What Does Machine Learning Mean for the Life Insurance Industry?

Doug Fullam, ASA

AIR

Agenda

State of the Life Insurance Industry

How Are We Evaluating Risk Today?

Can We "Amazon-ify" the Life Market?

Exploring Possible Applications of Machine Learning

- Artificial Neural Networks
- Genetic Algorithms
- How AIR Applies Machine Learning Techniques

Q&A

State of the Life Insurance Industry

The Race to Digitization and Automation

- Customer-centric focus on millennials
- Less invasive underwriting
- Processes becoming more digital

- Accelerated underwriting comes with more risk
- Heavy regulation of new technology
- Still dealing with siloed decision-making

Transformations in Process, Data, Society, and Analytics

Then:
- Full underwriting and holding pension risk

Now:
- Real-time policy quotes, PRT, and increased solvency demands

Then:
- MIB, medical data, driving, credit

Now:
- Avocation, voice, social media, marketing data

Then:
- Lower smoking rates

Now:
- Obesity, opioids, e-cigarettes

Then:
- Mortality tables, adjustment factors

Now:
- Probabilistic models and machine learning

Technology Shifts Require New Skill Sets

Underwriter

Pricing Actuary

Claims

ERM

Today

Underwriter

Pricing Actuary

ERM

Claims

Data Scientist: Advanced Analytics Expert

Data Scientist: Machine Learning Expert

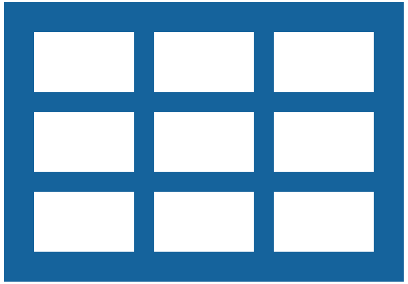
Data Analyst(s)

Tomorrow

AIR

How Are We Evaluating Risk Today?

How Are We Evaluating Risk Today?

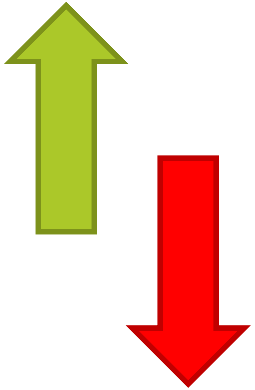


- Mortality tables built with limited information, e.g., age, sex, and smoking status and duration

- Underwriting often performed to meet minimum criteria for individual
- Designed to force firms to focus on an individual market and product
- Credit and debit systems require greater and greater expertise

AIR

How Current Methods Can Impact Your Business

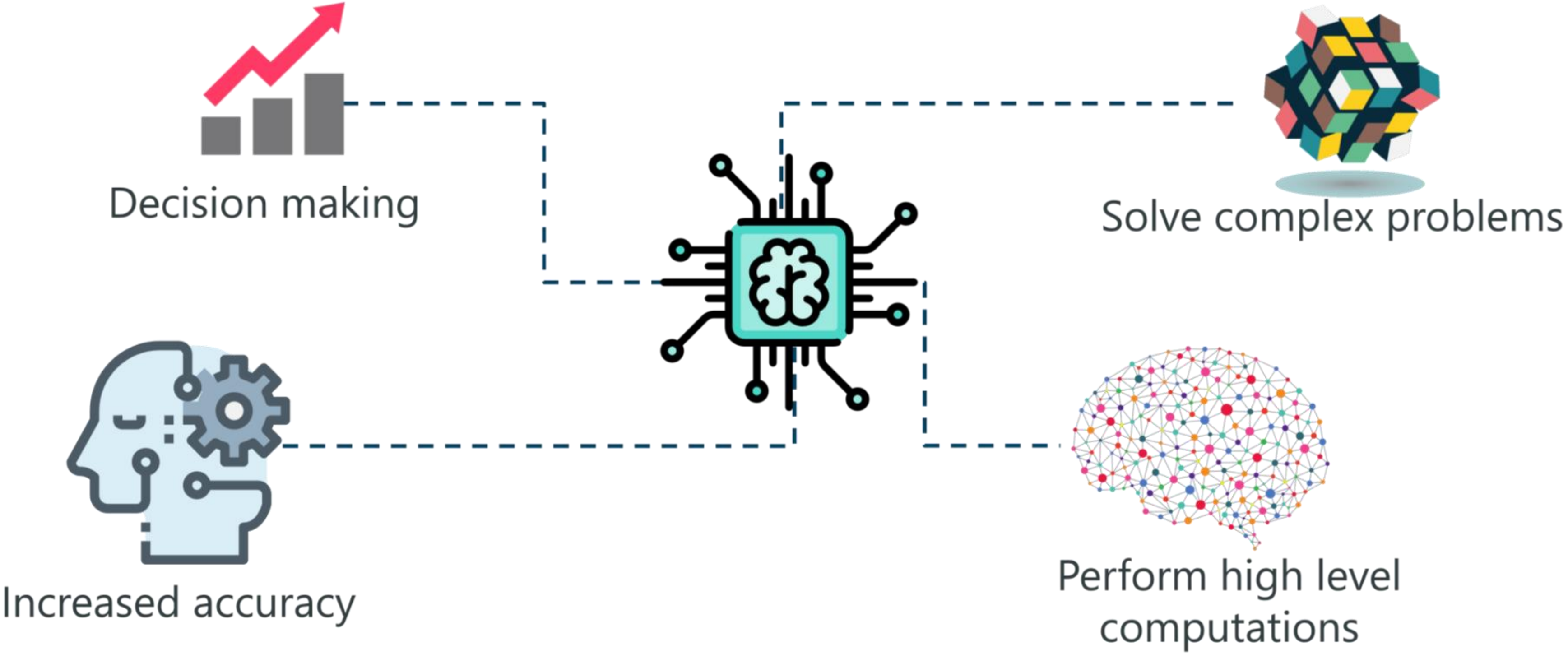


Over- or under-valuing risk

Anti-selection

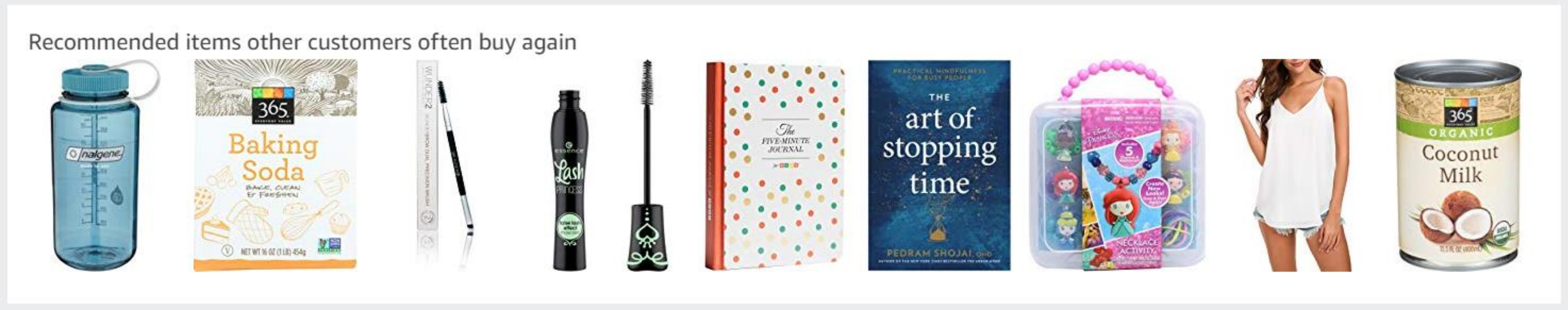
Lack of accounting for socioeconomic factors

