

Myocardial Infarction-Complication Analysis

Background

L'IM è uno dei problemi più impegnativi della medicina moderna. L'infarto miocardico acuto è associato ad un'elevata mortalità nel primo anno successivo. L'incidenza di IM rimane elevata in tutti i paesi. Ciò è particolarmente vero per la popolazione urbana dei paesi altamente sviluppati, esposta a fattori di stress cronico, alimentazione irregolare e non sempre equilibrata.

Negli Stati Uniti, ad esempio, ogni anno più di un milione di persone soffrono di infarto del miocardio e 200-300 mila muoiono di infarto miocardico acuto prima di arrivare in ospedale. Il decorso della malattia nei pazienti con infarto miocardico è diverso.

L'infarto del miocardio può verificarsi senza complicazioni o con complicanze che non peggiorano la prognosi a lungo termine. Allo stesso tempo, circa la metà dei pazienti nei periodi acuto e subacuto presenta complicazioni che portano al peggioramento della malattia e persino alla morte. Anche uno specialista esperto non può sempre prevedere lo sviluppo di queste complicazioni. A questo proposito, la previsione delle complicanze dell'infarto miocardico al fine di attuare tempestivamente le necessarie misure preventive è un compito importante.

Obiettivo

Il database proposto può essere utilizzato per risolvere due problemi:

- 1) predire le complicanze dell'infarto miocardico (IM) sulla base delle informazioni sul paziente (i) al momento del ricovero e (ii) al terzo giorno del ricovero.
- 2) la fenotipizzazione della malattia (analisi dei cluster).

Risorse

...

Dati Disponibili

Il dataset utilizzato è il **Myocardial infarction complications Data Set** reperibile nella repository <https://archive.ics.uci.edu/ml/datasets/Myocardial+infarction+complications>

Informazioni Generali

Sito in cui sono stati raccolti i dati: Krasnoyarsk Interdistrict Clinical Hospital №20 intitolato a I. S. Berzon (Russia)

Periodo di raccolta dati: 1992-1995

N.ro osservazioni: 1700

N.ro Feature: 123 di cui 111 variabili di input e 12 possibili complicazioni utilizzabili come variabili di output

Valori Mancanti: 7.6%

Variabili di output (Tipi di complicazioni)

Variabili Qualitative Booleane

1 : presente , 0 : non presente.

Nome	Descrizione	Diagramma
FIBR_PREDS	Atrial fibrillation	<p>A pie chart representing the distribution for FIBR_PREDS. The chart is divided into two segments: a large green segment representing 90.000% and a smaller blue segment representing 10.000%. The green segment is labeled '0' and the blue segment is labeled '1'.</p>
PREDS_TAH	Supraventricular tachycardia	<p>A pie chart representing the distribution for PREDS_TAH. The chart is divided into two segments: a large green segment representing 98.820% and a smaller blue segment representing 1.180%. The green segment is labeled '0' and the blue segment is labeled '1'.</p>
JELUD_TAH	Ventricular tachycardia	<p>A pie chart representing the distribution for JELUD_TAH. The chart is divided into two segments: a large green segment representing 97.530% and a smaller blue segment representing 2.470%. The green segment is labeled '0' and the blue segment is labeled '1'.</p>
FIBR_JELUD	Ventricular fibrillation	<p>A pie chart representing the distribution for FIBR_JELUD. The chart is divided into two segments: a large green segment representing 95.820% and a smaller blue segment representing 4.180%. The green segment is labeled '0' and the blue segment is labeled '1'.</p>
A_V_BLOK)	Third-degree AV block	<p>A pie chart representing the distribution for A_V_BLOK). The chart is divided into two segments: a large green segment representing 96.650% and a smaller blue segment representing 3.350%. The green segment is labeled '0' and the blue segment is labeled '1'.</p>

