

# CSE 575 : Final Report

Group Members: Chieh Yi, Devaraj Madhugondur, Gale Moore,  
Rahulrajan Karthikeyan, Yadvendra Naveen, Yuan-Yen Chang



An isometric illustration of a cityscape with various skyscrapers in shades of blue and teal. Some buildings have icons on top: a Wi-Fi symbol, a padlock, a dollar sign, and a helicopter landing pad. A small helicopter is flying in the center. The background is a solid light blue.

▶ 0 1

# OVERVIEW



# Introduction

## Factors

Race, gender, age, citizenship, and name

## BIAS



## To tackle this issue

The use of COMPAS

## Objective

Develop an advanced and interpretable ML model





# What is COMPAS?

## Correctional Offender Management Profiling for Alternative Sanctions



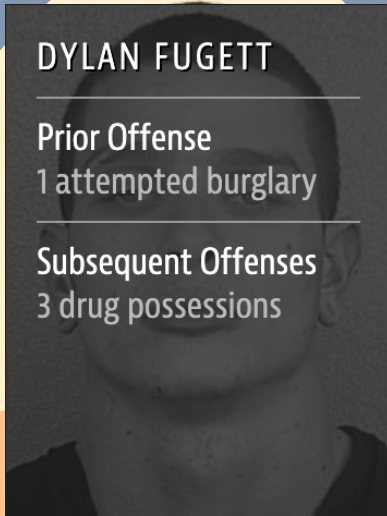


# MAIN PUBLICATION





# The Problem with COMPAS



**DYLAN FUGETT**

---

