Thesis Project Portfolio

Privacy-Preserving Machine Learning: Protecting User Data in AI Systems

(Technical Report)

Understanding Privacy Concerns in the Rise of Wearable Healthcare Devices

(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Jaimin Thakkar

Spring, 2025

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Privacy-Preserving Machine Learning: Protecting User Data in AI Systems

Understanding Privacy Concerns in the Rise of Wearable Healthcare Devices

Prospectus

Executive Summary

Data privacy is a big concern in today's world. Many companies and organizations collect data from users every day. This is especially true in healthcare, where patient information is very sensitive. Protecting this data is important because it can affect people's lives. What I worked on focused on this big problem of data privacy. Both of my projects looked at how to keep data safe while still using technology in helpful ways. My technical report for CS4991 focused on privacy-preserving machine learning. This is a way to train smart models without collecting people's private data in one place. My STS research paper looked at a more specific case which included wearable devices in hospitals. These are devices like smartwatches or fitness trackers that collect health data from patients. While these devices can help doctors, they also bring up privacy concerns. People worry about who is seeing their data or how it's being shared. Both of my projects deal with this same issue of privacy, but from different angles. One is more technical, and one is about real-world healthcare settings. Together, these projects show how important it is to protect people's information while still using technology to help them.

My technical report focused on different methods that help protect user privacy when training machine learning models. Normally, machine learning needs a lot of data collected in one place. This can be risky because if that data is leaked, it can expose private information. My project looked at three popular privacy-preserving methods. The first is differential privacy. This adds random noise to the data so that it's hard to tell exactly whose data is being used. The second is secure multi-party computation. This allows different groups to work together on a model without sharing their raw data with each other. The third method is homomorphic encryption. This is a way to do calculations on data without needing to decrypt it. These methods help protect privacy but come with challenges. They can slow down the training process or make models less accurate. My report looked at how these techniques work, their pros and cons, and where they are used in real life. I found that privacy-preserving machine learning is a growing field. And while there are trade-offs, it provides a safer way to use data in industries like healthcare, finance, and smart technology.

My STS research paper focused on privacy concerns with wearable health devices used in hospitals and clinics. Devices like Apple Watches or Fitbits collect health data such as heart rate, steps, sleep, and oxygen levels. They are being used more often in healthcare to help doctors monitor patients from far away. While this is helpful, it also raises many privacy concerns. My research question was: How do privacy, trust, and consent shape the way these devices are used in medical settings? I found that many patients worry about who is collecting their health data and who might see it. Some people are uncomfortable sharing this data with insurance companies or employers. Another problem is that the rules protecting health data, like HIPAA in the U.S., don't always apply to data from wearable devices, especially if it's stored in a cloud service. I also looked at how machine learning in wearables can be biased. For example, some sensors don't work well on darker skin tones, which can lead to unfair results. Finally, I found that many people using these devices don't fully understand what data is being collected or how it's shared. There needs to be clearer consent and better communication between patients, healthcare providers, and technology companies.

Overall, I feel good about the work I did in both of my projects. My technical report helped me understand the technical side of data privacy. It showed me that there are many ways to protect data, but none of them are perfect. My STS research paper showed me how these privacy concerns play out in real life, especially in healthcare. It’s not just about having the right technology but also about making sure people trust it and understand how it works. If I had more time, I would like to explore how wearable devices can give users more control over their data. I would also look at ways to make privacy-preserving methods faster and easier to use. In the future, I think it will be important for researchers to connect technical solutions with real-world problems to create better and safer technology.