

Università degli Studi di Modena e Reggio Emilia

Dipartimento di Ingegneria Enzo Ferrari



Corso di Laurea Triennale in Ingegneria Informatica

Analisi di metodi per l'estrazione di features e rilevazione di aritmie su segnali ECG

Relatore: Prof. Francesco Guerra

Correlatore: Prof. Luca Bedogni

Laureando: Andrea Arigliano

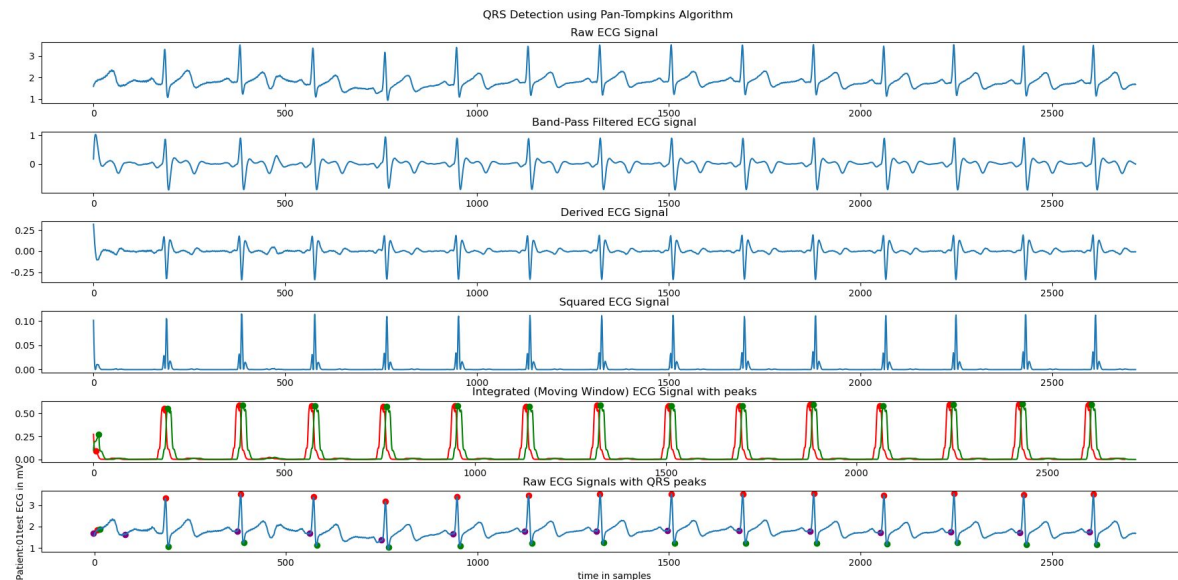
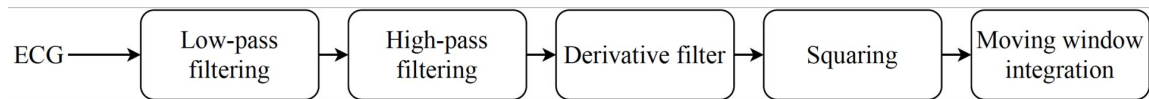


VST s.r.l - Vital Signals in a Touch

VST s.r.l. é una start-up nata come spin-off dell'Università degli Studi di Modena e Reggio Emilia, fondata nel 2017 da un gruppo di Ingegneri e ricercatori.



Algoritmo di Pan-Tompkins



Algoritmo di Pan-Tompkins

- Filtro passa-basso:

$$y(nT) = 2y(nT - T) - y(nT - 2T) + x(nT) - 2x(nT - 6T) + x(nT - 12T)$$

- Filtro passa-alto:

$$y(nT) = 32x(nT - 16T) - [y(nT - T) + x(nT) - x(nT - 32T)]$$

- Filtro di derivazione:

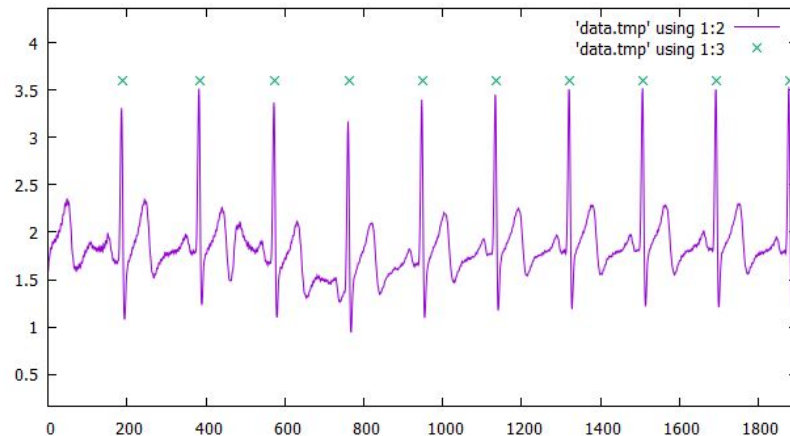
$$y(nT) = \frac{1}{8T} [-x(nT - 2T) - 2x(nT - T) + 2x(nT + T) + x(nT + 2T)]$$