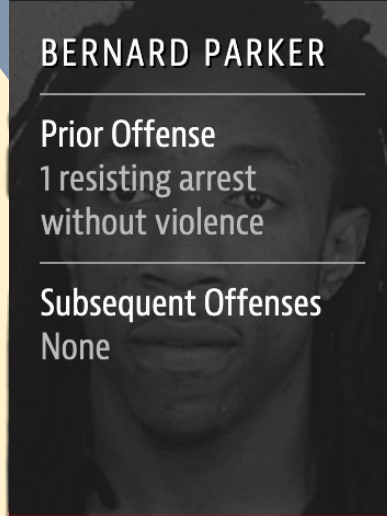
Prior Offense  
1 attempted burglary

---

Subsequent Offenses  
3 drug possessions

---

LOW RISK **3**



**BERNARD PARKER**

---

Prior Offense  
1 resisting arrest without violence

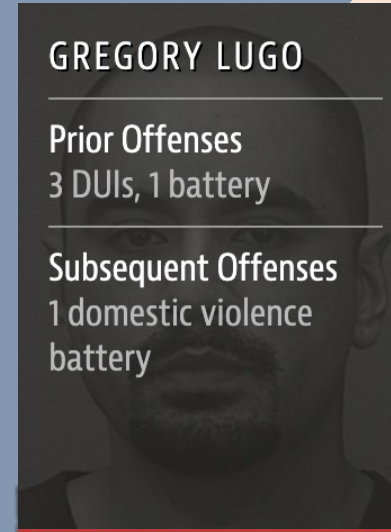
---

Subsequent Offenses  
None

---

HIGH RISK **10**

Two Drug Possession Arrests



**GREGORY LUGO**

---

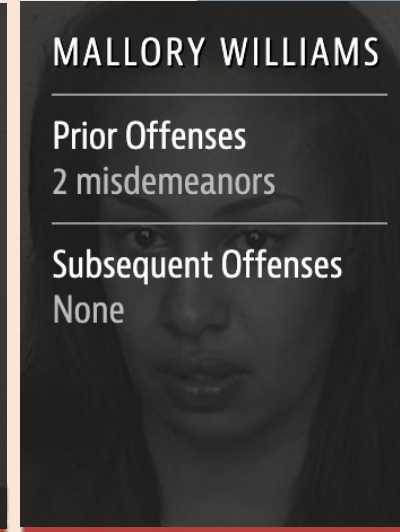
Prior Offenses  
3 DUIs, 1 battery

---

Subsequent Offenses  
1 domestic violence battery

---

LOW RISK **1**



**MALLORY WILLIAMS**

---

Prior Offenses  
2 misdemeanors

---

Subsequent Offenses  
None

---

MEDIUM RISK **6**

Two DUI Arrests



# Publication ML Techniques

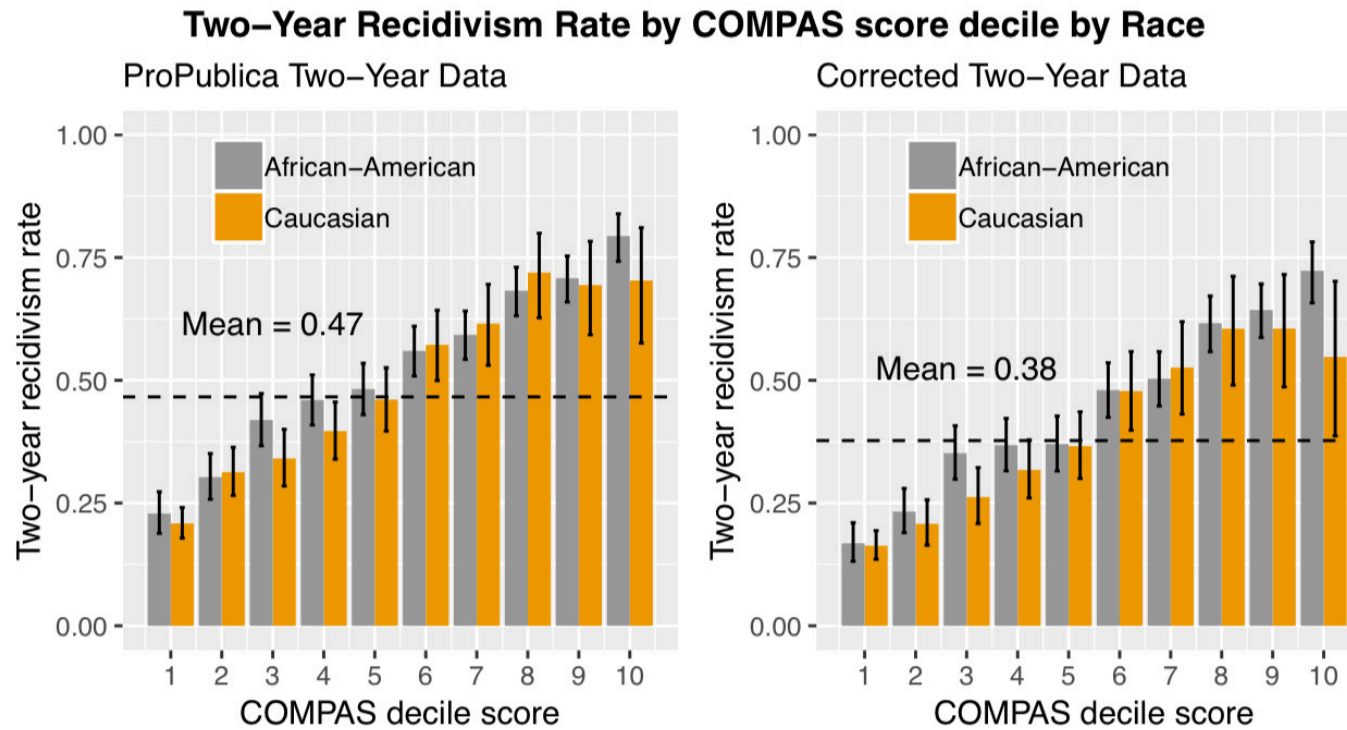
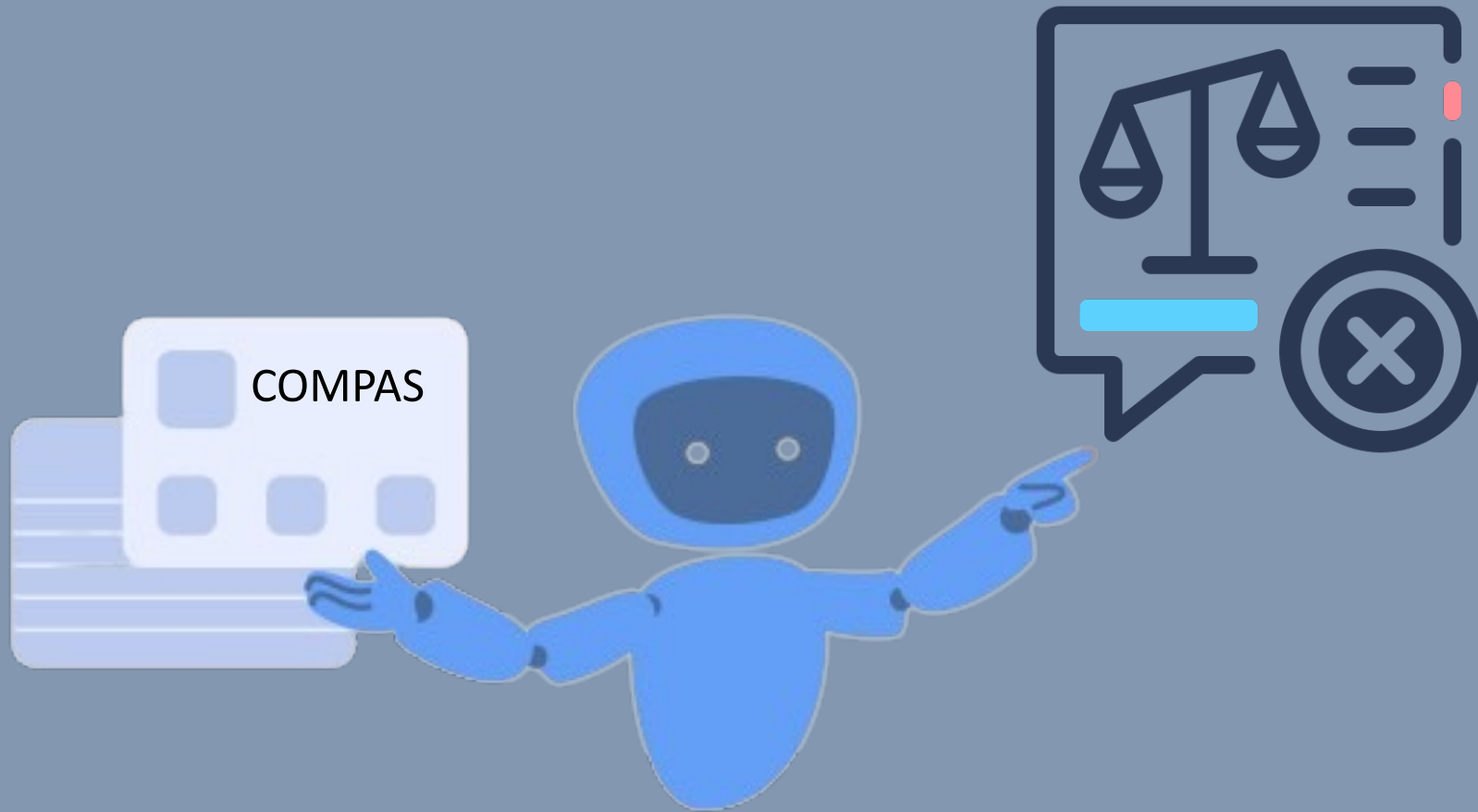


