Exploring Human-AI Collaboration for Fair Algorithmic Hiring

Hyun Joo Shin (hshin36@jh.edu) Anqi Liu (aliu@cs.jhu.edu)

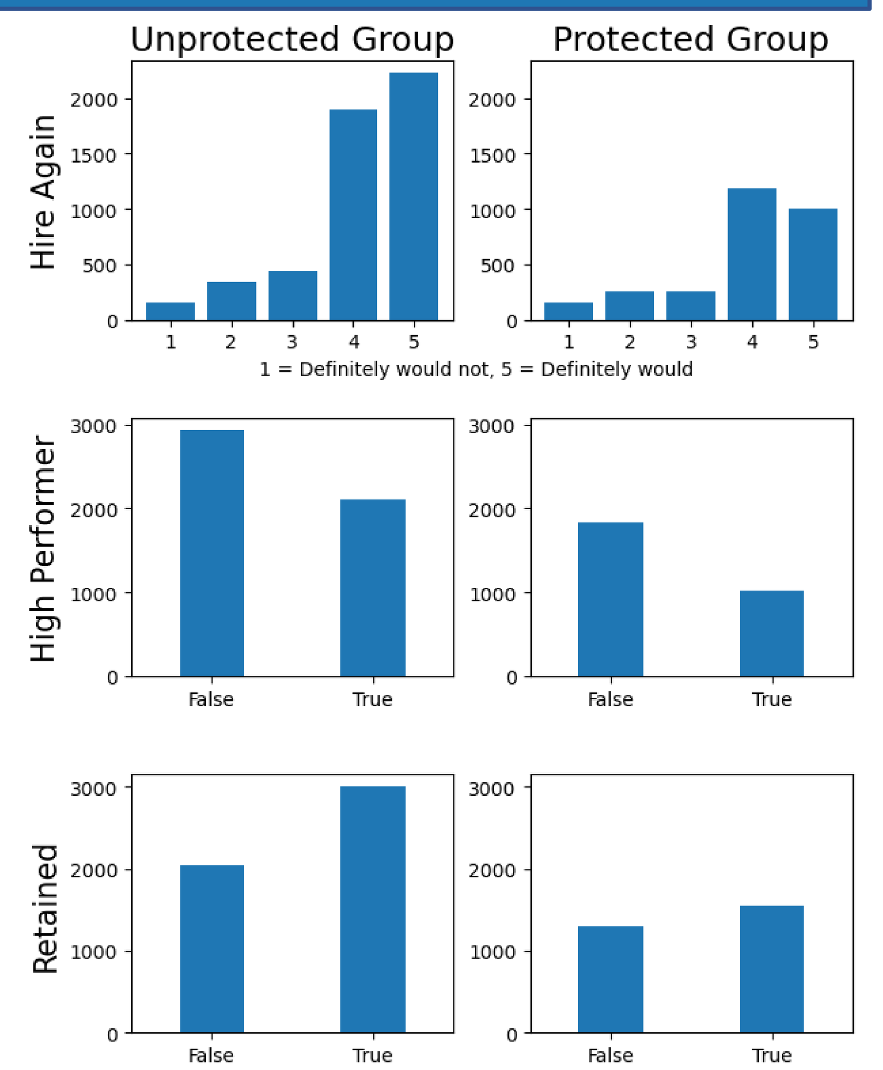
JOHNS HOPKINS UNIVERSITY

ML algorithms in the hiring process

- Increasing use of ML algorithms in hiring for greater efficiency, less human bias, and better quality of new hires
- Legal concerns about ML-induced discrimination against minority in algorithmic hiring processes, against Title VII, Affirmative Action and the Equal Employment Opportunity Commission (EEOC)

Data

- Walmart employee data [1]
  - 7890 employees (2846 in unprotected group*)
- **Input**: Three groups of features
  - Scenario Interpretation
  - Biodata / Work History Items
  - Personality / Work Style Items
- **Output**
  - Hire Again (Would you hire this employee again?)
  - High Performer (Is/Was employee a "high" performer?)
  - Retained (Was employee retained for a period of n days?)