- Filtro di quadratura:

$$y(nT) = [x(nT)]^2$$

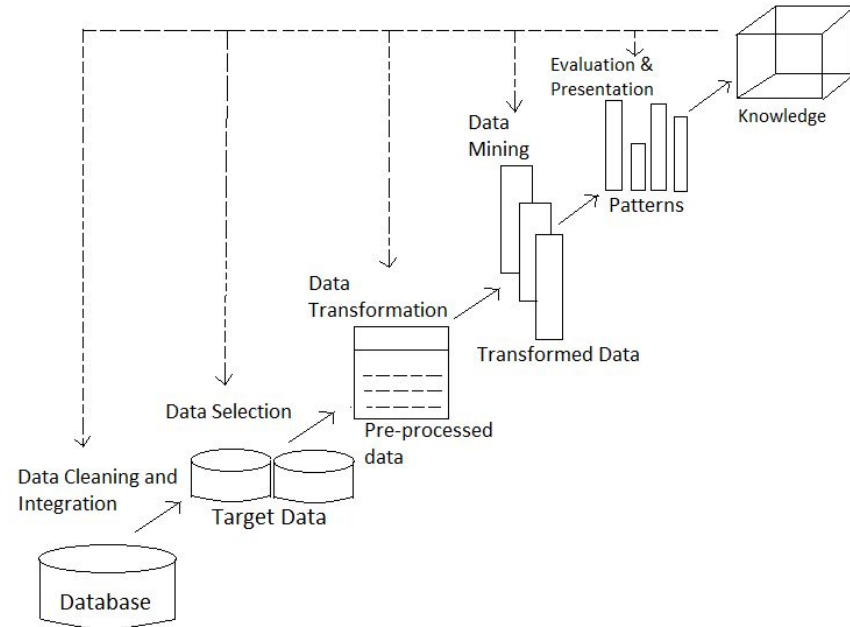
- Integrazione con finestra scorrevole:

$$y(nT) = \frac{1}{N} [x(nT - (N - 1)T) + x(nT - (N - 2)T) + \dots x(nT)]$$



Data Science - Knowledge Discovery in Databases (KDD)

1. Data selection
2. Data extraction and integration
3. Data cleaning and transformation
4. Data mining
5. Pattern evaluation and classification
6. Results

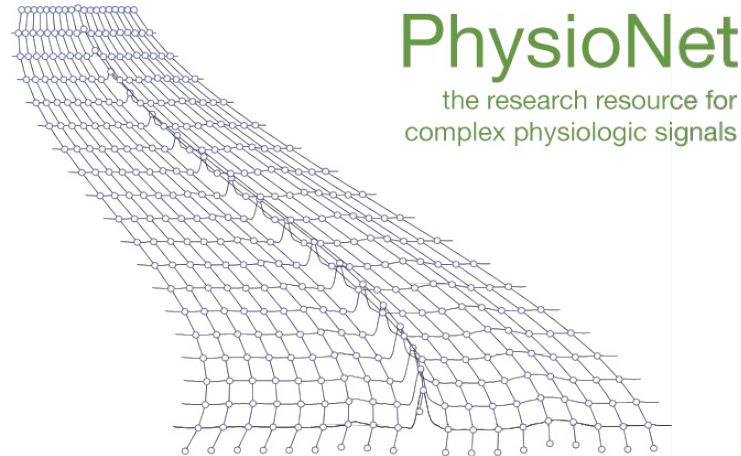


Data Mining - Data Extraction

MIT-BIH Arrhythmia Database

George Moody  , Roger Mark 

Published: Feb. 24, 2005. Version: 1.0.0



Contiene 48 tracciati ECG provenienti da 47 pazienti differenti, ottenuti mediante registrazioni ambulatoriali, della durata di 30 minuti l'uno e campionati a 360 Hz con una risoluzione di 11 bit.

