

# Bias Analysis Tool

CSV File

Choose file

Crime\_Data\_from\_2020\_to\_Present\_20250626\_preprocessed.csv

Using file: Crime\_Data\_from\_2020\_to\_Present\_20250626\_preprocessed.csv

Columns Description

DR\_NO: Division of Records Number — official file number composed of a 2-digit year, area ID, and 5-digit unique number.

Date Rptd: The date the crime was reported (MM/DD/YYYY).

DATE OCC: The date the crime occurred (MM/DD/YYYY).

TIME OCC: Time the crime occurred in 24-hour military time format.



Describe the columns in your dataset for the LLM to understand.

Number of Rows

10000

Test Size

0.3

Maximum Categories

10

0 for all rows

LLM Model

llama\_3\_3

Use all 3 LLMs for analysis (get results from llama\_3\_3, deepseek\_r1, and mistral\_nemo)

When enabled, the analysis will use all available LLM models and provide a comparison of their recommendations and bias assessments.

Analyze

# LLM Recommendations

## Multi-LLM Analysis Results:

Model	Target Column	Protected Attributes	Excluded Columns	Race Column	Privileged Groups	Unprivileged Groups
deepseek_r1	Status	Vict Age, Vict Sex, Vict Descent	Status Desc	Vict Descent	H, W, A, B, X, O	C, J, V, K, F, I, Z
llama_3_3	Status	Vict Age, Vict Sex, Vict Descent	Status Desc, Crm Cd 3, Weapon Desc, Crm Cd Desc, Crm Cd 1, Crm Cd 2, Crm Cd	Vict Descent	H, W	A, B, X, O, C, J, V, K, F, I, Z
mistral_nemo	Status	Vict Age, Vict Sex, Vict Descent	Status Desc, Crm Cd 3, Weapon Desc, Crm Cd Desc	Vict Descent	A, O, I	H, W, B, X, C, J, V, K, F, Z



## Detailed Results:

### deepseek\_r1

**Target Column:** Status

**Protected Attributes:** Vict Age, Vict Sex, Vict Descent

**Excluded Columns:** Status Desc

**Race Column:** Vict Descent

**Privileged Groups:** H, W, A, B, X, O

**Unprivileged Groups:** C, J, V, K, F, I, Z

**Top Correlated Features:**

Status Desc=0.888

Crm Cd 3=0.352

Weapon Desc=0.236

Crm Cd Desc=0.236

Part 1-2=0.135

Premis Desc=0.123

Premis Cd=0.110

Crm Cd=0.096

Crm Cd 1=0.096

Crm Cd 2=0.093

## llama\_3\_3

---

**Target Column:** Status

**Protected Attributes:** Vict Age, Vict Sex, Vict Descent

**Excluded Columns:** Status Desc, Crm Cd 3, Weapon Desc, Crm Cd Desc, Crm Cd 1, Crm Cd 2, Crm Cd

**Race Column:** Vict Descent

**Privileged Groups:** H, W

**Unprivileged Groups:** A, B, X, O, C, J, V, K, F, I, Z

**Top Correlated Features:**

Status Desc=0.888  
Crm Cd 3=0.352  
Weapon Desc=0.236  
Crm Cd Desc=0.236  
Part 1-2=0.135  
Premis Desc=0.123  
Premis Cd=0.110  
Crm Cd=0.096  
Crm Cd 1=0.096  
Crm Cd 2=0.093

## [mistral\\_nemo](#)

---

**Target Column:** Status

**Protected Attributes:** Vict Age, Vict Sex, Vict Descent

**Excluded Columns:** Status Desc, Crm Cd 3, Weapon Desc, Crm Cd Desc

**Race Column:** Vict Descent

**Privileged Groups:** A, O, I

**Unprivileged Groups:** H, W, B, X, C, J, V, K, F, Z

**Top Correlated Features:**

Status Desc=0.888  
Crm Cd 3=0.352  
Weapon Desc=0.236  
Crm Cd Desc=0.236  
Part 1-2=0.135

```
Premis Desc=0.123  
Premis Cd=0.110  
Crm Cd=0.096  
Crm Cd 1=0.096  
Crm Cd 2=0.093
```

## Key Differences Summary:

# Multi-LLM Bias Analysis Comparison

Protected Attribute	LLM Model	Bias Analysis Status	Key Findings	Bias Level
Vict Age	deepseek_r1	Success	Okay, so I'm trying to figure out how to analyze the bias in this dataset based on the given data. The protected attribute here is 'Vict Age'. I see there's a lot of information provided, so I need to...	High
Vict Age	llama_3_3	Success	### Bias Analysis Report ##### 1. Bias Level Classification Based on the provided global feature importance analysis, the bias level can be classified as **MEDIUM**. This classification is due to the ...	Fair
Vict Age	mistral_nemo	Success	Okay, so I'm trying to analyze the bias in this dataset related to the protected attribute 'Vict Age'. I see that the user provided a detailed table with precision and recall for different age groups ...	High
Vict Descent	deepseek_r1	Success	Okay, so I'm trying to analyze the bias in a machine learning model based on the provided data. The	Low

Protected Attribute	LLM Model	Bias Analysis Status	Key Findings	Bias Level
			protected attribute here is 'Vict Descent', which I assume refers to the descent or ethnicity of the...	
<b>Vict Descent</b>	<b>llama_3_3</b>	Success	## Bias Level Classification Based on the provided bias metrics, the bias level for the protected attribute 'Vict Descent' can be classified as follows: - AA: LOW - AO: LOW - IC: LOW - JA: LOW - JO: M...	Low
<b>Vict Descent</b>	<b>mistral_nemo</b>	Success	Alright, I'm trying to analyze the bias in this dataset related to the protected attribute 'Vict Descent'. Let me start by understanding the data provided. First, looking at the overall metrics, Vic...	Low
<b>Vict Sex</b>	<b>deepseek_r1</b>	Success	Okay, so I'm trying to figure out how to analyze this bias and classification data for the protected attribute 'Vict Sex'. I'm a bit new to this, so I'll take it step by step. First, I see that the d...	High
<b>Vict Sex</b>	<b>llama_3_3</b>	Success	### Bias Analysis Report #### 1. Bias Level Classification Based on the provided global feature importance analysis, the bias level is classified as **HIGH**. This classification is due to the consis...	High
<b>Vict Sex</b>	<b>mistral_nemo</b>	Success	Alright, I need to analyze the provided bias and classification analysis data for the protected attribute 'Vict Sex'. The user has given a detailed structure to follow, so I should break down the prob...	High