Machine Learning Offers Numerous Benefits



What Can We Learn from Amazon?

amazon

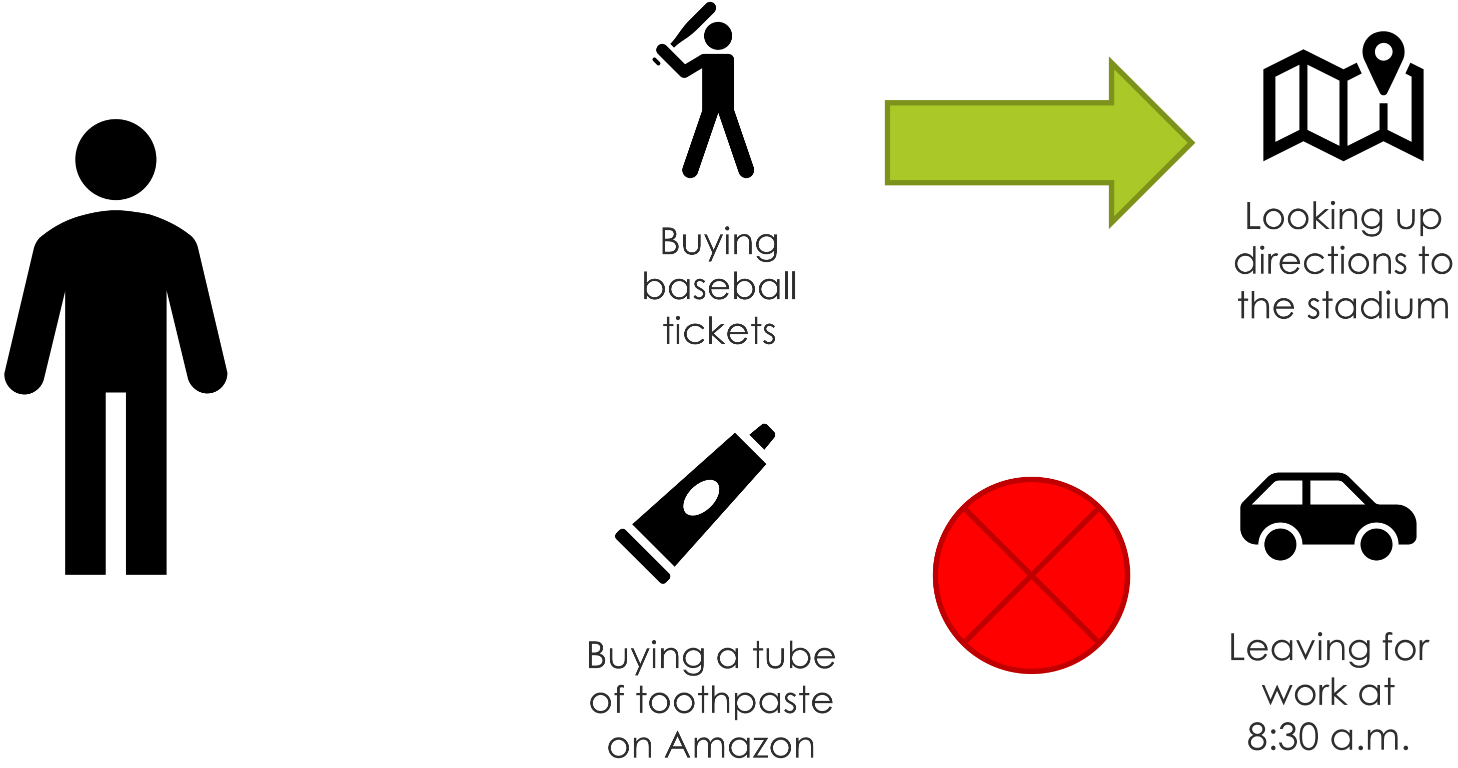


How do they know their customers—and why does it matter?

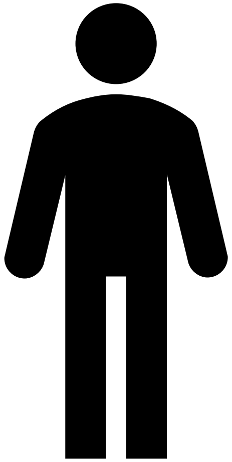
Can We "Amazon-ify" the Life Industry?

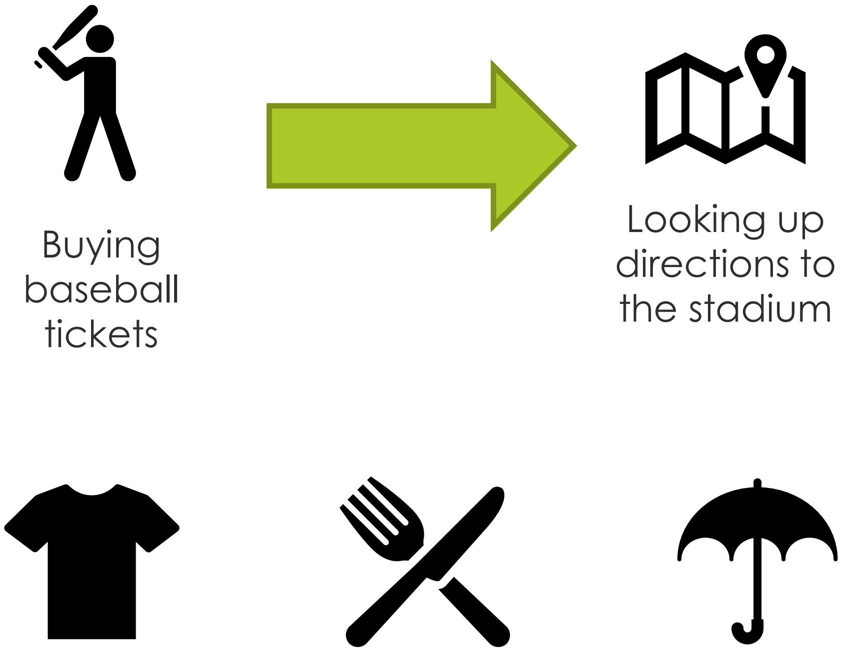

YOUTUBE GOOGLE+ FACEBOOK
COMMUNICATION
POPULAR WORDPRESS WEB
FRIENDS MEDIA GLOBAL FRIENDS MEDIA
MOBILE FLICKR SEARCH MOBILE FLICKR
COMMUNITY PEOPLE LINKEDIN
INFORMATION PINTEREST
SHARE CONTENT LIKE SHARE CONTENT
INTERNET MONEY LIKE
RADIO NETWORK ONLINE RADIO
DATA NEWS SOCIAL DATA NEWS SOCIAL
GOOGLE+ FACEBOOK
ONLINE
SOCIAL
YOUTUBE COMMUNICATION WEB
POPULAR WORDPRESS
MEDIA GLOBAL
FRIENDS FLICKR SEARCH MOBILE
LINKEDIN COMMUNITY PEOPLE
PINTEREST INFORMATION
LIKE SHARE CONTENT
FACEBOOK

