NATIONAL FAIR HOUSING ALLIANCE

The PPM Framework: A New Way of Auditing Algorithmic Systems

fairplay
FAIRNESS-AS-A-SERVICE

A Conceptual Introduction to Machine Learning

Kareem Saleh, Founder and CEO of FairPlay
January 19, 2022

What is Machine Learning?

Traditional Programming

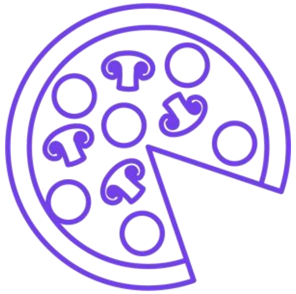
Ingredients

+

Recipe

MAKES

=



Outcome

But what if you don’t have a recipe?

Traditional / Programming

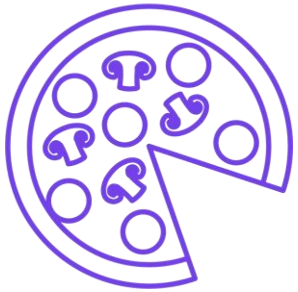
+

MAKES =

Ingredients | Recipe | Final Product

Machine Learning Algorithm

+



FIGURES OUT →

Ingredients | Outcome | Recipe

How is Machine Learning Being Used in Housing?

TENANT SCREENING

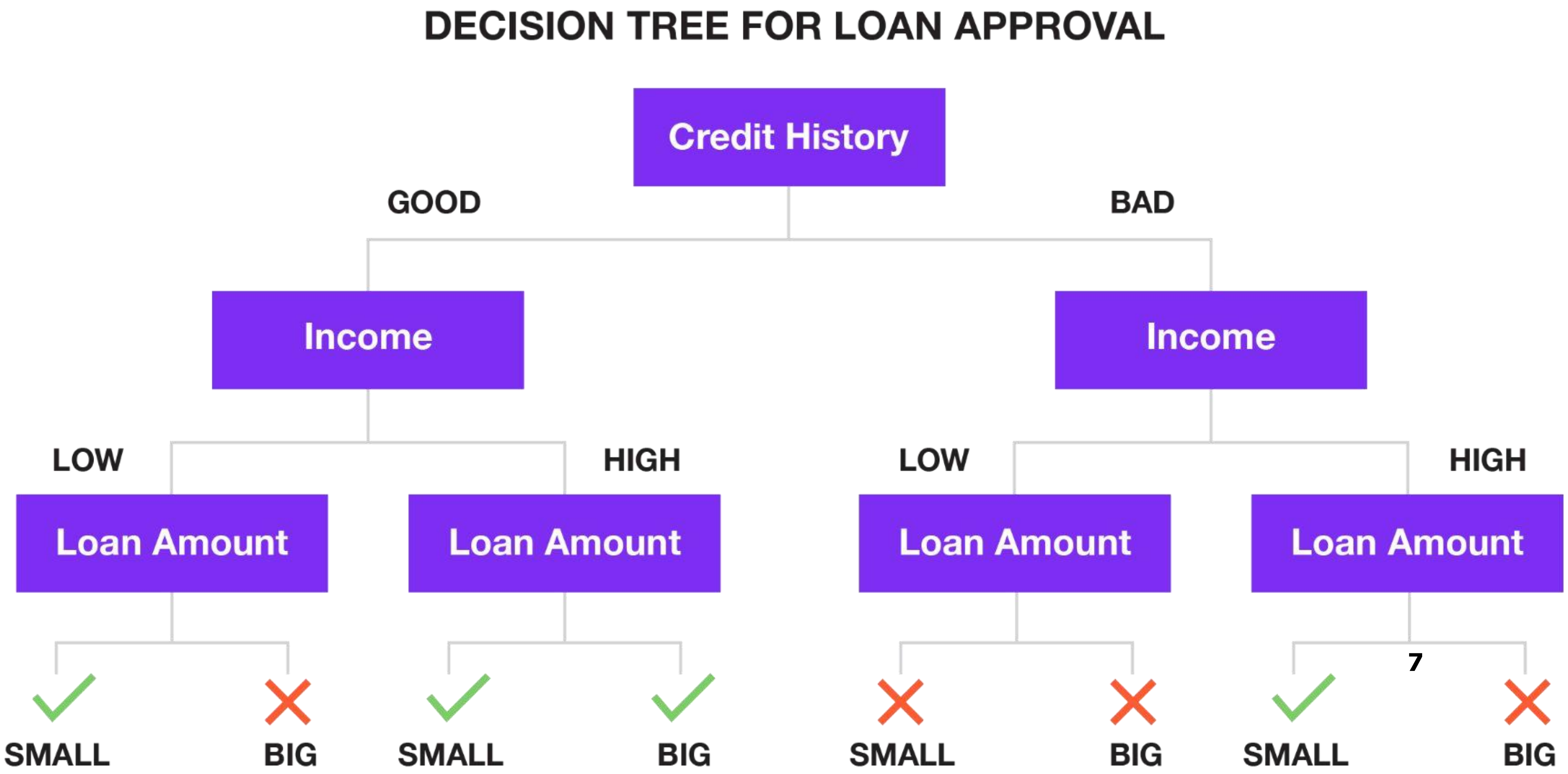
MORTGAGE UNDERWRITING

HOME APPRAISAL

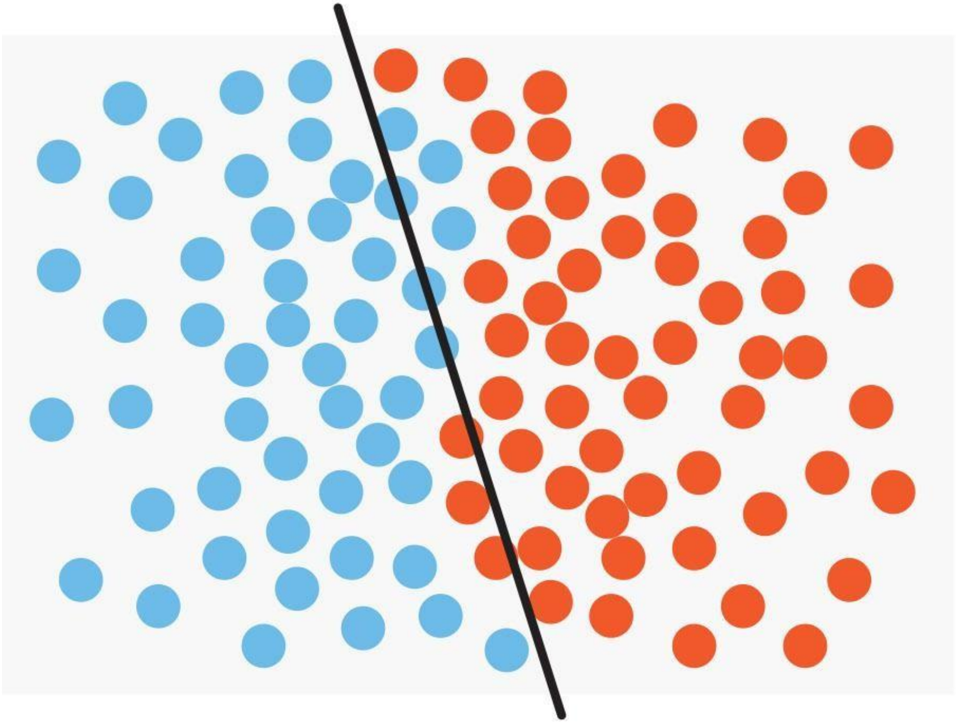
Lending decisions used to be made by humans who tried to assess an applicant's creditworthiness

