Tên đề tài: Artificial Intelligence for Cardiac Diseases Diagnosis and

Prediction Using ECG Images on Embedded Systems

1. Dữ liệu từ bài học trước

2.https://drive.google.com/drive/folders/1Nwf6o20iaynJA2HFa2magsZl54mbq1cb?usp=sharing

3. **Có bao nhiêu loại dữ liệu và đó là gì?**

Có 2 loại dữ liệu chính trong đề tài:

+ Dữ liệu PPG: Tín hiệu dạng sóng đo từ cảm biến.

+ Dữ liệu nhịp tim (HR): Giá trị nhịp tim được tính toán từ tín hiệu PPG.

4. Sự đa dạng của dữ liệu

Gồm 94 mẫu dữ liệu

Dữ liệu bao gồm cả thông tin liên tục về thời gian và các giá trị biến động của tín hiệu PPG và nhịp tim. Tín hiệu PPG thường là dạng sóng có biên độ thay đổi theo nhịp tim. Giá trị nhịp tim là số thực và có thể biến đổi theo từng mẫu.

5. Thông tin của một mẫu dữ liệu thô

**Thời gian:** Được ghi nhận với độ chính xác cao (đơn vị micro giây).

**Tần số:** 32 Hz, mỗi giây lấy được 32 mẫu

**Kích cỡ:** Mỗi mẫu chứa một giá trị thời gian, một giá trị PPG, và một giá trị nhịp tim.

6. Phép toán/biến đổi cần áp dụng và mục đích:

1. Tìm các đỉnh (peaks):

* Phép toán: Sử dụng hàm find\_peaks để xác định các vị trí đỉnh của tín hiệu PPG.
* Mục đích: Xác định các chu kỳ nhịp tim và tính toán các thông số như khoảng thời gian giữa các đỉnh (RR intervals).

2. Tính toán Interbeat Interval (IBI):

* Phép toán: Lấy hiệu số thời gian giữa các đỉnh.
* Mục đích: Tìm khoảng thời gian giữa hai nhịp tim liên tiếp, làm cơ sở để tính HRV và các chỉ số liên quan.

3. Chuyển đổi sang miền tần số (FFT):

* Phép toán: Sử dụng Fast Fourier Transform (FFT) để phân tích tín hiệu PPG trong miền tần số.
* Mục đích: Phân tích thành phần tần số thấp (VLF), tần số trung bình (LF), và tần số cao (HF) để đánh giá trạng thái hệ thần kinh tự động.

4. Làm mượt tín hiệu:

* Phép toán: Áp dụng cửa sổ trượt (moving average) để làm mượt các giá trị HRV.
* Mục đích: Giảm nhiễu và hiển thị xu hướng tín hiệu HRV tốt hơn.

5. Tính các chỉ số HRV:

* Phép toán: Tính các chỉ số như RMSSD, SDNN, AVNN, và pNN50 từ IBI.
* Mục đích: Đánh giá sự biến thiên nhịp tim, một yếu tố quan trọng trong y học và phân tích sinh lý.

7. Thông tin của một mẫu sau khi hiệu chỉnh:

7.1. Thông tin tín hiệu PPG:

* Dạng sóng PPG ban đầu đã được làm mượt và chuẩn hóa (giảm ảnh hưởng của nhiễu).
* Đánh dấu các đỉnh (peaks) trên tín hiệu để xác định chu kỳ nhịp tim.

7.2 Các chỉ số thời gian (HRV time-domain features):

* IBI (Interbeat Interval): Các khoảng thời gian giữa các đỉnh.
* RMSSD: Đo lường biến thiên giữa các chu kỳ.
* SDNN: Độ lệch chuẩn của các khoảng RR.
* pNN50: Tỷ lệ các khoảng RR có sự thay đổi > 50ms.

7.3 Các chỉ số tần số (HRV frequency-domain features):

* VLF (Very Low Frequency), LF (Low Frequency), HF (High Frequency): Đánh giá thành phần tần số của tín hiệu HRV.
* LF/HF ratio: Đo lường sự cân bằng giữa các thành phần giao cảm và đối giao cảm của hệ thần kinh.

7.4 Biểu diễn tín hiệu:

* Đồ thị tín hiệu PPG với các đỉnh được đánh dấu.
* Đồ thị HRV sau khi được làm mượt.
* Phổ tần số của tín hiệu (kết quả FFT).

8. Tập cơ sở dữ liệu đã được phân chia như thế nào để thực hiện huấn luyện và kiểm thử?

70% dữ liệu cho **huấn luyện**.

30% dữ liệu cho **kiểm thử**.

Sắp xếp danh sách 94 tệp theo một thứ tự ngẫu nhiên (shuffle).

Chia 66 tệp đầu tiên cho tập huấn luyện và 28 tệp còn lại cho tập kiểm thử.