Behaviors Can Be Independent or Correlated

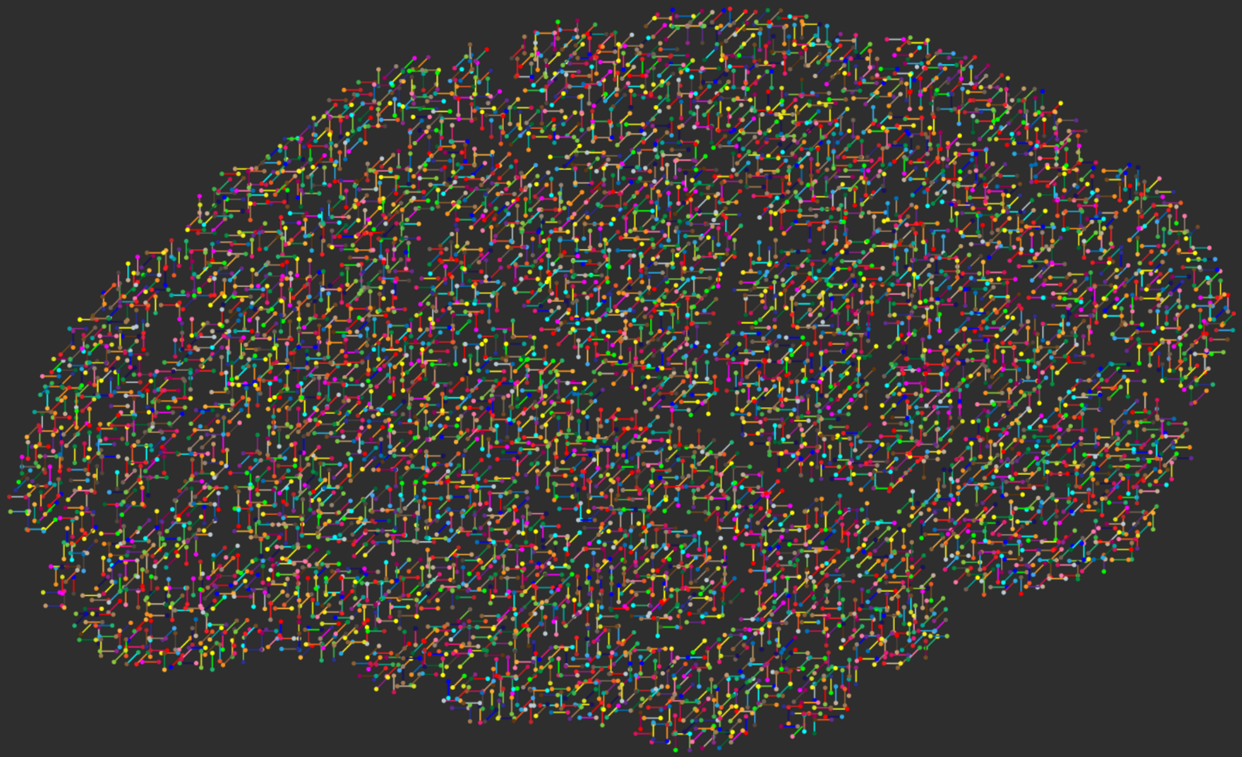


Behaviors Can Be Independent or Correlated





How Do We Find and Weigh Correlations in Mortality Risk Factors?

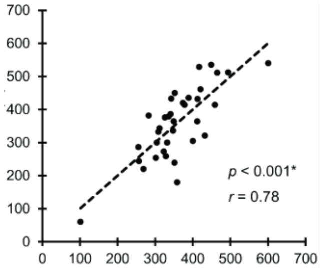


Exploring Possible Applications of Machine Learning

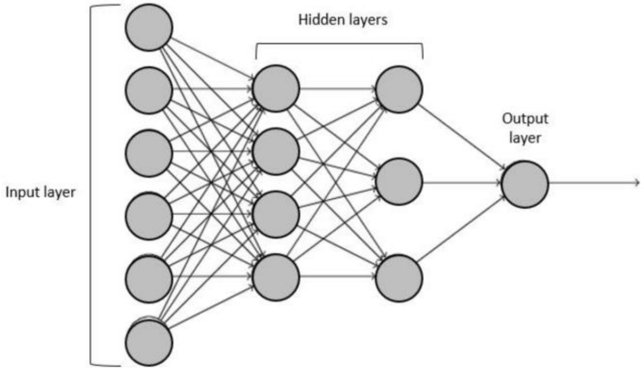
Comparing Modeling Methods

- **Traditional Statistical**
  - Regression, multi-variable regression, distribution fitting, etc.
  - Objective: understanding inference in the results

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + u_i.$$



- **Machine Learning**
  - Artificial neural networks (ANNs), deep neural networks, heuristic algorithms, etc.
  - Objective: predictability



Comparing Modeling Methods

	Traditional Statistical	Machine Learning
Advantages	• Easily assess parameters and relationships • Control model format • Compare against other known relationships	• Can smoothly take on very complex relationships • Often—not always—better at predicting outcomes
Disadvantages	• Must decouple any correlations • Model limited by modelers' knowledge • Complex relationships may be omitted, smoothed, or captured poorly	• Testing parameter output is complicated • Difficulty discerning some correlations • Subsequent testing can be problematic • Requires significant computational power

AIR

Multivariable Regression: Complexity Grows Quickly

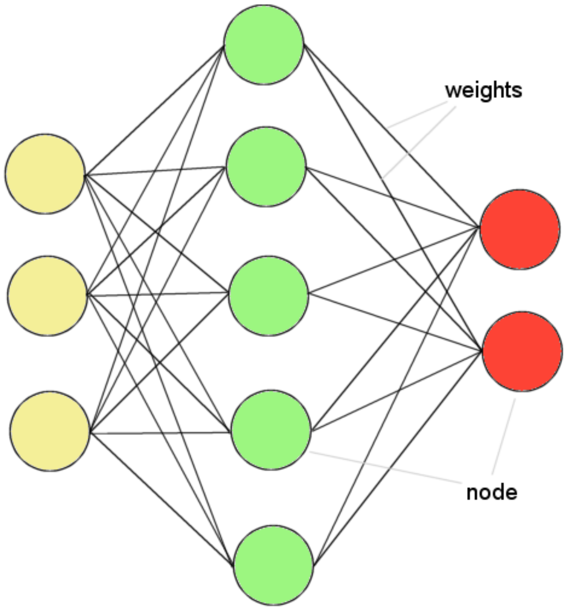
- Most simplistic form:
  - $mortality = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \cdots + \varepsilon$
- Alternatively:
  - $mortality = \beta_0 + \beta_1 *(x_1)^2 + \beta_2 * log(x_2) + \cdots + \varepsilon$
  - $mortality = \beta_0 + \beta_1 *(x_1)^2 + \beta_2 * (x_1) + \beta_3 * log(x_2) + \cdots + \varepsilon$
- What if x_1 and x_2 are related?
  - $f(x_2) = \beta_0 + \beta_1 *(x_1) + \varepsilon$
  - $mortality = \beta_0 + \beta_1 * (x_1) + \beta_2 * \varepsilon_{f(x_2)} + \beta_3 * (x_3) + \cdots + \varepsilon$

Artificial Neural Networks (ANN)

ANN: The Basics

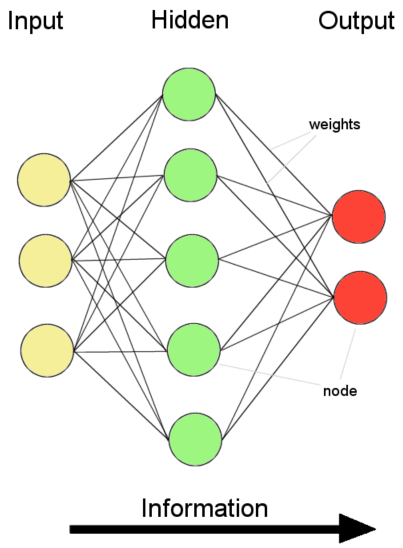
ANNs incorporate the **two fundamental components** of biological neural nets:

1. Neurons *(nodes)*
2. Synapses *(weights)*

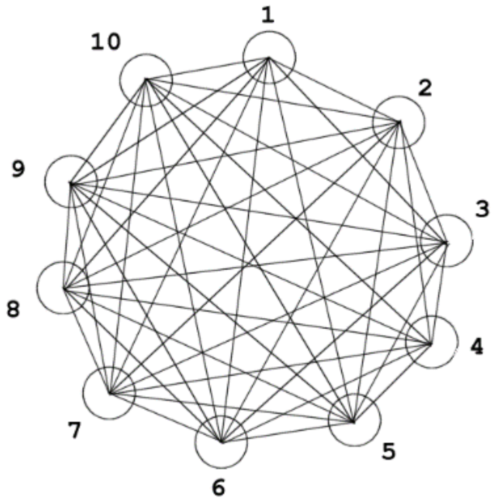


AIR

ANN: The Basics



Feed Forward Networks



Recurrent Networks

How Is a "Simple" ANN Model Applied?

