[COGS 9] Discussion
Reading 5, ML/DL Demos

Assignment 3 due on 31st July (Mon)

Quiz 5 on 1st Aug (Tue)

Accountability in Algorithmic Decision Making

Introduction

- Algorithms are increasingly used in various domains that have clean and well-structured data, such as finance, sports, weather, and education.
- The scale, speed, and cost-saving advantages of algorithms are undeniable.
- However, the trade-off appears to be nuance and accuracy, with thousands of automatically written articles requiring corrections.
- The paper argues for the need to consider the impact of automated decision-making in various sectors.

Algorithmic Decision Making

Algorithms make various types of decisions, including prioritizing, classifying, associating, and filtering.

Prioritizing information can lead to discrimination as certain things are emphasized at the expense of others.

Classification decisions can lead to bias, uncertainty, or outright mistakes.

Association decisions create relationships between entities, leading to potential misinterpretations.

Filtering decisions involve including or excluding information according to various rules or criteria

The Need for Accountability and Transparency

Important **or expensive errors, discrimination, unfair denials of public services, or censorship** can occur and algorithms need to be held accountable

Suggests **that computer science and engineering professionals have a role to play in ensuring accountability and transparency in algorithmic decision-making.**

The human influences in algorithms are many, including criteria choices, optimization functions, training data, and the semantics of categories.

Need **to incorporate ethical ideals throughout the engineering process**

An Algorithmic Transparency Standard

- Proposes five broad categories of information that might be disclosed for transparency: human involvement, data, the model, inferencing, and algorithmic presence.
- Suggests that transparency can be facilitated by disclosing certain key pieces of information, including aggregate results and benchmarks.
- Argues for regulations that compel information disclosure or at least routine audits around key algorithmically influenced decisions.
- Suggests the need for an adversarial approach to investigate black-box algorithms.
- However, every use case is different, and they cannot be held to the same level of transparency and disclosure

Machine Bias

Machine Bias in Criminal Sentencing

- The article discusses the use of software for predicting future criminals, highlighting its bias against black individuals.
- Presents the case of two individuals, Brisha Borden (black) and Vernon Prater (white), who committed similar petty thefts but were assessed differently by the software.
- Borden was rated a high risk while Prater was rated a low risk, despite Prater having a more severe criminal history.
- The software's prediction turned out to be incorrect as Borden did not reoffend, while Prater did.

Risk Assessments in Criminal Justice

Risk assessments are increasingly used in courtrooms to inform decisions about defendants' freedom.

They are used at every stage of the criminal justice system, from assigning bond amounts to sentencing.

The Justice Department encourages the use of such assessments at every stage of the criminal justice process.

However, concerns have been raised about the potential bias in these risk scores.

Racial Disparities in Risk Assessments

The algorithm used for risk assessments was found to have significant racial disparities.

Black defendants were more likely to be falsely flagged as future criminals at almost twice the rate as white defendants.

White defendants were mislabeled as low risk more often than black defendants.
James Rivelli had committed numerous crimes but still had a low score of 3/10

These disparities could not be explained by defendants' prior crimes or the type of crimes they were arrested for.

The Algorithm Behind Risk Assessments

The algorithm (called COMPAS) **used to create the risk scores is a product of a for-profit company, Northpointe.**

