# Processing Insurance Claims with Automated, Scalable and Fair AI

Kenneth Lim[1], Maxime Allard[1,2]

[1]IBM UK, [2]Imperial College London

{Kenneth.Lim, Maxime.Allard3}@ibm.com

*This work is not necessarily a representation of any past or current employers.*

# Inefficient claims processing is costly and a major priority for insurers this year.

Processing claims make up 30% of operating costs on average.[1]

Inefficient processing leads to claims leakage, costing the insurance industry more than $30B/year.

Insurers experience leakage up to 25% vs. 3% industry benchmark.[2]

This is becoming the #1 priority for insurers in 2021.[3]

1: https://www.mckinsey.com/industries/financial-services/our-insights/from-transparency-to-insights-mckinseys-insurance-cost-benchmarking-2016
2: https://www.pwc.com.au/industry/insurance/assets/stopping-the-leaks-jan15.pdf
3: https://go.forrester.com/blogs/europe-predictions-2021-financial-services/

# AI can help process claims more efficiently and save insurers substantial amounts of money.

Reducing claims leakage can save insurers 5%--10% of their costs, translating to millions of dollars.[4]

AI can minimize potential claims leakage prospectively vs. diagnosing retrospectively (status quo).

Up to 70% of claims can be handled automatically.[5]

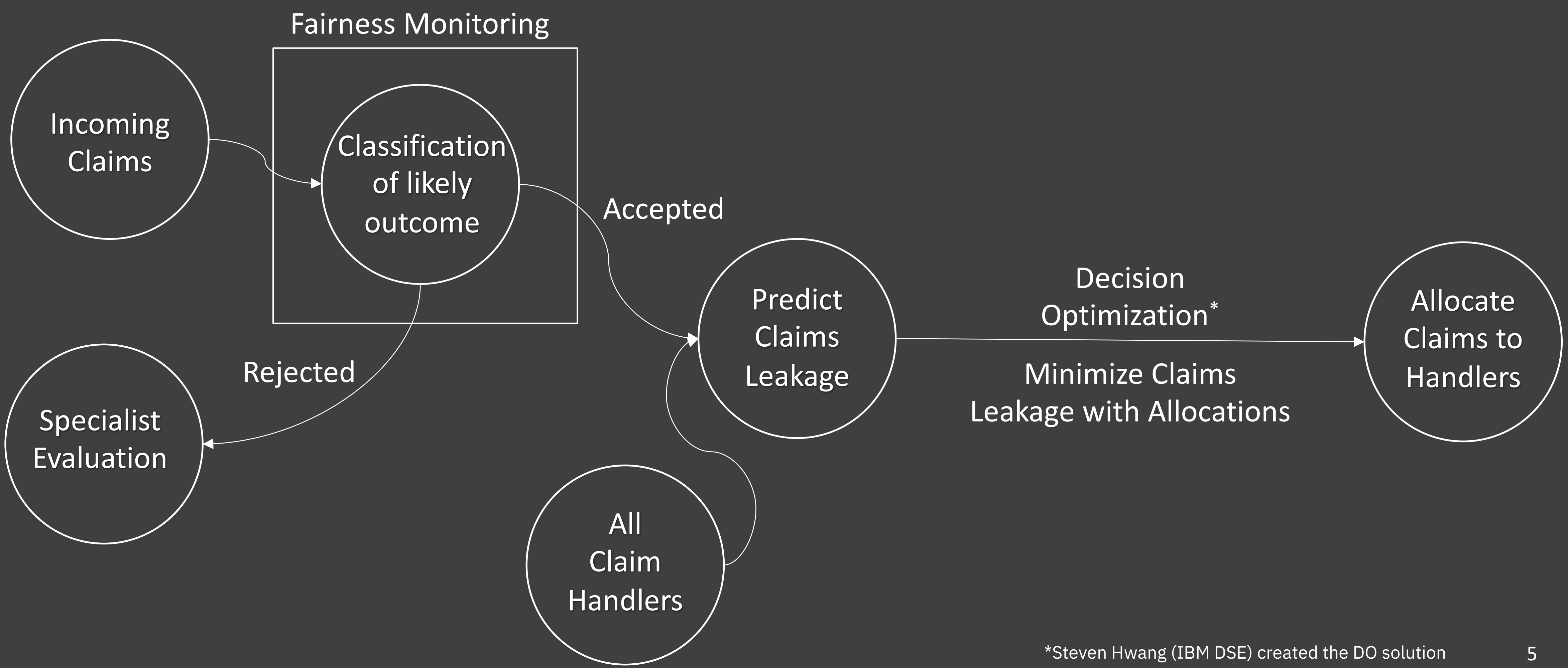
4: https://www.pwc.com.au/industry/insurance/assets/stopping-the-leaks-jan15.pdf
5: https://www.digitalistmag.com/customer-experience/2019/04/10/digitally-fully-automated-claims-is-not-2030-dream-its-possible-today-06197774

