provide insight on main drivers of fraud and identify high risk foreign death claims

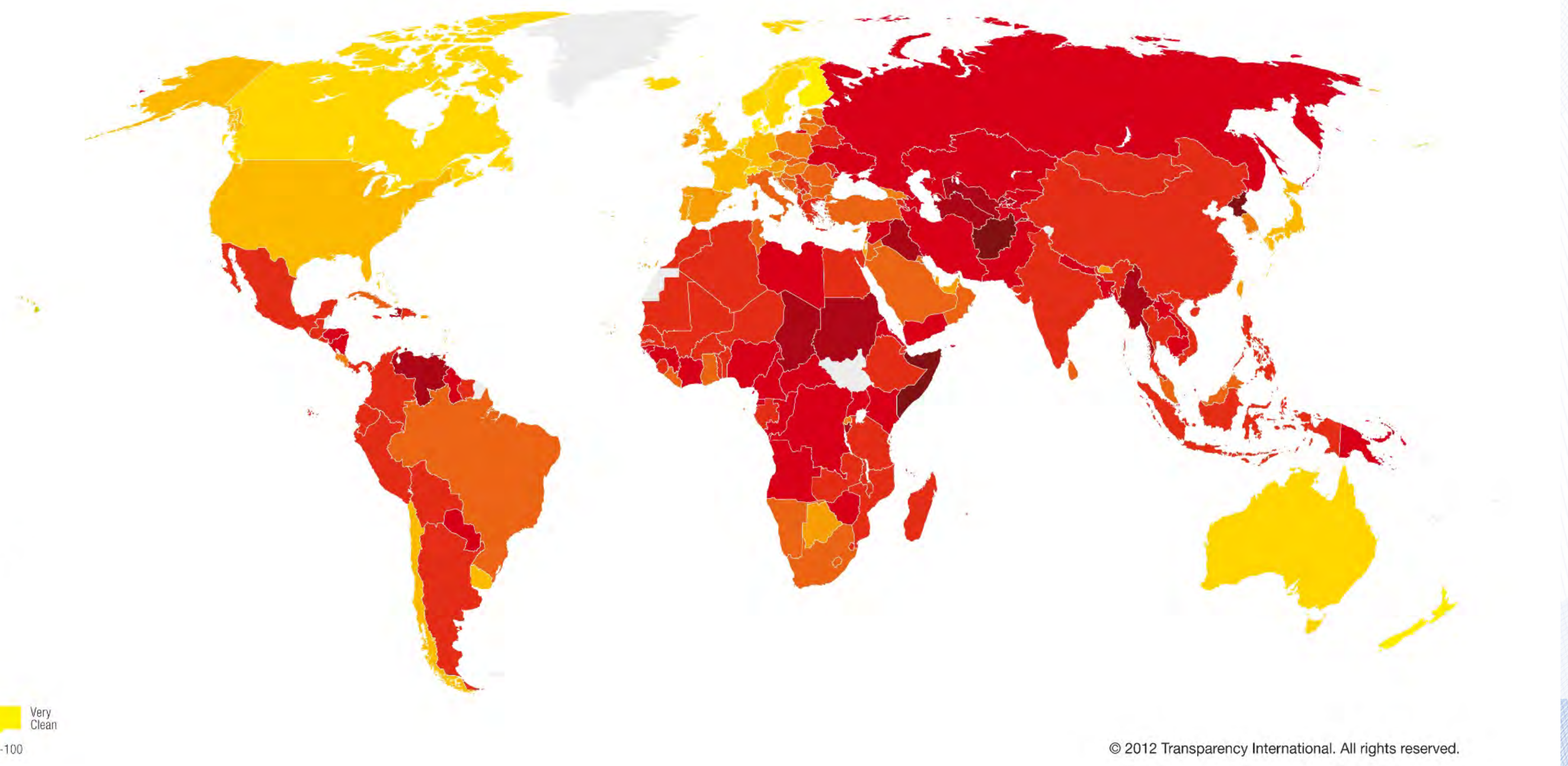
Data to consider

Demographics	Claim Data	Underwriting and policy data	Third Party/ External	Geospatial
Age / Date of birth	Country	Foreign travel, aviation, avocations	Corruption Perception Index	Population density
Gender	Cause	Tobacco, drug and alcohol usage	Credit/ public records	Unemployment rate
Income	Claim amount	Agent, face amount	Lifestyle	Tobacco tax
Occupation	Claim Age	Beneficiary	MIB	Home value
Income	Policy duration	Personal and family health history	MVR	Household size
Pre-existing diagnosis	Body found	Disability, bankruptcy, felony	Rx	Household composition
	Claim proofs	Product, simplified UW vs Full		Education

Corruption Perception Index (CPI)

Important continuous indicator that allows to smooth out underrepresentation of fraud cases in individual countries





Munich RE

Third-party data

Examples of third-party data include..