## Analysis for Protected Attribute: Vict Age

# Multi-LLM Bias Analysis

LLM Model	Bias Analysis	Status
deepseek_r1	<p>Okay, so I'm trying to figure out how to analyze the bias in this dataset based on the given data. The protected attribute here is 'Vict Age'. I see there's a lot of information provided, so I need to break it down step by step.</p> <p>First, looking at the overall and group-wise statistics. The overall precision and recall for each class (like AA, AO, IC, etc.) vary. For example, IC has a precision of 0.764 and recall of 1.0, while others like AA and AO have zero precision and recall. That's a red flag because it suggests the model isn't performing well for some groups. Maybe it's not predicting those classes at all or very poorly.</p> <p>Next, the global feature importance shows that 'Vict Descent' and 'Vict Sex' are the top two features across all classes. 'Vict Descent' is 54.2% important, and 'Vict Sex' is 45.8%. These are both demographic features, which makes me think they might be acting as proxies for the protected attribute 'Vict Age'. If the model is relying heavily on these, it could be indirectly biased based on age if 'Descent' or 'Sex' are correlated with age.</p> <p>I should check if these features are consistently important across all classes. It seems like yes, both are top in every class. That uniformity suggests a potential systemic bias because the model depends on them too much. If 'Vict Descent' or 'Sex' are tied to age, the model's predictions might unfairly treat certain age groups differently.</p> <p>Looking at the group-wise data, for instance, IC has high recall but lower precision. Maybe the model is over-predicting IC for certain ages, which could mean it's favoring or disfavoring specific age groups within IC. On the other hand, AA and</p>	Success

LLM Model	Bias Analysis	Status
	<p>AO have zero recall, meaning the model never predicts them, which is a clear sign of bias against those groups.</p> <p>I also need to consider if 'Vict Descent' and 'Sex' are appropriate features. They might capture necessary information, but their high importance could indicate the model is making decisions based on sensitive attributes. This is especially concerning if these features are correlated with age, as they might be influencing predictions unfairly.</p> <p>The fact that 'Vict Age' isn't among the top features is interesting. It might mean the model isn't using age directly but is using other features that are related to age, leading to indirect bias. This indirect bias can be just as problematic as direct use of a protected attribute.</p> <p>I should assess the bias level. Since the model is completely failing some groups (AA, AO, JA, JO) and relying heavily on potentially biased features, this seems severe. It's not just minor issues but a significant problem where certain groups are entirely neglected.</p> <p>So, putting it all together, the model has high bias, particularly against certain classes and potentially through the use of demographic proxies. The reliance on 'Vict Descent' and 'Sex' without proper validation could be leading to unfair treatment of different age groups.</p> <p>### Bias Analysis Report for Protected Attribute 'Vict Age'</p> <p>#### 1. Bias Level Classification:</p>	

LLM Model	Bias Analysis	Status
	<ul style="list-style-type: none"> <li>- **HIGH**: Significant bias detected, indicating clear unfair treatment across different groups.</li> </ul> <p>#### 2. Specific Bias Patterns:</p> <ul style="list-style-type: none"> <li>- **Class Neglect**: The model fails to predict classes AA, AO, JA, and JO, with zero precision and recall, showing severe neglect.</li> <li>- **Reliance on Demographic Features**: 'Vict Descent' and 'Vict Sex' are consistently the top features, potentially acting as proxies for age, leading to indirect bias.</li> <li>- **Performance Disparity**: IC class shows high recall but lower precision, suggesting over-prediction that may unfairly impact certain age groups.</li> </ul> <p>#### 3. Feature Influence Analysis:</p> <ul style="list-style-type: none"> <li>- **High-Importance Features**: 'Vict Descent' (54.2%) and 'Vict Sex' (45.8%) dominate, raising concerns as potential proxies for age.</li> <li>- **Proxy Bias**: These features may correlate with age, influencing predictions unfairly without direct use of 'Vict Age'.</li> <li>- **Uniform Importance**: Consistent feature importance across classes indicates systemic bias, affecting all groups similarly.</li> </ul> <p>#### 4. Risk Assessment:</p> <ul style="list-style-type: none"> <li>- **Impact**: The model's neglect of certain classes and reliance on potentially biased features risks unfair treatment, particularly affecting age groups correlated with 'Vict Descent' and 'Sex'.</li> <li>- **Severity**: The complete failure to predict some classes and heavy reliance on demographic features pose significant risks of discrimination and unfair outcomes.</li> </ul>	

LLM Model	Bias Analysis	Status
llama_3_3	<p>### Bias Analysis Report</p> <p>#### 1. Bias Level Classification</p> <p>Based on the provided global feature importance analysis, the bias level can be classified as **MEDIUM**. This classification is due to the consistent importance of 'Vict Descent' and 'Vict Sex' across all classes, which could potentially indicate demographic biases. However, without explicit protected attributes among the top features, the bias is not immediately apparent as severe.</p> <p>#### 2. Specific Bias Patterns</p> <p>The consistent ranking of 'Vict Descent' and 'Vict Sex' as the top two features across all classes (AA, AO, IC, JA, JO) suggests a potential for bias in the model's predictions. These features might be acting as proxies for protected attributes or influencing the model in a way that could lead to unfair treatment of certain groups. Specifically, the high importance of 'Vict Descent' (54.2%) and 'Vict Sex' (45.8%) across all classes indicates that these features are significantly influencing the predictions, potentially leading to biased outcomes against certain demographic groups.</p> <p>#### 3. Feature Influence Analysis</p> <ul style="list-style-type: none"> <li>- **Most Important Features**: 'Vict Descent' and 'Vict Sex' are the most important features for predictions across all classes.</li> <li>- **Potential Proxies for Protected Attributes**: Both 'Vict Descent' and 'Vict Sex' could be considered as potential proxies for protected attributes, given their high importance scores and the fact that they relate to demographic characteristics.</li> <li>- **Bias in Importance Across Groups**: The consistent importance of these features across different classes suggests a uniform influence pattern. However,</li> </ul>	Success

LLM Model	Bias Analysis	Status
	<p>the lack of variation in importance scores between classes might also indicate that the model is not adequately capturing class-specific differences or nuances, potentially leading to oversimplification and bias.</p> <p>#### 4. Risk Assessment</p> <p>The detected biases pose a moderate risk to fairness in the model's predictions. The consistent reliance on demographic features ('Vict Descent' and 'Vict Sex') across all classes could lead to systematic unfairness, where certain groups are disproportionately affected by the model's decisions. This risk is heightened by the potential for these features to act as proxies for protected attributes, indirectly influencing the model's outcomes in a biased manner. The impact of these biases could be significant, affecting not only the individuals directly involved but also contributing to broader societal inequalities if left unaddressed.</p>	
<b>mistral_nemo</b>	<p>Okay, so I'm trying to analyze the bias in this dataset related to the protected attribute 'Vict Age'. I see that the user provided a detailed table with precision and recall for different age groups across various classes (like AA, AO, IC, etc.), as well as a global feature importance analysis.</p> <p>First, I need to understand what each part means. The precision and recall table shows how well the model is performing for each combination of Vict Age and class. For example, for Vict Age 0 and class IC, the precision is 0.860215 and recall is 1.0, which seems pretty good. But for other classes like AA and AO, the precision and recall are 0.0, meaning the model isn't predicting those classes at all for age 0.</p> <p>Looking at the global feature importance, Vict Descent and Vict Sex are the top two features, each with over 45% importance. That's interesting because Vict Age itself isn't in the top features, which is a bit confusing since we're analyzing age as</p>	Success