1. **User**
   - Defines data for training and testing
   - Defines input data, number of neurons, their shape, and model output

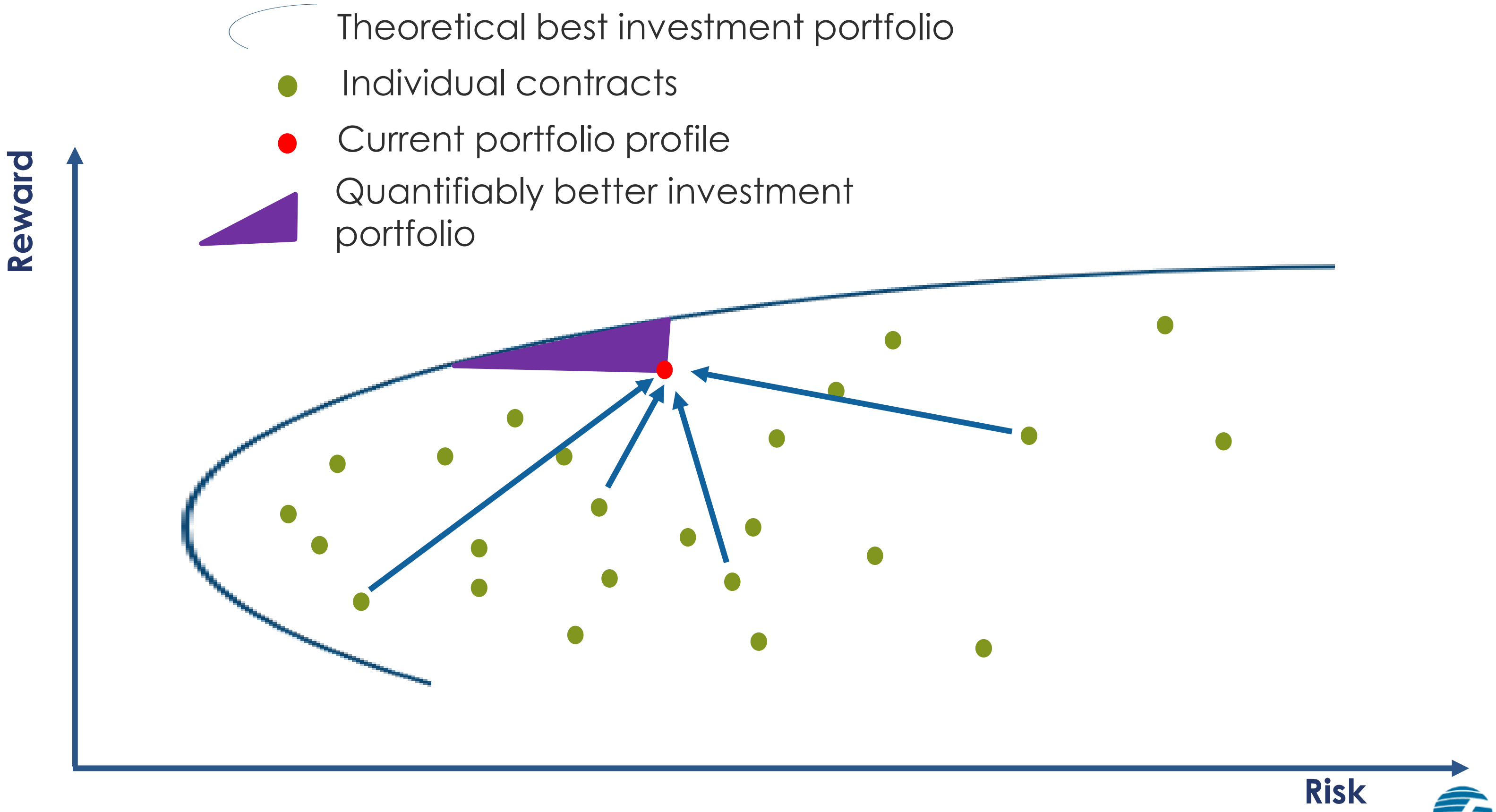
2. **Model**
   - Derives initial weights for each synapse
   - Compares the estimated output vs. the actual
   - Updates the weights for each synapse, with the updates being more refined over time
   - Iterates through this process until it achieves its overall objective

Testing an ANN Model

- Compare against testing data set to determine if the model is designed appropriately
- Run select data inputs to understand the output
  - Allows you to perform comparative analyses on specific iterations of data

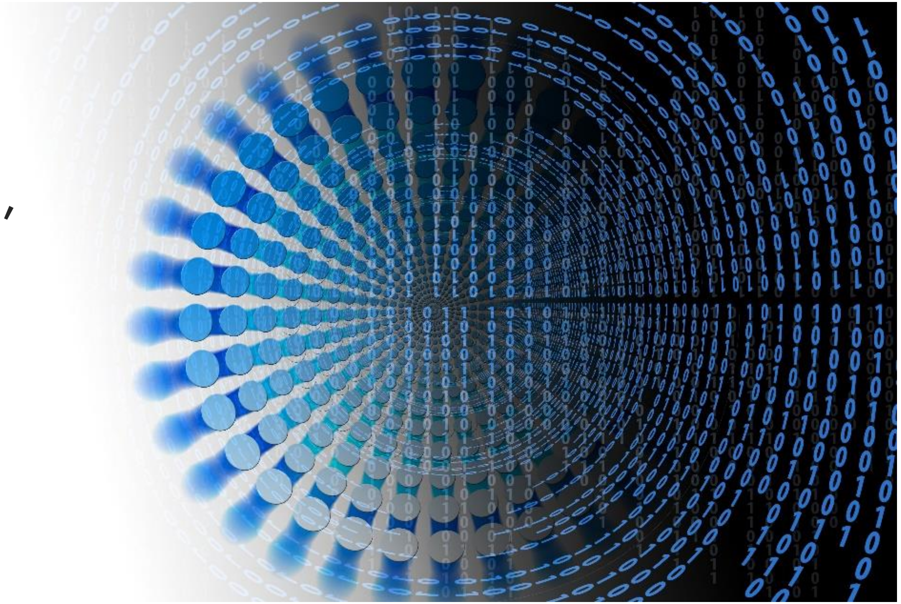
Genetic Algorithms (GA)