Figure 1: Two-Year Recidivism Rate by COMPAS score by Race (replicating Chouldechova 2016)



# Publication Conclusion







# Our Goal

The COMPAS  
dataset is  
notoriously biased

Showcase the  
potential of  
adversarial  
debiasing by using it  
with the COMPAS  
dataset to create a  
fair model



▶ 0 2

# MACHINE LEARNING TECHNIQUES





# Employing Fairness Techniques on Machine Learning Model

Fairness Metrics:

**Equalized Odds**

**Demographic Parity**

Adversarial Debiasing

Lime

Random Forest



# Dataset

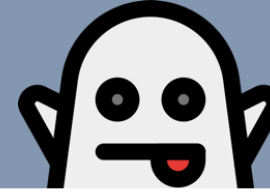
- COMPAS dataset is a collection of data on approximately 11,757 criminal defendants
- Dataset includes information on each defendant's criminal history, demographic information such as age, gender, and race, and the COMPAS scores assigned to them
- Dataset also includes information on each defendant's actual recidivism and re-arrest history, which can be used to assess the accuracy of the COMPAS scores





# ML techniques

## Dataset



A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	
id	name	first	last	compas_screen	sex	dob	age	age_cat	race	juv_fel_count	decile_score	juv_mid_count	juv_other_count	priors_count	days_b_screen	c_jail_in	c_jail_out	c_case_number	c_offense_date	c_arrest_date	c_days_from_arrest	c_charge_desc	c_charge_deg	is_recid	r_case_number	r_charge_desc	r_days
1	miguel hernandez	miguel	hernandez	2013-08-14	Male	1947-04-18	69	Greater than 45	Other	0	1	0	0	0	-1	2013-08-13	6:03	2013-08-14	5:41	13011352CF10A	2013-08-13	1	F	Aggravated Assault	0		
3	kevon dixon	kevon	dixon	2013-01-27	Male	1982-01-22	34	25 - 45	African-American	0	3	0	0	0	-1	2013-01-26	3:45	2013-02-05	5:36	13001275CF10A	2013-01-26	1	F	Felony Battery w/	1	13006779CF10/ (F3)	
4	ed philo	ed	philo	2013-04-14	Male	1991-05-21	24	Less than 25	African-American	0	4	0	1	4	-1	2013-04-13	4:58	2013-04-14	7:02	13005330CF10A	2013-04-13	1	F	Possession of C	1	13011511MM10/ (M1)	
5	marco brown	marco	brown	2013-01-13	Male	1993-01-21	23	Less than 25	African-American	0	8	1	0	1	0							1	F	Possession of C	0		
6	boutby pierre-louis	boutby	pierre-louis	2013-02-26	Male	1973-01-22	43	25 - 45	Other	0	1	0	0	2	0							1	F	arrest case no cl	0		
7	marsha millos	marsha	millos	2013-11-30	Male	1971-08-22	44	25 - 45	Other	0	1	0	0	0	0							2013-01-09	0	M	Battery	0	
8	edward riddle	edward	riddle	2014-02-19	Male	1974-07-23	41	25 - 45	Caucasian	0	6	0	0	14	-1	2014-02-18	5:08	2014-02-24	12:11	14002304CF10A	2014-02-18	1	F	Possession Burg	1	14004485CF10/ (F2)	
9	steven stewart	steven	stewart	2013-08-30	Male	1973-02-25	43	25 - 45	Other	0	4	0	0	3	-1	2013-08-29	8:55	2013-08-30	8:42	13012112CF10A	2014-02-18	1	F	arrest case no cl	0		
10	elizabeth thieme	elizabeth	thieme	2014-03-16	Female	1976-06-03	39	25 - 45	Caucasian	0	1	0	0	0	-1	2014-03-15	5:35	2014-03-18	4:28	14004524MM10	2014-03-15	1	M	Battery	0		
13	bo bradac	bo	bradac	2013-11-04	Male	1994-06-10	21	Less than 25	Caucasian	0	3	0	0	1	428	2015-01-06	3:56	2015-01-07	3:38	13000017CF10A	2012-12-31	308	F	Insurance Fraud	1	15002891MM10 (M1)	
15	elijaher lanza	elijaher	lanza	2013-10-03	Male	1992-08-18	23	Less than 25	African-American	0	6	0	0	3	0							0	M	Battery	1	14010414TC40/ (M2)	
16	kortney coleman	kortney	coleman	2013-01-01	Female	1978-06-22	37	25 - 45	Caucasian	0	1	0	0	0	0							0	M	Battery	0		
