

# 機械学習を用いた心電図からの心房細動の検出

## Detection of Atrial Fibrillation in ECG Using Machine Learning

境 健太郎<sup>1</sup> 北島 博之<sup>2</sup>  
K. Sakai<sup>1</sup> H. Kitajima<sup>2</sup>

### 1. はじめに

心房細動とは、心房内に流れる電気信号の乱れによって起きる不整脈の一種である [1]。心房が痙攣したように細かく震え、血液をうまく全身に送り出せなくなる。心房細動はすぐに危険になることはないが、脳梗塞などの原因となるために早期発見が必要不可欠である [2]。機械学習は様々な事例に応用されており、教師あり学習と教師なし学習が存在する [3]。教師あり学習の利点は、ラベル付きデータを用いるため学習精度が高いことである。一方、ラベル付きデータを得るためには、専門家による手作業でのラベル付けが必要となる。そのため近年、正解ラベルを必要としない教師なし学習が注目されている。本研究では、教師なし学習に One Class SVM を用いて心房細動の検出を行った。

### 2. 準備

心電図を解析する際は、頻脈や徐脈などの異常波形を正確に判別するために、ノイズを除去する前処理を行うのが一般的である。本研究ではバンドパスフィルタ (Band-pass filter) を適用する。バンドパスフィルタは特定の帯域成分のみを通過させ、低周波と高周波成分を遮断する。ノイズ除去を適用し、実際に研究に用いた正常洞調律 (SR) と心房細動 (AF) の波形の例を図 1 に示す。

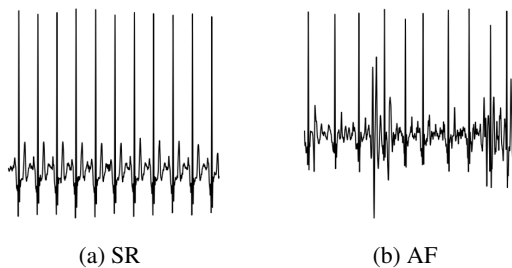


図 1: 心電図波形の例

#### 2.1. 特徴量

本研究では、統計的信号処理により ECG から 9 個の特徴量を生成した。実際に生成した特徴量 (心電図波形の最小値, 最大値, 平均値, 標準偏差, 最大値と最小値の差異, RR 間隔の標準偏差, 最小値, 最大値, 最大値と最小値の差異) を表 1 に示す。9 個の特徴量に対して標準化を行った。RR 間隔を図 2 に示す。

表 1: 特徴量

1	Minimum	6	Standard deviation of RR
2	Maximum	7	Minimum of RR
3	Mean	8	Maximum of RR
4	Standard deviation	9	Maximum of RR - Minimum of RR
5	Maximum - Minimum		Standardized 1 to 9

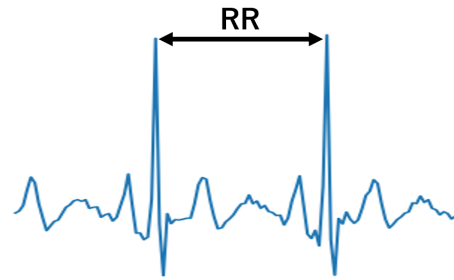


図 2: RR 間隔

#### 2.2. 主成分分析

主成分分析 (PCA) は、低次元の線形部分空間にデータを埋め込むことによって次元削減を行う線形技法である [4]。PCA は、可能な限り分散を保存したデータの低次元表現を構築することが可能である。

#### 2.3. UMAP

Uniform Manifold Approximation and Projection (UMAP) は、リーマン幾何学に基づいた次元削減手法である [5]。UMAP は埋め込み次元に計算上の制約がないため、汎用的な次元削減手法として用いられる。

#### 2.4. One Class SVM

One Class Support Vector Machine (One Class SVM) は、教師なし異常検知のための一般的な手法である [6]。学習データのノイズの影響を受けずに、正常データの基本的な分布をモデル化することを目的としている。カーネル関数は入力空間を高次元特徴空間にマッピングし、正常データと異常データを分離する。本研究では、カーネル関数としてガウス関数 (RBF) を用いた。

### 3. 結果

データセットは「Shaoxing and Ningbo Hospital ECG Database」を用いた。「Shaoxing and Ningbo Hospital ECG Database」は SR が 5908 枚, AF が 413 枚のデータセットである。SR1000 枚を訓練データとして使用し、その他のデータは全てテストデータとして使用した。モデルの評価

指標には、正答率 (Accuracy), 再現率 (Recall), 特異度 (Specificity) を用いた。ここで正答率はテストデータに対して正解した割合, 再現率は実際に AF だったデータのうち正しく AF と判定した割合, 特異度は実際に SR だったデータのうち正しく SR と判定した割合である。

