**AI Risk Report Template**

**Filename:** ai\_risk\_report.docx or ai\_risk\_report.pdf **Length:** 3–6 pages recommended

## **Project Title**

*Your team name and project name*

## **1. Problem Overview**

* What was the task? (e.g., Predict loan approval)
* Why does it matter in a real-world or ethical context?
* What dataset were you given, and what were the known sensitive attributes?

## **2. Model Summary**

* What model(s) did you use and why?
* Key preprocessing, feature engineering, or hyperparameter choices
* Performance on internal validation data (accuracy, precision, recall, etc.)

## **3. Bias Detection Process**

* Methods used to detect bias (e.g., SHAP, Disparity Impact Ratio, False Positive Rate comparisons, Grouped AUC, Fairlearn audit, etc.)
* Did you audit raw data, model output, or both?
* Were audits performed at the individual or group level?

## **4. 📉 Identified Bias Patterns**

Summarize the biases your model exhibited. Aim for at least 4 clear findings. Include charts or tables if needed.

| Bias Type | Affected Group | Evidence | Metric | Comment |
| --- | --- | --- | --- | --- |

## **5. Visual Evidence**

Embed key visuals here or reference attached files. Examples:

* Approval rate heatmaps
* SHAP/LIME feature breakdowns
* Confusion matrix by group

## **6. Real-World Implications**

* Who is most at risk if your model were deployed as-is?
* What are the ethical or social consequences?
* Would your model pass a fairness audit in a regulated setting?

## **7. Limitations & Reflections**

* What didn’t work?
* What would you try next time with more time or data?
* Any lessons learned on fairness or auditing?