17	craig gilbert	craig	gilbert	2013-10-30	Female	1968-06-14	47	Greater than 45	Caucasian	0	1	0	0	1	-20	2013-10-10	5:12	2013-10-24	11:3	12000212CF10A	2013-10-03	490	F	arrest case no cl	1	14007704CF10/ (F3)	
20	samuel seraphin	samuel	seraphin	2014-03-24	Male	1985-03-25	31	25 - 45	African-American	0	3	0	0	7	22	2014-06-25	2:15	2014-06-28	5:02	14004186CF10A	2014-03-24	71	F	Felony Driving V	1	14009921MM10 (M1)	
21	mario hernandez	mario	hernandez	2014-03-24	Male	1979-01-25	37	25 - 45	Hispanic	0	1	0	0	0	0							0	M	Battery	0		
22	darrious davis	darrious	davis	2013-12-22	Male	1990-06-22	25	25 - 45	African-American	0	10	0	0	0	-1	2013-12-21	5:42	2013-12-22	9:11	13017598CF10A	2013-12-21	1	F	Grand Theft in Tr	0		
23	neil heckart	neil	heckart	2013-11-17	Male	1984-12-24	31	25 - 45	Caucasian	0	5	0	0	6	-1	2013-11-16	7:12	2013-11-17	8:28	13015994CF10A	2013-11-16	1	F	Driving While Lic	1	14010409CF10/ (F3)	
24	michael lux	michael	lux	2014-11-15	Male	1985-01-08	31	25 - 45	Caucasian	0	3	0	0	5	-1	2014-11-14	1:37	2014-11-15	4:57	14015458CF10A	2014-11-14	1	F	Possession Of F	1	15005595CF10/ (F3)	
25	columbus wilson	columbus	wilson	2014-06-02	Male	1991-08-28	64	Greater than 45	African-American	0	9	0	0	13	-1	2014-05-01	6:49	2014-05-02	9:28	14000363CF10A	2014-05-01	1	F	arrest case no cl	1	14008198CF10/ (F2)	
26	vandivuet william	vandivuet	williams	2013-04-18	Male	1994-11-29	21	Less than 25	African-American	0	1	0	0	1	-1	2013-04-17	1:40	2013-04-20	4:19	13005492CF10A	2013-04-17	1	F	Battery on Law E	1	13009581MM10 (M2)	
27	nelson avolio	nelson	avolio	2014-10-16	Male	1988-06-06	27	25 - 45	Caucasian	0	2	0	0	0	0							0	M	Possession Of L	1	15003574TC30/ (M2)	
28	janel denicola	janel	denicola	2013-11-22	Female	1995-03-22	21	Less than 25	Caucasian	0	4	0	0	0	-2	2013-11-20	4:12	2013-11-21	7:53	13016112CF10A	2013-11-20	2	F	Introduce Contr	0		
30	dominic pabon	dominic	pabon	2013-02-08	Male	1992-01-23	24	Less than 25	Hispanic	0	4	0	0	1	0							1	F	Lewd/Lasc Batte	1	14001650MM40 (M1)	
32	russell sottle	russell	sottle	2013-01-10	Male	1973-01-10	43	25 - 45	Caucasian	0	1	0	0	1	-1	2013-01-24	3:43	2013-01-25	9:12	13003773TC10A	2013-01-24	1	M	Susp Drivers Lic	0		
33	andre ashley	andre	ashley	2013-05-11	Male	1983-08-24	32	25 - 45	Other	0	3	0	0	0	-1	2013-05-10	8:38	2013-05-12	3:33	13009088MM10	2013-05-10	1	M	Battery	0		
37	deandra counts	deandra	counts	2013-05-06	Female	1989-02-08	27	25 - 45	African-American	0	3	0	0	8	-1	2013-05-05	9:07	2013-05-22	9:08	13009419CF10A	2013-05-05	1	F	Carrying Conceal	1	14001039TC10/ (M2)	
38	victoria soltau	victoria	soltau	2013-03-18	Female	1979-09-03	36	25 - 45	Caucasian	0	3	0	0	3	53	2013-05-10	11:1	2013-05-11	7:01	12018170CF10A	2012-12-13	96	F	arrest case no cl	0		
39	najee sapp	najee	sapp	2013-02-20	Male	1989-04-27	26	25 - 45	African-American	0	7	0	1	5	-1	2013-02-19	7:47	2013-02-20	8:33	13002542CF10A	2013-02-19	1	F	Pos Cannabis W	1	14011400MM10 (M1)	
40	victor moreno	victor	moreno	2014-10-24	Male	1983-02-03	33	25 - 45	African-American	0	10	0	0	0	0							1	F	Tampering With	1	15010523CF10/ (F3)	
41	william pinkie	william	pinkie	2013-01-28	Male	1985-10-17	30	25 - 45	Caucasian	0	9	0	0	9	0							0	F	Felony Battery w	1	13009605CF10/ (F3)	
42	maslin brutus	maslin	brutus	2014-11-15	Male	1984-03-16	32	25 - 45	African-American	0	8	0	0	4	-1	2014-11-14	3:07	2014-11-15	4:56	14015347CF10A	2014-11-14	1	F	Att Tamper w/Pl	1	14015680CF10/ (M1)	
45	mark friedland	mark	friedland	2013-12-30	Male	1980-07-27	55	Greater than 45	Caucasian	0	1	0	0	0	-1	2013-12-29	6:37	2013-12-31	10:2	13017948CF10A	2013-12-29	1	F	arrest case no cl	0		