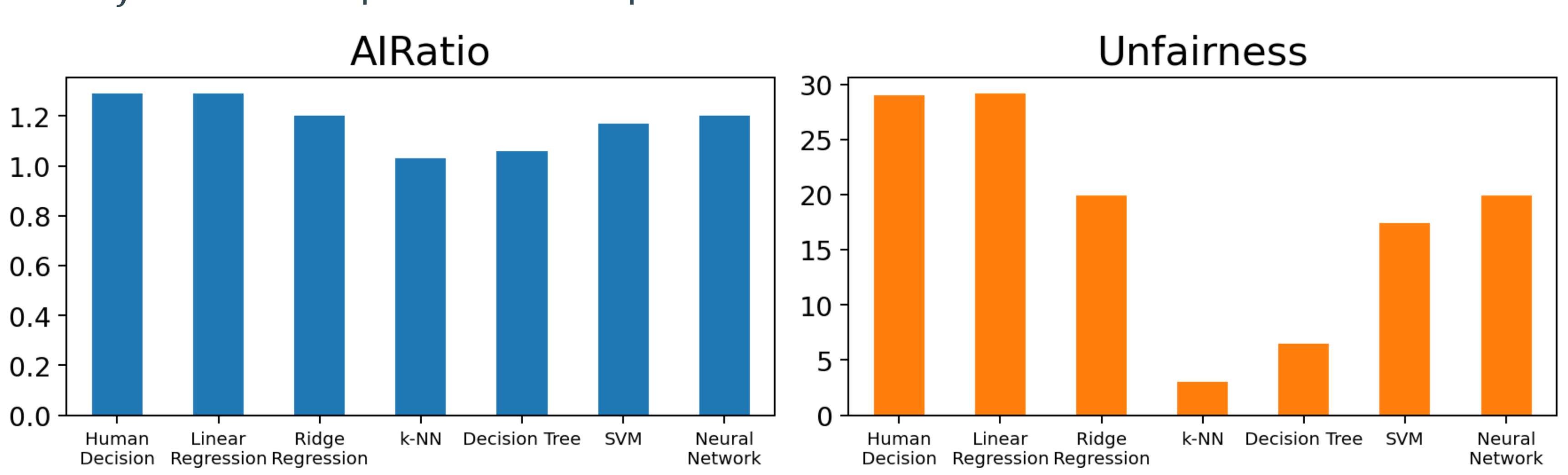


Fairness of ML decisions

- Adverse impact (AI) ratio with a selection ratio, 0.5

$$A = \frac{Protected\ Hired}{Protected\ Applicants} \quad B = \frac{Unprotected\ Hired}{Unprotected\ Applicants} \quad AI\ Ratio = \frac{A}{B}$$

- $Unfiarness = |1 - AI\ Ratio| * 100$



- Human decision AI ratio = 1.29 → Reverse discrimination?
- ML decision AI ratio = min (103), max (1.29), average (1.16)