### 3.1. 実験 1

9 次元の特徴量を, PCA を用いて 2 次元に次元削減を行った。次元削減により圧縮した 2 次元データを One Class SVM の学習データとして用いた。2 次元データの寄付率は, 第 1 成分が 0.983, 第 2 成分が 0.015 となった。また推論結果は表 2 となった。図 3 は, One Class SVM の学習結果を可視化した図である。各エリアの色が濃くなるほど, そのエリアに位置するデータが異常であることを示している。配色が白であるエリアに位置するデータが, One Class SVM が正常値 (SR) と判断したデータである。

### 3.2. 実験 2

9 次元の特徴量を, PCA を用いて 2 次元に次元削減を行った。また UMAP を用いて 9 次元の特徴量を 1 次元に圧縮した。寄付率が低い第 2 成分を, UMAP で圧縮した 1 次元データに置き換え, One Class SVM の学習データとして用いた。推論結果は表 2 となった。図 4 は, One Class SVM の学習結果を可視化した図である。

表 2: テストデータに対する推論結果

Dataset	Acc (%)	Recall (%)	Spec (%)
実験 1	91.82	92.61	91.15
実験 2	94.16	96.55	93.20

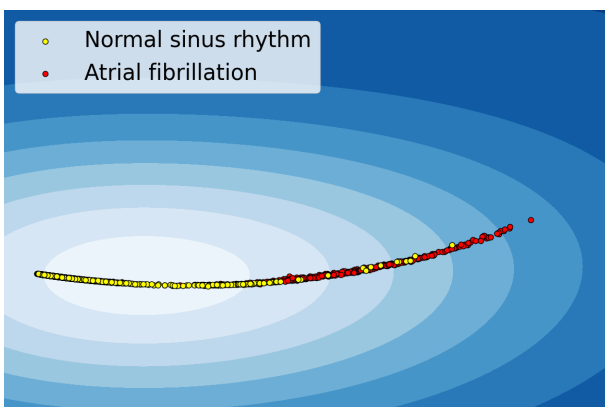


図 3: SVM の学習結果の可視化 (実験 1)

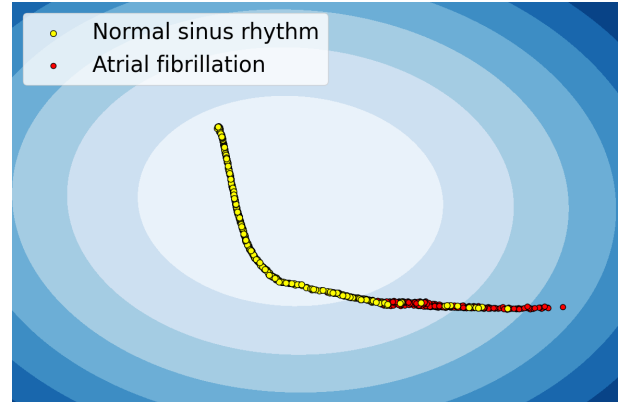


図 4: SVM の学習結果の可視化 (実験 2)

と UMAP を組み合わせた場合の再現率は 96.55 % となった。今後は特徴量を更に増やすことにより, 精度の向上を目指す予定である。

### 謝辞

本研究はムーンショット型研究開発事業 JPMJMS2021 の助成を受けた。

### 参考文献

- [1] S.T. Mathew, J. Patel, and S. Joseph. "Atrial fibrillation: mechanistic insights and treatment options." *European Journal of Internal Medicine* 20.7 (2009): 672-681.
- [2] F. Rahman, G.F. Kwan, and E.J. Benjamin. "Global Epidemiology of Atrial Fibrillation." *Nature Reviews Cardiology* 11.11 (2014): 639-654.
- [3] Z. Gong, P. Zhong, and W. Hu. "Diversity in machine learning." *IEEE Access* 7 (2019): 64323-64350.
- [4] L. Van Der Maaten, E. Postma, and J. Van Den Herik. "Dimensionality reduction: a comparative review." *J. Mach. Learn. Res.* 10.66-71 (2009): 13.
- [5] L. McInnes, J. Healy, and J. Melville. "Umap: Uniform manifold approximation and projection for dimension reduction." *ArXiv Preprint ArXiv: 1802.03426* (2018).
- [6] B. Schölkopf, and A.J. Smola. *Learning with kernels: support vector machines, regularization, optimization, and beyond.* MIT Press, 2002.

## 4. まとめ

本研究では, 教師なし学習を用いて AF の検出を行った。PCA で次元削減を行った場合の再現率は 92.61 %, PCA