50	thaurica watson	thaurica	watson	2013-04-04	Male	1986-07-27	49	Greater than 45	Other	0	3	0	0	7	-1	2013-04-03	4:47	2013-04-07	2:44	13004783CF10A	2013-04-04	1	F	Agg Fleeting and	1	15000545MM20 (F3)	
51	kurt towks	kurt	towks	2013-04-09	Male	1990-02-11	26	25 - 45	Caucasian	0	8	0	0	2	6							382	M	Carrying W/O I	0		
52	michael trisch	michael	trisch	2013-12-03	Female	1983-02-02	33	25 - 45	Caucasian	0	5	0	0	4	-1	2013-12-02	1:48	2013-12-13	5:25	13022428MM10	2013-12-02	1	M	Battery	1	14012292CF10/ (M1)	
53	brooks nunez	brooks	nunez	2013-09-15	Male	1991-05-14	24	Less than 25	Caucasian	0	2	0	0	1	0							0	M	Battery	1	14006651CF10/ (F3)	
54	walter atwell	walter	atwell	2013-01-11	Male	1981-06-01	34	25 - 45	African-American	2	9	1	3	21	0							0	F	Poss Wep Conv	1	14009119MM10 (M1)	
55	darling madrano	darling	madrano	2013-01-16	Male	1987-04-10	29	25 - 45	Caucasian	0	2	0	0	0	0							1	M	Possess Cannab	0		
56	kiantle stocum	kiantle	stocum	2013-08-24	Female	1994-08-17	21	Less than 25	African-American	0	8	0	0	2	-1	2013-08-23	10:32	2013-11-11	4:30	13011710CF10A	2013-08-23	1	F	arrest case no cl	0		
57	porfirio zamot	porfirio	zamot	2013-02-28	Male	1964-11-24	51	Greater than 45	African-American	0	1	0	0	0	-1	2013-02-27	4:44	2013-02-28	8:01	13004098MM10	2013-02-27	1	M	Unlaw Use Falx	0		
59	moises miranda	moises	miranda	2014-04-06	Male	1986-06-27	29	25 - 45	Caucasian	0	2	0	0	2	-1	2014-04-05	8:45	2014-04-06	1:46	14004060MM10	2014-04-05	1	M	Viol Injunct Dom	1	14009508MM10 (M1)	
61	brenda plummer	brenda	plummer	2013-02-14	Female	1964-11-11	51	Greater than 45	African-American	0	7	0	0	7	-1	2013-02-13	9:29	2013-02-15	1:35	13002244CF10A	2012-03-13	1	F	Possession of C	0		
64	derrick nims	derrick	nims	2014-03-03	Male	1987-03-03	29	25 - 45	African-American	0	2	0	0	0	0							0	F	Defrauding Inck	1	14007343MM10 (M1)	
66	jeffery dowdy	jeffery	dowdy	2013-03-28	Male	1990-05-28	25	25 - 45	Caucasian	0	10	0	0	9	-1	2013-03-27	10:12	2014-04-03	9:27	14005313MM10	2014-03-27	1	M	Battery	1	14004719CF10/ (F2)	
67	eddie jones	eddie	jones	2014-06-19	Male	1981-01-19	35	25 - 45	African-American	1	8	4	13	-1	2014-05-18	4:05	2014-05-19	1:47	14008989CF10A	2014-06-19	1	F	Uttering a Forge	1	14008544MM10 (M3)		
68	michael harper	michael	harper	2013-08-02	Male	1987-02-19	49	Greater than 45	Caucasian	0	1	0	0	0	0							0	M	DUI Level 0.15 C	0		
69	anthony bennett	anthony	bennett	2014-12-29	Male	1982-02-27	34	25 - 45	African-American	0	3	0	0	4	-59	2014-10-28	4:07	2014-12-18	9:44	14075628TC20A	2014-12-29	59	M	Driving License i	1	15051590TC20/ (M2)	
70	jeffery piens	jeffery	piens	2013-04-24	Male	1986-10-06	29	25 - 45	African-American	0	4	0	0	0	0							1	F	Tampering With	0		
72	joshel marcoran	joshel	marcoran	2013-06-03	Male	1986-07-30	29	25 - 45	Caucasian	0	2	0	0	1	-24	2013-05-10	5:18	2013-05-10	7:43	13006728CF10A	2013-05-10	24	F	Possession of O	1	14029973MU10/ (M1)	
72	shinell badstreck	shinell	badstreck	2013-11-18	Female	1987-09-29	28	25 - 45	Other	0	2	0	0	0	-1	2013-11-17	9:43	2013-11-18	8:03	13015985CF10A	2013-11-17	1	F	Grand Theft in Tr	0		
75	elliott knauff	elliott	knauff	2013-06-02	Male	1987-05-01	48	Greater than 45	Caucasian	0	6	0	0	20	302	2014-02-28	11:1	2014-03-11	2:46	11010574ACF10A	2011-10-18	562	F	arrest case no cl	1	14007040MM20 (M2)	
77	graciela quevedo	graciela	quevedo	2014-01-22	Female	1952-08-15	63	Greater than 45	Hispanic	0																	