- **ML algorithms makes more fair hiring decisions across two groups**

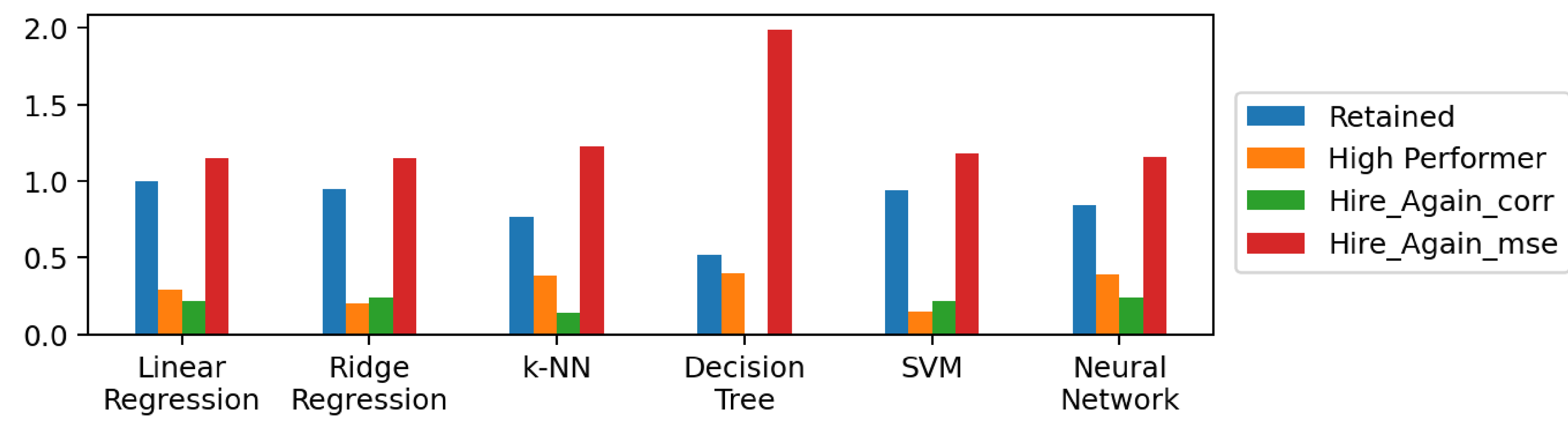
* An artificially contrived variable intended to be used surrogate for protected class variables (e.g., race, gender, sex, age)

Performance of ML algorithms

- ML algorithms performance
  - Retained and High Performer:

$$\frac{\#\ of\ same\ prediction\ as\ human\ decision\ makers}{total\ \#\ of\ predictions}$$

  - Hire Again: MSE & Pearson R on human decision and machine decision



- Good performance on **Retention** (average = **0.84**)
- Poor performance on **High Performer** (average = **0.30**)
- Poor performance on **Hire Again** (avg. R = **0.18**, avg. MSE = **1.31**)
- **Why are the performances different?**
  - Retention is a factual information
  - Performance evaluation and hiring decision involve 3rd person's evaluation and decision

- **ML algorithms fails to mimic human decision makers**

Human decision vs. Machine decision

- To understand why algorithms fail, two-fold Blinder-Oaxaca decomposition was used comparing the characteristics of Human and ML decisions across protected and unprotected groups on Hire Again

	Human decision		Machine decision	
	average	std. dev.	average	std. dev.
unprotected group	4.12	1.08	4.13	0.23
protected group	3.85	1.15	4.06	0.24
difference	0.27		0.06	
	coefficient	std. err.	coefficient	std. err.
explained	-0.09	0.15	0.05	0.03
unexplained	0.36	0.15	0.01	0.02