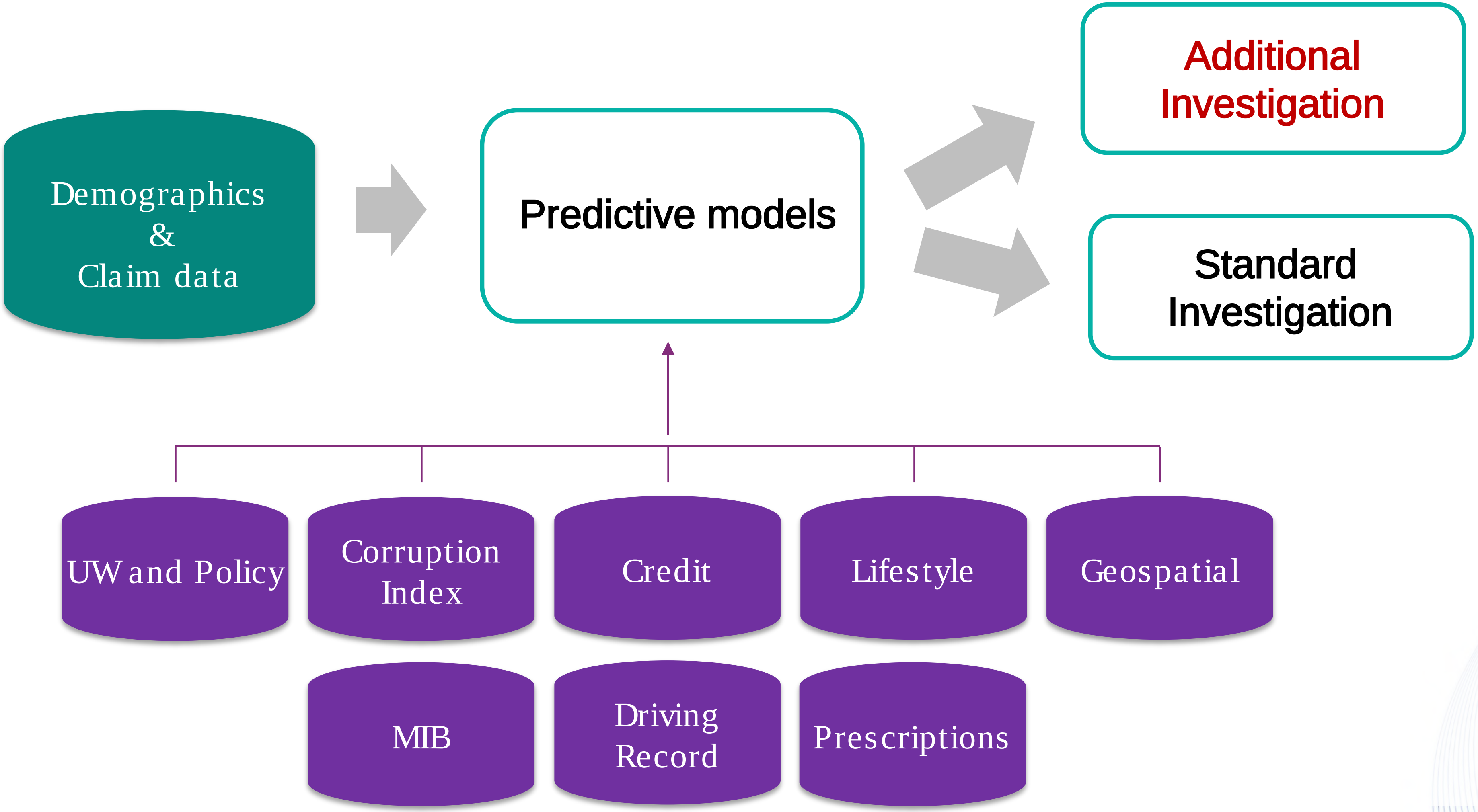
Epsilon
Individual & household level

- Demographics:
  Age, income, children, occupation
- Property:
  Home value, rent/own, living area
- Financial:
  Credit activity, number of tradelines, loans
- Lifestyle/interests:
  Pets, golf, travel, music
- Ailments:
  Asthma, diabetes, heart condition