Envisioning the Efficient Frontier

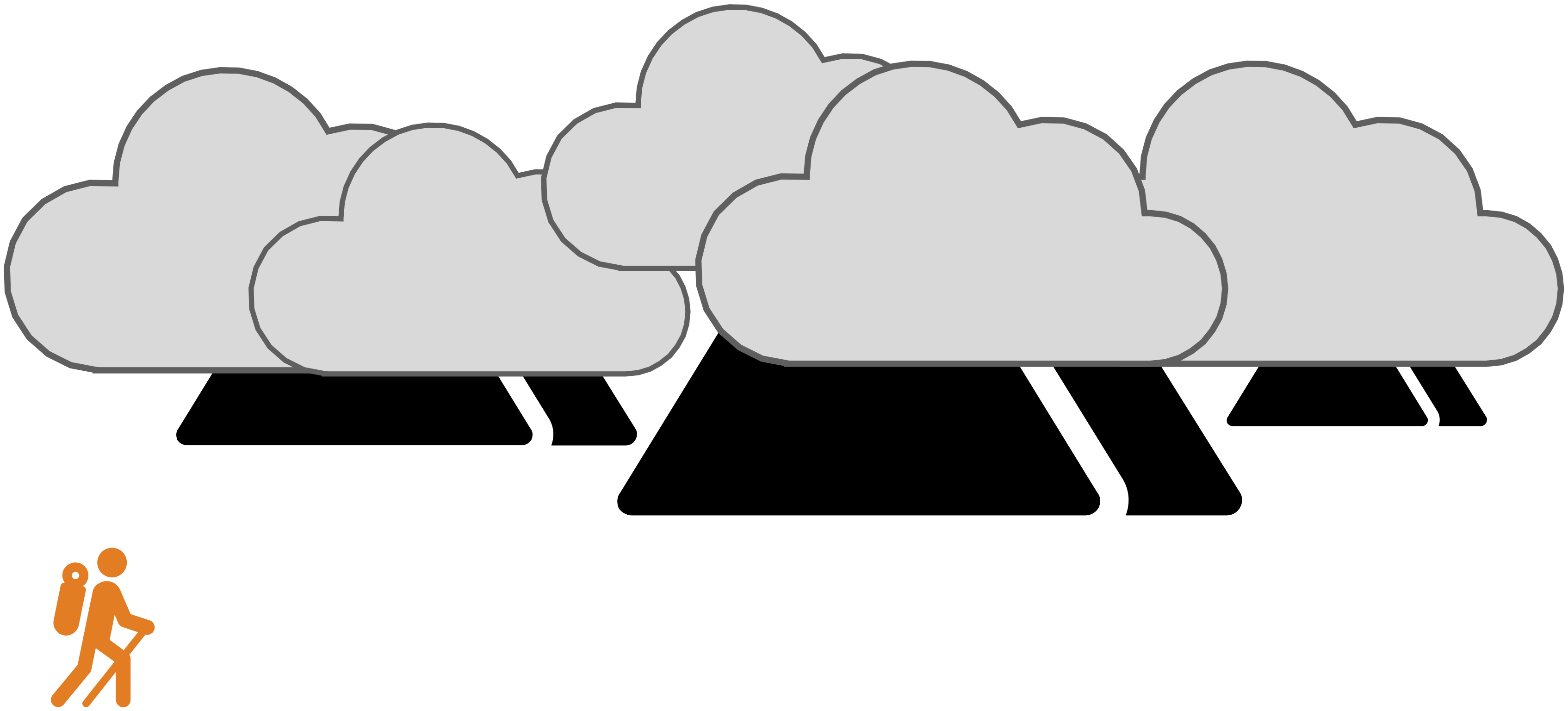


Quantifying the Size of the Problem

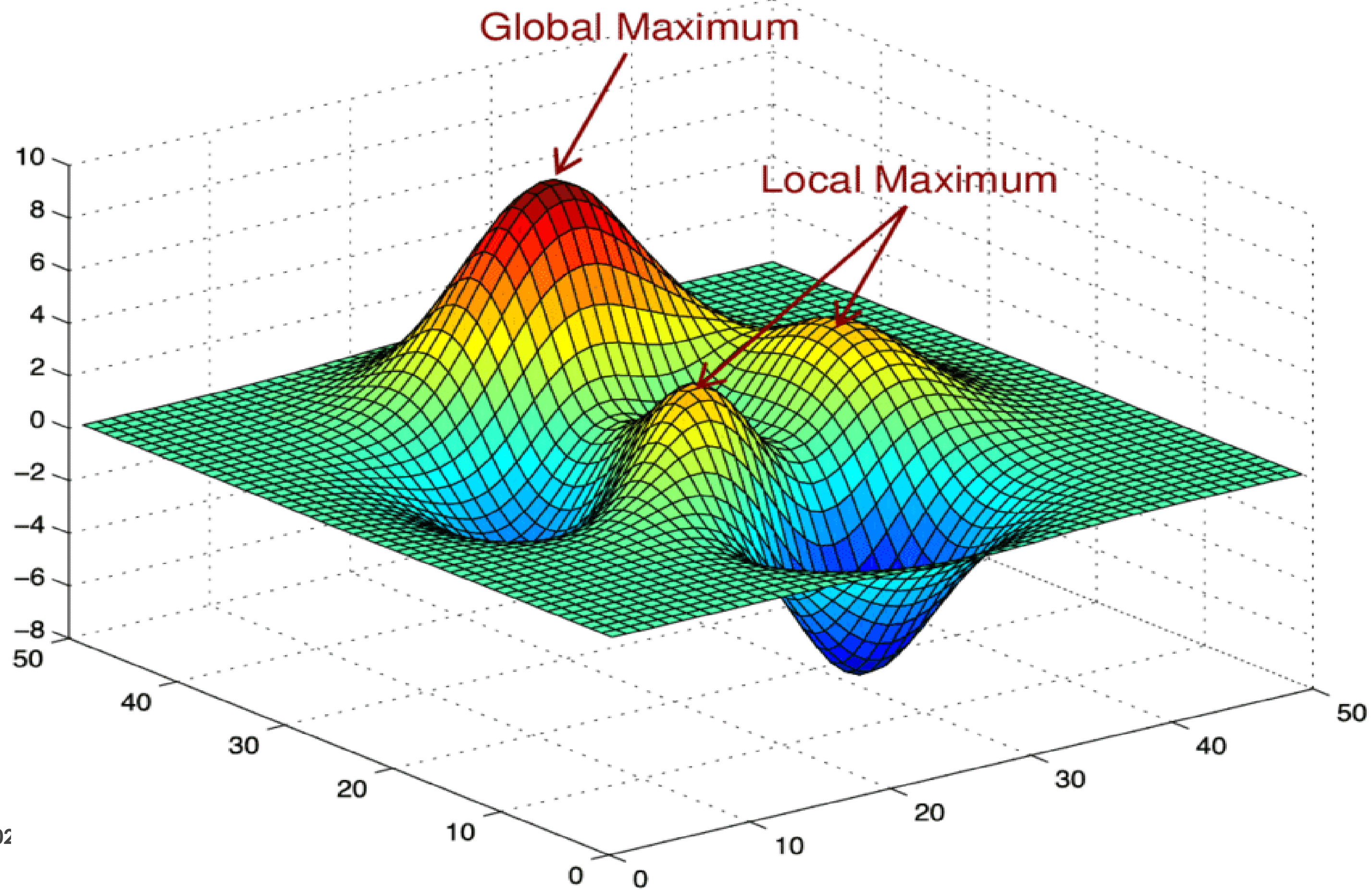
- With a fixed number of contracts, if each contract can only be sold as a whole, the resulting theoretical combinations will be:
  - 5 contracts = 32
  - 10 contracts = 1,024
  - 100 contracts = $1.27 * 10^{30}$



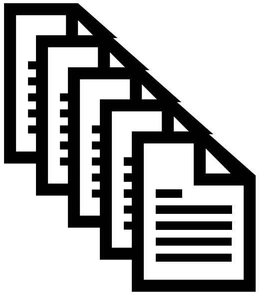
Can We Deduce Which Is the Highest Peak?