ML techniques

# Dataset

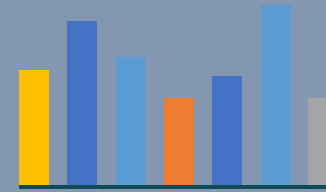


A	B	C	D	E	F	G	H	I	J
id	name	first	last	sex	age	race	priors_count	c_charge_degree	two_year_recid
1	miguel hernandez	miguel	hernandez	Male	69	Other	0	F	0
3	kevon dixon	kevon	dixon	Male	34	African-American	0	F	1
4	ed philo	ed	philo	Male	24	African-American	4	F	1
5	marcu brown	marcu	brown	Male	23	African-American	1	F	0
6	bouthy pierrelouis	bouthy	pierrelouis	Male	43	Other	2	F	0
7	marsha miles	marsha	miles	Male	44	Other	0	M	0
8	edward riddle	edward	riddle	Male	41	Caucasian	14	F	1
9	steven stewart	steven	stewart	Male	43	Other	3	F	0
10	elizabeth thieme	elizabeth	thieme	Female	39	Caucasian	0	M	0
13	bo bradac	bo	bradac	Male	21	Caucasian	1	F	1
15	ellyaher lanza	ellyaher	lanza	Male	23	African-American	3	M	1
16	kortney coleman	kortney	coleman	Female	37	Caucasian	0	M	0
19	craig gilbert	craig	gilbert	Female	47	Caucasian	1	F	1
20	samuel seraphin	samuel	seraphin	Male	31	African-American	7	F	1
21	mario hernandez	mario	hernandez	Male	37	Hispanic	0	M	0
22	darrious davis	darrious	davis	Male	25	African-American	3	F	0
23	neil heckart	neil	heckart	Male	31	Caucasian	6	F	1
24	michael lux	michael	lux	Male	31	Caucasian	5	F	1
25	columbus wilson	columbus	wilson	Male	64	African-American	13	F	1
26	vandivuiet william	vandivuiet	williams	Male	21	African-American	1	F	1
27	nelson avalo	nelson	avallo	Male	27	Caucasian	0	M	1
28	janel denicola	janel	denicola	Female	21	Caucasian	0	F	0
30	dominic pabon	dominic	pabon	Male	24	Hispanic	1	F	1
32	russell sottile	russell	sottile	Male	43	Caucasian	1	M	0
33	andre ashley	andre	ashley	Male	32	Other	0	M	0
37	deandrae counts	deandrae	counts	Male	27	African-American	8	F	1
38	victoria soltau	victoria	soltau	Female	36	Caucasian	3	F	0
39	najee sapp	najee	sapp	Male	26	African-American	5	F	1
40	victor moreno	victor	moreno	Male	33	African-American	0	F	1
41	william pirkle	william	pirkle	Male	30	Caucasian	9	F	1
42	maslin brutus	maslin	brutus	Male	32	African-American	4	F	1
45	mark friedland	mark	friedland	Male	55	Caucasian	0	F	0
50	maurice watson	maurice	watson	Male	49	Other	7	F	1
51	kurt fowks	kurt	fowks	Male	26	Caucasian	6	M	0
52	michael tritsch	michael	tritsch	Female	33	Caucasian	4	M	1
53	brooks nunez	brooks	nunez	Male	24	Caucasian	1	M	1
54	walter atwell	walter	atwell	Male	34	African-American	21	F	1
55	darling madrano	darling	madrano	Male	29	Caucasian	0	M	0
56	kianté slocum	kianté	slocum	Female	21	African-American	2	F	0
57	porfirio zamot	porfirio	zamot	Male	51	African-American	2	M	0