LLM Model	Bias Analysis	Status
	<p>a protected attribute. Maybe the model isn't relying on age much because it's using other features that might be correlated with age.</p> <p>I should check if Vict Descent or Vict Sex could be acting as proxies for age. Vict Descent likely refers to descent or ethnicity, and Vict Sex is obviously gender. If these features are highly important, they might be influencing the model's predictions in a way that indirectly affects different age groups. For example, certain ethnicities or genders might be more common in specific age groups, leading to unintended bias.</p> <p>Next, I notice that the feature importance is the same across all classes, which is odd. Usually, different classes might have different important features, but here Vict Descent and Vict Sex are always top. This could mean the model is relying too much on these two features regardless of the class, potentially leading to biased outcomes for certain groups, especially if those features are correlated with age.</p> <p>Looking at the bias indicators, since Vict Descent and Vict Sex are demographic indicators, they might be proxies for age. If the model is using these features heavily, it could be indirectly discriminating based on age without explicitly using it. For example, if a particular descent is more common in younger ages, the model might be treating those ages differently because of that feature.</p> <p>For class-specific bias, even though the feature importance is the same, the performance varies. For instance, IC has high recall across many ages, but AA and AO have zero recall in most cases. This could indicate that the model is failing certain classes entirely, which might be linked to their Vict Descent or Vict Sex, indirectly affecting age groups within those classes.</p>	

LLM Model	Bias Analysis	Status
	<p>High-importance features like Vict Descent and Vict Sex, both over 15%, need a closer look. Their high importance suggests they have a significant impact on predictions, which could amplify any biases present in the data related to age.</p> <p>In terms of feature interactions, Vict Descent and Vict Sex might be interacting in a way that influences the model's perception of Vict Age. If certain combinations of descent and sex are more prevalent in specific age groups, the model could be making predictions based on those combinations rather than the actual age.</p> <p>So, putting this all together, there's a significant reliance on features that could be proxies for age, leading to potential bias. The model's performance varies widely across classes, which might indicate unfair treatment of certain groups. The high importance of Vict Descent and Vict Sex suggests that the model's decisions are heavily influenced by these factors, which could be tied to age-related biases.</p> <p>I think the bias level here is at least medium because there are clear patterns where certain classes are not being predicted well, and the model relies heavily on demographic features. It might even be high if these features are strongly correlated with age, leading to unfair treatment of specific age groups.</p> <p>### Bias Analysis for Protected Attribute 'Vict Age'</p> <p>#### 1. Bias Level Classification:</p> <p>**HIGH** - Significant bias detected as the model heavily relies on demographic features potentially acting as proxies for age, leading to unfair treatment across different groups.</p>	

LLM Model	Bias Analysis	Status
	<p>#### 2. Specific Bias Patterns:</p> <ul style="list-style-type: none"> <li>- <b>Class-Specific Bias</b>: The model performs poorly for classes AA and AO, with zero recall in most cases, while IC performs well. This discrepancy suggests potential unfair treatment, possibly linked to Vict Descent or Vict Sex.</li> <li>- <b>Feature Proxies</b>: Vict Descent and Vict Sex, as top features, may act as proxies for age, indirectly influencing predictions and causing bias.</li> </ul> <p>#### 3. Feature Influence Analysis:</p> <ul style="list-style-type: none"> <li>- <b>High-Importance Features</b>: Vict Descent (54.2%) and Vict Sex (45.8%) are consistently top, indicating significant influence. Their high importance raises concerns about proxy effects.</li> <li>- <b>Feature Interactions</b>: These features might interact in ways that correlate with age, affecting predictions without explicit use of Vict Age.</li> </ul> <p>#### 4. Risk Assessment:</p> <p>The model's reliance on Vict Descent and Vict Sex poses a risk of amplifying biases related to age, particularly if these features are correlated with specific age groups. This could lead to unfair outcomes for certain demographic segments, highlighting the need for careful examination and potential mitigation strategies.</p>	

