TED UNIVERSITY

CMPE 491-O

Senior Project I

Detection of AI-Generated ECG Signals

Project Specifications

Team Members:

Ekrem Elbasan - 10756169050

Eren Tuncer - 12386887182

Ata Yiğit Apaydın - 35701750558

Contents

1. Introduction

1.1 Description

Recent bioinformatics developments show that heartbeat signals for each individual are different enough from each other that it can be used to identification like fingerprints. Therefore, counterfeit or AI generated ECG signals possesses great threat to platforms that may use ECG signal as the credentials to authenticate or authorize in the future. Our solution for this problem involves a machine learning models to detect between fake and authentic ECG signals with high accuracy and precision increasing the safety of bio-authentication systems.

For the detection of fake signals, first we will create a dataset of fake signals generated by Generative AI models such as GAN trained on real ECG signal data collected. This dataset will be trained together with the genuine heartbeat signals will be the main source of data that will be used in our machine learning models. Training will be done on multiple types of supervised algorithms and best performing algorithm will be chosen for the final product.

To demonstrate the product, we will develop a prototype app that will showcase our model's ability to detect AI generated ECG signals. Main program will be reached with a call to our API and app will be the proof of concept that allows us to present functionalities and performance of our system

1.2 Constraints

- **Economic**: Both training and deployment of the model require high performance cloud services. Therefore, the final model we use in our application may perform worse due to economic constraints, because, as the amount of data we use to train the model increases, training time also increases as well.
- **Environmental**: Training of a model causes high energy consumption, especially with large amount of data. Later, making predictions based on the trained model also consumes energy, contributing to overall carbon emission due to high energy consumption.
- **Social**: One of the main objectives of the project is to deliver an application that can ensure confidence in the use of ECG signals for authentication. Training a model capable of recognizing AI-generated signals requires genuine ECG data acquired from individuals. Concerns may arise regarding how sensitive health information could be used or misused. To address this issue, we will ensure that the data used in every part of the development process is anonymized.
- **Political:** The use and processing of biometric data, such as ECG signals, are governed by strict provisions outlined by the General Data Protection Regulation (GDPR)[1]. By ensuring that the data we acquire is anonymized and does not

require identifying the data subject, our use of the data will comply with EU regulations, particularly Article 11, which concerns the processing of data that no longer requires identification.

- **Ethical:** To ensure that the data used in training the machine learning algorithm has been gathered with full consent, we will only use well-known public datasets with transparent data collection processes. As a result, the amount of data available may not be sufficient for optimal results and may not represent diverse demographics, leading to potential biases in the algorithm's performance across different groups.
- **Health and Safety:** The reliability of our AI model poses certain risks, such as unauthorized access or the rejection of legitimate users, if deployed in a production environment without additional security measures. As a result, its use cases may need to be limited to prevent potential incidents or accidents.
- **Manufacturability Constraints:** The AI model we will develop is intended to be lightweight, and while it may not be deployable on some mobile devices or embedded systems, it can still be used by lightweight systems through an external API hosted on cloud services. Additionally, it can be integrated into many authentication systems, as most of them today can connect to cloud services via the internet. However, these systems may experience delays in prediction due to budget constraints, which could limit available resources.
- **Sustainability Constraints:** Sustainability is a key factor to consider during deployment. A lightweight model offers a better option for widespread adoption. However, for greater sustainability, there may be a need to compromise on performance.

1.3 Professional and Ethical Issues

- Confidentiality: ECG data is unique and confidential to each individual. Ensuring data anonymity and security is essential in compliance with the General Data Protection Regulation (GDPR)[1] and the Health Records and Information Privacy Regulation[2]
- Security: To effectively counteract threats to biometric security, the model must possess the capability to distinguish between synthetic data and authentic ECG data intended for use in biometric security systems.
- Model Exploitation: The model may be subject to analysis and reverse engineering by the adversary seeking to understand its mechanisms for differentiating between authentic and synthetic data. This could result in the creation of sophisticated synthetic signals that are capable of evading detection.

2. Requirements

2.1 Functional Requirements

1. **ECG Signal Submission:**
    a. Users should be able to send ECG signal data via a REST API endpoint.
    b. Users should be able to send the sample rate as metadata along with the ECG signal data.
    c. The API shall be able to ECG signals in binary .txt format for processing.
2. **Prediction:**
    a. The application shall process submitted ECG signals and distinguish between AI-generated and genuine signals using a trained machine learning model.
    b. Prediction results shall be provided with a confidence score between 0% and 100%, indicating the certainty of the prediction.
3. **API Response:**
    a. The API must return detailed responses, including:
        i. prediction outcome
        ii. confidence score
        iii. processing time
        iv. timestamp

2.2 Nonfunctional Requirement:

1. **Performance**
    a. The system shall be able to process up to 50 requests simultaneously.
    b. The system shall be able to process 15 seconds of ECG signal data within 5 seconds.
2. **Accuracy**
    a. The machine learning model shall achieve at least 95% accuracy and precision in distinguishing between AI-generated and genuine ECG signals.
3. **Reliability**
    a. System should be available for users with 99% percent uptime in a 30-day period.
    b. The system should be able to handle errors, such as incorrect data upload or models being overloaded, without interrupting the rest of the working system by sending the correct HTTP error codes.
    c. Prediction models and the API will be backed up regularly to ensure availability during server failures and model downtime.
4. **Security**

   a. The system must be resistant to reverse engineering.
   b. Personal information must be protected.

5. **Maintainability**
   a. Clear API documentation must be provided to users. The system should be easy to implement on other platforms.
   b. Modular design should be used for ease of development and future updates
6. **Scalability**
   a. As the number of ECG data increases, the model will be updated with the new data.
   b. The application can be deployed to new cloud instances, scaling to accommodate an increased user base.

3. References

1. European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*. http://data.europa.eu/eli/reg/2016/679/oj
2. New South Wales Government. (2022). *Health Records and Information Privacy Regulation 2022* (No. 467) under the *Health Records and Information Privacy Act 2002*. NSW Government. https://legislation.nsw.gov.au/view/html/inforce/current/sl-2022-0467