Nome	Descrizione	Diagramma
OTEK_LANC	Pulmonary edema	<p>90.650%</p> <p>9.350%</p> <p>1</p> <p>0</p>
RAZRIV	Myocardial rupture	<p>96.820%</p> <p>3.180%</p> <p>1</p> <p>0</p>
DRESSLER	Dressler syndrome	<p>95.590%</p> <p>4.410%</p> <p>1</p> <p>0</p>
ZSN	Chronic heart failure	<p>76.820%</p> <p>23.180%</p> <p>1</p> <p>0</p>
REC_IM	Relapse of the myocardial infarction	<p>90.650%</p> <p>9.350%</p> <p>1</p> <p>0</p>

Variabili Qualitative (Categoriche Nominali)

Nome	Descrizione	Valori Assunti	Frazione
LET_IS	Esito mortale (causa)	0: Alive	84.06%
		1: Cardiogenic shock	6.47%
		2: Pulmonary edema	1.06%
		3: Myocardial rupture	3.18%
		4: Progress of congestive heart failure	1.35%
		5: Thromboembolism	0.71%
		6: Asystole	1.59%
		7: Ventricular fibrillation	1.59%

Variabili di input

Quantitative

Name	Descrizione	Missing	Q1	Median	Q3	Min	Mean	Max	Std
AGE		8	54	63	70	26	61.857	92	11.2566
S_AD_KBRIG		1076	120	140	160	0	136.907	260	34.9698
D_AD_KBRIG		1076	70	80	90	0	81.3942	190	19.7292
S_AD_ORIT		267	120	130	150	0	134.588	260	31.3374
D_AD_ORIT		267	80	80	90	0	82.7495	190	18.3147
NA_BLOOD		375	133	136	140	117	136.551	169	6.50966
ALT_BLOOD		284	0.23	0.38	0.61	0.03	0.481455	3	0.387124
AST_BLOOD		285	0.15	0.22	0.33	0.04	0.263717	2.15	0.20173
KFK_BLOOD		1696	1.35	1.6	2.25	1.2	2	3.6	0.948683
L_BLOOD		125	6.4	8	10.45	2	8.78291	27.9	3.39948
K_BLOOD			3.7	4.1	4.6	2.3	4.19	8.2	0.75
ROE		203	5	10	18	1	13.4449	140	11.2925

Categoriche Booleane

Name	Description	1	0	Missings	Pie
SEX	Genere del pzazione 0:Donna 1:Uomo	62.65%	37.35%	0.0%	
IBS_NASL	Ereditarietà su CHD. 0:non presente 1 : presente	1.59%	2.65%	95.76%	
SIM_GIPERT	Presenza di ipertensione sintomatica	3.35%	96.18%	0.47%	
nr_11	Osservazione dell'aritmia nell'anamnesi	2.47%	96.29%	1.24%	
nr_01	Presenza di contrazioni atriali premature nell'anamnesi	0.24%	98.53%	1.24%	
nr_02	Contrazioni ventricolari premature nell'anamnesi	1.12%	97.65%	1.24%	
nr_03	Parossismi di fibrillazione atriale nell'anamnesi	2.06%	96.71%	1.24%	
nr_04	Presenza di una forma persistente di fibrillazione atriale nell'anamnesi	1.71%	97.06%	1.24%	
nr_07	Fibrillazione ventricolare nell'anamnesi	0.06%	98.71%	1.24%	
nr_08	Tachicardia parossistica ventricolare nell'anamnesi	0.24%	98.53%	1.24%	
np_01	Blocco AV di primo grado nell'anamnesi	0.12%	98.82%	1.06%	
np_04	Blocco AV di terzo grado nell'anamnesi	0.18%	98.76%	1.06%	
np_05	LBBB (ramo anteriore) nell'anamnesi	0.65%	98.29%	1.06%	
np_07	LBBB incompleto nell'anamnesi	0.06%	98.88%	1.06%	
np_08	LBBB completo nell'anamnesi	0.35%	98.59%	1.06%	
np_09	RBBB incompleto nell'anamnesi	0.12%	98.82%	1.06%	