# Flow Chart

1

1 Biased Data



2

2 Adversarial Debiasing

3

3 Model (Random Forest + Fairness Metric)

4

4 LIME

5

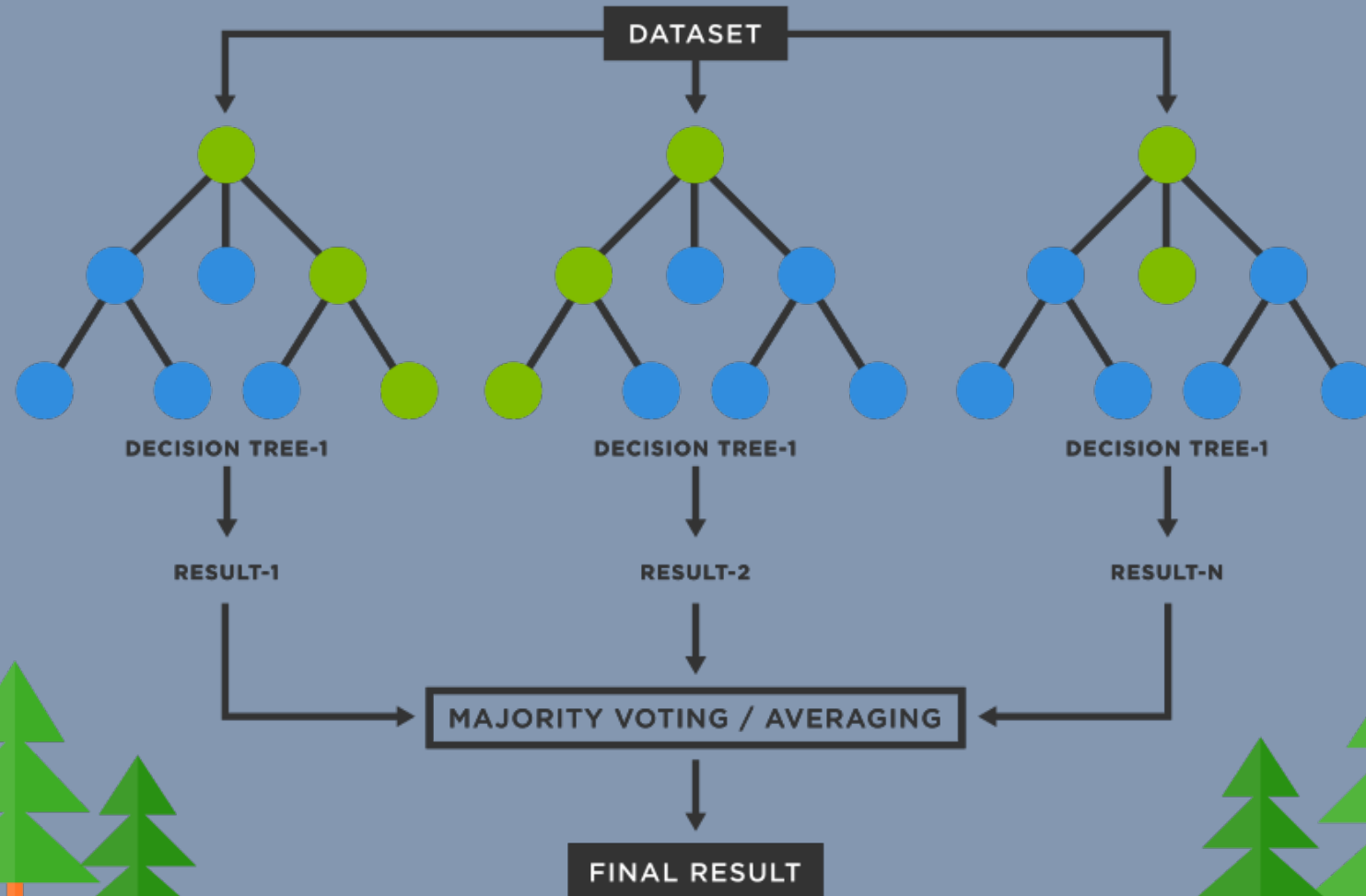
RESULT 5







# Random Forest





# Random Forest

```
# Train the random forest classifier model  
rf_model = RandomForestClassifier(n_estimators=100, max_depth=6, random_state=42)  
rf_model.fit(train_dataset.features, train_dataset.labels.ravel())
```



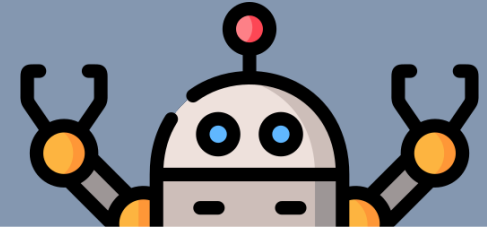


# Adversarial Debiasing





# Adversarial Debiasing

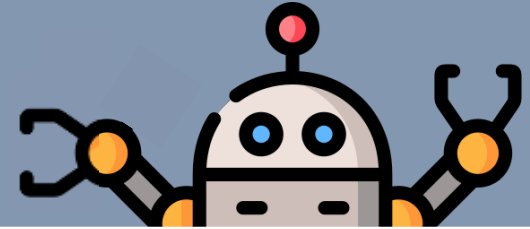


```
# Define the Adversarial Debiasing model
model = Sequential([
    Dense(64, activation='relu', input_dim=X_train.shape[1]),
    Dense(32, activation='relu'),
    Dense(1, activation='sigmoid')
])

# Use Adversarial Debiasing to mitigate bias
adv_debiasing = AdversarialDebiasing(sensitive_features=sensitive_features, random_state=42)
adv_debiasing.fit(train_dataset)
test_pred_debias = adv_debiasing.predict(test_dataset)
```