Lexis Nexis
Individual level

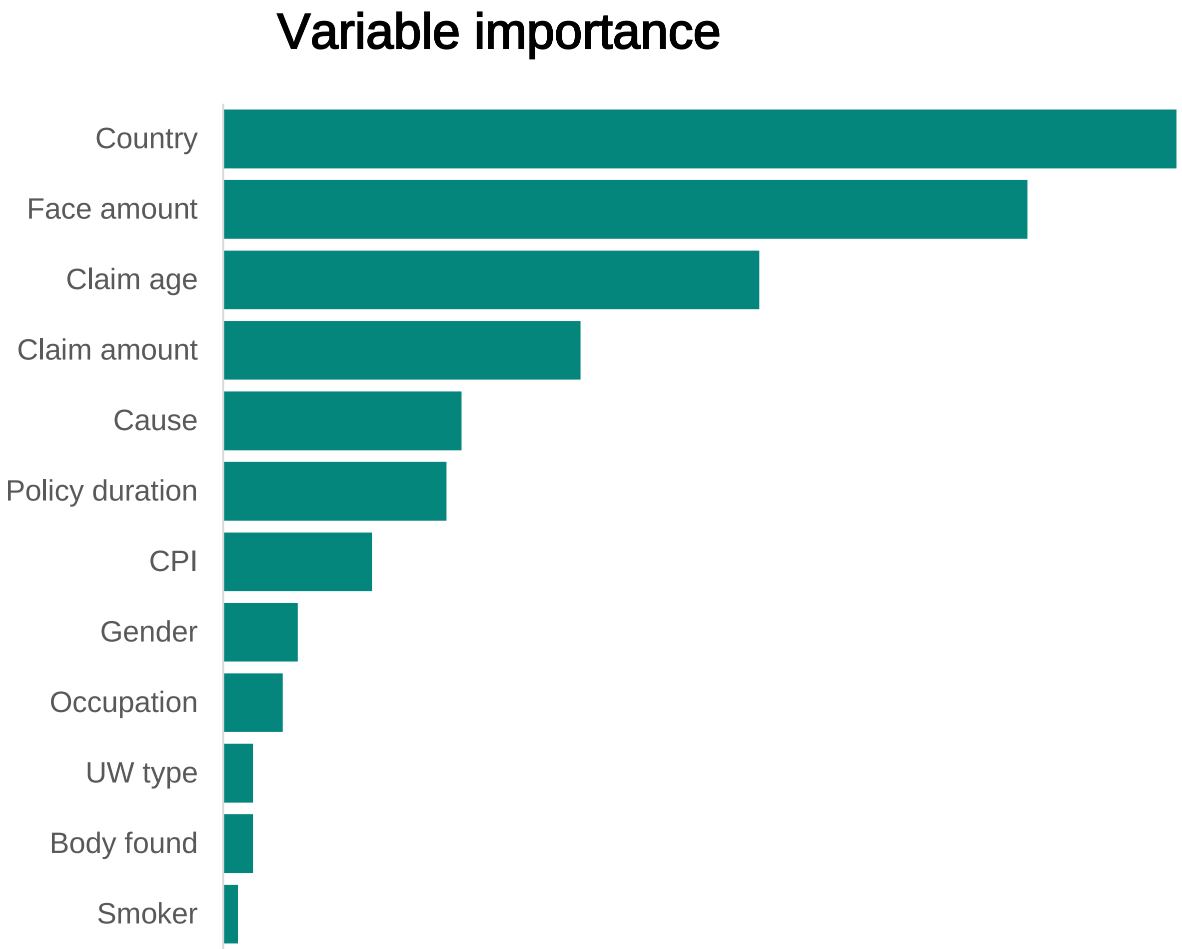
- Demographics:
  Age, education, income
- Court records
- Property:
  Property ownership, value
- Police records:
  Bankruptcy, evictions
- Credit records:
  Auto finance, mortgage, retail credit accounts

