PHIẾU HỌC TẬP NHÓM ĐÔI: BÓC TRẦN "SỰ NHẦM LẪN" CỦA AI

Giới thiệu: Ma trận nhầm lẫn (Confusion Matrix) không phải là một công thức tính toán phức tạp. Nó thực chất là một chiếc bảng thống kê để chúng ta nhìn thấu xem AI đã làm tốt ở đâu và "ngốc nghếch" (nhầm lẫn) ở những chỗ nào.

Trong ma trận này, chúng ta luôn so sánh 2 điều: **Thực tế sự việc** và **Dự đoán của AI**.

HOẠT ĐỘNG 1: AI MỞ KHÓA ĐIỆN THOẠI (FACEID)

Giả sử bạn cài đặt AI nhận diện khuôn mặt để mở khóa chiếc điện thoại của mình.

- **Mục tiêu tìm kiếm (Positive):** Là CHÍNH BẠN (Chủ nhân).
- **Không phải mục tiêu (Negative):** Là NGƯỜI LẠ.

Nhiệm vụ: Hai bạn hãy cùng thảo luận và điền các tình huống (A, B, C, D) vào 4 ô tương ứng của Ma trận nhầm lẫn dưới đây:

- **Tình huống A:** Bạn đang ngáp rách miệng, AI không nhận ra và từ chối mở khóa.
- **Tình huống B:** Đứa em ruột có nét hao hao giống bạn cầm máy lên, AI tự động mở khóa.
- **Tình huống C:** Kẻ gian nhặt được điện thoại của bạn, đưa mặt vào và máy báo khóa vĩnh viễn.
- **Tình huống D:** Bạn cầm điện thoại lên, máy lập tức mở khóa.

	AI dự đoán: LÀ CHỦ NHÂN (Mở khóa)	**AI dự đoán: LÀ NGƯỜI LẠ (Khóa máy)**
Thực tế: Bạn cầm máy	**True Positive (TP)** Tình huống:	**False Negative (FN)** - *AI bỏ lọt* Tình huống: .....
Thực tế: Người khác cầm máy	**False Positive (FP)** - *AI báo động giả* Tình huống:	**True Negative (TN)** Tình huống: .....

LIVEWORKSHEETS

HOẠT ĐỘNG 2: THỬ TÀI LÀM "GIÁM KHẢO AI"

Trường học áp dụng hệ thống "Camera AI Giám thị" để phát hiện học sinh gian lận trong phòng thi.

- **Positive (Đối tượng cần bắt):** Học sinh ĐANG quay cóp.
- **Negative (Người bình thường):** Học sinh KHÔNG quay cóp.

Nhiệm vụ: Hãy đánh dấu (X) vào cột phân loại lỗi đúng nhất cho các sự việc sau mà AI đã báo cáo:

Sự việc xảy ra trong phòng thi	Lỗi False Positive(Báo động giả)	Lỗi False Negative(Bỏ lọt tội phạm)	Không có lỗi(TP hoặc TN)
1. Bạn Nam làm bài đàng hoàng nhưng ngước lên nhìn đồng hồ treo tường, AI réo còi báo gian lận.			
2. Bạn Hoa mở tài liệu dưới gầm bàn, AI khoanh đỏ và gửi ảnh cho giám thị.			
3. Bạn Tuấn chép phao thu nhỏ trong lòng bàn tay, AI không phát hiện ra điều gì bất thường.			
4. Cả phòng đang cúi xuống viết bài, AI im lặng quan sát, không có báo động nào.			

LIVEWORKSHEETS

HOẠT ĐỘNG 3: TƯ DUY PHẢN BIỆN - NHẦM LẪN NÀO TAI HẠI HƠN?

Nhóm hãy thảo luận và ghi lại ý kiến thống nhất của nhóm:

Câu 1: Trong "AI Giám thị" ở Hoạt động 2, theo em, lỗi **False Positive (FP - Báo động giả)** hay lỗi **False Negative (FN - Bỏ lọt tội phạm)** sẽ gây hậu quả xấu hơn? Tại sao?

.....

.....

Câu 2: Cho bài toán y tế: **AI chẩn đoán bệnh ung thư qua ảnh chụp X-quang.** (Positive = Có bệnh; Negative = Khỏe mạnh).

Trong trường hợp sinh tử này, lỗi **FP** (Người khỏe mạnh nhưng AI bảo mắc bệnh) hay lỗi **FN** (Người đang mắc bệnh nhưng AI bảo khỏe mạnh), lỗi nào sẽ gây nguy hiểm hơn? Tại sao?

.....

.....

LIVEWORKSHEETS