Name _____

When Smart Machines Make Dumb Decisions

AI systems are only as fair and smart as the data and humans behind them. Sometimes, algorithms make mistakes that seem logical to machines - but unfair to people. In this activity, you'll read a few short stories about how an AI reached a wrong or biased conclusion. For each one, identify **where the bias or error might have entered the system**, then **suggest a practical improvement** that could make the AI's decision more fair, accurate, or human-aware.

1. The Job Application Filter - A company used an AI system to screen résumés for "strong leadership potential." The program learned from ten years of hiring data. Within a month, it started rejecting most female applicants. When engineers checked, they found the algorithm had noticed that most past "successful hires" were men - and treated gender as a signal for success.

Where might the bias or error have entered?

How could this system be improved?

2. The Predictive Policing Program - A city's police department deployed an AI to predict where crimes were likely to occur. The tool based its predictions on arrest records from the past decade. Soon, the AI kept flagging the same neighborhoods - areas that already had heavier policing. As a result, officers were sent there more often, creating more arrests and feeding the cycle.

Where might the bias or error have entered?

How could this system be improved?

3. The Loan Approval Assistant - A bank used an algorithm to decide who qualified for home loans. It trained the model using decades of data from past borrowers. But many applicants from certain ZIP codes were consistently rejected - even when they had good credit. The algorithm had learned to associate those areas (historically minority neighborhoods) with "higher risk."

Where might the bias or error have entered?

How could this system be improved?