LOAN OFFICER

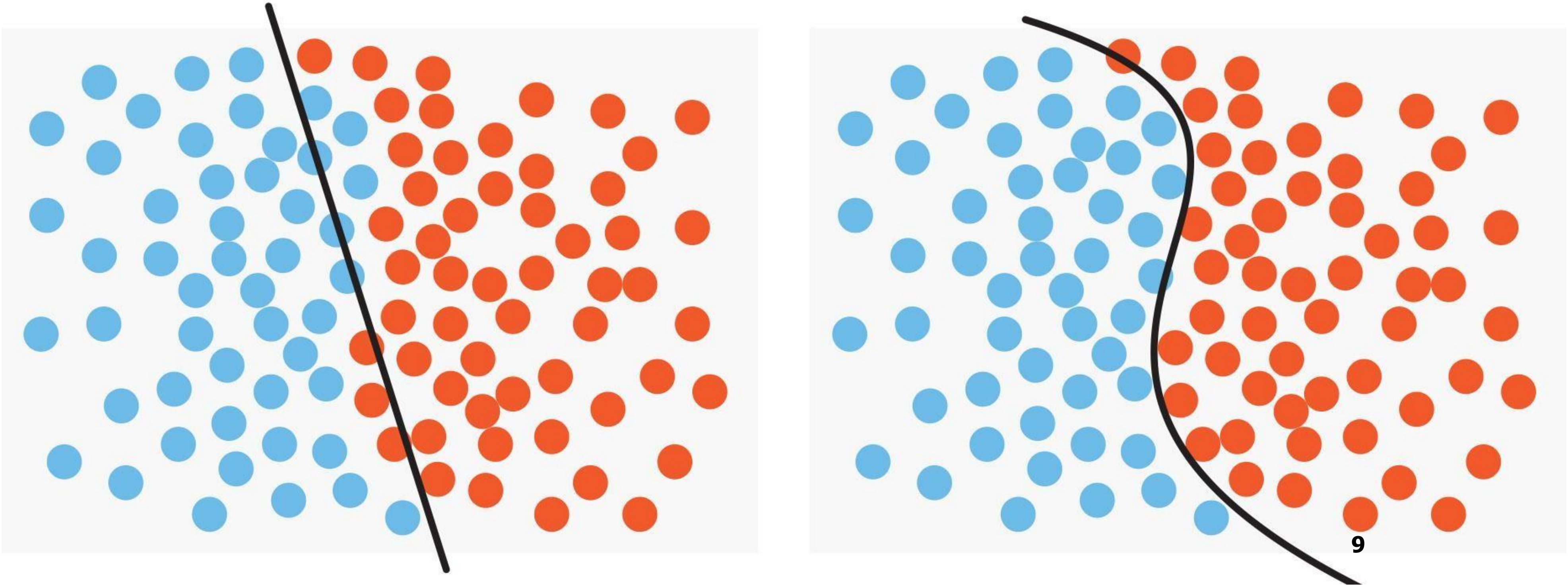
In the 1980s we started using math instead of humans to make lending decisions because it was seemingly more “neutral” and “objective”



The old math assumes relationships between ingredients (variables) are straight-forward…



The old math assumes relationships between ingredients (variables) are straight-forward, but the world is actually complex and nonlinear



Imagine we tried to build a model that predicts sex

CAN WE DETERMINE GENDER USING HEIGHT?

+ Men, on average, are taller than women.

fairplay

Height is somewhat but not perfectly predictive of sex

CAN WE DETERMINE SEX USING HEIGHT?

+ Men, on average, are taller than women.
+ But there are tall women and short men, and many people that are the same height.
+ ***Using height alone isn't very accurate.***

Including weight adds predictive power, but the model still isn't perfect

CAN WE DETERMINE SEX USING HEIGHT + WEIGHT?

+ Because men are, on average, heavier than women of the same height, accuracy would improve.

Using height and weight to predict gender causes kids to be classified as women

CAN WE DETERMINE SEX USING HEIGHT + WEIGHT?

- Because men are, on average, heavier than women of the same height, accuracy would improve.
- ***But children would mostly be misclassified as women.***

Is birthdate predictive of Sex?

CAN WE DETERMINE SEX USING HEIGHT + WEIGHT + BIRTHDATE?

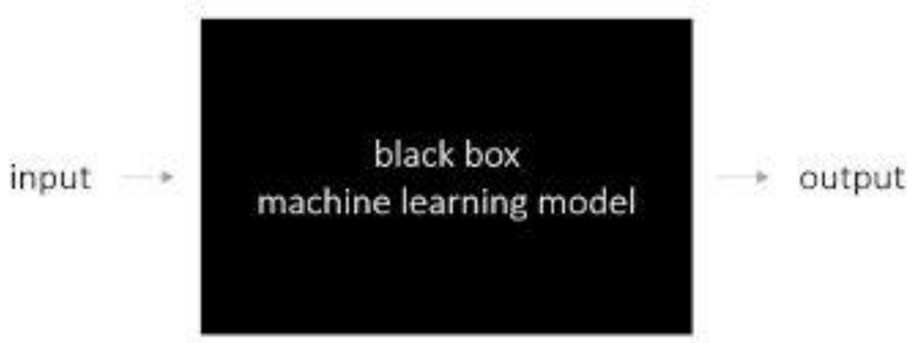
- Knowing people's age eliminates the misclassification of children, and improves the model's accuracy of determining gender across all age groups.

- ***But the idea that birthdate can help determine sex is not obvious***

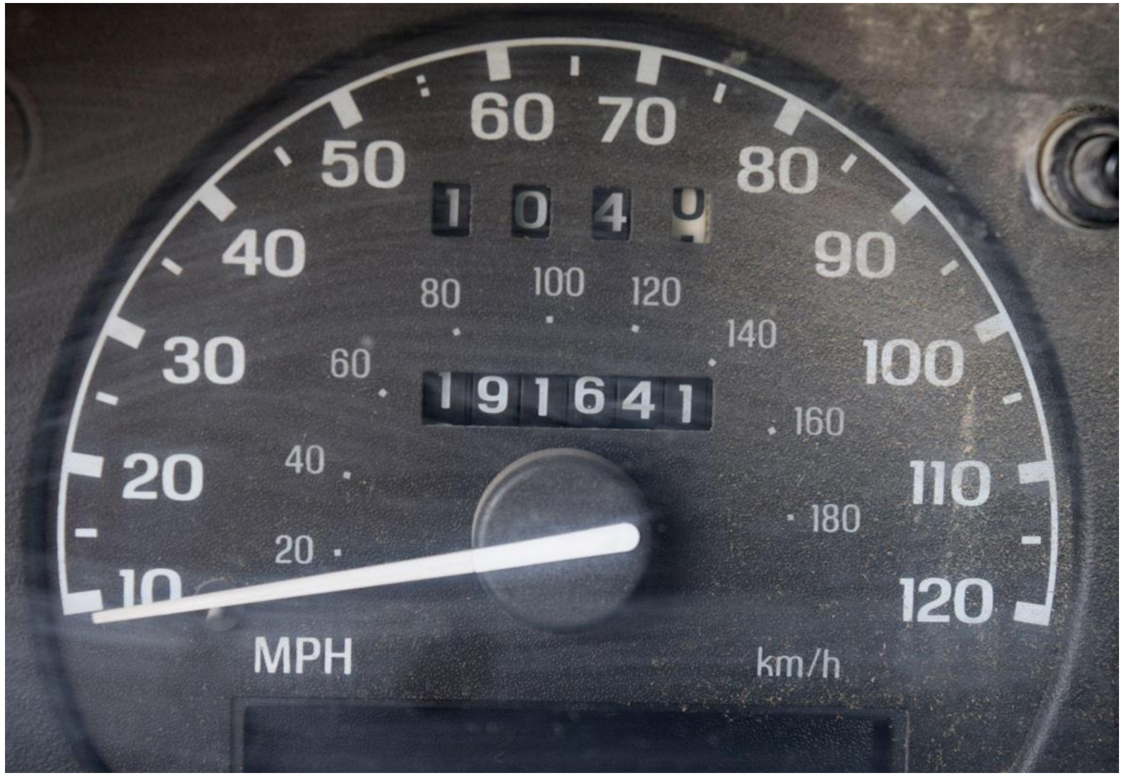
fairplay

When there are a lot variables, the recipe connecting them to an outcome can be so complex no one can understand it:

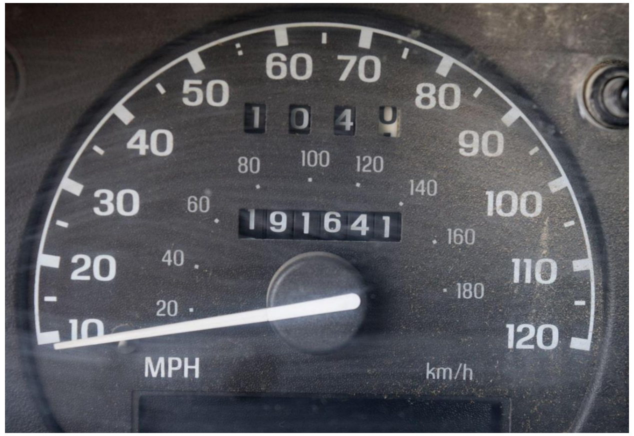
When there are a lot variables, the recipe connecting them to an outcome can be so complex no one can understand it: a black box



The problem with the black box is the risk it will be biased in ways we don’t understand.



