ETHICAL REFLECTION

When deploying the predictive model from Task 3 in a real company for issue/resource prioritization:

Potential biases

1. Dataset imbalance: Malignant (high priority) cases are underrepresented in some populations; model may under-prioritize critical issues from minority groups or rare projects.
2. Historical labeling bias: Past triage decisions often reflect senior engineer availability rather than true urgency → model perpetuates delayed handling for certain teams/products.
3. Feature proxy bias: Texture, perimeter etc. analogs in software issues (e.g. log volume, affected users) can indirectly encode team size or department, leading to discriminatory resource allocation.

Mitigation with fairness tools

IBM AI Fairness 360 (AIF360) or Fairlearn can be applied as follows:

- Measure disparity metrics (e.g., equalized odds, demographic parity) across protected groups (team, region).
- Apply pre-processing (re-weighting), in-processing (adversarial debiasing), or post-processing (equalized odds post-hoc) to reduce disparity while maintaining ≥90% overall accuracy. Regular auditing with AIF360's interactive dashboard ensures transparency before production rollout.