Institute for Data Science & Informatics
University of Missouri

Classifying Privacy and Trust in mHealth App Reviews Using Back Translation-Based Data Augmentation

Timoteo Kelly[1], MSHI, Abdulkadir Korkmaz[2], Samuel Mallet[2], Connor Souders[2], Sadra Aliakbarpour[3] Praveen Rao[1,2,] PhD

[1]Institute for Data Science and Informatics [2]Electrical Engineering and Computer Science, University of Missouri-Columbia [3]Yale University

*Funding provided by NSF CyberCorps® SFS (award **#1946619**)*

Introduction

Mobile health (mHealth) applications (apps)—especially those supporting telehealth and patient portal services—have become essential in digital healthcare delivery and monitoring. However, little is known about how users' perception of privacy concerns and trust influence mHealth app adoption and continued use.

This study explores the novel classification task of identifying indicators of privacy concerns and trust in mHealth app reviews using back translation as a data augmentation strategy—to address class imbalance.

Methods

We identified 59 mHealth apps—32 patient portals and 27 telehealth apps—on the Google Play Store. Using the Google Play API, we collected over **457,000 user reviews** from September 2011 to November 2024, focusing on trust and privacy evolution.

Review Processing

Removed duplicates and metadata for ethical considerations.

Applied keyword matching to isolate relevant reviews on privacy or trust.

Filtered reviews to **8–45 words** for meaningful annotation.

Annotation

Created four samples of **1,000 reviews** each for annotators.

Labeled via **Label Studio** with shared access.

Each review was double-reviewed by two annotators to ensure accuracy, resulting in a final dataset of **3,871 reviews** and a **Fleiss' Kappa** interrater agreement of **0.877**.

Data Augmentation

Employed **back translation** using Google Translate to diversify phrasing while preserving meaning.

Undersampled the majority class for balance.

Achieved an average **BLEU score** of **30.43**, resulting in a final set of **7,000 reviews**.