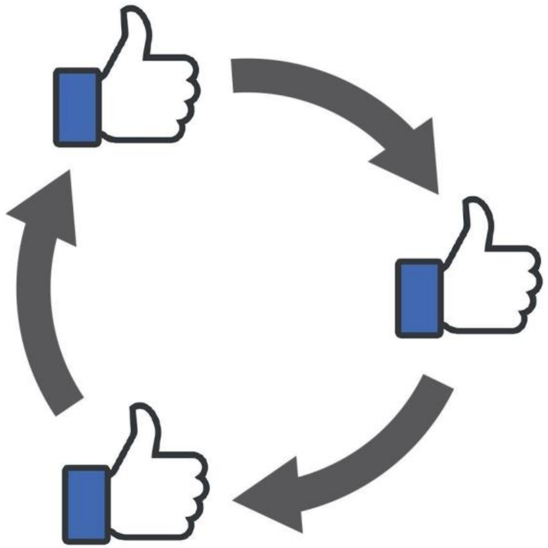
If you're buying a high mileage car in Nevada there's a big probability you're a person of color



ML algorithms relentlessly refine their "recipe" to achieve the best outcome (aka they adjust to better accomplish their target or objective)

Every algorithm must be given a target

Social media algorithms seek to maximize their target: engagement

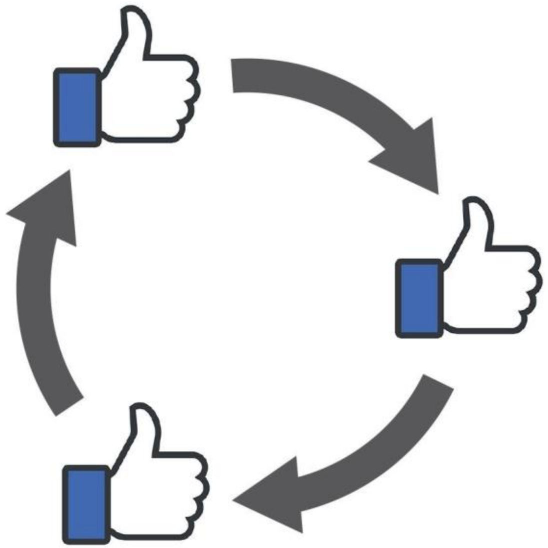


Every algorithm must be given a target

Social media target: Maximize engagement

The algorithm single mindedly focuses on engagement regardless of whether it's good for your health or good for society

Every algorithm must be given a target



Social media target: Maximize engagement

Without regard for societal harm

Giving an algorithm one target is problematic: imagine a self-driving car whose only target was to get you from point (a) to point (b)

Target: Get from point (a) to point (b)

Self-driving cars have a second target: Safety

Target: Get from point (a) to point (b)

Second Target: Safety (obey traffic laws; avoid accidents with cars, pedestrians, cyclists, etc)

We can do this in financial services: Target a low risk of default . . .

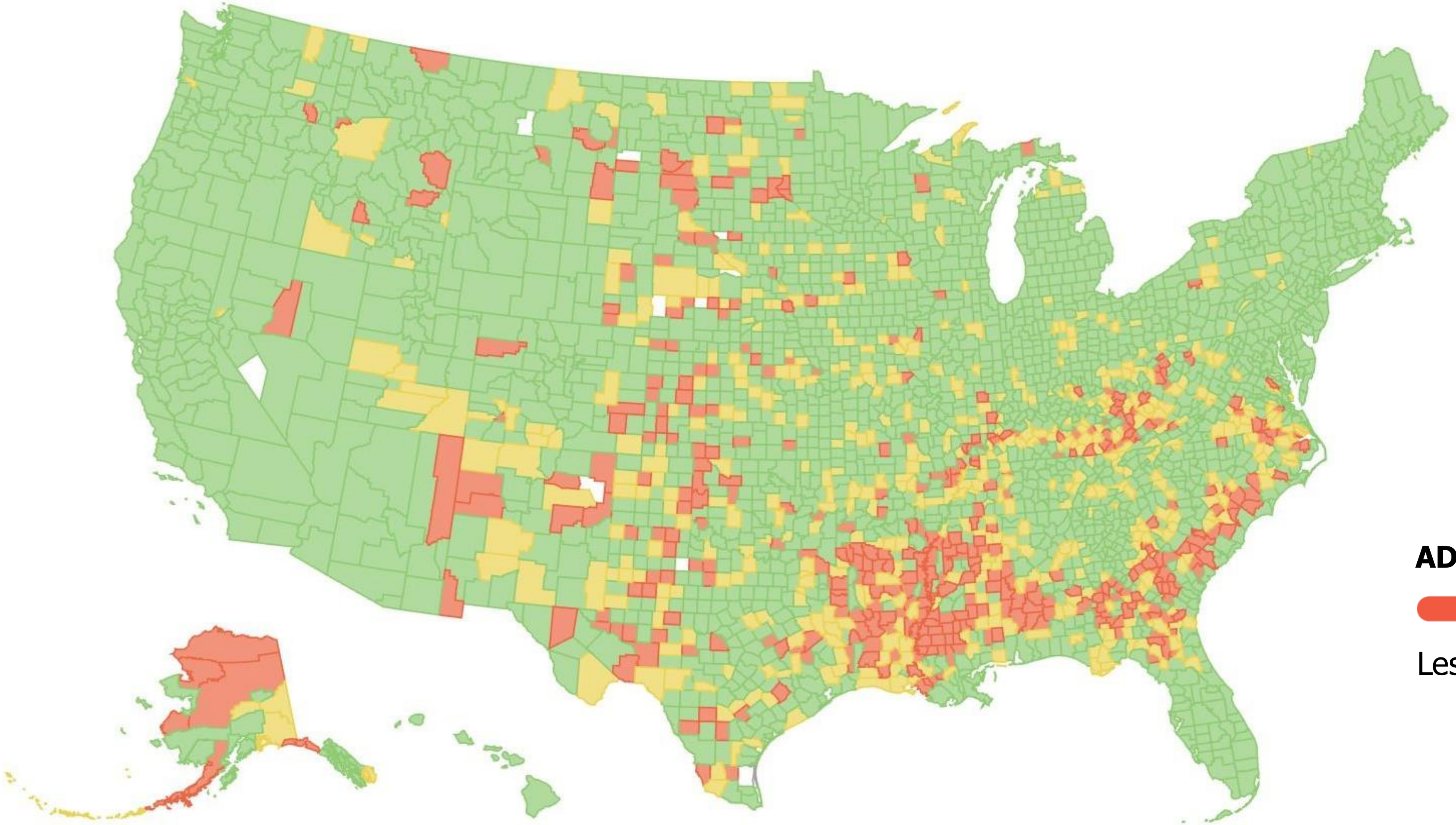
Target: Low risk of default

We can do this in financial services: Target a low risk of default while also targeting fairness