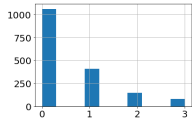
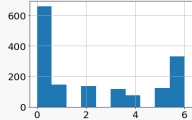
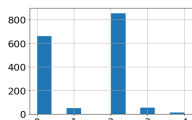
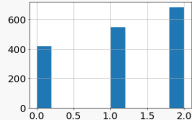
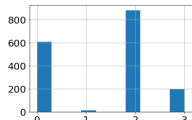
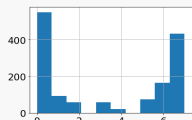
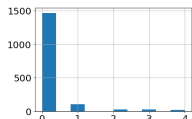
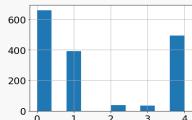
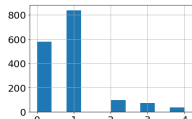
Name	Description	1	0	Missings	Pie
np_10	RBBB completo nell'anamnesi	0.18%	98.76%	1.06%	
endocr_01	Diabete mellito nell'anamnesi	13.41%	85.94%	0.65%	
endocr_02	L'obesità nell'anamnesi	2.47%	96.94%	0.59%	
endocr_03	Tireotossicosi nell'anamnesi	0.76%	98.65%	0.59%	
zab_leg_01	Bronchite cronica nell'anamnesi	7.88%	91.71%	0.41%	
zab_leg_02	Bronchite cronica ostruttiva nell'anamnesi	7.12%	92.47%	0.41%	
zab_leg_03	Asma bronchiale nell'anamnesi	2.18%	97.41%	0.41%	
zab_leg_04		0.53%	99.06%	0.41%	
zab_leg_06	Tubercolosi polmonare nell'anamnesi	1.29%	98.29%	0.41%	
O_L_POST	Edema polmonare al momento dell'ammissione all'unità di terapia intensiva	6.47%	92.82%	0.71%	
K_SH_POST	Shock cardiogeno al momento dell'ammissione all'unità di terapia intensiva	2.71%	96.41%	0.88%	
MP_TP_POST	Parossismi di fibrillazione atriale al momento dell'ammissione all'unità di terapia intensiva unità di cura, (o in una fase pre-ospedaliera)	6.71%	92.47%	0.82%	
SVT_POST	Parossismi di tachicardia sopraventricolare al momento dell'ammissione all'unità di terapia intensiva, (o in una fase pre-ospedaliera)	0.47%	98.82%	0.71%	
GT_POST	Parossismi di tachicardia ventricolare al momento del ricovero in unità di terapia intensiva, (o in una fase pre-ospedaliera)	0.47%	98.82%	0.71%	
FIB_G_POST	Fibrillazione ventricolare al momento dell'ammissione all'unità di terapia intensiva (o in una fase pre-ospedaliera)	0.88%	98.41%	0.71%	
IM_PG_P	Presenza di un infarto miocardico del ventricolo destro	2.94%	97.0%	0.06%	