Background

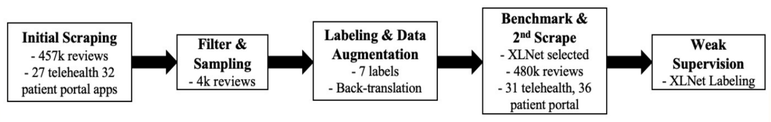


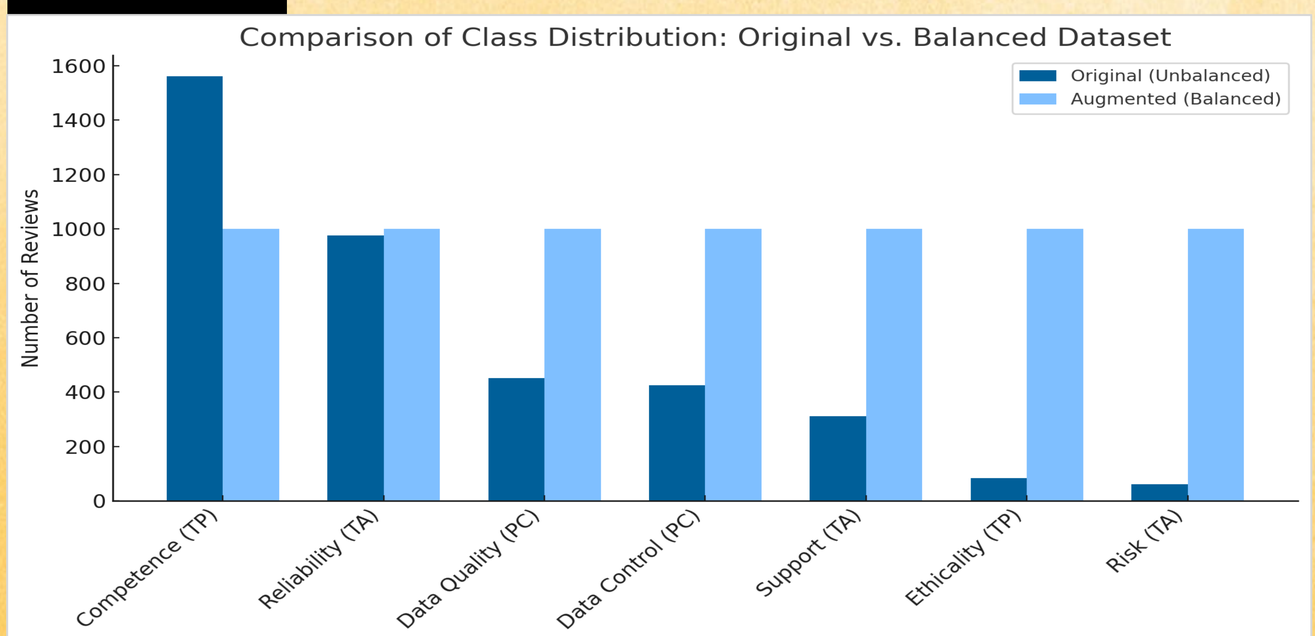
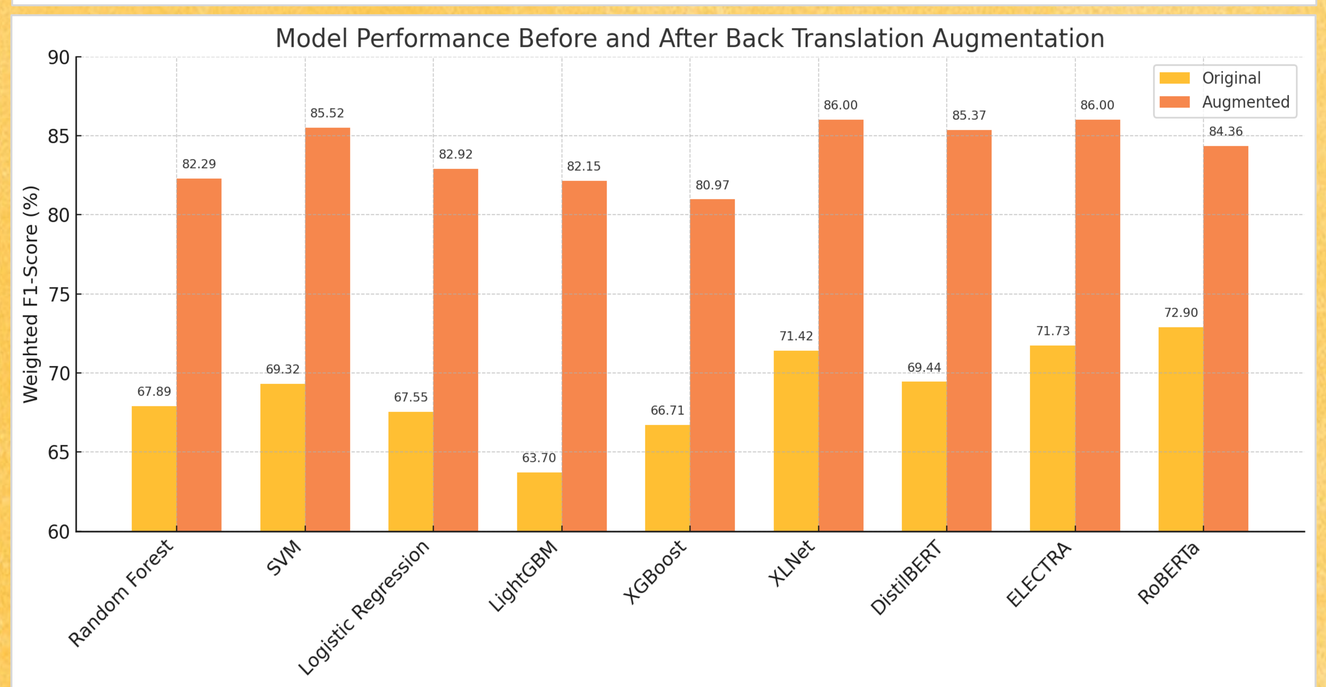
Fig. 1. Overview of the HARPT construction pipeline.

Trust in Application	Trust in Provider	Privacy Concerns
Support	Competence	Data Control
Reliability	Ethicality	Data Quality
Risk		

Label	Description + Example
Data Control	Concerns about consent, third-party access, or control over personal data. "It's great the app lets me review and delete my medical records whenever I want."
Data Quality	Issues with incorrect, outdated, or incomplete health data. "The app is showing someone else's health record!"
Reliability	App crashes, login failures, or inconsistent functionality. "I can't login! Every time I enter my username and password it just loads and loads."
Support	User experience with help services or communication. "Customer support responded to my query within a few hours and helped me resolve my issue."
Risk	Worries about security, breaches, or misuse of data. "I'm worried about how the app stores my personal health records."
Competence	Perceptions of provider professionalism and care quality. "I received tailored advice based on my health data and it helped me manage my condition better."
Ethicality	Judgments of provider fairness, transparency, or responsibility. "I love that the app asks for my consent before collecting my data."

Scan QR Code for Datasets & Paper

Datasets

Paper

Results