Implementation



Essential Data Sources
Nice to have's

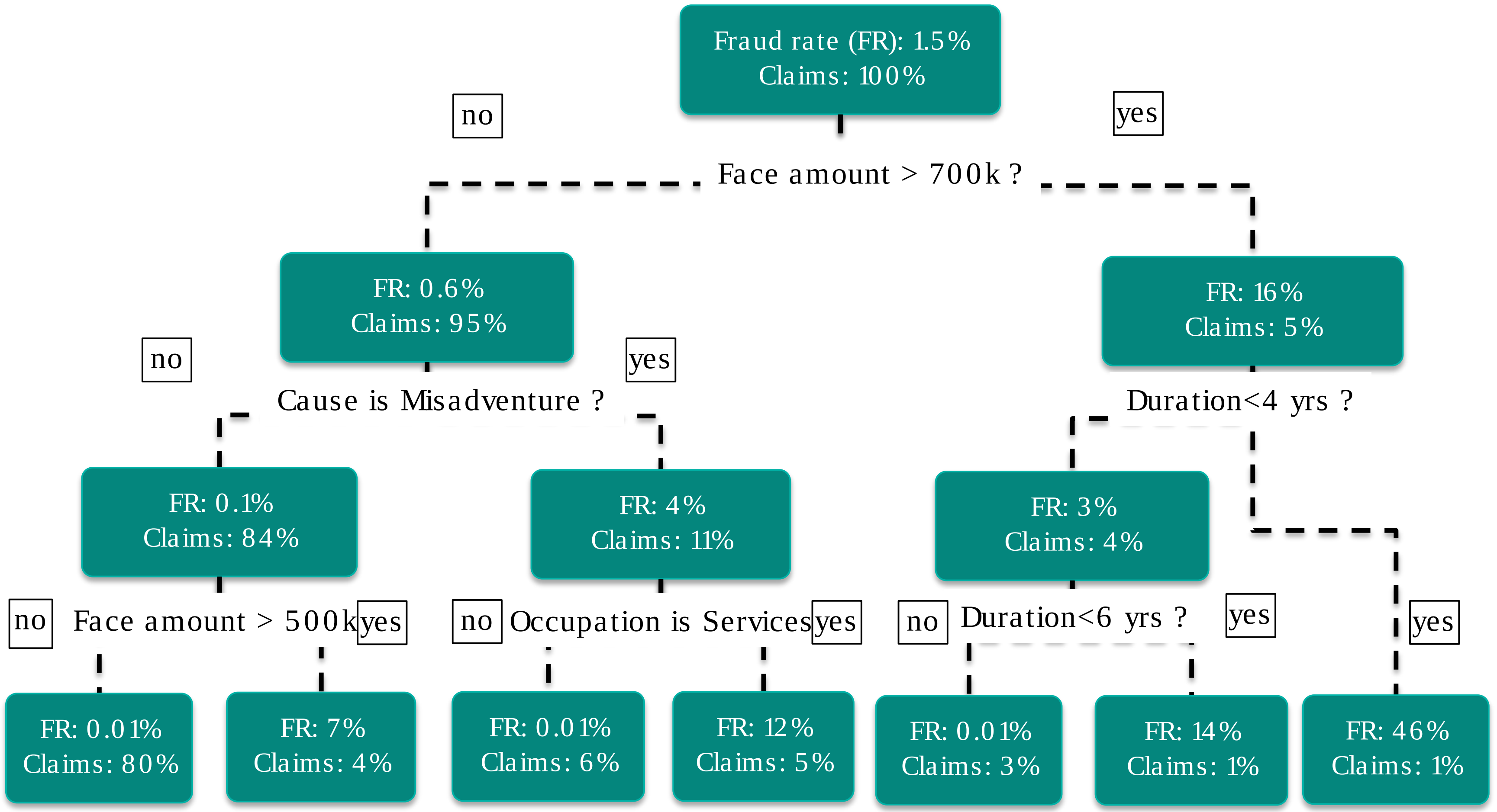
Machine learning to detect key fraud drivers



- Data:
  - ❑ Demographics
  - ❑ Policy data
  - ❑ Claims
  - ❑ Corruption Perception Index
- Average fraud rate: 1.5%
- Machine learning methods allow to estimate marginal impacts of all variables on fraud likelihood and thus detect **key fraud drivers**