Target: Low risk of default

Second Target: Fairness

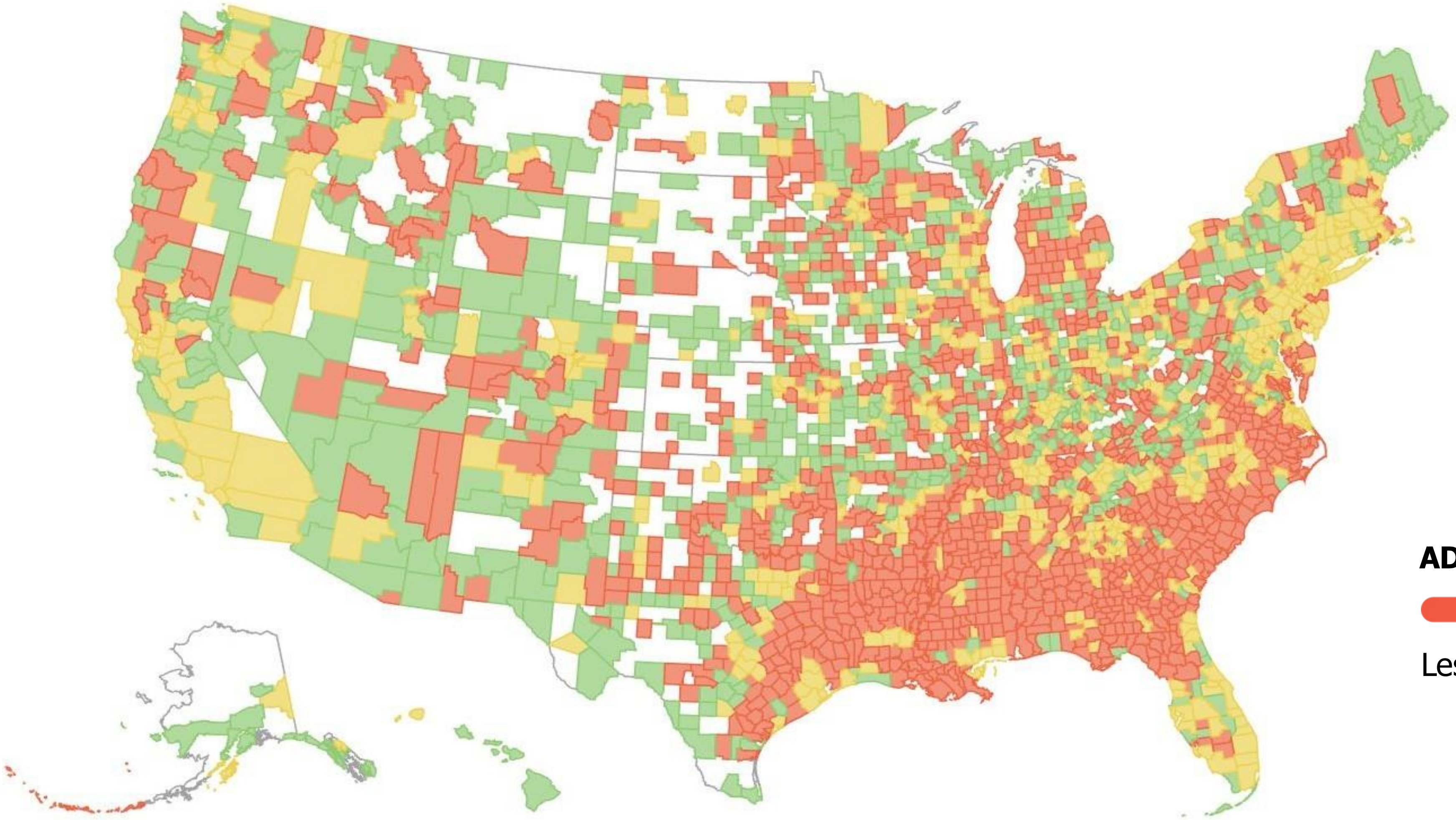
U.S. Mortgage Fairness in 2020: Female



ADVERSE IMPACT RATIO:

Less than 80% | 80 to 90% | Over 90%

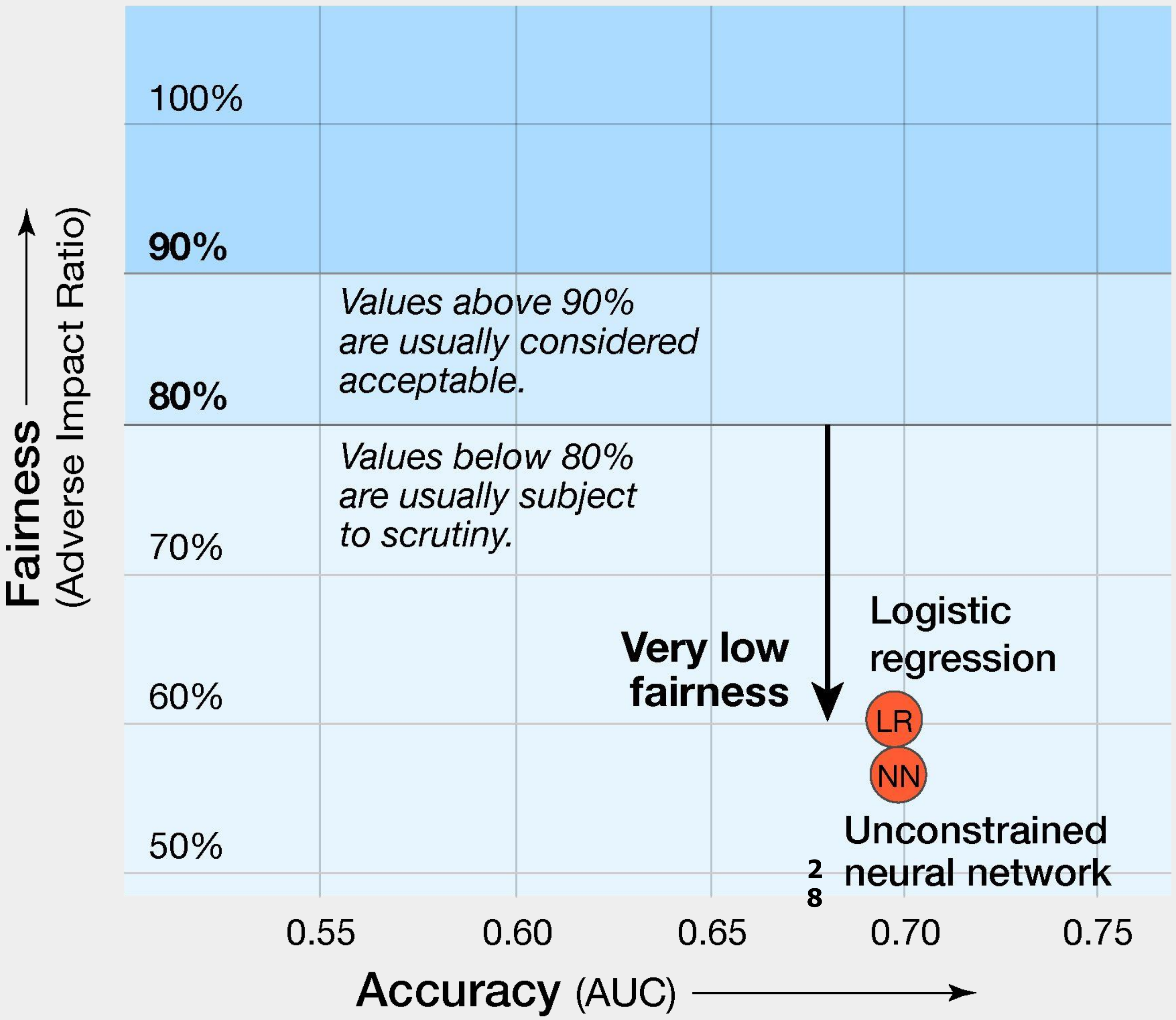
U.S. Mortgage Fairness in 2020: Black



ADVERSE IMPACT RATIO:

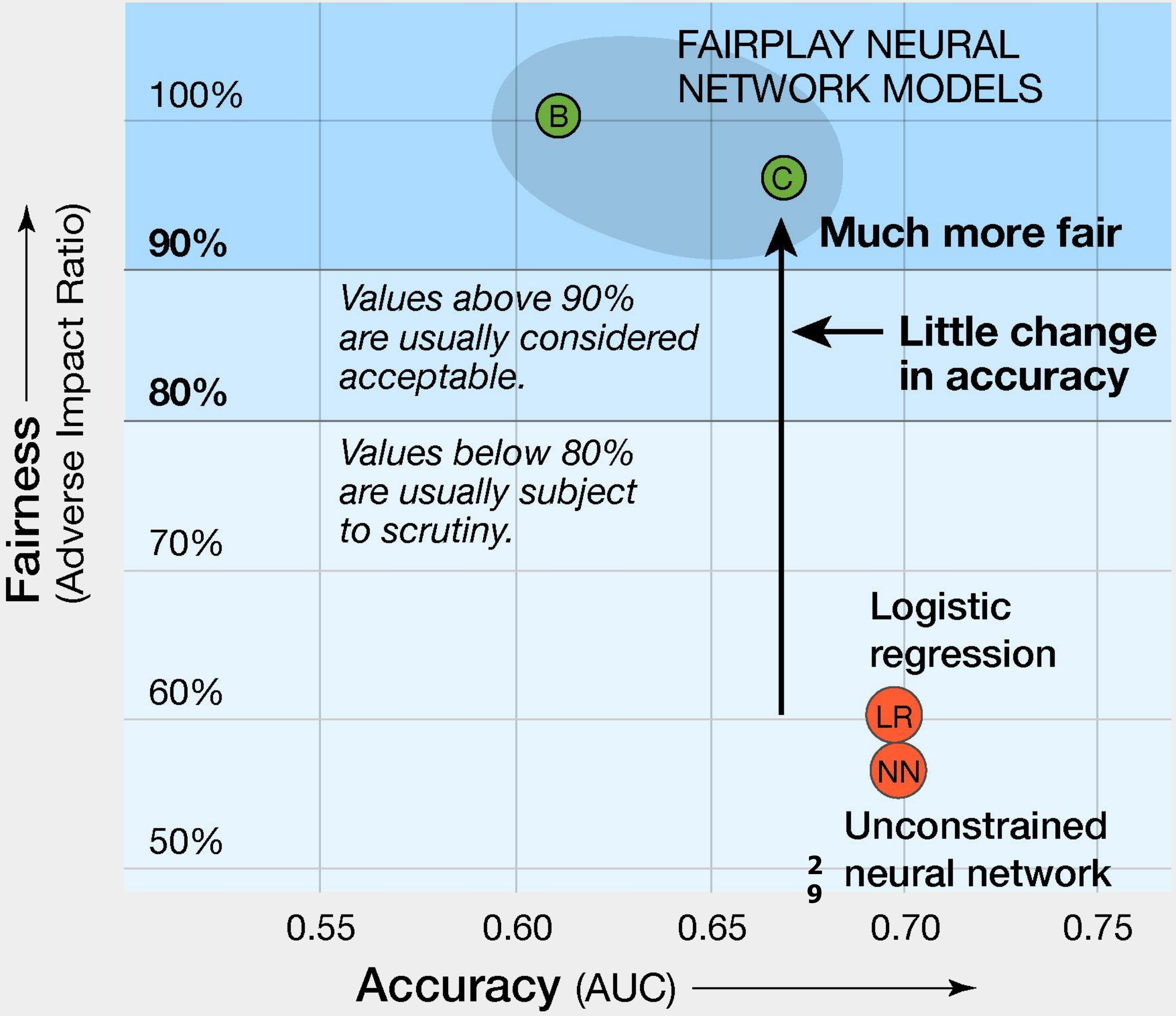
Less than 80% | 80 to 90% | Over 90%

A fairness-optimized neural network can produce very accurate and considerably fairer mortgage outcomes than a logistic regression or unconstrained neural network



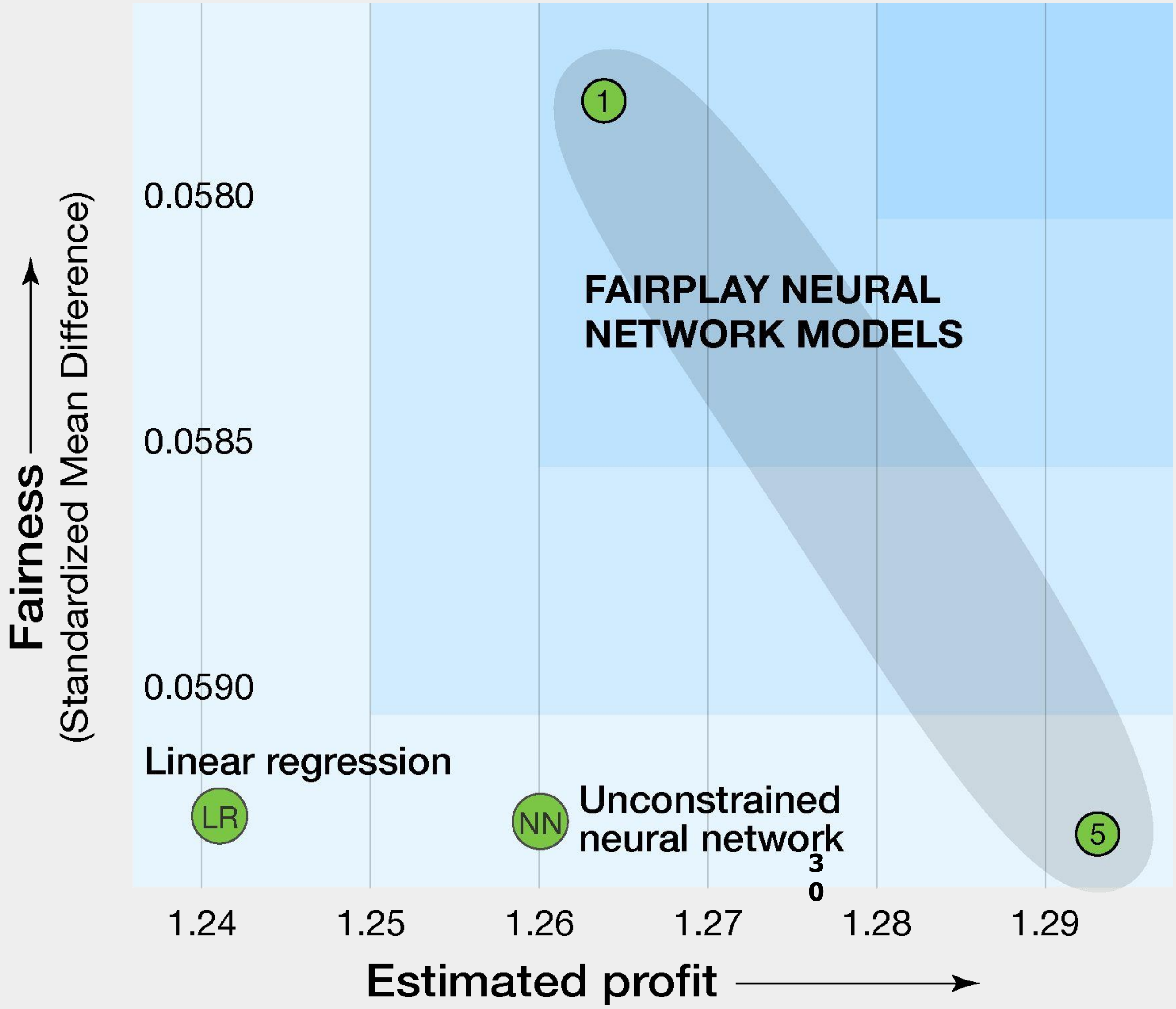
fairplay

A fairness-optimized neural network can produce very accurate and considerably fairer mortgage outcomes than a logistic regression or unconstrained neural network



fairplay

A fairness-optimized neural network can produce fairer mortgage pricing than a linear regression or unconstrained neural network



SMD values less than 0.2 are generally considered acceptable; all of the models shown here are well within that range.

fairplay

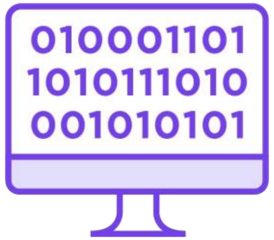
The fairness constrained neural network generates fairer results using the same variables

	Logistic Regression	Unconstrained Neural Network	Fair Neural Network
ADVERSE IMPACT RATIO	60% (90%)	56% (90%)	96% (90%)

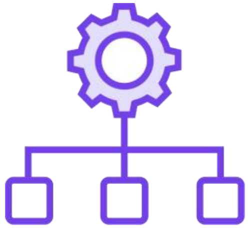
RANK	REASON	REASON	REASON
1	**Borrower's score** Previous December	**Borrower's score** Previous December	**Borrower's score** Previous December
2	**Borrower's score** Previous September	**Borrower's score** Previous September	**Borrower's score** Previous June
3	**Borrower's score** Previous June	**Borrower's score** Previous June	**Co-borrower's score** Previous December
4	**Borrower's score** Previous March	**Borrower's score** Previous March	**Co-borrower's score** At origination
5	**Is this the borrower's first mortgage?**	**Co-borrower's score Previous December**	**Loan-to-value**

fairplay

Catalysts driving AI’s takeover of high stakes decisions



Data



Advanced Algorithms

Cloud Computing

fairplay

FAIRNESS-AS-A-SERVICE

CONTACT

kareem@fairplay.ai

BROOKINGS | QUALITY. INDEPENDENCE. IMPACT.

Algorithmic auditing – why and how

Alex C. Engler
Twitter: @aengler
Email: aengler@brookings.edu

Proliferation of algorithms in important services

- "Over 75% of colleges and universities use analytics for enrollment management" – EduCause Survey, 2015
- "55% of human resources leaders in the United States use predictive algorithms in hiring, at many stages" – Mercer Global Talent Trends
- "98 percent of mortgages originated by Quicken Loans, the country's largest lender, used the company's digital platform, Rocket Mortgage" writes the NYTimes
- "We found that a category of algorithms that influences health care decisions for over a hundred million Americans shows significant racial bias," said Sendhil Mullainathan