Global Maximum vs. Local Maximum



How It Works: Reinsurance Example



- Problem
  - Investor has USD 1 billion to invest in the reinsurance market
  - There are 1,000 contracts, each with their own price, expected return, risk profile, and correlation with other contracts

- Objective
  - What is the best investment portfolio I can derive?

Genetic Algorithm Method

Available Investments

A	B	C	D	E	F	G	H	I	J	K	L	M	●●●

Each has its own **expected return**, **risk profile**,
and **correlation** to other investments

Simulated Investments

Sim. 1	1	0	0	0	1	1	0	1	0	0	0	1	1	●●●
Sim. 2	1	0	1	1	1	0	0	0	0	1	0	0	1	●●●
Sim. 3	0	1	1	1	0	1	0	0	1	1	0	0	0	●●●
Sim. 4	●●●													

Genetic Algorithm Method: Rank and Order

Iteration 1		**Rank Iteration 1**	
Sim. 1	→	Sim. 111	Keep
Sim. 2		Sim. 12	
Sim. 3		Sim. 405	
Sim. 4		Sim. 8	
●		●	
●		●	Discard
●		●	
Sim. n		Sim. #	

AIR

Genetic Algorithm Method: Offspring

Available Investments

	A	B	C	D	E	F	G	H	I	J	K	•••
Sim. 111	1	0	0	0	1	1	0	1	0	0	1	•••
Sim. 12	0	1	1	1	1	1	0	1	0	1	0	•••
Offspring 1	1	0	0	0	1	1	0	1	0	0	1	•••
Offspring 2	0	1	1	1	1	1	0	1	0	1	0	•••

Genetic Algorithm Method: Offspring

Available Investments

	A	B	C	D	E	F	G	H	I	J	K		
Sim. 111	1	0	0	0	1	1	0	1	0	0	1	•••	Parents
Sim. 12	0	1	1	1	1	1	0	1	0	1	0	•••	
Offspring 1	1	1	1	0	1	1	0	1	0	1	1	•••	Children
Offspring 2	0	0	0	1	1	1	0	1	0	0	0	•••	

Swapped

AIR

Repeat with Subsequent Iterations

Iter. 2

Parents

Children

New Simulations

Rank Iter. 2

Sim. 198
Sim. 111
Sim. 20
Sim. 88

Keep

Discard

Sim. #

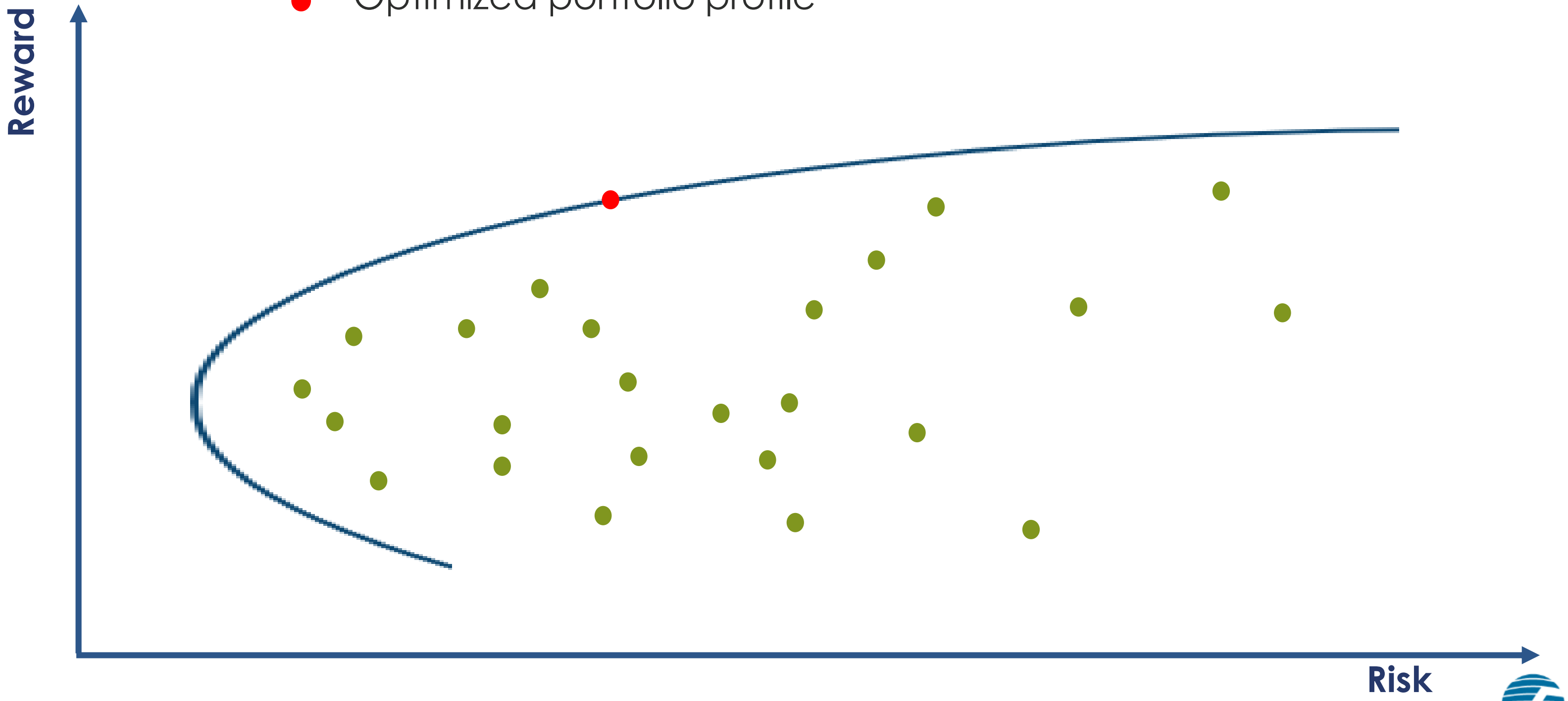
This iterative process can consider **constraints** as well as **objectives**, e.g., life insurance limit, age restrictions, disease restrictions, broker caveats, percent change from current investments, etc.

Reaching the Efficient Frontier

Theoretical best investment portfolio

Individual contracts

Optimized portfolio profile



How AIR Applies Machine Learning Techniques

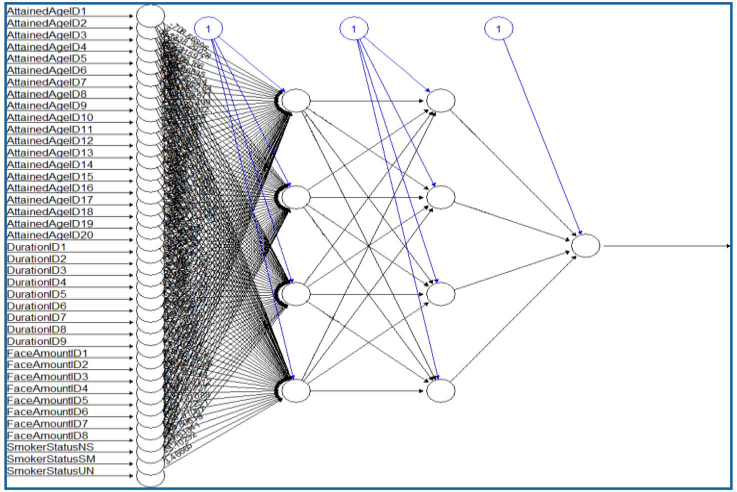
AIR

How AIR Applies Machine Learning Techniques

INPUT	METHODOLOGY

- Age
- Sex
- Duration of policy
- Face amount/annuity amount
  (income level predictor)
- Health status
  - Smoking status
- Socioeconomic status
- Post-term period
- Geographical location
- Policy options
- …