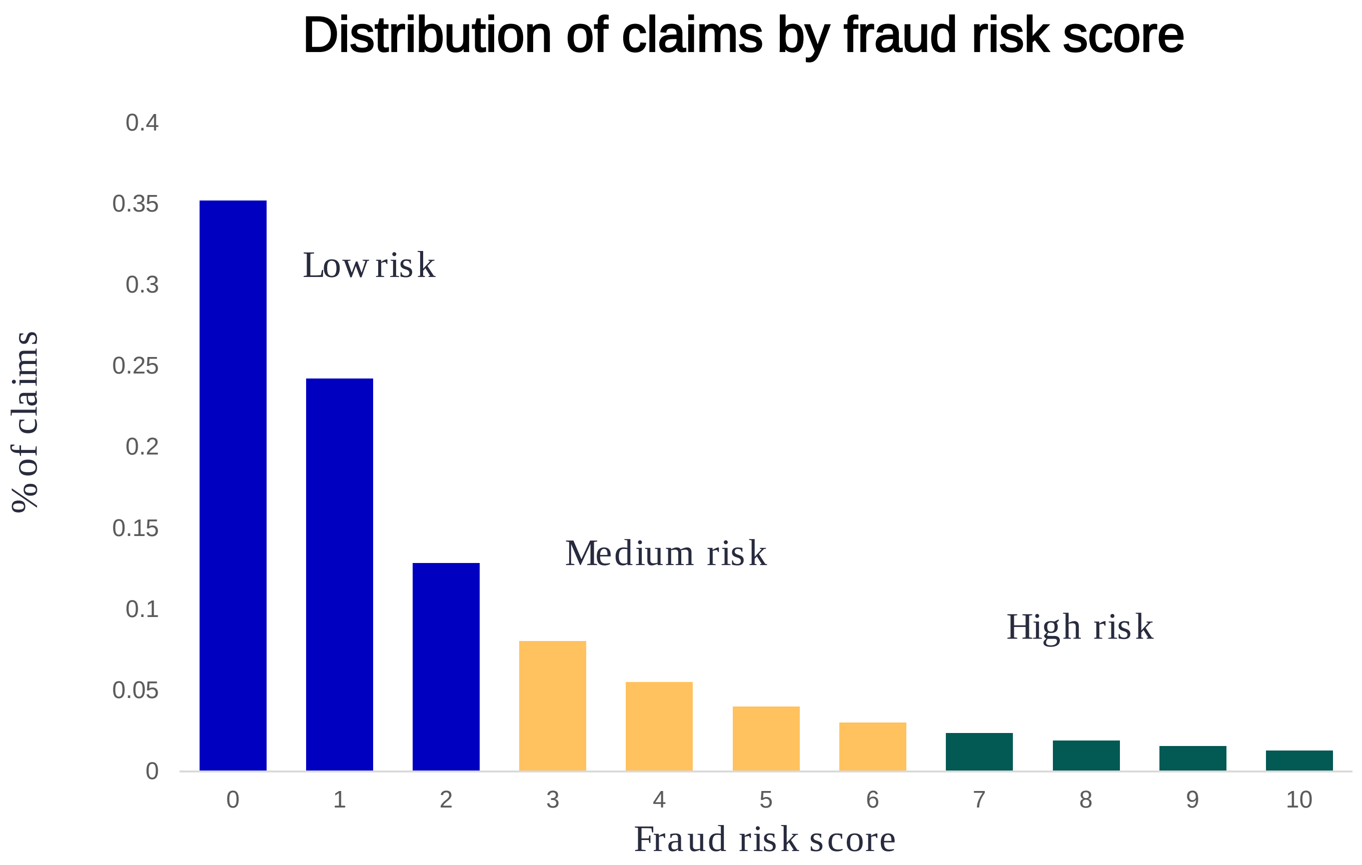
Munich RE

Machine learning to explore fraud cases



- Machine learning methods can be applied to segment claims from least risky to most risky classes
- The decision tree identifies claims with face amounts higher than $700K and policy duration under 4 years as high risk with fraud likelihood of 46%

Machine learning to produce fraud risk score



- Given sufficient data, more sophisticated machine learning methods can be applied to produce a continuous risk score
- The score can then be used to determine optimal strategy for claim investigations
- Machine learning methods produce 0-10 fraud score for each claim. Claims can then be segmented into low, medium and high risk categories for targeted investigations

Partnership is the key to fraud detection

Data is essential

CLAIMS RISK ASSESSMENT

Munich RE

Claims approval/denial and misrepresentation

Motivation

- Claims adjudication is a complex process that requires significant amount of human intervention
- Average claim processing time can exceed several weeks