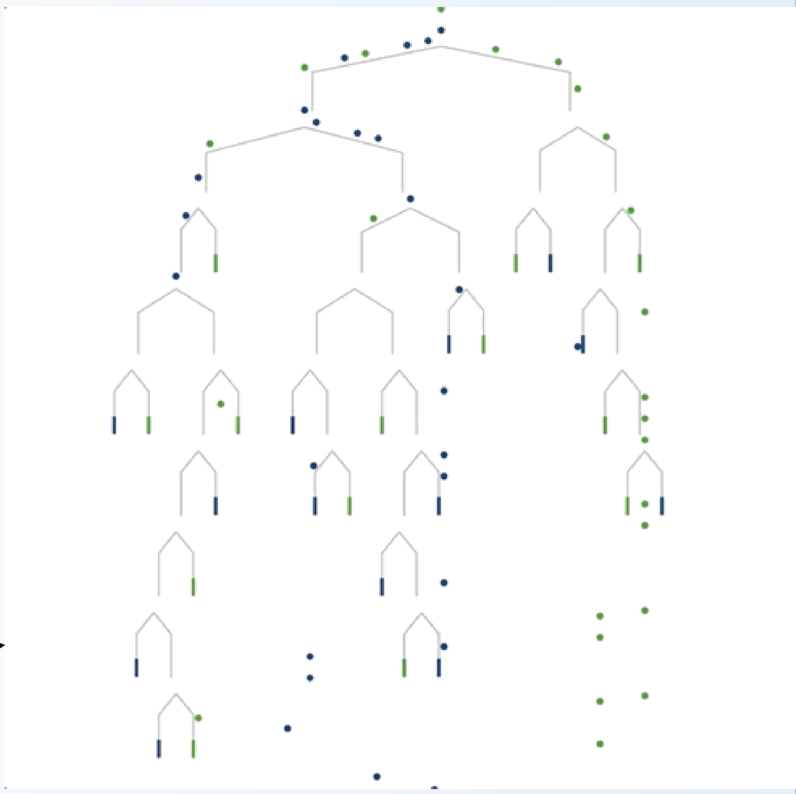
Characteristics of algorithms in important services

- Scaled – they often affect many people

- Automated – they often replace an existing process that was done by a person (but not always)

- "Black box" – They may obscure the decision-making process & reduce efficacy of civil liability

- Might be consistent in treatment (but often not in impact)



- "Learning" means algorithmic behavior may change over time

Concerns with algorithms in important services

- Power centralizing towards institutions (e.g. tuition, ride-sharing)
- Incentivizes privacy violations (e.g. Clearview AI, EverAlbum)
- Opaque process may offer no feedback (e.g. mortgage denial)
- Fraud (different than efficacy)
- Discrimination
- Systemic Risks (e.g. flash crashes)

Proposed solution: Algorithmic audits!

- An algorithmic audit is a review of the environment, inputs, functions, outputs, and results of an algorithmic system
- This is what the **NFHA's Purpose, Process, Monitoring Framework** describes, but there are many other examples from health, hiring, education, social media, etc.
- Can be paired with a fine for illegal activity, possibly algorithmic deletion in extreme cases

Do algorithmic audits help with these problems?

- Power centralizing towards institutions (e.g. tuition, ride-sharing)
  - Mostly no, may create public awareness
- Incentivizes privacy violations (e.g. Clearview AI, EverAlbum)
  - Yes, if the privacy violations are illegal
- Opaque process may offer no feedback (e.g. mortgage denial)
  - Mostly no – audits offer global, not individual, explanations.
- Fraud
  - Yes!
- Discrimination
  - Yes!
- Systemic Risks
  - Maybe

What should an audit be?

Process considerations:
- Independent
- Adversarial
- Representative (if there are many models)
- Access to data/code/models

Dependencies & Docs
- Data collection process
- Training data representativeness
- Dependency analysis
- Accurate documentation

Model & outcome review
- Look at outcome/training/hypothetical data
- Consider problem definition
- Consider algorithmic drift, frequency of audit

BROOKINGS

Who needs to do audits?

- Criteria
  - Audits must be **independent** and **adversarial**.
  - The require extensive access to the data and algorithms
  - This will often mean government action

- Audits do not automatically mean accountability. In accounting, it takes:
  - Market incentives (these don't apply to algorithmic discrimination)
  - Financial auditors (the people!) are legally liable if they commit fraud
  - Plus direct government oversight (SEC)

FAST COMPANY LOGIN

01-26-21 | POV

Independent auditors are struggling to hold AI companies accountable

Controversial AI company HireVue implied that an external audit showed its algorithms had no bias. But a look at the audit itself tells a different story.

So, Government Audits

- Why:
  - They can be genuinely independent.
  - Usefully, fraud & discrimination are often already against the law.
- How:
  - Develop specific best practices/guidance by model/application type
  - Some legal changes may be necessary
    - Expand administrative subpoenas to include data sets
    - E.g. remove ‘predictive validity’ from EEOC Uniform Guidelines
  - Hire regulatory data scientists (soon to be a booming field)
  - Start information gathering, enable consumer complaints, and encourage industry whistleblowing

Government Agency Action

- HHS was only agency to detail its algorithmic regulatory authority
- EEOC launched Initiative on AI Fairness
- Federal Trade Commission
  - Threatening blog post on fraud and discrimination
  - Fines and algorithmic deletions for privacy violations (EverAlbum)
  - Started rulemaking consideration

- Five financial regulators started information gathering on algorithms, citing Fair Housing + Equal Credit Opportunity Acy
  - (DOT OCC, Fed Reserve, FDIC, CFPB, NCUA)
- Housing and Urban Development
  - Reversing the Trump admin change to the 'discriminatory effects' rule
  - Has not responded to the two E.O.s on algorithms (E.O. 13960 and E.O. 13859)

Aside: Agencies have algorithms too

- Federal agencies can show the way by documenting and auditing their own algorithms
- Algorithmic documentation required under E.O. 13960
- Eventually, standards for algorithmic auditing will be the norm

Complications of algorithms in important services

- Many algorithmic stages of "funnel processes;" consider hiring:
  - Job recommendation -> Resume analysis -> Questionaries/Games -> Automated interviews
- Different types of algorithms combined, consider college tuition:
  - First: predict what one student would pay to attend
  - Second: assign scholarships across many prospective students
- Separate algorithmic systems affecting one another (e.g. online markets)
- Human + algorithmic decisions can have unexpected outcomes
  - Felony sentencing by judge with algorithmic risk assessment

BROOKINGS

Work supported by:

STIFTUNG MERCATOR

FULBRIGHT Schuman Program

BELGIUM / LUXEMBOURG / SCHUMAN

Get in touch:

Twitter: @alexcengler

aengler@brookings.edu

NFHA NATIONAL FAIR HOUSING ALLIANCE

PPM AUDITING FRAMEWORK

A COMPREHENSIVE NEW FRAMEWORK FOR AUDITING ALGORITHMIC SYSTEMS

AI AUDITING IN HOUSING

LISA RICE, PRESIDENT AND CEO OF NFHA
MARCH 22, 2022

ALGORITHMIC BIAS

Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women

Isobel Asher Hamilton Oct 10, 2018, 5:47 AM

Employment

Health

Racial bias in a medical algorithm favors white patients over sicker black patients

+ Add to list

Scientists discovered racial bias in a widely used medical algorithm that predicts which patients will have complex health needs. (iStock)

Healthcare

Credit scores in America perpetuate racial injustice. Here's how

Sarah Ludwig

Credit reports and scores reflect existing racial inequities in our credit system and economy

Tue 13 Oct 2015 10.14 EDT

503 200

▲ A bad credit score can lead to discrimination in housing and employment. Photograph: Alamy Stock Photo

Credit Scoring

NFHA

ALGORITHMIC BIAS

APPLE \ POLICY \ TECH

Apple's credit card is being investigated for discriminating against women

Customers say the card offers less credit to women than men

By James Vincent | Nov 11, 2019, 5:57am EST

SHARE

Image: Apple

Credit

The Markup

Big Tech Is Watching You. We're Watching Big Tech.

Locked Out

Can Algorithms Violate Fair Housing Laws?

