

Failure Case Studies

📌 Case 1: Racial Bias in Skin Lesion Diagnosis AI

Problem: Trained primarily on data from white patients, resulting in lower accuracy for patients of color

Result: 20% higher misdiagnosis rate for skin cancer in Black patients

Lesson: Need for balanced data collection across diverse racial groups

📌 Case 2: Racial Discrimination in Healthcare Resource Allocation Algorithm

Problem: Used healthcare costs as a proxy variable for health status

Result: Black patients incorrectly assessed as being less sick than white patients

Lesson: Must consider social inequalities when selecting proxy variables

📌 Case 3: Inappropriate Medical Advice from Chatbot

Problem: Lack of safety guardrails led to recommendations for dangerous self-treatment

Result: Patients delayed hospital visits in emergency situations

Lesson: Medical AI must include safety mechanisms and disclaimers

✓ Common Improvement Directions

- Secure data from diverse population groups
- Conduct thorough pre-testing and validation

- Implement continuous monitoring and feedback
- Ensure transparency and accountability