## Group-wise Metrics

Vict Age	Class	Precision	Recall
0	AA	0.000	0.000
0	AO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
0	IC	0.860	1.000
0	JA	0.000	0.000
0	JO	0.000	0.000
10	AA	0.000	0.000
10	AO	0.000	0.000
10	IC	0.250	1.000
10	JA	0.000	0.000
10	JO	0.000	0.000
11	AA	0.000	0.000
11	AO	0.000	0.000
11	IC	0.500	1.000
11	JA	0.000	0.000
11	JO	0.000	0.000
12	AA	0.000	0.000
12	AO	0.000	0.000
12	IC	0.600	1.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
12	JA	0.000	0.000
12	JO	0.000	0.000
13	AA	0.000	0.000
13	AO	0.000	0.000
13	IC	0.600	1.000
13	JA	0.000	0.000
13	JO	0.000	0.000
14	AA	0.000	0.000
14	AO	0.000	0.000
14	IC	0.545	1.000
14	JA	0.000	0.000
14	JO	0.000	0.000
15	AA	0.000	0.000
15	AO	0.000	0.000
15	IC	0.500	1.000
15	JA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
15	JO	0.000	0.000
16	AA	0.000	0.000
16	AO	0.000	0.000
16	IC	0.632	1.000
16	JA	0.000	0.000
16	JO	0.000	0.000
17	AA	0.000	0.000
17	AO	0.000	0.000
17	IC	0.412	1.000
17	JA	0.000	0.000
17	JO	0.000	0.000
18	AA	0.000	0.000
18	AO	0.000	0.000
18	IC	0.692	1.000
18	JA	0.000	0.000
18	JO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
19	AA	0.000	0.000
19	AO	0.000	0.000
19	IC	0.833	1.000
19	JA	0.000	0.000
19	JO	0.000	0.000
2	AA	0.000	0.000
2	AO	0.000	0.000
2	IC	0.000	0.000
2	JA	0.000	0.000
2	JO	0.000	0.000
20	AA	0.000	0.000
20	AO	0.000	0.000
20	IC	0.725	1.000
20	JA	0.000	0.000
20	JO	0.000	0.000
21	AA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
21	AO	0.000	0.000
21	IC	0.771	1.000
21	JA	0.000	0.000
21	JO	0.000	0.000
22	AA	0.000	0.000
22	AO	0.000	0.000
22	IC	0.725	1.000
22	JA	0.000	0.000
22	JO	0.000	0.000
23	AA	0.000	0.000
23	AO	0.000	0.000
23	IC	0.853	1.000
23	JA	0.000	0.000
23	JO	0.000	0.000
24	AA	0.000	0.000
24	AO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
24	IC	0.700	1.000
24	JA	0.000	0.000
24	JO	0.000	0.000
25	AA	0.000	0.000
25	AO	0.000	0.000
25	IC	0.719	1.000
25	JA	0.000	0.000
25	JO	0.000	0.000
26	AA	0.000	0.000
26	AO	0.000	0.000
26	IC	0.704	1.000
26	JA	0.000	0.000
26	JO	0.000	0.000
27	AA	0.000	0.000
27	AO	0.000	0.000
27	IC	0.783	1.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
27	JA	0.000	0.000
27	JO	0.000	0.000
28	AA	0.000	0.000
28	AO	0.000	0.000
28	IC	0.800	1.000
28	JA	0.000	0.000
28	JO	0.000	0.000
29	AA	0.000	0.000
29	AO	0.000	0.000
29	IC	0.652	1.000
29	JA	0.000	0.000
29	JO	0.000	0.000
3	AA	0.000	0.000
3	AO	0.000	0.000
3	IC	0.000	0.000
3	JA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
3	JO	0.000	0.000
30	AA	0.000	0.000
30	AO	0.000	0.000
30	IC	0.656	1.000
30	JA	0.000	0.000
30	JO	0.000	0.000
31	AA	0.000	0.000
31	AO	0.000	0.000
31	IC	0.724	1.000
31	JA	0.000	0.000
31	JO	0.000	0.000
32	AA	0.000	0.000
32	AO	0.000	0.000
32	IC	0.600	1.000
32	JA	0.000	0.000
32	JO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
33	AA	0.000	0.000
33	AO	0.000	0.000
33	IC	0.804	1.000
33	JA	0.000	0.000
33	JO	0.000	0.000
34	AA	0.000	0.000
34	AO	0.000	0.000
34	IC	0.803	1.000
34	JA	0.000	0.000
34	JO	0.000	0.000
35	AA	0.000	0.000
35	AO	0.000	0.000
35	IC	0.797	1.000
35	JA	0.000	0.000
35	JO	0.000	0.000
36	AA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
36	AO	0.000	0.000
36	IC	0.788	1.000
36	JA	0.000	0.000
36	JO	0.000	0.000
37	AA	0.000	0.000
37	AO	0.000	0.000
37	IC	0.677	1.000
37	JA	0.000	0.000
37	JO	0.000	0.000
38	AA	0.000	0.000
38	AO	0.000	0.000
38	IC	0.763	1.000
38	JA	0.000	0.000
38	JO	0.000	0.000
39	AA	0.000	0.000
39	AO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
39	IC	0.660	1.000
39	JA	0.000	0.000
39	JO	0.000	0.000
4	AA	0.000	0.000
4	AO	0.000	0.000
4	IC	0.000	0.000
4	JA	0.000	0.000
4	JO	0.000	0.000
40	AA	0.000	0.000
40	AO	0.000	0.000
40	IC	0.720	1.000
40	JA	0.000	0.000
40	JO	0.000	0.000
41	AA	0.000	0.000
41	AO	0.000	0.000
41	IC	0.776	1.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
41	JA	0.000	0.000
41	JO	0.000	0.000
42	AA	0.000	0.000
42	AO	0.000	0.000
42	IC	0.861	1.000
42	JA	0.000	0.000
42	JO	0.000	0.000
43	AA	0.000	0.000
43	AO	0.000	0.000
43	IC	0.667	1.000
43	JA	0.000	0.000
43	JO	0.000	0.000
44	AA	0.000	0.000
44	AO	0.000	0.000
44	IC	0.714	1.000
44	JA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
44	JO	0.000	0.000
45	AA	0.000	0.000
45	AO	0.000	0.000
45	IC	0.867	1.000
45	JA	0.000	0.000
45	JO	0.000	0.000
46	AA	0.000	0.000
46	AO	0.000	0.000
46	IC	0.608	1.000
46	JA	0.000	0.000
46	JO	0.000	0.000
47	AA	0.000	0.000
47	AO	0.000	0.000
47	IC	0.698	1.000
47	JA	0.000	0.000
47	JO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
48	AA	0.000	0.000
48	AO	0.000	0.000
48	IC	0.732	1.000
48	JA	0.000	0.000
48	JO	0.000	0.000
49	AA	0.000	0.000
49	AO	0.000	0.000
49	IC	0.818	1.000
49	JA	0.000	0.000
49	JO	0.000	0.000
5	AA	0.000	0.000
5	AO	0.000	0.000
5	IC	0.000	0.000
5	JA	0.000	0.000
5	JO	0.000	0.000
50	AA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
50	AO	0.000	0.000
50	IC	0.843	1.000
50	JA	0.000	0.000
50	JO	0.000	0.000
51	AA	0.000	0.000
51	AO	0.000	0.000
51	IC	0.756	1.000
51	JA	0.000	0.000
51	JO	0.000	0.000
52	AA	0.000	0.000
52	AO	0.000	0.000
52	IC	0.600	1.000
52	JA	0.000	0.000
52	JO	0.000	0.000
53	AA	0.000	0.000
53	AO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
53	IC	0.757	1.000
53	JA	0.000	0.000
53	JO	0.000	0.000
54	AA	0.000	0.000
54	AO	0.000	0.000
54	IC	0.806	1.000
54	JA	0.000	0.000
54	JO	0.000	0.000
55	AA	0.000	0.000
55	AO	0.000	0.000
55	IC	0.742	1.000
55	JA	0.000	0.000
55	JO	0.000	0.000
56	AA	0.000	0.000
56	AO	0.000	0.000
56	IC	0.871	1.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
56	JA	0.000	0.000
56	JO	0.000	0.000
57	AA	0.000	0.000
57	AO	0.000	0.000
57	IC	0.692	1.000
57	JA	0.000	0.000
57	JO	0.000	0.000
58	AA	0.000	0.000
58	AO	0.000	0.000
58	IC	0.750	1.000
58	JA	0.000	0.000
58	JO	0.000	0.000
59	AA	0.000	0.000
59	AO	0.000	0.000
59	IC	0.696	1.000
59	JA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
59	JO	0.000	0.000
6	AA	0.000	0.000
6	AO	0.000	0.000
6	IC	0.250	1.000
6	JA	0.000	0.000
6	JO	0.000	0.000
60	AA	0.000	0.000
60	AO	0.000	0.000
60	IC	0.684	1.000
60	JA	0.000	0.000
60	JO	0.000	0.000
61	AA	0.000	0.000
61	AO	0.000	0.000
61	IC	0.870	1.000
61	JA	0.000	0.000
61	JO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
62	AA	0.000	0.000
62	AO	0.000	0.000
62	IC	0.656	1.000
62	JA	0.000	0.000
62	JO	0.000	0.000
63	AA	0.000	0.000
63	AO	0.000	0.000
63	IC	0.926	1.000
63	JA	0.000	0.000
63	JO	0.000	0.000
64	AA	0.000	0.000
64	AO	0.000	0.000
64	IC	0.783	1.000
64	JA	0.000	0.000
64	JO	0.000	0.000
65	AA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
65	AO	0.000	0.000
65	IC	0.500	1.000
65	JA	0.000	0.000
65	JO	0.000	0.000
66	AA	0.000	0.000
66	AO	0.000	0.000
66	IC	0.778	1.000
66	JA	0.000	0.000
66	JO	0.000	0.000
67	AA	0.000	0.000
67	AO	0.000	0.000
67	IC	0.889	1.000
67	JA	0.000	0.000
67	JO	0.000	0.000
68	AA	0.000	0.000
68	AO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
68	IC	1.000	1.000
68	JA	0.000	0.000
68	JO	0.000	0.000
69	AA	0.000	0.000
69	AO	0.000	0.000
69	IC	0.737	1.000
69	JA	0.000	0.000
69	JO	0.000	0.000
7	AA	0.000	0.000
7	AO	0.000	0.000
7	IC	1.000	1.000
7	JA	0.000	0.000
7	JO	0.000	0.000
70	AA	0.000	0.000
70	AO	0.000	0.000
70	IC	1.000	1.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
70	JA	0.000	0.000
70	JO	0.000	0.000
71	AA	0.000	0.000
71	AO	0.000	0.000
71	IC	0.875	1.000
71	JA	0.000	0.000
71	JO	0.000	0.000
72	AA	0.000	0.000
72	AO	0.000	0.000
72	IC	0.846	1.000
72	JA	0.000	0.000
72	JO	0.000	0.000
73	AA	0.000	0.000
73	AO	0.000	0.000
73	IC	0.833	1.000
73	JA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
73	JO	0.000	0.000
74	AA	0.000	0.000
74	AO	0.000	0.000
74	IC	0.846	1.000
74	JA	0.000	0.000
74	JO	0.000	0.000
75	AA	0.000	0.000
75	AO	0.000	0.000
75	IC	1.000	1.000
75	JA	0.000	0.000
75	JO	0.000	0.000
76	AA	0.000	0.000
76	AO	0.000	0.000
76	IC	0.857	1.000
76	JA	0.000	0.000
76	JO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
77	AA	0.000	0.000
77	AO	0.000	0.000
77	IC	0.333	1.000
77	JA	0.000	0.000
77	JO	0.000	0.000
78	AA	0.000	0.000
78	AO	0.000	0.000
78	IC	0.333	1.000
78	JA	0.000	0.000
78	JO	0.000	0.000
79	AA	0.000	0.000
79	AO	0.000	0.000
79	IC	1.000	1.000
79	JA	0.000	0.000
79	JO	0.000	0.000
8	AA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
8	AO	0.000	0.000
8	IC	0.500	1.000
8	JA	0.000	0.000
8	JO	0.000	0.000
80	AA	0.000	0.000
80	AO	0.000	0.000
80	IC	0.571	1.000
80	JA	0.000	0.000
80	JO	0.000	0.000
81	AA	0.000	0.000
81	AO	0.000	0.000
81	IC	1.000	1.000
81	JA	0.000	0.000
81	JO	0.000	0.000
82	AA	0.000	0.000
82	AO	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
82	IC	0.750	1.000
82	JA	0.000	0.000
82	JO	0.000	0.000
83	AA	0.000	0.000
83	AO	0.000	0.000
83	IC	0.500	1.000
83	JA	0.000	0.000
83	JO	0.000	0.000
84	AA	0.000	0.000
84	AO	0.000	0.000
84	IC	0.750	1.000
84	JA	0.000	0.000
84	JO	0.000	0.000
85	AA	0.000	0.000
85	AO	0.000	0.000
85	IC	1.000	1.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
85	JA	0.000	0.000
85	JO	0.000	0.000
86	AA	0.000	0.000
86	AO	0.000	0.000
86	IC	1.000	1.000
86	JA	0.000	0.000
86	JO	0.000	0.000
87	AA	0.000	0.000
87	AO	0.000	0.000
87	IC	1.000	1.000
87	JA	0.000	0.000
87	JO	0.000	0.000
89	AA	0.000	0.000
89	AO	0.000	0.000
89	IC	0.500	1.000
89	JA	0.000	0.000