Landlords increasingly use screening services to weed out renters. Advocates say both landlords and the algorithms should be accountable when things go wrong

By Lauren Kirchner

September 24, 2020 08:00 ET

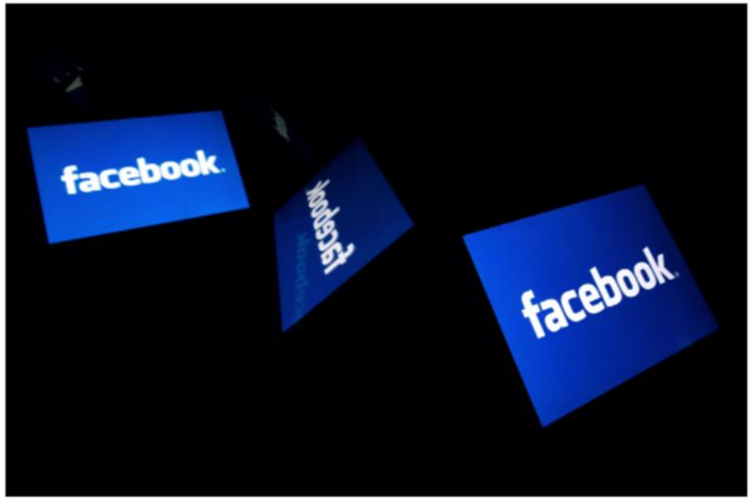
Buśa Photography/Getty Images

Tenant Screening Services

After Lawsuits, Facebook Announces Changes To Alleged Discriminatory Ad Targeting

npr By Brakkton Booker

Published March 19, 2019 at 1:32 PM CDT



Marketing Platforms

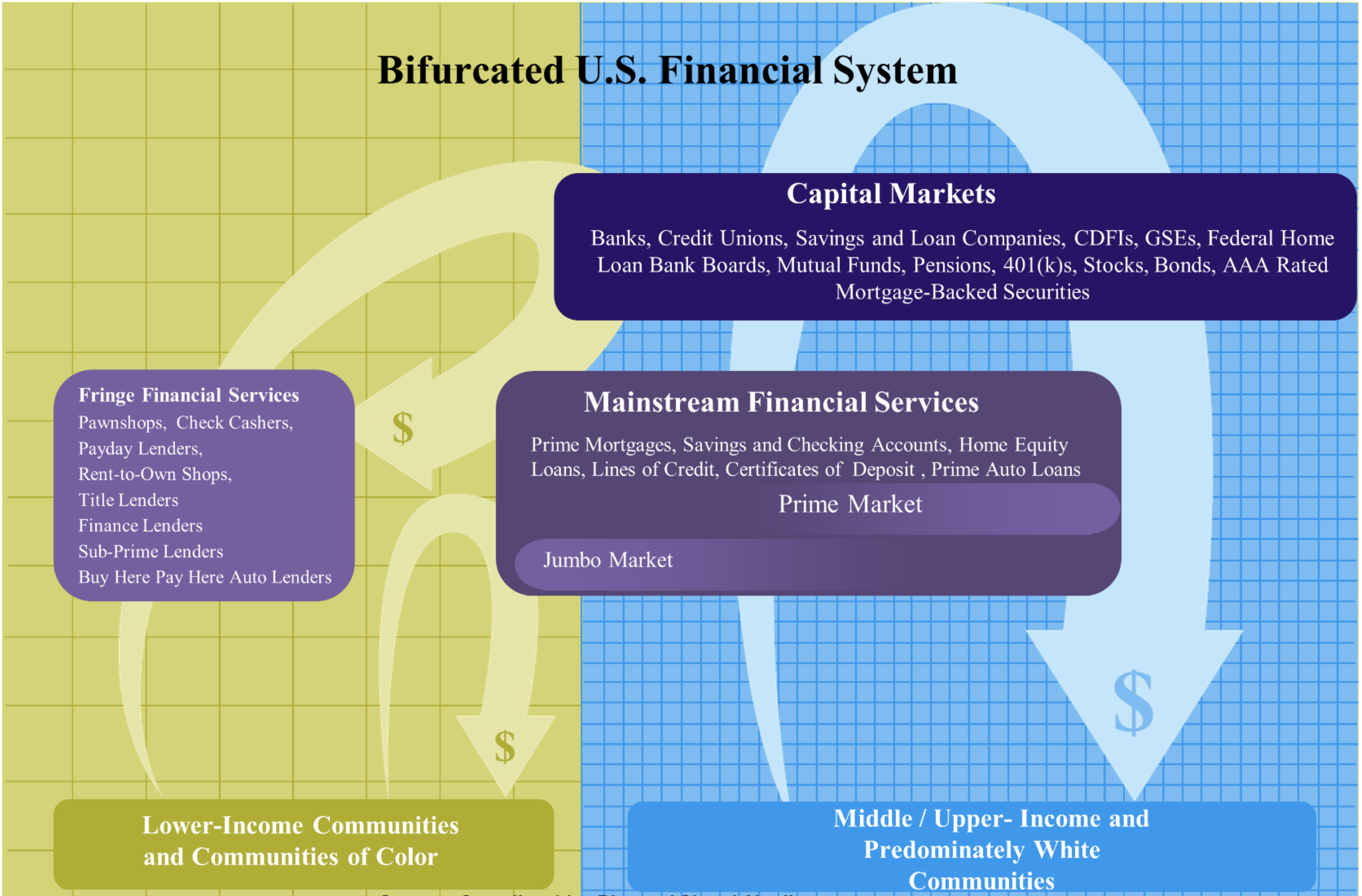
@ItsLisaRice

NFHA

TECH IS THE NEW CIVIL RIGHTS FRONTIER

- Credit Scoring Systems
- Automated Underwriting Systems
  - Automized Verifications Processing
- Risk-based Pricing Systems
- Advertising Systems/Platforms
- Revenue Management Software
- E-loans Processing
  - Video-Conferencing
  - Digital Fingerprinting
- Facial Recognition Systems
- Automated Property Valuation Systems
- Tenant Selection Systems

Bifurcated U.S. Financial System



Source: Carr, Jim, Lisa Rice and Shanti Abedin

How Technologies Manifest Bias

- Unrepresentative, Insufficient or Flawed Data
- Design Flaws
- Biased Feedback Loops
- Insufficient or No Testing for Bias
- Untrained Designers
- Lack of Diversity

NFHA

EFFECTIVE AUDITING OF SYSTEMS CAN REDUCE DISPARITIES

- Compelling models to be fairer can lead to the development of more accurate and safer models that minimize or eliminate disparities in housing and lending.
- The adoption of fairer models would expand access to credit for underserved groups and result in less disruption for consumers, families, communities, and industry.
- Fairer models can help reduce the racial wealth and homeownership gaps.
- The development of state-of-the-art policies and protocols would mark the U.S. as a leader in advancing responsible technologies that serve markets, consumers, and the greater society.

NATIONAL FAIR HOUSING ALLIANCE

Overview and Purpose Element of the PPM Framework

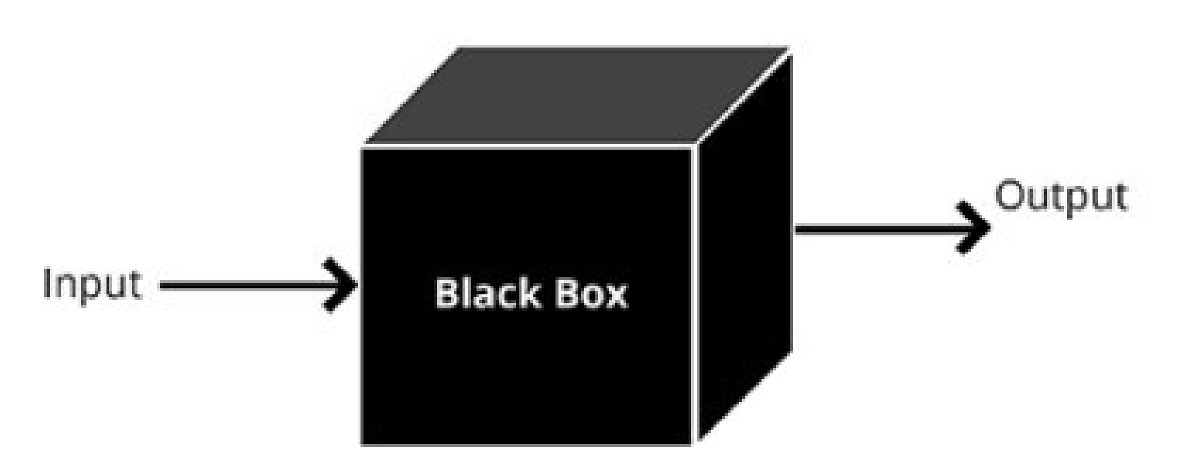
Snigdha Sharma, FairOps Team Lead & Tech Equity Analyst, NFHA

Outline

- What job does PPM – Purpose, Process, Monitoring - do?
- The Purpose element of the PPM Framework?

What Job does PPM Try to Do?

- Examines every stage of algorithm solution development
- A trust tool regulators can use to gather information that could be used to decide how fair, transparent and explainable an algorithmic system is.
- Helps policymakers see what aspects of algorithmic systems need mitigation and can lead to algorithmic risks to consumers
- Helps policymakers hold those developing algorithms accountable for any decisions they made while building the system



Quora.com, 2022

The Purpose in PPM: an application of algorithm in renting

- Business Problem: Would this person meet their rent obligation or not?

- Input: application information, employment, bank statements etc. of previous or current tenants

- Black box: a classification algorithm like Random Forest

- Output: tenant score that rates an applicant and is used to decide if an individual is given a lease or not

Laila Milevski/ProPublica

The Purpose in PPM: what Questions does it answer?