Blinder-Oaxaca Decomposition Result

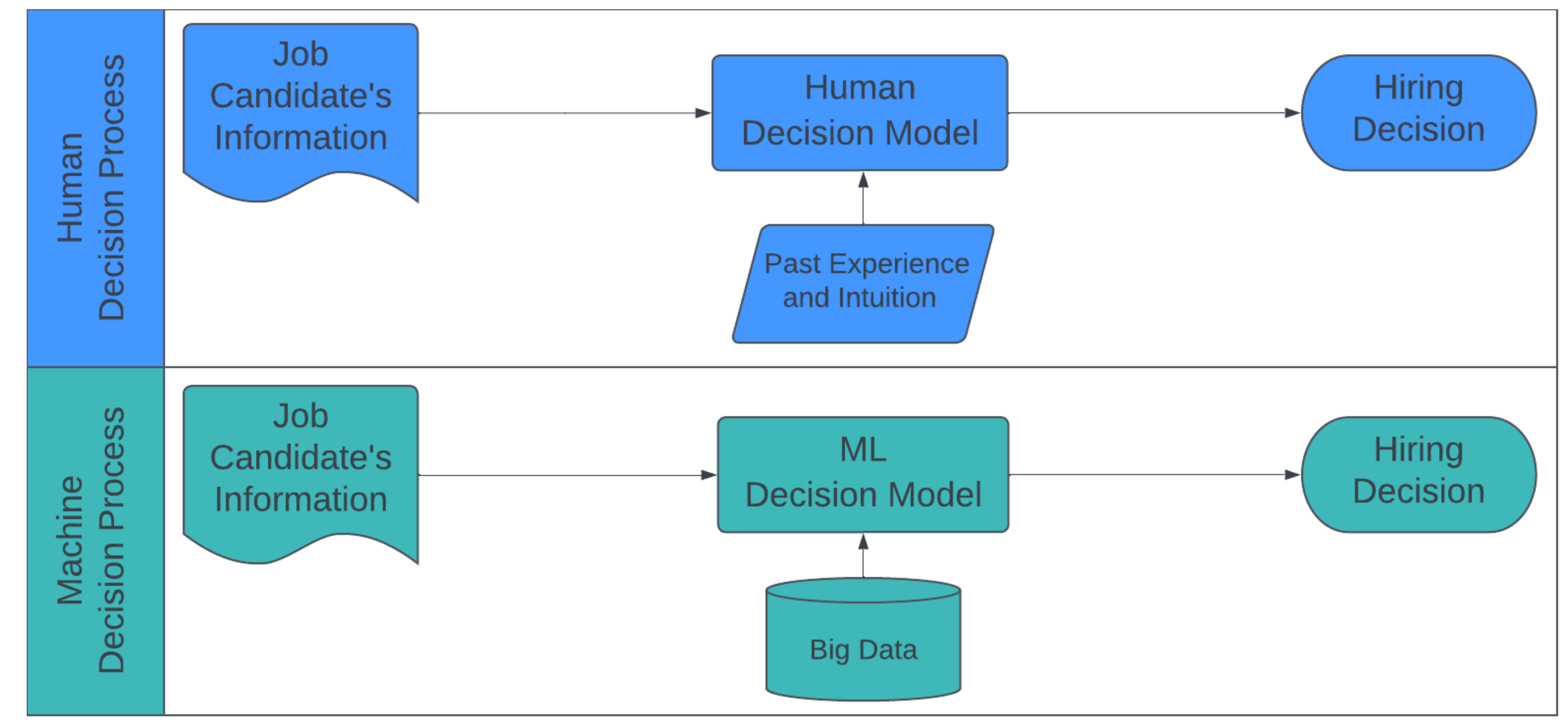
- **Unexplained** components are often…
  - attributed to discrimination
  - resulted from the influence of unobserved features
- Human: 133% (=0.36/0.27) ML: 17% (=0.01/0.06)

→ Human decisions are greatly influenced by many factors, such as …
  - labor market discrimination
  - unobserved features, such as decision makers' past experience and intuition