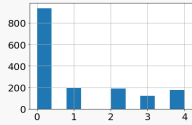
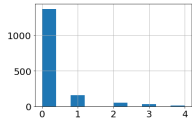
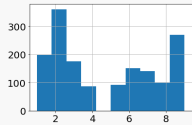
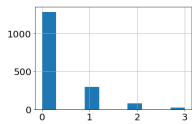
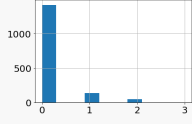
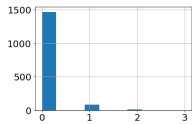
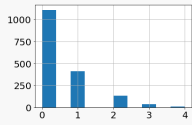
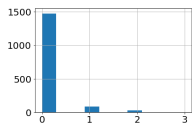
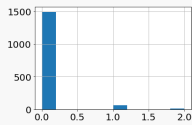
Name	Description	1	0	Missings	Pie
ritm_ecg_p_01	Ritmo ECG al momento dell'ammissione in ospedale - sinusale (con una frequenza cardiaca 60-90)	60.53%	30.53%	8.94%	
ritm_ecg_p_02	Ritmo ECG al momento dell'ammissione in ospedale - fibrillazione atriale	5.59%	85.47%	8.94%	
ritm_ecg_p_04	Ritmo ECG al momento dell'ammissione in ospedale - atriale	1.35%	89.71%	8.94%	
ritm_ecg_p_06	Ritmo ECG al momento dell'ammissione in ospedale - idioventricolare	0.06%	91.0%	8.94%	
ritm_ecg_p_07	Ritmo ECG al momento dell'ammissione in ospedale - seno con una frequenza cardiaca superiore a 90 (tachicardia)	20.76%	70.29%	8.94%	
ritm_ecg_p_08	Ritmo ECG al momento dell'ammissione in ospedale - seno con una frequenza cardiaca inferiore a 60 (bradicardia)	2.71%	88.35%	8.94%	
n_r_ecg_p_01	Contrazioni atriali premature su ECG al momento del ricovero in ospedale	3.41%	89.82%	6.76%	
n_r_ecg_p_02	Frequenti contrazioni atriali premature su ECG al momento del ricovero in ospedale	0.47%	92.76%	6.76%	
n_r_ecg_p_03	Contrazioni ventricolari premature su ECG al momento dell'ammissione in ospedale	12.0%	81.24%	6.76%	
n_r_ecg_p_04	Frequenti contrazioni ventricolari premature sull'ECG al momento dell'ammissione in ospedale	4.06%	89.18%	6.76%	
n_r_ecg_p_05	Parossismi di fibrillazione atriale su ECG al momento del ricovero in ospedale	4.12%	89.12%	6.76%	
n_r_ecg_p_06	Forma persistente di fibrillazione atriale su ECG al momento dell'ammissione in ospedale	1.88%	91.35%	6.76%	
n_r_ecg_p_08	Parossismi di tachicardia sopraventricolare su ECG al momento dell'ammissione in ospedale	0.24%	93.0%	6.76%	
n_r_ecg_p_09	Parossismi di tachicardia ventricolare su ECG al momento del ricovero in ospedale	0.12%	93.12%	6.76%	
n_r_ecg_p_10	Fibrillazione ventricolare su ECG al momento dell'ammissione in ospedale	0.12%	93.12%	6.76%	
n_p_ecg_p_01	Blocco sinoatriale su ECG al momento dell'ammissione in ospedale	0.12%	93.12%	6.76%	

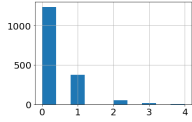
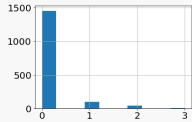
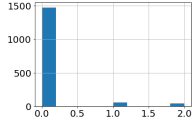
Name	Description	1	0	Missings	Pie
n_p_ecg_p_03	Blocco AV di primo grado su ECG al momento dell'ammissione in ospedale	1.88%	91.35%	6.76%	
n_p_ecg_p_04	Blocco AV di secondo grado di tipo 1 (Mobitz I/Wenckebach) su ECG al al momento del ricovero in ospedale	0.29%	92.94%	6.76%	
n_p_ecg_p_05	Blocco AV di secondo grado di tipo 2 (Mobitz II/Hay) su ECG al momento dell'ammissione all'ospedale	0.12%	93.12%	6.76%	
n_p_ecg_p_06	Blocco AV di terzo grado su ECG al momento dell'ammissione in ospedale	1.59%	91.65%	6.76%	
n_p_ecg_p_07	LBBB (branca anteriore) sull'ECG al momento dell'ammissione in ospedale	6.0%	87.24%	6.76%	
n_p_ecg_p_08	LBBB (branca posteriore) su ECG al momento del ricovero in ospedale	0.41%	92.82%	6.76%	
n_p_ecg_p_09	LBBB incompleto su ECG al momento del ricovero in ospedale	0.59%	92.65%	6.76%	
n_p_ecg_p_10	LBBB completo su ECG al momento dell'ammissione in ospedale	2.0%	91.24%	6.76%	
n_p_ecg_p_11	RBBB incompleto su ECG al momento dell'ammissione in ospedale	1.65%	91.59%	6.76%	
n_p_ecg_p_12	RBBB completo su ECG al momento dell'ammissione in ospedale	4.59%	88.65%	6.76%	
fibr_ter_01	Terapia fibrinolitica con Celiasum 750k IU	0.76%	98.65%	0.59%	
fibr_ter_02	Terapia fibrinolitica con Celiasum 1m IU	0.94%	98.47%	0.59%	
fibr_ter_03	Terapia fibrinolitica con Celiasum 3m IU	4.0%	95.41%	0.59%	
fibr_ter_05	Terapia fibrinolitica con Streptase	0.24%	99.18%	0.59%	
fibr_ter_06	Terapia fibrinolitica con Celiasum 500k IU	0.53%	98.88%	0.59%	
fibr_ter_07	Terapia fibrinolitica con Celiasum 250k IU	0.35%	99.06%	0.59%	