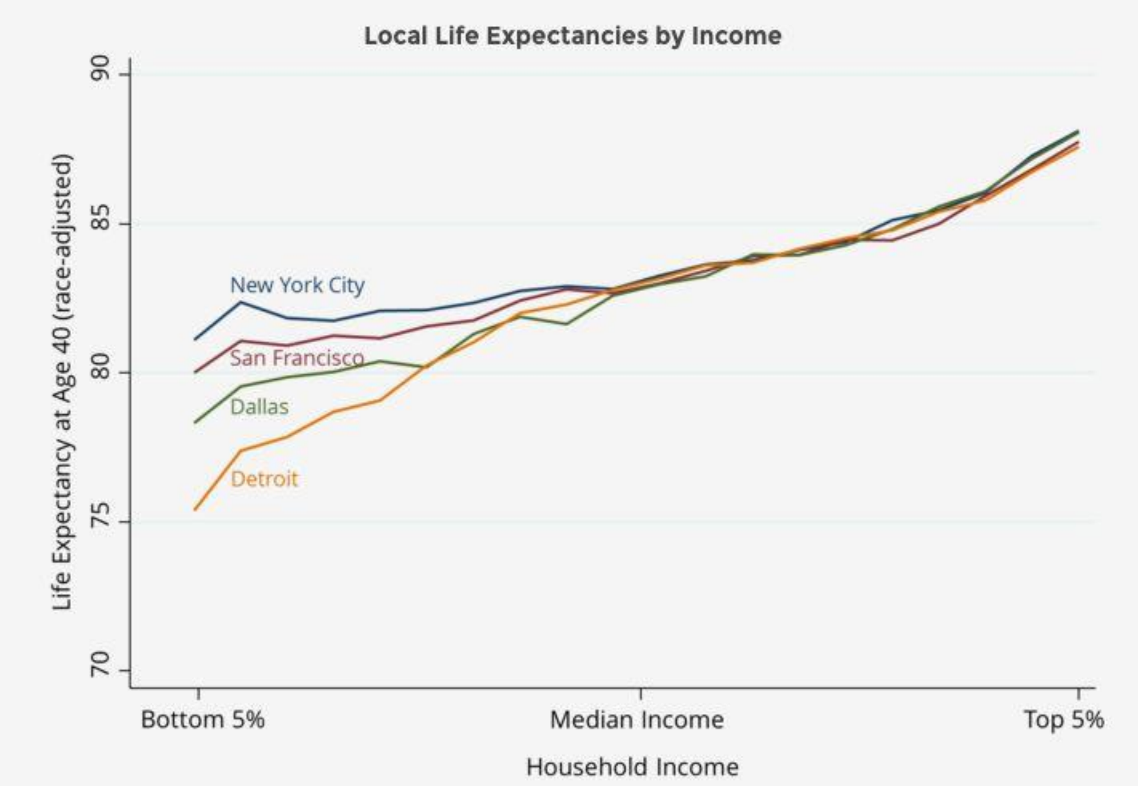
Artificial Neural Network



Logistic Regression

$$\begin{aligned}\mathrm{Log}(\mathrm{mort}_{x,gender}) = {} & \beta_0^{\,x,gender} + \beta_1^{\,x,gender}\mathrm{Age} + \beta_2^{\,x,gender,tems}\mathrm{Policy\ Terms} \\ & + \beta_3^{\,x,gender}\mathrm{Duration} + \beta_4^{\,x,gender}\mathrm{Duration}^2 \\ & + \beta_5^{\,x,gender}\mathrm{PostTerm} + \beta_6^{\,x,gender}\mathrm{PostTerm}^2 \\ & + \beta_7^{\,x,gender,group}\mathrm{Age}_{\mathrm{Smk}} + \beta_8^{\,x,gender,InsGroup}\mathrm{InsuranceUnderwriting} \\ & + \beta_9^{\,x,gender,Face}\mathrm{FaceAmount}\end{aligned}$$

Improvement Rates Vary Across Demographics



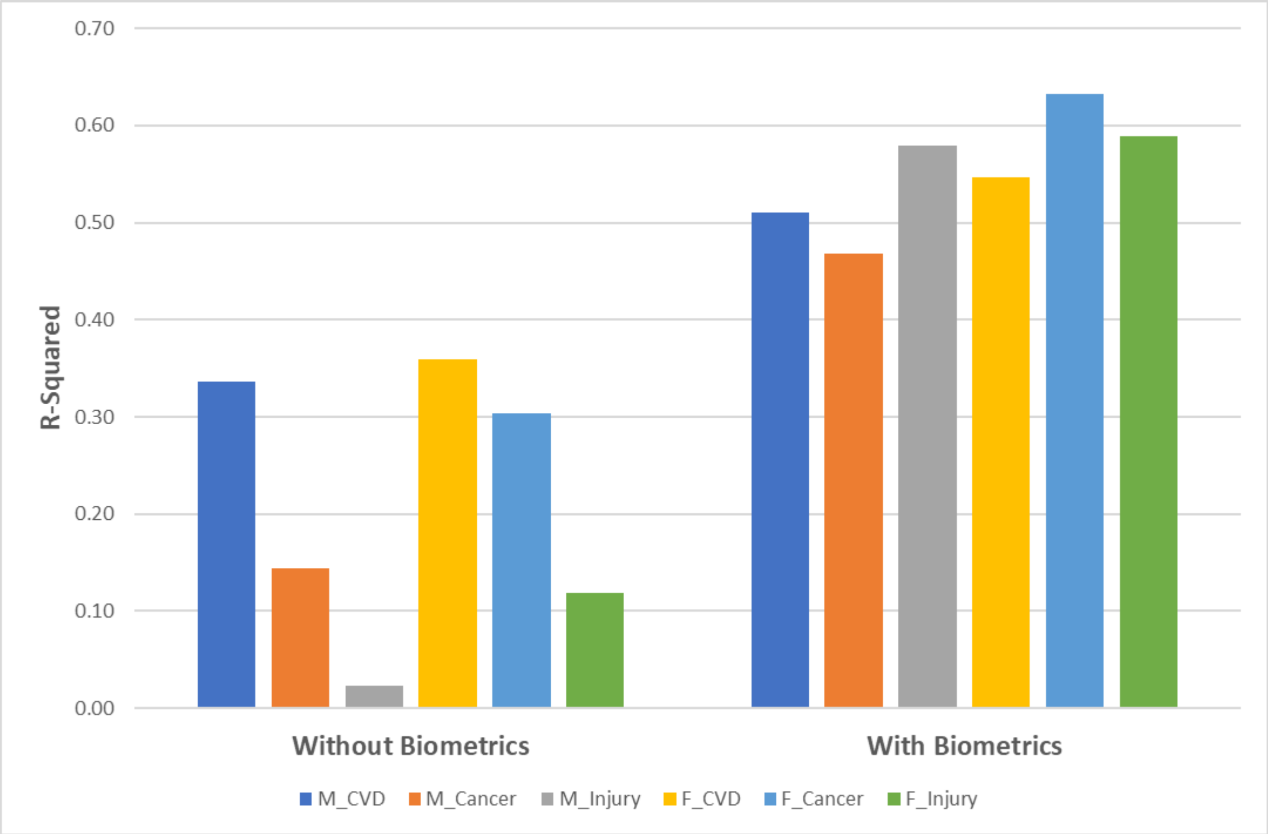
Source: The Health Inequality Project

Mortality Improvements Are Not Simply a Function of Income

- Income is related to mortality, but is also correlated with other factors

- Breaking down biomedical information and using a multi-variable dynamic model provides a better view of risk

