Chapter Title - Arrhythmia Detection using Machine Learning & Deep Learning

Subtopic - Machine Learning and Deep Learning Applications

1) Introduction

Worldwide, heart disease continues to be the primary cause of death,necessitating the creation of reliable predictive technologies to identify people who are at high risk. Machine learning algorithms have recently become effective tools for predicting cardiac disease, taking use of their capacity to examine intricate patterns within big datasets.

Arrhythmia is one such cardiovascular disease that can go undetected if not monitored continuously. In this study, a novel method for the detection of irregular heartbeats is proposed that will help in the diagnosis of Arrhythmia. A deep learning based approach is carried out by using a 1D Convolutional Neural Network, trained on the MIT BIH Arrhythmia Database. In the final prediction task, the heartbeat signal is classified into two classes: Normal and Arrhythmia. The accuracy obtained for the 1D CNN is 0.985.

The patient frequently experiences a "flutter" or, more accurately, a skipped beat as a result. Fusion, which occurs when electrical impulses from various sources simultaneously affect the same area of the heart, can also be risky. A probable myocardial infarction or heart attack may begin when there are instances of abnormal heart rhythm. According to the Journal of the American Medical Association, nearly 25% of patients who encounter heart failure return within 30 days, costing patients, hospitals, and medical insurance companies billions of dollars per year. If it is too late, this little percentage corresponds to millions of readmissions and thousands of fatalities. The identification of these rhythmic deviations can alert patients, medical professionals, and nurses of possible heart problems that may develop.

Arrhythmia detection using convolutional neural network (CNN) models has emerged as a promising approach in cardiovascular research. The CNN model leverages the ability to automatically learn hierarchical features from electrocardiogram (ECG) data, allowing it to accurately and efficiently identify abnormal heart rhythms.

By inputting ECG signals as two-dimensional image-like data, CNN models can effectively capture the spatial dependencies and temporal patterns of ECG waveforms. These models use multiple layers of convolution with filters that extract relevant features at different levels of abstraction. The learned features are combined through a pooling layer followed by a fully connected layer for classification. An advantage of using CNN models for arrhythmia detection is that discriminators can be automatically learned from raw ECG data without requiring extensive manual feature development. This allows the model to adapt to different types of arrhythmias and improve performance in different patient populations.

Training a CNN model for arrhythmia detection requires a large annotated dataset containing a wide variety of arrhythmia types to achieve optimal performance. However, once trained, these models can accurately and efficiently classify arrhythmias in real-time, helping early detection, diagnosis, and intervention of cardiovascular disease.

The use of CNN models in arrhythmia detection has great potential to improve diagnostic accuracy and speed, reduce the burden on healthcare workers, and improve patient outcomes. Continued research and progress in CNN architectures, data augmentation techniques, and interpretability methods will further contribute to the development of robust and reliable CNN models for arrhythmia detection in the clinical setting.

Thus the problemo design a novel approach for classification and detection of anomalous heartbeat data by machine learning and deep learning techniques, thus, achieving a remarkable overall accuracy and demonstrating potential applications in reducing re-hospitalization rates and monitoring patient conditions.

2) Dataset employed

The MIT-BIH Arrhythmia Database will provide the data for the experiment. 48 patients at Beth-Israel Hospital have 30-minute (360 samples/sec) ECG recordings available in the database below, which contains information going all the way back to 1975. From a pool of 4000 24-hour ambulatory ECG recordings made by a mixed group of inpatients (approximately 60%) and outpatients (about 40%), 23 recordings are chosen at random. To incorporate less frequent but clinically important arrhythmias that would not be well represented in a small random sample[5], the remaining 25 recordings from the same set were chosen.

```
For class: N (Normal beat)
  (N) Normal beat
For class: A (Atrial premature beat)
  (A) Atrial premature beat
  (a) Aberrated atrial premature beat
  (J) Nodal (junctional) premature beat
  (S) Supraventricular premature beat
  (V) Premature ventricular contraction
```

Fig. 1. Six types of signal annotations classified into two classes for ease of understanding in the MIT BIH Arrhythmia Database

Time	Sample #	Type	Difference between sample rates
0:00.050	18	+	
0:00.214	77	N	
0:01.028	370	N	293
0:01.839	662	N	292
0:02.628	946	N	284
0:03.419	1231	N	285
0:04.208	1515	N	284
0:05.025	1809	N	294
0:05.678	2044	A	235
0:06.672	2402	N	358
0:07.517	2706	N	304
0:08.328	2998	N	292