# But ethical and scalability concerns remain barriers to AI adoption.

| Concern | Solution |
|---|---|
| Can we trust ‘black box’ AI models?[6] | Explainable and fair results. |
| Can we scale AI models across an organization? | Automated predictions and optimized actions. |

6: https://emerj.com/partner-content/what-do-insurance-experts-think-about-ai-in-claims-processing/

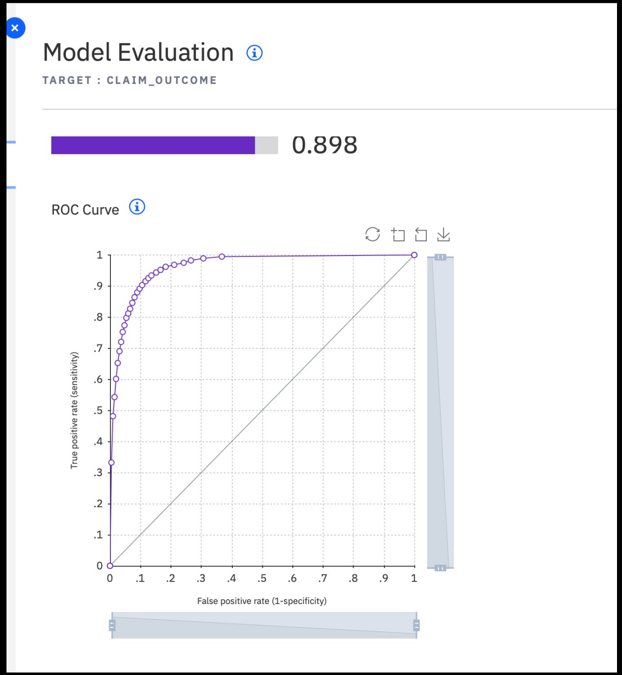
# Solution overview




*Steven Hwang (IBM DSE) created the DO solution

# 1. Automating Claims Classifications and Regression

- Tabular Claims data, containing both categorical and numerical features (e.g. age, dates, covered amounts)
- 2 **Predictive Steps** in the solution:
  - Predicting likely **outcome of claim** (i.e. Settled or Rejected)
  - Predicting possible **claim leakage** for assigned claim handlers
- We let AutoAI handle imputations and model choice as well as **Feature engineering** through Cognito[1] and ADMM[2]
- Cognito finds **relevant features** (e.g. through PCA or other simple transformations)



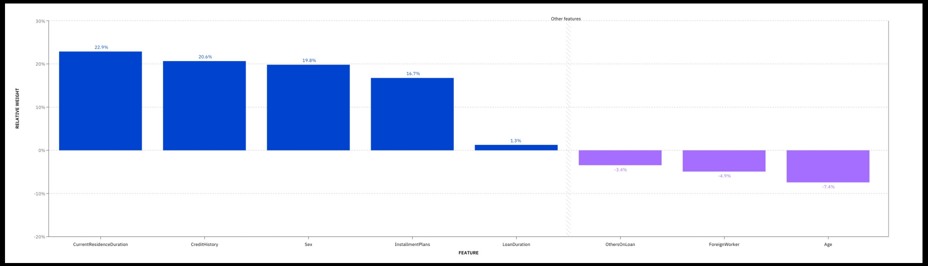

1: https://ieeexplore.ieee.org/abstract/document/7836821

2: https://arxiv.org/abs/1905.00424



(c) IBM Corporation 2021

# 2. Fairness in Claims Predictions




- Certain features induce bias against a certain group of people (**group fairness**) or individuals (**individual fairness**)
- Both **human specialist** and Openscale monitor this bias
- For our example, bias against **younger and older age-groups** has been detected
- Bias mitigation can be performed to assist the Claims Handler (by using for example Reject option classification[1])
- In addition to LIME[2], we use Contrastive Explanations[3] to explain the model’s decisions to help Claim Handler understand the decisions

1: https://ieeexplore.ieee.org/document/6413831

2: https://arxiv.org/abs/1602.04938

3: https://www.ibm.com/downloads/cas/0ZRZNR8E



(c) IBM Corporation 2021

# 3. Optimization of scheduling

Based on predictions for the claims leakage, we solve a problem to find the **optimal scheduling among claim handlers**

