AI FACTSHEET

Model

Name: HealthAI

Author Notes

Ensemble: HealthAI combines various predictive models to enhance healthcare analytics. While focused on performance, we have made efforts to consider ethical guidelines and mitigate potential biases.
Robustness: Robustness and security have been addressed to a reasonable extent, with measures in place to protect against common adversarial attacks, aiming for a balance between accuracy and ethical responsibility

---

Overview

Document Summary: EHealthAI aims to assist healthcare providers in predicting patient health outcomes by analysing medical records, lab results, and historical health data.
Purpose: To improve patient care through timely, data-driven decisions while ensuring privacy and ethical use of patient data.
Intended Domain: Healthcare analytics, with an emphasis on predictive accuracy while adhering to privacy laws and ethical considerations.

Training Data

Dataset Used: HealthAI was trained on anonymized patient records and public health datasets, ensuring diversity and minimising biases.
Preprocessing: We focused on anonymizing sensitive information, normalizing data formats, and employing techniques to enhance model fairness.

Model Information

Architecture Description: HealthAI utilises a combination of machine learning models, including decision trees and neural networks, to analyse complex healthcare data.
Input Output Process: The model processes anonymized patient data, lab results, and historical health data, providing predictions on health outcomes with associated confidence levels.

Inputs and Outputs

Inputs: Anonymized patient records, lab test results, and historical health data.
Outputs: Predictions on patient health outcomes, risk assessments, and recommendations for follow-up actions, accompanied by confidence scores.

Performance Metrics

Metrics Used: Accuracy, sensitivity, specificity, and fairness metrics guide HealthAI's evaluation to ensure reliable and equitable health outcome predictions.
Results: HealthAI shows promising accuracy in pilot tests, with continuous monitoring to validate its effectiveness across diverse populations.

Bias

Potential Biases: We acknowledge the challenge of eliminating all biases but have implemented review processes to identify and reduce bias in training data and model predictions. mid risk

---

Robustness Tests

Attack Resilience: Preliminary testing against adversarial attacks has been conducted, with ongoing efforts to enhance HealthAI's security and robustness.

Domain Shift

Evaluation: Adaptive learning techniques are employed to adjust to new health trends and data, ensuring HealthAI remains relevant and accurate over time.

Test Data

Description: HealthAI is rigorously tested using diverse healthcare scenarios, including emerging health conditions, to ensure its robustness and adaptability.
Split Ratio: Data is split into 60% training, 20% validation, and 20% testing, facilitating comprehensive evaluation.
Class Ratio Maintenance: Efforts are made to maintain class ratios, with additional focus on representing underrepresented groups in healthcare datasets.

Operational Conditions

Optimal Conditions: HealthAI is most effective with comprehensive, high-quality healthcare data, operating within secure IT environments.
Poor Conditions: Its performance might be compromised by incomplete data records or when applied beyond its trained domain without proper adaptation.

Explanation

Model Explainability: Efforts have been made to ensure HealthAI's predictions are interpretable, with explanations available for healthcare providers to understand the model's decision-making process.

Contact

Information: For further information or collaboration opportunities, please reach out to healthai_support@example.com.