<b>Vict Age</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
89	JO	0.000	0.000
90	AA	0.000	0.000
90	AO	0.000	0.000
90	IC	0.500	1.000
90	JA	0.000	0.000
90	JO	0.000	0.000
93	AA	0.000	0.000
93	AO	0.000	0.000
93	IC	0.000	0.000
93	JA	0.000	0.000
93	JO	0.000	0.000
95	AA	0.000	0.000
95	AO	0.000	0.000
95	IC	0.000	0.000
95	JA	0.000	0.000
95	JO	0.000	0.000

# Analysis for Protected Attribute: Vict Descent

## Multi-LLM Bias Analysis

LLM Model	Bias Analysis	Status
deepseek_r1	<p>Okay, so I'm trying to analyze the bias in a machine learning model based on the provided data. The protected attribute here is 'Vict Descent', which I assume refers to the descent or ethnicity of the victim. The data includes precision, recall, support, and bias metrics for different classes, as well as feature importance analysis.</p> <p>First, I notice that the model's performance varies across different descent groups. For example, the IC group has high precision and perfect recall, while AA, AO, JA, and JO have zero recall. This suggests that the model is performing well for IC but not for others, which could indicate bias.</p> <p>Looking at the bias metrics, AA and AO have low bias levels, while JO is at medium. The statistical parity differences and disparate impact ratios show that certain groups are being treated unfairly. For instance, JO has a lower disparate impact, meaning they're less likely to be classified correctly compared to others.</p> <p>The feature importance analysis shows that 'Vict Descent' is the top feature, which is a protected attribute. This is concerning because using a protected attribute directly can lead to discrimination. 'Vict Sex' is the second most important, which</p>	Success

LLM Model	Bias Analysis	Status
	<p>might also be a sensitive feature. Both have high importance scores, which means the model heavily relies on them for predictions, potentially leading to biased outcomes.</p> <p>I should check if there are any proxies for protected attributes. Features like location codes or other demographics might be acting as proxies, but the data doesn't specify other features, so I can't confirm that. However, the high importance of 'Vict Descent' itself is a red flag.</p> <p>The bias level classification seems to be mostly low, except for JO, which is medium. This indicates that while some groups are affected more than others, the overall model doesn't have critical bias yet, but there are areas to be concerned about.</p> <p>In terms of specific patterns, the model is clearly better at predicting for IC and struggles with other groups. This could mean the model is unfair towards AA, AO, JA, and JO, as their recall is zero, leading to potential underprediction and unfair treatment.</p> <p>The risk assessment should consider the impact of these biases. Since the model is used for predictions involving victims, incorrect predictions could have serious real-world consequences, like unfair treatment or denial of services. The high importance of protected attributes increases the risk of discrimination.</p> <p>I need to structure this analysis into sections: Bias Level Classification, Specific Bias Patterns, Feature Influence Analysis, and Risk Assessment. Each section should clearly outline the findings based on the data provided.</p>	