Name	Description	1	0	Missings	Pie
fibr_ter_08	Terapia fibrinolitica con Streptodecase 1,5m IU	0.12%	99.29%	0.59%	
GIPO_K	Ipopotassiemia (< 4 mmol/L)	31.41%	46.88%	21.71%	
GIPER_NA	Aumento del sodio nel siero	1.76%	76.18%	22.06%	
NA_KB	Uso di farmaci oppioidi da parte del team di cardiologia d'emergenza	36.35%	25.0%	38.65%	
NOT_NA_KB	Uso dei NSAID da parte del team di cardiologia d'urgenza	41.24%	18.41%	40.35%	
LID_KB	Uso della lidocaina da parte del team di cardiologia d'emergenza	23.29%	36.88%	39.82%	
NITR_S	Uso di nitrati liquidi in terapia intensiva	11.47%	88.0%	0.53%	
LID_S_n	Uso della lidocaina in terapia intensiva	28.18%	71.24%	0.59%	
B_BLOK_S_n	Uso dei beta-bloccanti in terapia intensiva	12.65%	86.71%	0.65%	
ANT_CA_S_n	Uso di calcio-antagonisti in terapia intensiva	66.18%	33.06%	0.76%	
GEPAR_S_n	Uso di a anticoagulanti (eparina) in terapia intensiva	70.76%	28.24%	1.0%	
ASP_S_n	Uso dell'acido acetilsalicilico in terapia intensiva	73.65%	25.35%	1.0%	
TIKL_S_n	Uso di Ticlid in terapia intensiva	1.76%	97.29%	0.94%	
TRENT_S_n	Uso di Trental in terapia intensiva	20.06%	79.0%	0.94%	