Objective

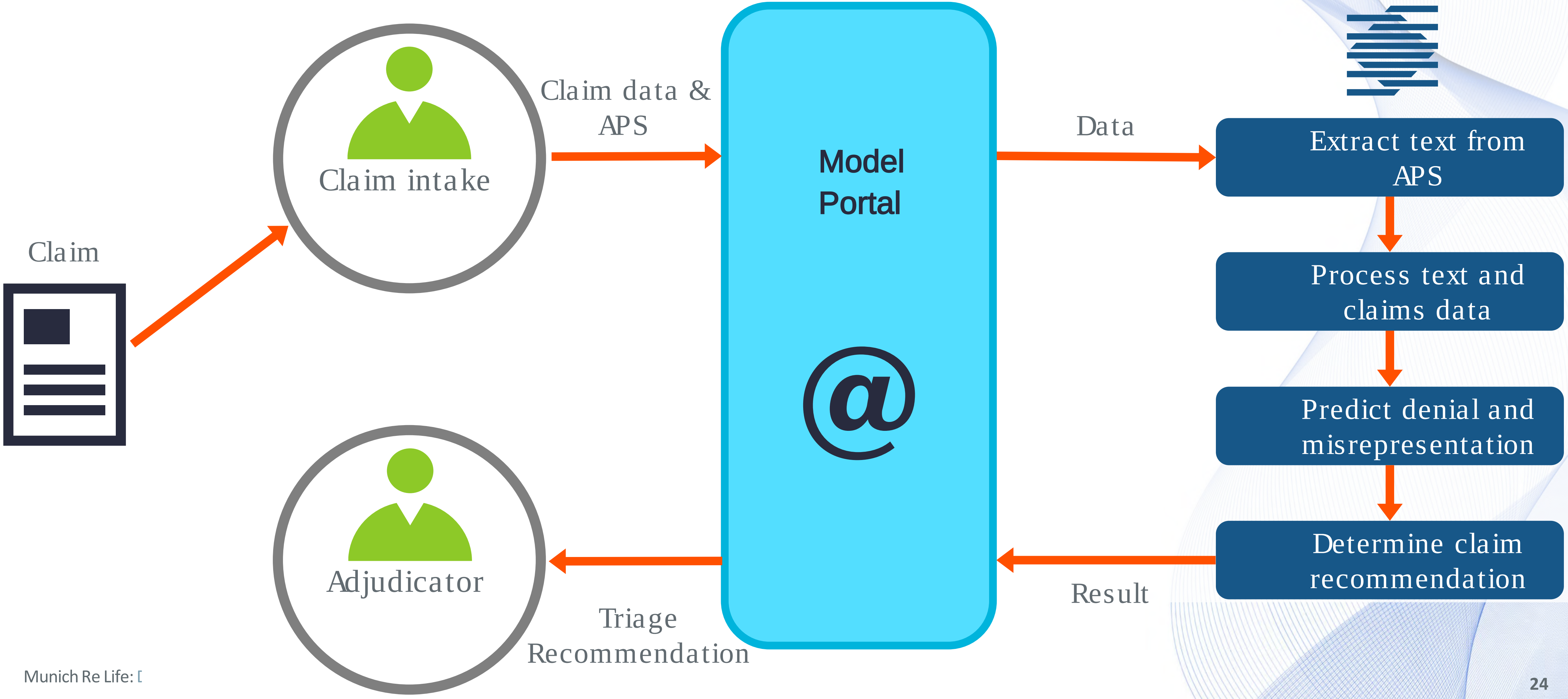
- Use machine learning to identify critical illness insurance (CI) claims that have a high risk of being denied
- Triage claims based on adjudication difficulty and likely outcome

Our approach

Two model approach is proposed to classify CI claims based on:

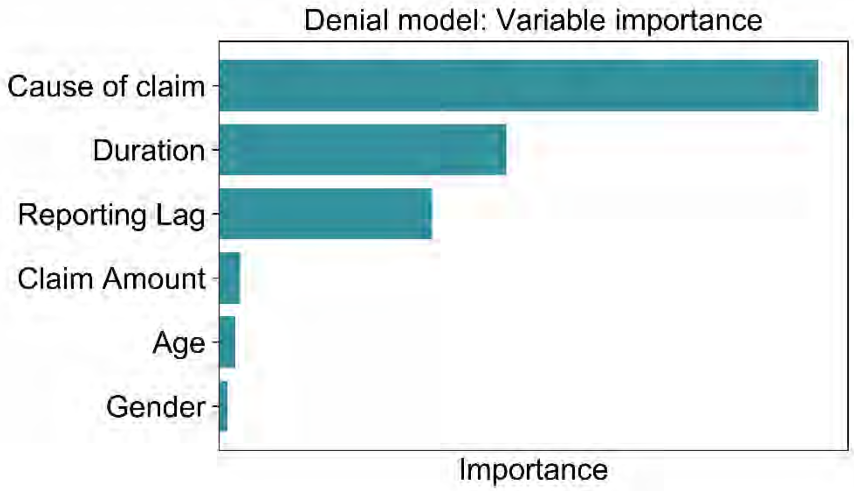
- Risk of claim denial
- Risk of claim involving misrepresentation, non-disclosure or fraud