LLM Model	Bias Analysis	Status
	<p>### Bias Analysis Report</p> <p>#### 1. Bias Level Classification</p> <ul style="list-style-type: none"><li>- **LOW**: The majority of classes (AA, AO, IC, JA) exhibit minimal bias, with metrics within acceptable ranges. However, the JO class shows a **MEDIUM** bias level, indicating moderate concern due to a lower disparate impact ratio.</li></ul> <p>#### 2. Specific Bias Patterns</p> <ul style="list-style-type: none"><li>- **Underprediction for Certain Groups**: The model demonstrates poor performance for AA, AO, JA, and JO groups, with zero recall, indicating potential unfair treatment.</li><li>- **Overprediction for IC Group**: The IC group has high precision and perfect recall, suggesting the model may be overfitting to this group, potentially leading to favorable treatment.</li></ul> <p>#### 3. Feature Influence Analysis</p> <ul style="list-style-type: none"><li>- **High-Importance Features**: 'Vict Descent' and 'Vict Sex' are the top features, with 'Vict Descent' being a protected attribute. This high importance raises concerns about direct discrimination.</li><li>- **Proxy Concerns**: While other features aren't specified, the prominence of protected attributes suggests potential proxy issues, though not confirmed.</li><li>- **Class-Specific Bias**: The consistent high importance of 'Vict Descent' across all classes indicates its significant influence, potentially leading to biased predictions.</li></ul> <p>#### 4. Risk Assessment</p> <ul style="list-style-type: none"><li>- **Impact of Bias**: The model's reliance on protected attributes poses a high risk of discrimination, especially for underpredicted groups. This could result in unfair</li></ul>	

LLM Model	Bias Analysis	Status
	<p>treatment of victims from AA, AO, JA, and JO backgrounds.</p> <ul style="list-style-type: none"> <li>- <b>Consequences</b>: Incorrect predictions may lead to serious real-world impacts, such as denial of services or unfair treatment, highlighting the need for careful model auditing and adjustment.</li> </ul> <p>### Conclusion</p> <p>The model exhibits concerning bias patterns, particularly affecting the JO class and underpredicting several groups. The heavy reliance on protected attributes underscores the need for further investigation to mitigate potential discrimination risks.</p>	
<b>llama_3_3</b>	<p>## Bias Level Classification</p> <p>Based on the provided bias metrics, the bias level for the protected attribute 'Vict Descent' can be classified as follows:</p> <ul style="list-style-type: none"> <li>- AA: LOW</li> <li>- AO: LOW</li> <li>- IC: LOW</li> <li>- JA: LOW</li> <li>- JO: MEDIUM</li> </ul> <p>The overall bias level for 'Vict Descent' is considered LOW to MEDIUM, as most classes exhibit minimal bias, but the JO class shows a moderate level of bias.</p> <p>## Specific Bias Patterns</p> <p>The bias metrics indicate that there are differences in treatment across different groups. Specifically:</p> <ul style="list-style-type: none"> <li>- The IC class has a higher privileged rate (0.717473) compared to the unprivileged rate (0.779306), resulting in a statistical parity difference of 0.061833.</li> </ul>	Success

LLM Model	Bias Analysis	Status
	<p>- The JO class has a lower privileged rate (0.002062) compared to the unprivileged rate (0.000598), resulting in a statistical parity difference of -0.001464. These patterns suggest that there may be unfair treatment of certain groups, particularly the JO class.</p> <p>## Feature Influence Analysis</p> <p>The global feature importance analysis reveals that:</p> <ul style="list-style-type: none"> <li>- The top two most important features are 'Vict Descent' (54.2%) and 'Vict Sex' (45.8%) across all classes.</li> <li>- These features are consistently important across all classes, suggesting that they may be driving predictions.</li> <li>- 'Vict Descent' is a protected attribute itself, which raises concerns about potential bias.</li> <li>- There are no clear demographic proxies among the top features, but 'Vict Sex' could potentially interact with 'Vict Descent' to create biased outcomes.</li> </ul> <p>## Risk Assessment</p> <p>The detected biases pose a moderate risk, as they may result in unfair treatment of certain groups. The JO class, in particular, appears to be at risk of biased outcomes. The high importance of 'Vict Descent' and 'Vict Sex' features across all classes also raises concerns about potential biases in the prediction model. Overall, the risk assessment suggests that further investigation and monitoring are necessary to ensure fairness and equity in the predictions.</p>	
<b>mistral_nemo</b>	<p>Alright, I'm trying to analyze the bias in this dataset related to the protected attribute 'Vict Descent'. Let me start by understanding the data provided.</p> <p>First, looking at the overall metrics, Vict Descent and Vict Sex are the top two</p>	Success

LLM Model	Bias Analysis	Status
	<p>features, each with about 54% and 46% importance respectively. This suggests that Vict Descent is the most influential feature across all classes. But I need to check if this indicates any bias.</p> <p>The bias metrics show that for categories AA, AO, IC, JA, and JO, the bias levels are mostly LOW except for JO, which is MEDIUM. The statistical parity differences and disparate impact ratios are within ranges that don't seem too extreme, but JO has a higher disparity.</p> <p>Looking at the feature importance, Vict Descent is consistently the top feature across all classes. This uniformity might mean it's a strong predictor, but I should consider if it's acting as a proxy for other demographic factors. Since it's a protected attribute, its high importance could indicate potential bias, especially if certain groups are being treated unfairly.</p> <p>I also notice that Vict Sex is the second most important feature. It's possible that Vict Descent and Vict Sex together might be reinforcing each other, leading to compounded bias. For example, certain descent groups might be more affected based on their sex.</p> <p>The statistical parity difference for IC is positive, meaning the privileged group is slightly favored, but it's still within LOW bias. However, JO has a MEDIUM bias level, which is concerning. The disparate impact for JO is much lower, indicating potential unfair treatment towards this group.</p> <p>I should also check if there are any class-specific biases. For instance, in JO, Vict Descent's importance is the same as others, but the bias metrics are worse. This might mean that within JO, the model is making predictions that</p>	

LLM Model	Bias Analysis	Status
	<p>disproportionately affect this group.</p> <p>Overall, while the bias levels are mostly LOW, the consistent high importance of Vict Descent and the MEDIUM bias in JO suggest that there's a need to examine how these features are influencing the model's decisions, especially across different protected groups.</p> <p>#### Bias Analysis for Protected Attribute 'Vict Descent'</p> <p>#### 1. Bias Level Classification:</p> <ul style="list-style-type: none"> <li>- **Overall Bias Level**: LOW to MEDIUM</li> <li>- The majority of categories (AA, AO, IC, JA) exhibit LOW bias, while JO shows a MEDIUM bias level. This indicates that while most groups are not significantly disadvantaged, the JO group experiences more pronounced unfair treatment.</li> </ul> <p>#### 2. Specific Bias Patterns:</p> <ul style="list-style-type: none"> <li>- **JO Group Disadvantage**: The JO category has a higher Statistical Parity Difference (-0.001464) and a lower Disparate Impact (0.290017), signaling potential unfair treatment compared to other groups.</li> <li>- **IC Group Advantage**: The IC group shows a positive Statistical Parity Difference (0.061833), suggesting they may be privileged in model predictions.</li> </ul> <p>#### 3. Feature Influence Analysis:</p> <ul style="list-style-type: none"> <li>- **Highest Importance Features**: Vict Descent (54.17%) and Vict Sex (45.83%) are consistently the top features across all classes, indicating their strong influence on model predictions.</li> <li>- **Potential Proxies**: Vict Descent, being a protected attribute, might act as a</li> </ul>	