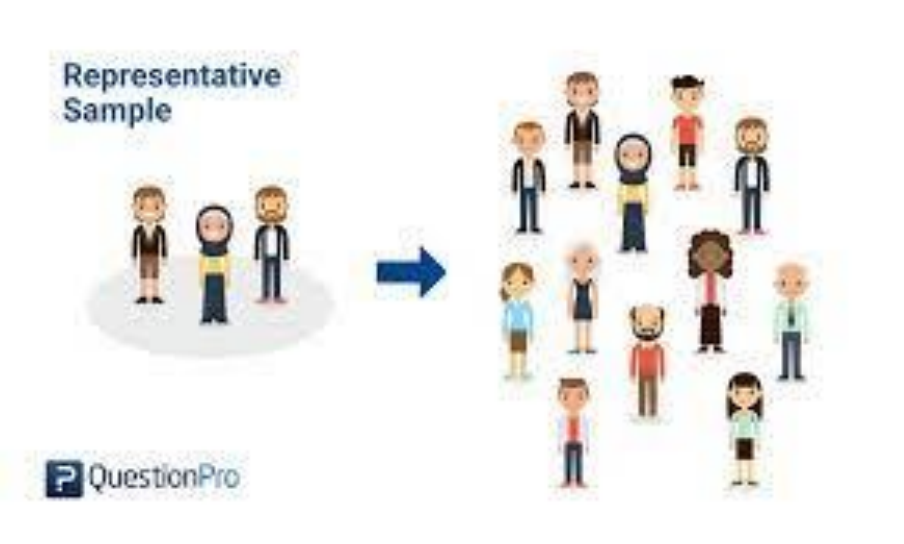
Business Understanding:

- The Purpose section asks why the model is being built and what are the business objectives.
- Auditors are guided to seek information to make informed decisions about risks the business problem may pose to consumers, institutions, and society at large.
- Does the system pose too great a harm that it should not be built or put into production?
- What exactly is the business problem being solved?

The Purpose in PPM: what Questions does it answer?

Data Understanding :

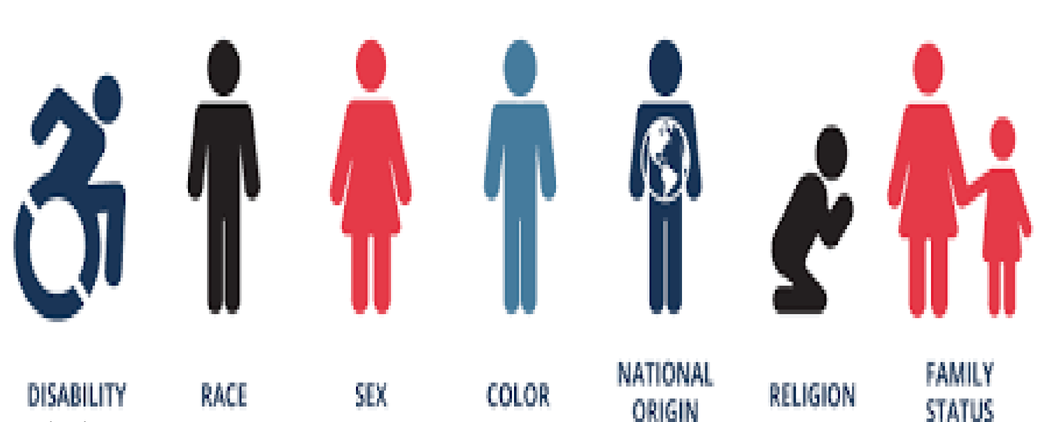
- Is there data to represent the business problem?
- Were any techniques used to mitigate risks associated with data paucity or data quality.



The Purpose in PPM: what Questions does it answer?

Data Understanding :

- Is there proxy testing for protected classes?
- What metrics are being used to test model accuracy and fairness?
- Are model features or variables representative across protected class data?



EqualRightsCenter.org

NFHA PPM Auditing Framework: Process

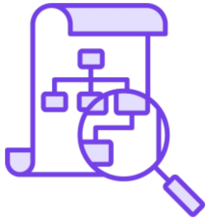
John Merrill, Chief Technology Officer of FairPlay

March 22, 2022

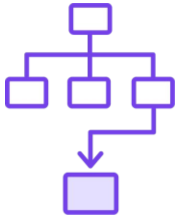
The Process Stage of the PPM Framework Includes Five Elements

STAFF PROFILE

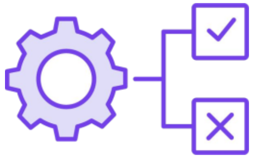
DATA ASSESSMENT



MODEL ASSESSMENT



OUTCOME ASSESSMENT



MODEL USE AND LIMITATIONS

Staff Profile

Diverse Team

Do the members represent a broad class of possible stakeholders?

Functional Team

Can the members work together?

Communication

Can the team explain its results to others outside of the team?

Data Assessment

Collection

How was data gathered?

Appropriate

Is the data well-suited to the model's objective?

Representative

Does the data reflect the characteristics and qualities of the subject being modeled?

Accurate

Are the values in the data correct and consistent?

Well-behaved

Are the data targets and values well-behaved? Are any computed features reasonable?

Model Assessment

Model Type

Is the model type appropriate to the data and problem?

Parameters

Which variables are included in the model? How is each one treated?

Hyper-Parameters

How did you set the parameters? For example, did you control how fast the algorithm learned? When did you cut off its learning?

Fairness Constraints & Less Discriminatory Alternatives

What steps did you take to ensure model was fair? If a model exhibits disparities for a group, did you search for fairer alternatives? If so, how?

Outcomes Assessment: Did the model meet its objectives?

Accurate

Are the model's predictions correct?

Fair

Do the model outcomes favor one group over another?

Explainable

Are the model's decisions known and sensible?

Stable & Robust

Do the model's predictions vary widely when things change?

Model Use and Limitations

Limitations and Assumptions

What does the model assume about the data?

Other Applications

To what other problems can the model outputs be or nor be applied?

Disaster Response

Are there procedures and processes in place to handle disasters?

fairplay

FAIRNESS-AS-A-SERVICE

CONTACT

john@fairplay.ai

NATIONAL FAIR HOUSING ALLIANCE

The Monitoring Element of the PPM Framework

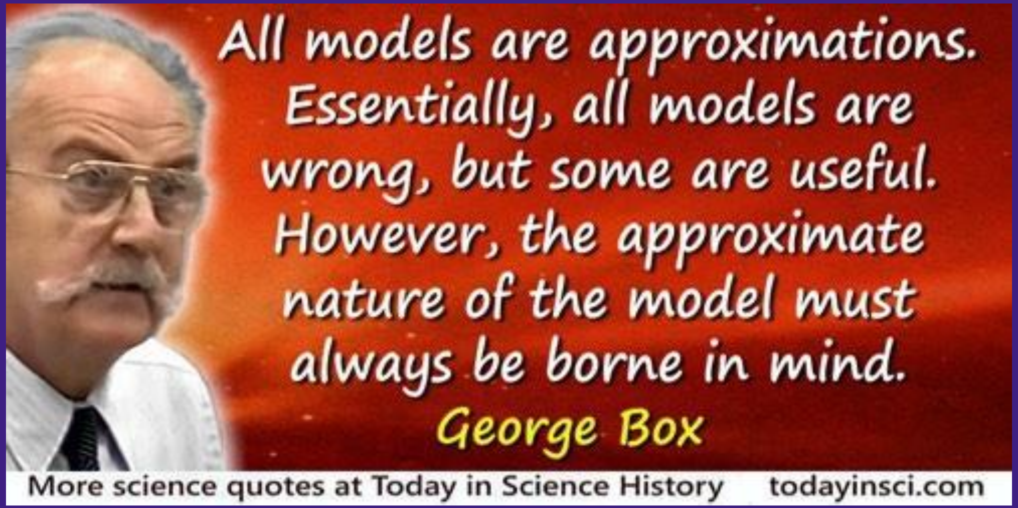
Michael Akinwumi, Chief Tech Equity Officer, NFHA

Outline

- Model Use and Model Limitations
- An algorithmic solution needs monitoring
- Q & A (All)

Model Use and Limitations

- Developer needs to be transparent about blind spots in their model. Every model has a limitation.
- What business problem can the model be applied to?



An algorithmic solution needs monitoring

- “All models are wrong, but some are useful.”
- Document what metrics developers plan to use for monitoring
  - The usefulness of the model and
  - Potential harms it may cause to consumers while in production
- Integrity of the model should be protected, and consumer privacy should be protected in the event of a hack

NFHA NATIONAL FAIR HOUSING ALLIANCE

Question & Answer