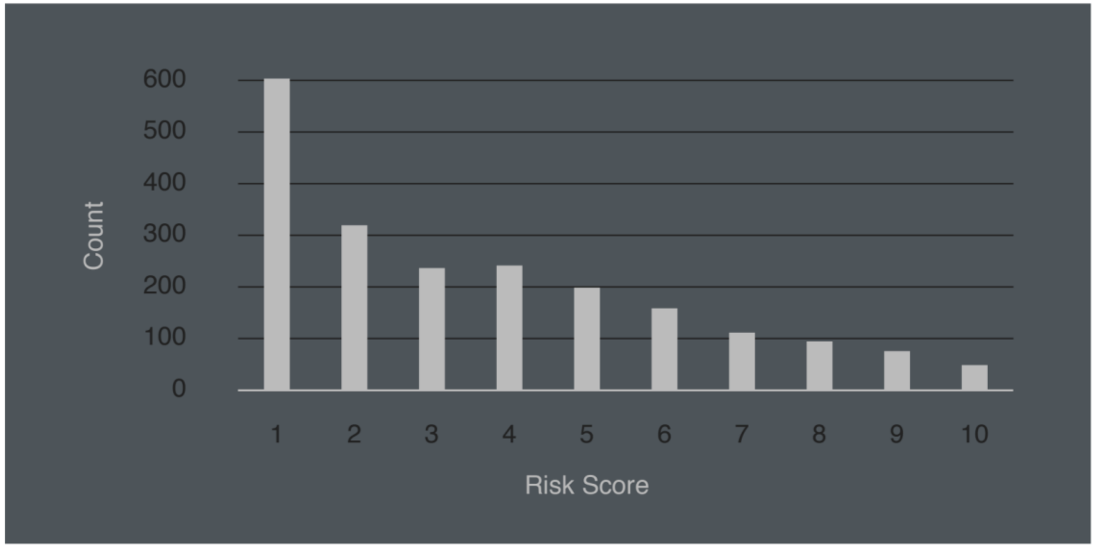
The company does not publicly disclose the calculations used to arrive at defendants' risk scores.

They put out a validation study where COMPAS had a 68% accuracy to predict reoffense and it was "less predictive" for black men than white men

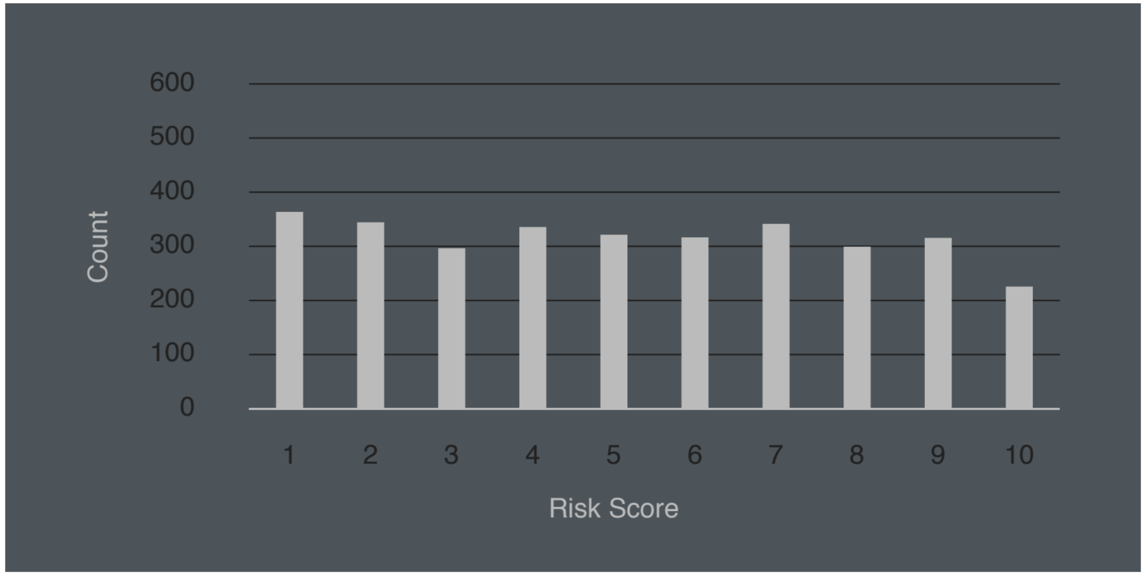
There were no studies done by neutral parties on the risk scores put out by COMPAS

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

White Defendants' Risk Scores



Black Defendants' Risk Scores



These charts show that scores for white defendants were skewed toward lower-risk categories. Scores for black defendants were not. (Source: ProPublica analysis of data from Broward County, Fla.)

The Impact and Future of Risk Assessments

Proponents argue that risk scores can be used to reduce the rate of incarceration and make the criminal justice system fairer.

However, the scores can be inaccurate or biased due to the underlying data

There is a need for more transparency and scrutiny in the use of these algorithms

The debate around the use of risk assessments in criminal sentencing continues, with ongoing legal challenges and discussions about their appropriateness and accuracy.

Machine Learning/Deep Learning Demo