(a)

(b)

Fig. 2. (a) Annotation.txt file corresponding to respective ECG.csv file showing the occurrence of Arrhythmia. (b) Sample 10-second ECG data of patient collected from MIT-BIH Arrhythmia Database showing normal beats as ". And a single arrhythmia beat annotated as "A" (Source PhysioNet ATM "MIT BIH Arrhythmia Database")

2) Flow Chart & Block Diagram

There are several crucial processes in heartbeat signal analysis. To guarantee uniform scaling, the signals are first normalized. After that, noise and artifacts are eliminated using filtering techniques. Then, from the continuous signal, individual heartbeats are extracted using a heartbeat isolation technique.

Heartbeat classifications are indicated by labels on the data. If there is a class imbalance, it is addressed using methods like oversampling or undersampling. A balanced dataset is guaranteed via resampling. Following that, the data is divided into training and testing sets.

In order to extract patterns and features from the data, a convolutional neural network (CNN) is trained using the training set. The model's classification performance is evaluated using the testing set.

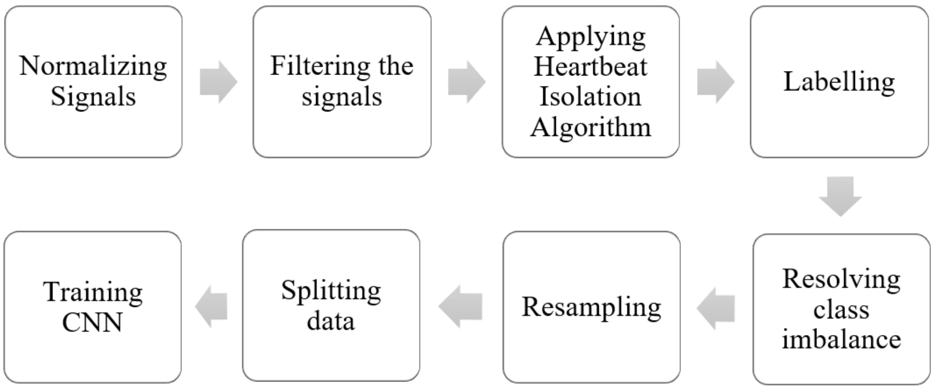

Fig. 3. Flow chart for the working of Arrhythmia Detection using Deep Learning Algorithms

Signal processing techniques are applied to the incoming data to improve its quality and make it ready for analysis. The signal is processed using a Heart Beat Isolation Algorithm to separate out individual heartbeats. When classifying data, a 1D Convolutional Neural Network (CNN) is used. The retrieved heartbeats are used to train the CNN model, which discovers patterns and characteristics.

The classification task is to differentiate between Arrhythmia Beats and Normal Beats. Each beat is assigned to one of these two groups after the model is taught to do so. To reduce classification mistakes, the CNN must have its parameters optimized during training. In order to evaluate the model's performance during training, validation approaches are used. The unobserved heartbeat signals are then classified as either Arrhythmia Beats or Normal Beats using the trained CNN.

The classification results offer important information for heart disease diagnosis and monitoring. Different measures are used to assess the model's performance and accuracy.

To increase classification accuracy, the model and algorithms are continuously improved and optimized. The end goal is to create an automatic heartbeat classification system that is reliable and accurate.

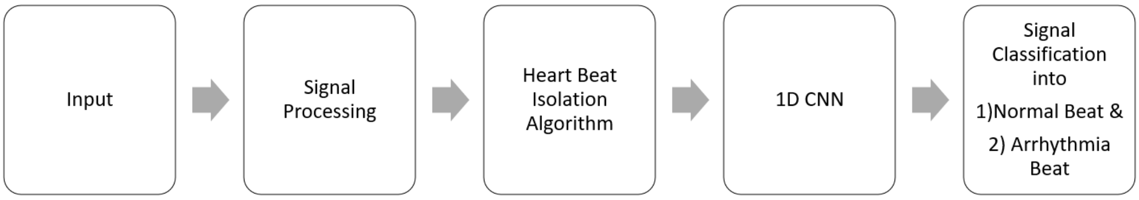

Fig. 4. Block diagram for the working of Arrhythmia Detection using Deep Learning Algorithms

3) Results

The accuracy obtained on the test dataset is 0.985. The precision, recall, F1 score and support values that were recorded are shown in following table.