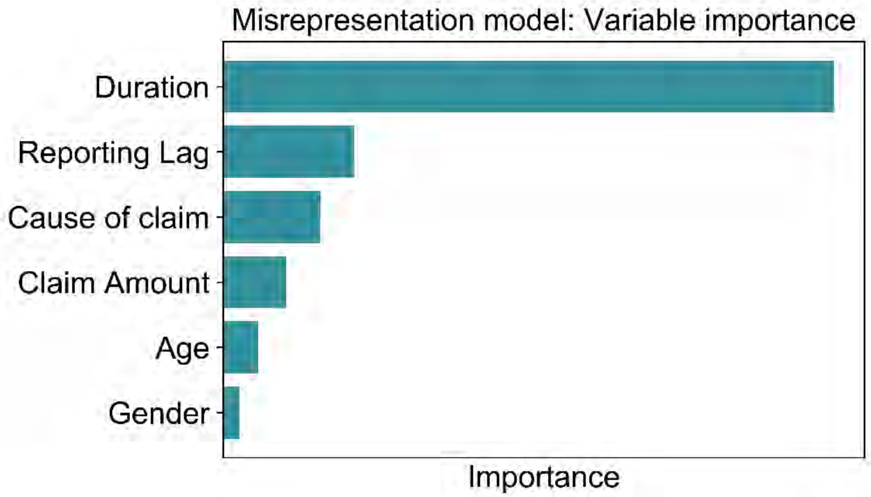
How does it work?