Claim handler features and their workload are used as inputs to the claims leakage classifier

**Optimal distribution** between experience of claim handler, their case load and difficulty of claim case

**CPLEX** Optimizer allows for **large-scale optimization** across thousands of claims and handlers

| CLAIM_NBR | Incurred Claim Cost | Handler Case Load at Allocation | Notification Delay | Claim Complexity | Claim Type | Days from Policy Inception to Date of Claim | Claim Handler Level | Number of Previous Claims | Handler Allocation | Pred_Leakage |
|---|---|---|---|---|---|---|---|---|---|---|
| PE34481 | 1455 | 185 | 19 | 2 | Personal Effects | 211 | 4 | 2 | ClaimHandler1 | 0.051046 |
| PE34481 | 1455 | 304 | 19 | 2 | Personal Effects | 211 | 1 | 2 | ClaimHandler2 | 0.159686 |
| PE34481 | 1455 | 288 | 19 | 2 | Personal Effects | 211 | 1 | 2 | ClaimHandler3 | 0.156615 |
| PE34481 | 1455 | 312 | 19 | 2 | Personal Effects | 211 | 4 | 2 | ClaimHandler4 | 0.071500 |
| PE34481 | 1455 | 232 | 19 | 2 | Personal Effects | 211 | 1 | 2 | ClaimHandler5 | 0.156163 |
| PE34481 | 1455 | 689 | 19 | 2 | Personal Effects | 211 | 2 | 2 | ClaimHandler6 | 0.197361 |
| PE34481 | 1455 | 146 | 19 | 2 | Personal Effects | 211 | 3 | 2 | ClaimHandler7 | 0.078544 |
| PE34481 | 1455 | 80 | 19 | 2 | Personal Effects | 211 | 2 | 2 | ClaimHandler8 | 0.116680 |
| PE34481 | 1455 | 232 | 19 | 2 | Personal Effects | 211 | 1 | 2 | ClaimHandler9 | 0.154322 |

# 4. Transparency is the Key

- It is not enough to model claims automatically but the whole data and modelling process needs to be **documented**
- AI Fact Sheets are important for any business using ML models
- Communicating what steps are used makes it easier for the **consumer to understand** how this could impact any decision

AI FACTSHEET

| | |
|---|---|
| **Model Name** | Audio Classifier |
| **Overview** | This document is a FactSheet accompanying the Audio Classifier model on IBM Developer Model Asset eXchange. |
| **Purpose** | This model classifies an input audio clip. |
| **Intended Domain** | This model is intended for use in the audio processing and classification domain. |
| **Training Data** | The model is trained on the AudioSet dataset by Google. |
| **Model Information** | The audio classifier is a two-stage model:<br>• The first model (MAX-Audio-Embedding-Generator) converts each second of input raw audio into vectors or embeddings of size 128 where each element of the vector is a float between 0 and 1.<br>• Once the vectors are generated, there is a second deep neural network that performs classification. |
| **Inputs and Outputs** | **Input**: a 10 second clip of audio in signed 16-bit PCM wavfile format.<br>**Output**: a JSON with the top 5 predicted classes and probabilities. |

**Performance Metrics**

| Metric | Value |
|---|---|
| Mean Average Precision | 0.357 |
| Area Under the Curve | 0.968 |
| d-prime | 2.621 |

**Bias**

The majority of audio samples in the training data set represent voice and music content. Potential bias caused by this over-representation has not been evaluated. Careful attention should be paid if this model is to be incorporated in an application where bias in voice type or music genre is potentially sensitive or harmful.

**Robustness**

This audio classifier is not robust to the L-infinity and L2 norms for the HopSkipJump attack.

| | L2 | L-Infinity |
|---|---|---|
| 5th Percentile | 887.0 (200.9) | 5.5 (4.9) |
| 10th Percentile | 1496.6 (720.6) | 7.53 (5.73) |
| 15th Percentile | 3723.1 (4707.2) | 52.8 (41.8) |
| 25th Percentile | 7187.9 (---) | 187.6 (198.1) |
| 50th Percentile | 11538.6 (---) | 502.8 (---) |

The susceptibility of the model to the two attacks. The parenthetical values in the table above represent the fitted curve evaluated at 11 iterations. (When we are unable to fit a curve, or the result is negative, we denote by ---.)


(c) IBM Corporation 2021

# Conclusion

Inefficient claims processing is costly.

AI can reduce inefficiencies and costs but face ethical and scalability concerns.

We developed a solution that automates the creation of predictive and prescriptive models with explainable and fair results.