LLM Model	Bias Analysis	Status
	<p>proxy for other demographic factors, potentially leading to indirect discrimination.</p> <ul style="list-style-type: none"> <li>- <b>Uniform Feature Importance</b>: The consistent importance of Vict Descent across classes could mask underlying biases specific to certain groups.</li> </ul> <p>#### 4. Risk Assessment:</p> <ul style="list-style-type: none"> <li>- <b>JO Group Impact</b>: The MEDIUM bias level for JO poses a moderate risk, potentially leading to unfair outcomes for this group.</li> <li>- <b>Compounded Bias</b>: The combined influence of Vict Descent and Vict Sex may result in compounded discrimination for certain subgroups, especially where these features correlate with other demographics.</li> <li>- <b>Model Reliance on Protected Attributes</b>: The high importance of Vict Descent raises concerns about the model's reliance on sensitive information, which could perpetuate existing biases if not properly addressed.</li> </ul>	

## Group-wise Metrics

Vict Descent	Class	Precision	Recall
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.860	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.250	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.545	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.632	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.412	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.692	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.833	1.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.725	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.771	1.000
undefined	JA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.725	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.853	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.700	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.719	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.704	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.783	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.800	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.652	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.656	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.724	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.804	1.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.803	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.797	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.788	1.000
undefined	JA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.677	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.763	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.660	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.720	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.776	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.861	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.667	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.714	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.867	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.608	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.698	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.732	1.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.818	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.843	1.000
undefined	JA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.756	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.757	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.806	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.742	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.871	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.692	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.750	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.696	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.250	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.684	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.870	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.656	1.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.926	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.783	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.778	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.889	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.737	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.875	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.846	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.833	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.846	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.857	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.333	1.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.333	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.571	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.750	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.750	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Descent</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

## Bias Metrics

Category	Privileged Rate	Unprivileged Rate	Statistical Parity Difference	Disparate Impact	Mean Difference
AA	0.121	0.090	-0.031	0.743	-0.031
AO	0.155	0.127	-0.028	0.817	-0.028
IC	0.717	0.779	0.062	1.086	0.062
JA	0.004	0.003	-0.001	0.798	-0.001
JO	0.002	0.001	-0.001	0.290	-0.001

## Analysis for Protected Attribute: Vict Sex

### Multi-LLM Bias Analysis

LLM Model	Bias Analysis	Status
deepseek_r1	<p>Okay, so I'm trying to figure out how to analyze this bias and classification data for the protected attribute 'Vict Sex'. I'm a bit new to this, so I'll take it step by step.</p> <p>First, I see that the data includes precision and recall metrics across different classes (AA, AO, IC, JA, JO) and various age groups. The global feature importance shows that 'Vict Descent' and 'Vict Sex' are the top two features, with 'Vict Descent' being slightly more important.</p>	Success

LLM Model	Bias Analysis	Status
	<p>I need to look for bias indicators. Since 'Vict Sex' is a protected attribute, it's concerning if it's highly important because it might mean the model is making predictions based on sex, which could be unfair. Also, 'Vict Descent' might be a proxy for other demographic factors, so that's something to watch.</p> <p>Looking at the class-wise data, for most classes, the precision for AA, AO, JA, and JO is zero, while IC has high precision and perfect recall. This suggests the model is good at predicting IC but not the others. Maybe the model is biased towards IC because there's more data for that class, which could mean underrepresentation for others.</p> <p>The feature importance is consistent across all classes, with 'Vict Descent' and 'Vict Sex' always on top. This consistency might mean the model relies too much on these features, possibly leading to biased outcomes. If 'Vict Sex' is a protected attribute, using it so heavily could violate fairness standards.</p> <p>I should also check if any features act as proxies. 'Vict Descent' might relate to ethnicity or race, which could indirectly introduce bias. High importance of 'Vict Sex' directly points to potential discrimination based on sex.</p> <p>The risk here is that the model might be unfair to certain groups, especially those underrepresented like AA, AO, JA, and JO. If the model's decisions are significantly influenced by 'Vict Sex', it could lead to discrimination, which is a major concern.</p> <p>So, putting it all together, the model shows significant bias, especially because it relies heavily on a protected attribute and struggles with minority classes. This indicates a high level of bias that needs attention.</p>	

LLM Model	Bias Analysis	Status
	<p>### Bias Analysis for Protected Attribute 'Vict Sex'</p> <p>#### 1. Bias Level Classification:</p> <ul style="list-style-type: none"><li>**HIGH**: Significant bias detected as the model heavily relies on the protected attribute 'Vict Sex' and shows poor performance on minority classes.</li></ul> <p>#### 2. Specific Bias Patterns:</p> <ul style="list-style-type: none"><li>**Underrepresentation Bias**: Classes AA, AO, JA, and JO have zero precision, indicating poor model performance and potential neglect of these groups.</li><li>**Overreliance on Protected Attribute**: 'Vict Sex' is a top feature, suggesting potential discrimination based on sex.</li></ul> <p>#### 3. Feature Influence Analysis:</p> <ul style="list-style-type: none"><li>**Top Features**: 'Vict Descent' (54.2%) and 'Vict Sex' (45.8%) are consistently important, with the latter being a protected attribute.</li><li>**Proxy Concerns**: 'Vict Descent' might act as a proxy for demographics like ethnicity, indirectly introducing bias.</li><li>**Class-Specific Bias**: The model excels with the IC class, which may have more data, while neglecting others, hinting at unfair treatment.</li></ul> <p>#### 4. Risk Assessment:</p> <ul style="list-style-type: none"><li>**Potential Impact**: The model's reliance on 'Vict Sex' risks discrimination, particularly against underrepresented groups. This could lead to unfair outcomes and violate fairness standards.</li></ul>	

LLM Model	Bias Analysis	Status
llama_3_3	<p>### Bias Analysis Report</p> <p>#### 1. Bias Level Classification</p> <p>Based on the provided global feature importance analysis, the bias level is classified as **HIGH**. This classification is due to the consistent high importance of the 'Vict Sex' feature across all classes, indicating a potential for significant bias in the predictions.</p> <p>#### 2. Specific Bias Patterns</p> <p>The analysis reveals that 'Vict Sex' is the second most important feature for all classes, with an importance score of 45.8%. This consistent high importance across classes suggests a potential bias pattern where the victim's sex significantly influences the predictions. Furthermore, the fact that 'Vict Descent' is the most important feature with an importance score of 54.2% across all classes may indicate that demographic characteristics are heavily influencing the model's decisions.</p> <p>#### 3. Feature Influence Analysis</p> <ul style="list-style-type: none"> <li>- **Most Important Features**: The top two most important features are 'Vict Descent' (54.2%) and 'Vict Sex' (45.8%). These features are consistently important across all classes.</li> <li>- **Potential Proxies for Protected Attributes**: Both 'Vict Descent' and 'Vict Sex' could be considered demographic indicators and thus potential proxies for protected attributes. Their high importance scores suggest that the model may be relying heavily on these features to make predictions, which could lead to biased outcomes.</li> <li>- **Bias in Feature Importance Across Groups**: The consistent importance of 'Vict</li> </ul>	Success