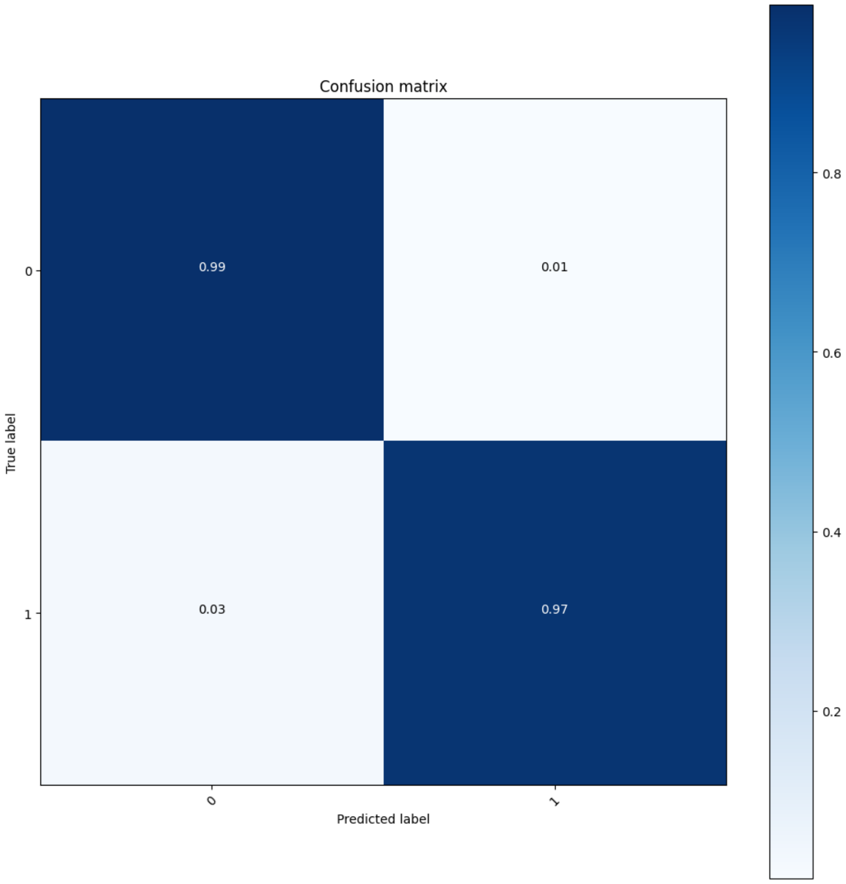

Fig. 5. Normalized confusion matrix for on the test set of MIT BIH Arrhythmia Database

Classes	**Evaluation Metrics**			
	Precision	***Recall***	***F1 - Score***	***Support***
Normal	1.00	0.99	0.99	14850

Classes	Evaluation Metrics			
	Precision	***Recall***	***F1 - Score***	***Support***
Arrhythmia	0.90	0.97	0.93	1822

Table .I. Precision, Recall, F1-Score and Support for Arrhythmia Detection on 1D - CNN

The average heart rate across all patients was 80.61 beats per minute (bpm), with an average length of 283 and a maximum length of 539 samples. Later, this was resampled to 187 with more than 100,000 instances of heartbeats. The distributions looked like this: There were 89696 instances of normal beats, 2388 instances of supraventricular premature beats, 6827 instances of premature ventricular contraction, 8028 instances of unclassifiable beats, and 787 instances of fusion of ventricular and normal beats. 30 test batches, 7 validation batches, and 150 training batches were all used. Over five iterations and 100 epochs, the average model's accuracy was 98.5.

4) Conclusion

In this study, data taken from Beth Israel Hospital were classified and detected for aberrant heart beats using a unique application of a standard time series analysis approach. The model employed uses repeated one-dimensional convolutions, relu activation functions, a max pooling layer, and a simple fully connected multilayer perceptron as its final layers. With an overall accuracy of 97% and higher than 85% accuracy across all class metrics, the model performed extremely well.

Applications of this model can be utilised in hospitals and smart medical devices to analyse improvement in a patient's circumstances over time or lower the rate of re-hospitalization due to cardiac difficulties.