Munich RE

Approval/denial vs misrepresentation





Claims denials

- Denied (any reason) , rescinded, death, litigation

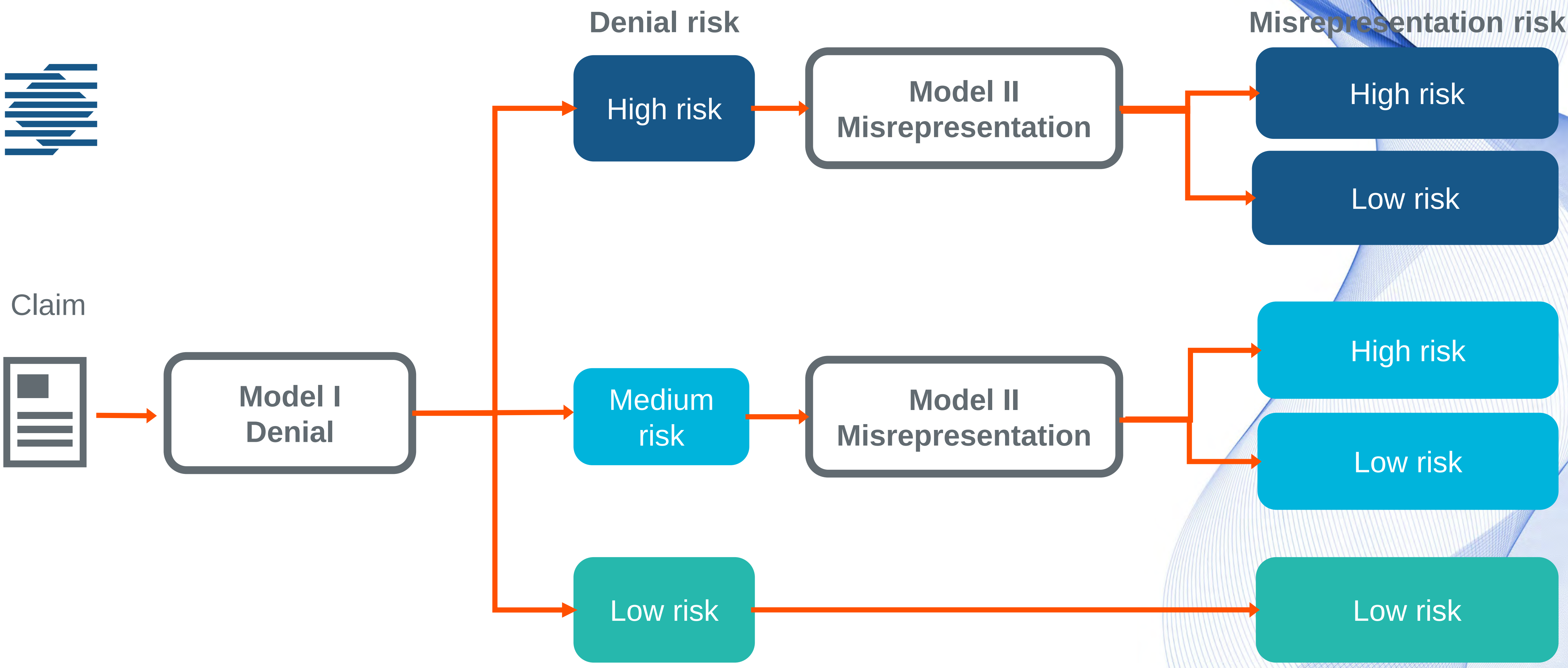
Claims misrepresentations

- Medical misrepresentation, fraud, non-med nondisclosure, rescission identity fraud, rescission family history, denial family history, etc

Key drivers

- Cause of claim, duration and reporting lag are the key drivers of claim denial and misrepresentation predictions
- Cause of claim is the most important feature predicting claim denial whereas policy duration is the most important feature for predicting likelihood of claim misrepresentation

Implementation



Munich RE

Dashboard

Enter Claim age

20

Enter gender

Female

Enter policy duration (years)

10

Enter Claim amount

5000

Enter reporting lag (days)

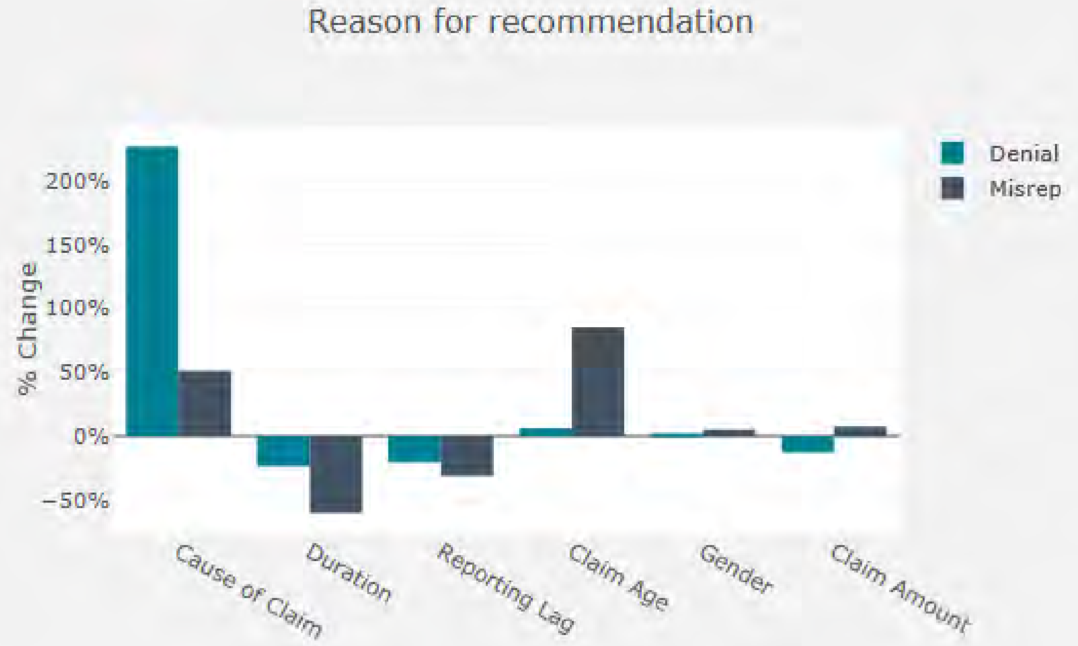
0100

Enter cause of claim

Gastro-Intestinal Acute Pancreatitis

SUBMIT

Case has a high risk of denial, with a low risk of misrepresentation.



Key findings

- Approval/denial and misrepresentation models provide an efficient risk segmentation of all CI claims allowing for claims to be triaged based on quantitative risk assessment
- The most impactful drivers of denials and misrepresentation are identified during risk assessment
- Easy to use through a user friendly interface

Thank you for your attention!