Biometric Data Enhances Understanding of Trends



Life Risk Models

Longevity

- Cause of death modeling
- Stochastic mortality projections
- Individual risk factors

Excess Mortality

- Pandemic, earthquake, terrorism
- Antibiotic resistance

Insurance Mortality Adjustment

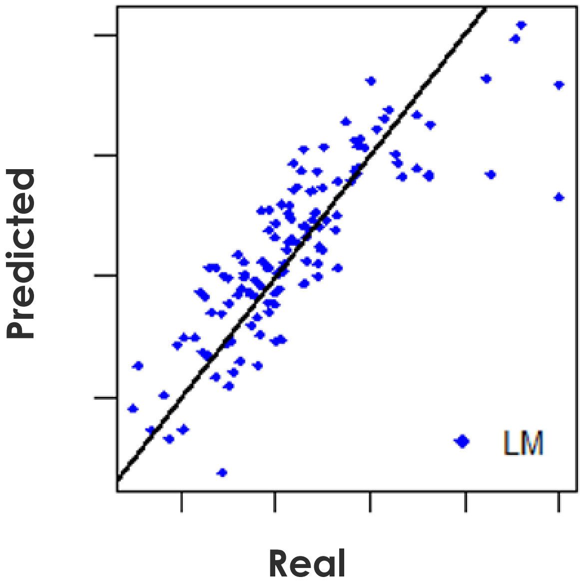
- Reported policyholder data
- Inferences from reported data

Individual Adjusted Mortality Factor

Hypothetical Fits

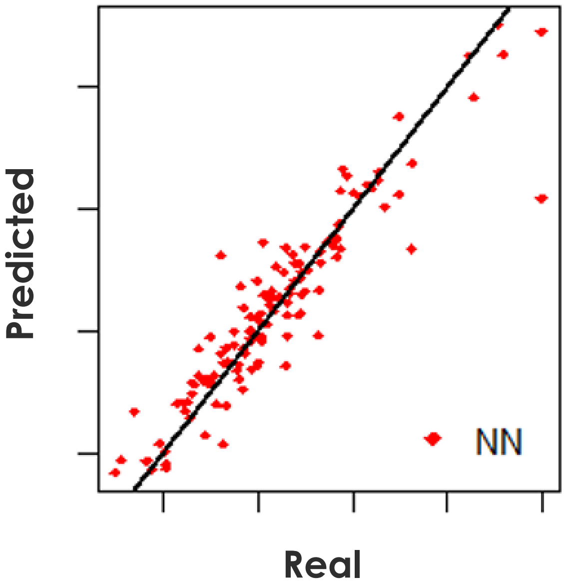
Linear Model (LM)

σ = 4.66



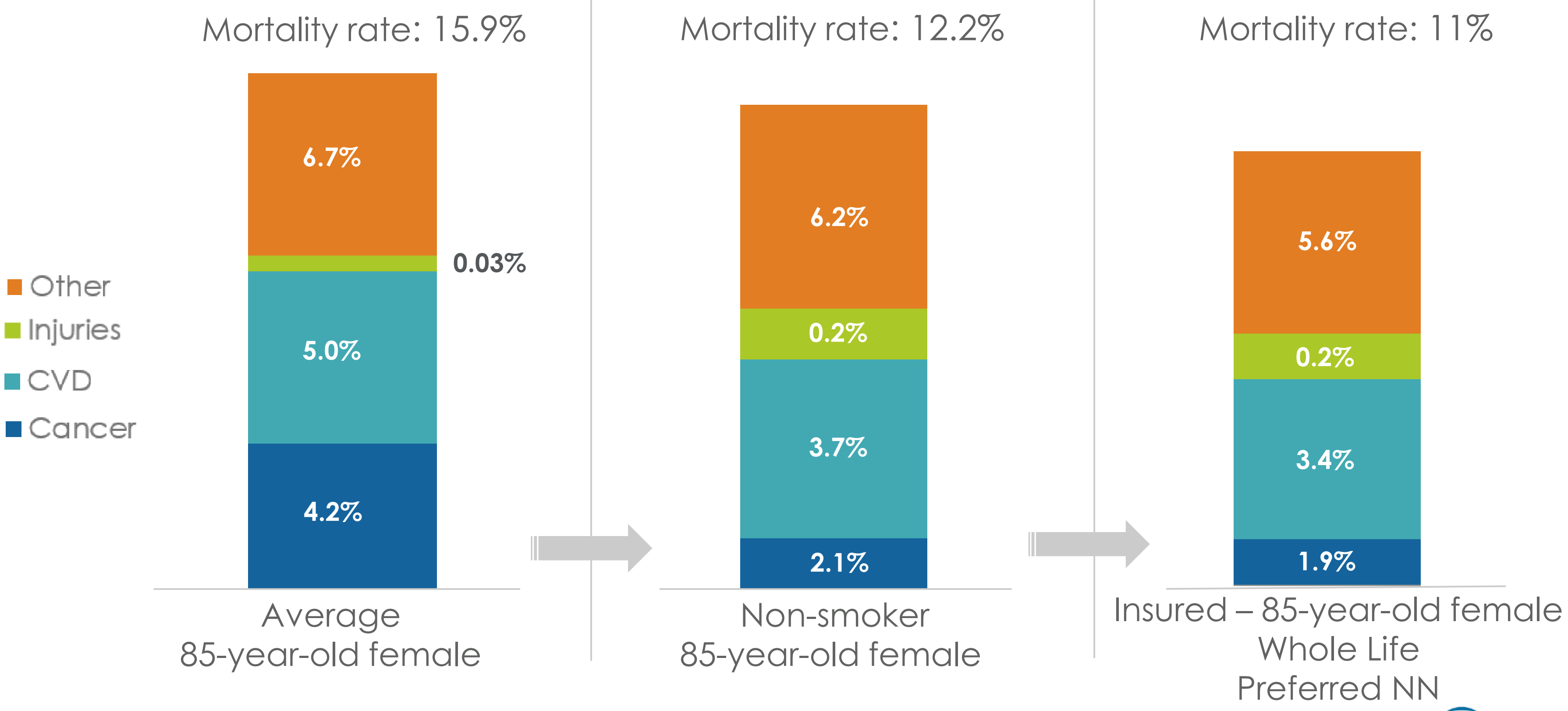
Neural Network (NN)

σ = 3.97



AIR

Mortality Adjustment for Insured Individuals



Key Takeaways

1. Machine learning is already changing our world—and it shows no signs of slowing down
2. Machine learning applications in the life insurance industry are helping us rethink risk
3. The value of AI underwriting will surpass USD 20 billion by 2024 from USD 1.3 billion in 2019 (Juniper Research)


Verisk Life Seminar
The Westin New York at Times Square
March 10, 2020
· Industry perspectives on an evolving life insurance landscape
· Innovations using machine learning and predictive analytics
· Capturing the impact of shifting life expectancy trends
· Leveraging emerging data sources to better understand and manage risk
· Enhancing underwriting, claims, and portfolio management workflows
$125 Early Bird Rate available now

Thank you!
Questions?

Learn more at Verisk.com/Life