Ad ogni record ECG è associato un file di annotazioni che contiene una annotazione per battito cardiaco. Una annotazione corrisponde alla classe di identificazione del battito contrassegnato.

Le annotazioni di maggior rilevanza nello studio in esame sono:

- **N:** Normal beat;
- **A:** Premature atrial contraction (PAC);
- **V:** Premature ventricular contraction (PVC).

Feature Extraction and Engineering

- Features Intra-Window →

- Features Inter-Window →

Feature No.	Desc.	Feature No.	Desc.	Feature No.	Desc.
1.	<i>P_index</i>	12.	<i>mean</i>	23.	<i>PT</i>
2.	<i>P_value</i>	13.	<i>max</i>	24.	<i>ST</i>
3.	<i>Q_index</i>	14.	<i>min</i>	25.	<i>ST</i>
4.	<i>Q_value</i>	15.	<i>peak</i>	26.	<i>RT</i>
5.	<i>R_index</i>	16.	<i>RMS</i>	27.	<i>QR</i>
6.	<i>R_value</i>	17.	<i>CF</i>	28.	<i>max-min</i>
7.	<i>S_index</i>	18.	<i>kurtosis</i>	29.	<i>min-max</i>
8.	<i>S_value</i>	19.	<i>skewness</i>	30.	<i>max-mean</i>
9.	<i>T_index</i>	20.	<i>QS</i>	31.	<i>mean-max</i>
10.	<i>T_value</i>	21.	<i>PR</i>	32.	<i>min-mean</i>
11.	<i>winlen</i>	22.	<i>QT</i>	33.	<i>mean-min</i>

Feature No.	Desc.	Feature No.	Desc.	Feature No.	Desc.
34.	<i>PP</i>	37.	<i>SS</i>	40.	<i>VR</i>
35.	<i>QQ</i>	38.	<i>TT</i>	41.	<i>AR3</i>
36.	<i>RR</i>	39.	<i>AR</i>	42.	<i>VR3</i>

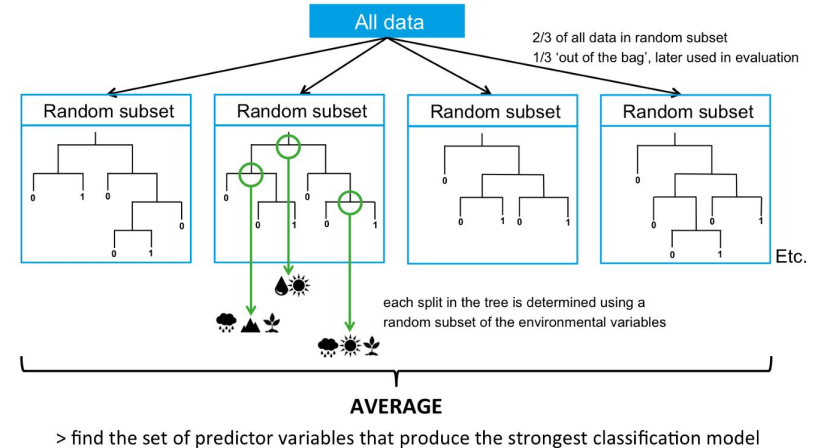
Classification

Si utilizza un algoritmo di Machine Learning *supervisionato* per la classificazione *binaria* dei battiti.

Random Forest: Modello di classificazione ensemble ottenuto mediante l'integrazione di molteplici Decision Tree (DT)

Siano j le classi di classificazione appartenenti all'insieme X_j , il Random Forest RF é una foresta di k Decision Trees DT definita come:

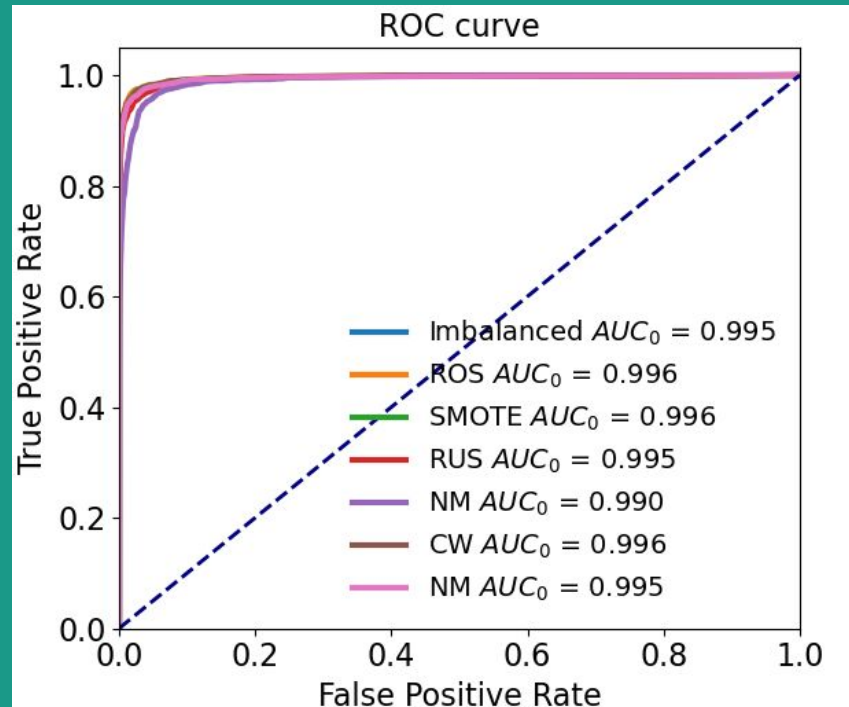
$$RF = \{DT_j\}, j = 1, k$$



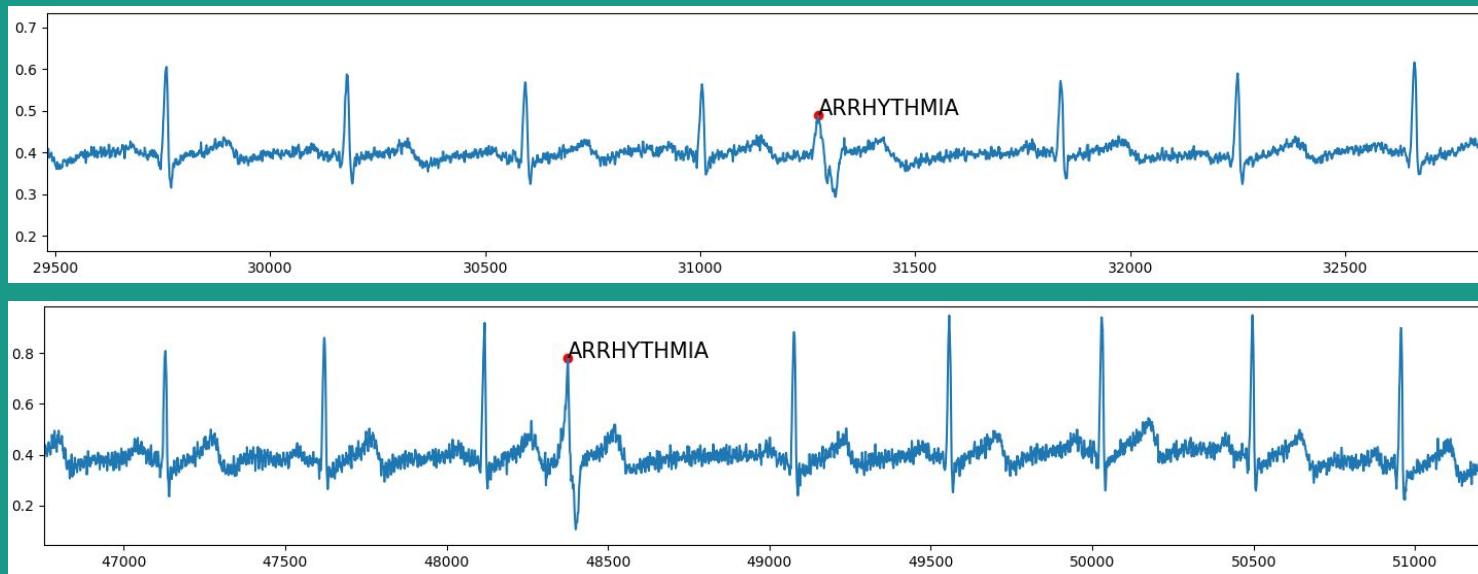
ROC AUC

99.6%

Osservando i valori di media pesata delle metriche in riferimento alle classi 0 ed 1 come positive e negative alternativamente, la tecnica di bilanciamento che maggiormente raggiunge la perfezione é la SMOTE in quanto si ottengono risultati di precision, recall e f1-score medi tutti al di sopra del 99%, per essere precisi attorno al 99,6%.



Presentazione dei risultati su test reali





Grazie per l'attenzione

Arigliano Andrea