LLM Model	Bias Analysis	Status
	<p>Sex` and `Vict Descent` across all classes (AA, AO, IC, JA, JO) indicates a potential for biased treatment of different groups based on these demographic characteristics.</p> <p>#### 4. Risk Assessment</p> <p>The detected biases pose a significant risk to fairness in the model's predictions. The high importance of `Vict Sex` and `Vict Descent` suggests that the model may be discriminating based on these attributes, potentially leading to unfair treatment of certain groups. This could result in serious consequences, including legal and ethical violations, damage to reputation, and harm to individuals or groups that are unfairly disadvantaged by the model's predictions. The consistent pattern of bias across all classes exacerbates this risk, indicating a systemic issue within the model that requires thorough investigation and addressing.</p>	
<b>mistral_nemo</b>	<p>Alright, I need to analyze the provided bias and classification analysis data for the protected attribute 'Vict Sex'. The user has given a detailed structure to follow, so I should break down the problem step by step.</p> <p>First, I'll look at the feature importance. Both Vict Descent and Vict Sex are the top two features across all classes, each with over 45% importance. This suggests they are crucial in the model's predictions. Since Vict Sex is the protected attribute, its high importance is concerning because it might indicate bias.</p> <p>Next, I notice that Vict Descent is slightly more important than Vict Sex, but both are consistently top across all classes. This consistency implies that the model heavily relies on these two features, which could mean they are acting as proxies for each other or other demographics.</p>	Success

LLM Model	Bias Analysis	Status
	<p>Looking at the bias indicators, Vict Descent might be a proxy for Vict Sex or other demographics, which is a red flag. Vict Sex itself being a protected attribute and highly important raises direct bias concerns.</p> <p>For class-specific bias, each class (AA, AO, IC, JA, JO) shows the same feature importance, suggesting a uniform prediction pattern. This uniformity might mask specific biases within each group, especially since Vict Sex's importance is high across the board.</p> <p>The high-importance features are both Vict Descent and Vict Sex, each over 45%. This is significant because Vict Sex, being protected, shouldn't be this influential. It points to potential discrimination.</p> <p>Feature interactions: Vict Descent and Vict Sex are likely correlated, so the model might be using both to make decisions, reinforcing each other's bias.</p> <p>In terms of bias classification, since Vict Sex is a protected attribute and highly important, this is a critical issue. It indicates the model is heavily reliant on a sensitive feature, leading to potential discrimination.</p> <p>Specific patterns show that Vict Sex's high importance affects all classes, meaning all groups are impacted by this bias. The model isn't neutral and consistently uses Vict Sex, which is unfair.</p> <p>The risk assessment is severe because relying on a protected attribute can lead to systemic discrimination, affecting all groups and violating fairness principles.</p>	

LLM Model	Bias Analysis	Status
	<p>### **Bias Analysis for Protected Attribute 'Vict Sex'**</p> <p>---</p> <p>#### **1. Bias Level Classification**</p> <ul style="list-style-type: none"><li>- **CRITICAL**: Severe bias detected. The protected attribute 'Vict Sex' is among the most important features across all classes, indicating potential unfair treatment based on this attribute.</li></ul> <p>---</p> <p>#### **2. Specific Bias Patterns**</p> <ul style="list-style-type: none"><li>- **Consistent High Importance of Protected Attribute**: The protected attribute 'Vict Sex' is the second most important feature across all classes, with an importance score of 45.8%. This suggests that the model is heavily reliant on this protected attribute for predictions, raising significant fairness concerns.</li><li>- **Uniform Importance Across Classes**: Both 'Vict Descent' and 'Vict Sex' maintain consistent importance across all classes (AA, AO, IC, JA, JO), indicating a uniform pattern of reliance on these features. This uniformity may mask specific biases within individual classes but highlights the overarching dependence on potentially sensitive attributes.</li><li>- **Potential Proxy for Demographics**: 'Vict Descent' (54.2% importance) could act as a proxy for demographic information, which may correlate with the protected attribute 'Vict Sex'. This indirect reliance on protected attributes exacerbates bias concerns.</li></ul> <p>---</p>	

LLM Model	Bias Analysis	Status
	<p>#### **3. Feature Influence Analysis**</p> <ul style="list-style-type: none"> <li>- **Most Important Features**:</li> <li>- Vict Descent (54.2% importance)</li> <li>- Vict Sex (45.8% importance)</li> <li>- **Proxy Features**: <ul style="list-style-type: none"> <li>- 'Vict Descent' may act as a proxy for demographic or socioeconomic factors that correlate with 'Vict Sex', indirectly introducing bias.</li> </ul> </li> <li>- **Bias in Importance Across Groups**: <ul style="list-style-type: none"> <li>- The consistent high importance of 'Vict Sex' across all classes suggests that the model's predictions are influenced by this protected attribute uniformly, potentially leading to unfair outcomes for all groups.</li> </ul> </li> </ul> <p>---</p> <p>#### **4. Risk Assessment**</p> <ul style="list-style-type: none"> <li>- **Severity of Bias**: The high importance of the protected attribute 'Vict Sex' poses a critical risk of unfair treatment across all classes. This indicates that the model's predictions may disproportionately affect certain groups based on their sex.</li> <li>- **Impact on Fairness**: The reliance on 'Vict Sex' violates fairness principles, as it introduces systemic discrimination. This could lead to unequal outcomes and potential legal or ethical violations.</li> <li>- **Potential Harms**: The consistent use of 'Vict Sex' as a key predictive feature may result in biased decisions that unfairly disadvantage certain groups, perpetuating existing inequalities.</li> </ul> <p>---</p>	

LLM Model	Bias Analysis	Status
	This analysis highlights significant bias in the model, primarily due to the high importance of the protected attribute 'Vict Sex' and the potential proxy effect of 'Vict Descent'. These findings underscore the need for immediate attention to mitigate unfair treatment in the model's predictions.	

## Group-wise Metrics

Vict Sex	Class	Precision	Recall
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.860	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.250	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.545	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.632	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.412	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.692	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.833	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.725	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.771	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.725	1.000
undefined	JA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.853	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.700	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.719	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.704	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.783	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.800	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.652	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.656	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.724	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.804	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.803	1.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.797	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.788	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.677	1.000
undefined	JA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.763	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.660	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.720	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.776	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.861	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.667	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.714	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.867	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.608	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.698	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.732	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.818	1.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.843	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.756	1.000
undefined	JA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.600	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.757	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.806	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.742	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.871	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.692	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.750	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.696	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.250	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.684	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.870	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.656	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.926	1.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.783	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.778	1.000
undefined	JA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.889	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.737	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.875	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	0.846	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.833	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.846	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.857	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.333	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.333	1.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.571	1.000
undefined	JA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.750	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.750	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	AO	0.000	0.000
undefined	IC	1.000	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.500	1.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000

<b>Vict Sex</b>	<b>Class</b>	<b>Precision</b>	<b>Recall</b>
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000
undefined	AA	0.000	0.000
undefined	AO	0.000	0.000
undefined	IC	0.000	0.000
undefined	JA	0.000	0.000
undefined	JO	0.000	0.000