# Fairness Metrics

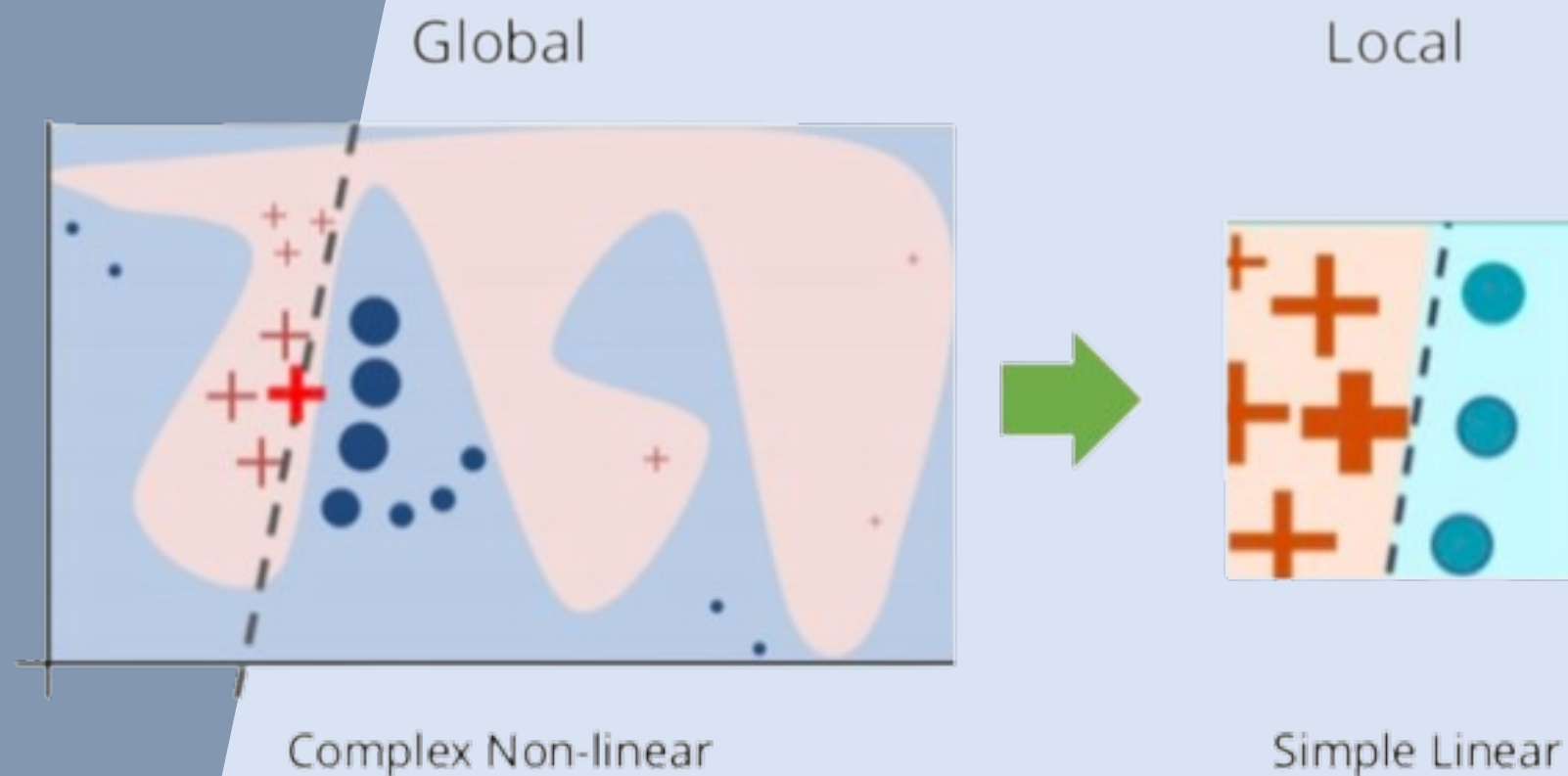


```
# Calculate fairness metrics
cm = ClassificationMetric(test_dataset, test_pred_debias, privileged_groups=[{'race': 1}], unprivileged_groups=[{'race': 0}])
print("Accuracy: ", accuracy)
print("Confusion Matrix: \n", conf_matrix)
print("Equalized Odds difference: ", cm.equalized_odds_difference())
print("Disparate Impact: ", cm.disparate_impact())
```



# Lime

Local  
Interpretable  
Model-agnostic  
Explanations





🔍 ML techniques

# Lime







# Result

The current model uses a random forest algorithm that achieves high accuracy where the data is less biased.

01

The model provides unbiased life sentences for criminals based on their crimes.

02

The model predicts an unbiased likelihood of recidivism for the criminal.

03

The model provides valid explanations for the length of the sentence given to the criminal.

04

# Conclusion

- The project aims to develop a fair machine learning model.
- This project addresses human biases in the data to offer a fair solution for all offenders.
- The goal is to counter latent bias, maximize accuracy, and ensure fairness for first-timers, those with a pre-existing criminal history, and recidivism offenders.
- The use of this model would help to overcome the problem of discrimination in the judiciary system.
- The model would provide a fair solution to any offender, regardless of factors such as race, gender, name, age, citizenship, etc.



# EXTENSION

- Debiasing the data by employing alternative methods:
  - Gradient Penalty
  - Disparate Impact Remover
  - Synthentic Minority Over-sampling Technique (SMOTE)
- Shapely Additive Explanations can be used instead of LIME.
- Decision trees or rule-based systems, could be used instead of Adversarial Debiasing to provide more transparency and understanding of the decision-making process





Application

# Application



Mortgage credit



Education



Job hiring process



Health Insurance



Criminal justice

# REFERENCES

- Barenstein, M. (2019, June 11). *[1906.04711] ProPublica's COMPAS Data Revisited*. arXiv. Retrieved April 17, 2023, from <https://arxiv.org/abs/1906.04711>
- CHRISTIN, A., ROSENBLAT, A., & BOYD, D. (2015, October 27). *Courts and Predictive Algorithms*. Data & Civil Rights. Retrieved April 17, 2023, from [http://www.datacivilrights.org/pubs/2015-1027/Courts\\_and\\_Predictive\\_Algorithms.pdf](http://www.datacivilrights.org/pubs/2015-1027/Courts_and_Predictive_Algorithms.pdf)
- Larson, J., Roswell, M., & Atlidakis, V. (2016, May 22). *propublica/compas-analysis: Data and analysis for 'Machine Bias'*. GitHub. Retrieved April 17, 2023, from <https://github.com/propublica/compas-analysis/>
- Badr, W. (2019, March 13). *Evaluating Machine Learning Models Fairness and Bias*. | by Will Badr. Towards Data Science. Retrieved April 17, 2023, from <https://towardsdatascience.com/evaluating-machine-learning-models-fairness-and-bias-4ec82512f7c3>