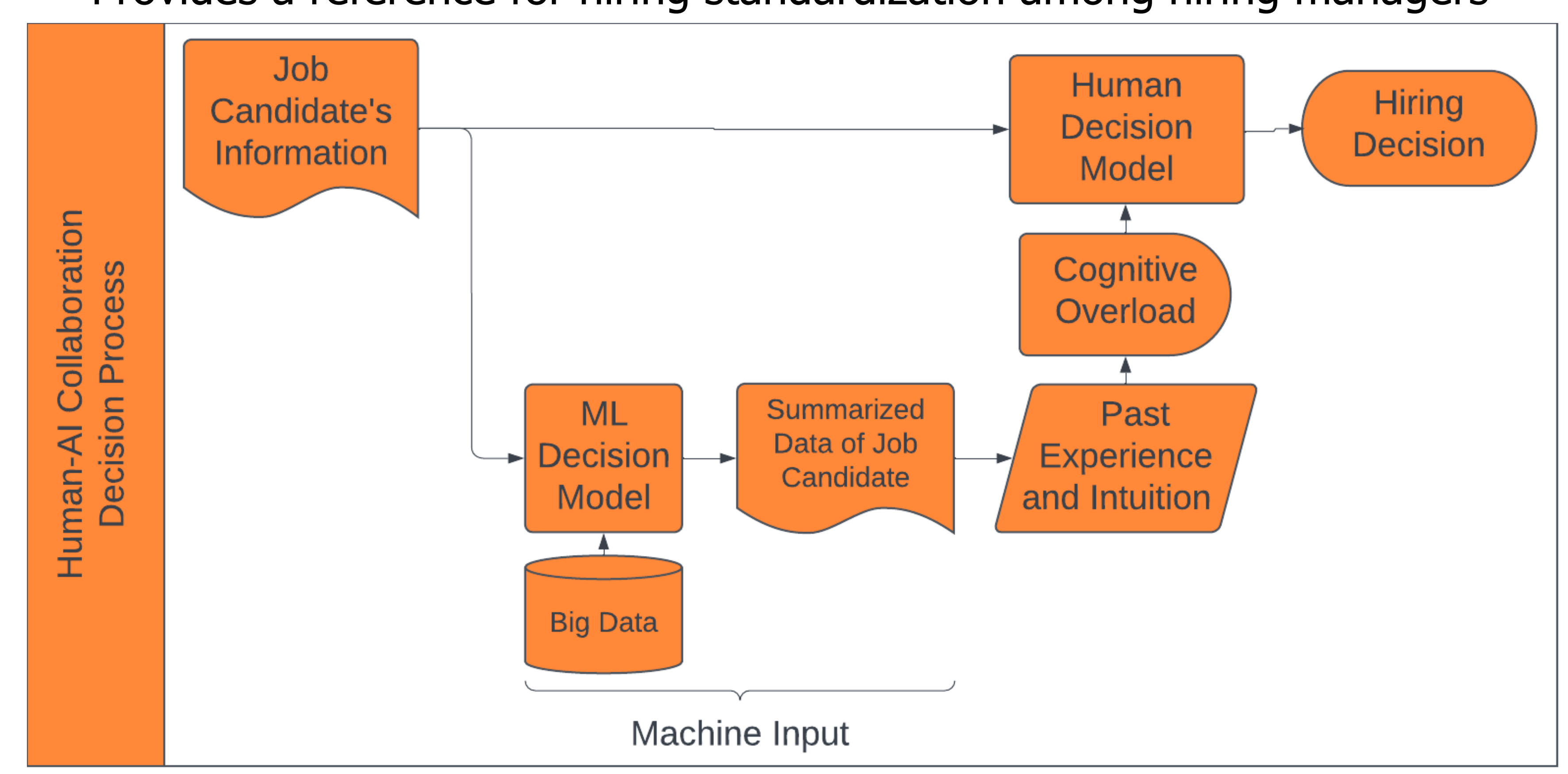
- **ML algorithms fails to mimic human decisions because humans use external data not available to algorithms**

Human-AI collaboration during hiring

Each decision model has its own benefit:
- Human hiring managers use their past experience and intuition that are not available to algorithms
- ML algorithms purely makes data-driven decisions using past employee data



Combining two models has potential benefits of …
- Enforce **cognitive overload** providing a chance to confirm human decisions
- **Subjective evaluation** of a candidate given past hires
- **Mitigate** humans' **implicit bias** by slowing down the process
- Provides a reference for hiring standardization among hiring managers



- **Human-AI collaboration has a potential to improve both hiring accuracy and fairness during hiring processes**

- Human subject study should follow to measure the impact of the cognitive overload introduced by the summarized data of job candidate from ML algorithms

References

[1] Koenig, N., and Thompson, I. 2021. The 2020-2021 SIOP Machine Learning Competition. In Presented at the 36th annual Society for Industrial and Organizational Psychology. SIOP, New Orleans, LA. https://github.com/izk8/2021_SIOP_Machine_Learning_Winners