Ordinali

Name	Description	Values	Hist
INF_ANAM	Quantità di infarti miocardici nell'anamnesi	0,1,2,3 (missing: 4)	
STENOK_AN	Angina pectoris da sforzo nell'anamnesi	0-6 (missing:106)	
FK_STENOK	Classe funzionale (FC) dell'angina pectoris nell'ultimo anno	0-nessun angina pectoris, 1-I FC, 2- II FC, 3-III FC, 4-IV FC (missing : 73)	
IBS_POST	Malattia coronarica (CHD) nelle ultime settimane, giorni prima ricovero in ospedale	0-Nessun CHD, 1 -angina pectoris da sforzo, 2- angina pectoris instabile (missing :51)	
GB	Presenza di un'ipertensione essenziale	0-nessuna ipertensione essenziale, 1- Stadio 1, 2-Stadio 2, 3- Stadio 3 (missing : 9)	
DLIT_AG	Durata dell'ipertensione arteriosa	0- nessun ipertensione arteriosa, 1,2,4,5 (anni), 6 (da 6-10 anni), 7 (più di 10 anni) (missings : 248)	
ZSN_A	Presenza di insufficienza cardiaca cronica (HF) nell'anamnesi, Attributo parzialmente ordinato: ci sono due linee di gravità: 0<1<2<4, 0<1<3<4. Lo stato 4 significa gli stati 2 e 3 simultanei	0 : - non c'è insufficienza cardiaca cronica 1: Stadio I 2:IIA stadio(insufficienza cardiaca dovuta a disfunzione sistolica del ventricolo destro) Stadio IIA (insufficienza cardiaca dovuta a disfunzione sistolica ventricolare sinistra) 4:Stadio IIB (insufficienza cardiaca dovuta a disfunzione sistolica del ventricolo destro e sinistro disfunzione) (missing: 54)	
ant_im	Presenza di un infarto miocardico anteriore (ventricolare sinistro) (cambiamenti ECG nelle derivazioni V1 - V4)	0: Nessuna presenza 1:Il QRS non ha cambiamenti 2:- Il QRS è come il complesso QR 3:- Il QRS è come il complesso Qe 4- Il QRS è come il complesso QS (missings: 83)	
lat_im	Presenza di un infarto miocardico laterale (ventricolare sinistro) (cambiamenti ECG nelle derivazioni V5 - V6, I, AVL)	0: Nessuna presenza 1:Il QRS non ha cambiamenti 2:- Il QRS è come il complesso QR 3:- Il QRS è come il complesso Qe 4- Il QRS è come il complesso QS (missings: 80)	

Name	Description	Values	Hist
inf_im	Presenza di un infarto miocardico inferiore (ventricolare sinistro) (cambiamenti ECG nelle derivazioni III, AVF, II)	0: Nessuna presenza 1: Il QRS non ha cambiamenti 2:- Il QRS è come il complesso QR 3:- Il QRS è come il complesso Qe 4- Il QRS è come il complesso QS (missings: 80)	
post_im	Presenza di un infarto miocardico posteriore (ventricolare sinistro) (cambiamenti ECG in V7 - V9, cambiamenti di reciprocità nelle derivazioni V1 - V3)	0: Nessuna presenza 1: Il QRS non ha cambiamenti 2:- Il QRS è come il complesso QR 3:- Il QRS è come il complesso Qe 4- Il QRS è come il complesso QS (missings: 72)	
TIME_B_S	Tempo trascorso dall'inizio dell'attacco di CHD al ospedale	1: meno di 2 ore 2: 2-4 ore 3: 4-6 ore 4: 6-8 ore 5 : 8-12 ore 6 : 12-24 ore 7 : più di un giorno 8: Più di 2 giorni 9: oltre 3 giorni (missings: 126)	
R_AB_1_n	Ricaduta del dolore nelle prime ore del periodo di ricovero	0: Nessuna ricaduta 1:solo una 2: 2 volte 3: 3 o più volte (missings : 16)	
R_AB_2_n	Ricaduta del dolore nel secondo giorno del periodo di ricovero	0: Nessuna ricaduta 1:solo una 2: 2 volte 3: 3 o più volte (missings: 108)	
R_AB_3_n	Ricaduta del dolore nel terzo giorno del periodo ospedaliero	0: Nessuna ricaduta 1:solo una 2: 2 volte 3: 3 o più volte (missings: 128)	
NA_R_1_n	Uso di farmaci oppioidi in terapia intensiva nelle prime ore del periodo	0 : No 1: Una volta 2: 2 Volte 3: 3volte 4: 4 volte (missings: 5)	
NA_R_2_n	Uso di farmaci oppioidi in terapia intensiva nel secondo giorno di ricovero periodo	0 : No 1: Una volta 2: 2 Volte 3: 3volte (missings:108)	
NA_R_3_n	Uso di farmaci oppioidi in terapia intensiva nel terzo giorno di ricovero periodo	0 : No 1: Una volta 2: 2 Volte 3: 3volte (missings: 131)	

Name	Description	Values	Hist
NOT_NA_1_n	Uso di NSAID in terapia intensiva nelle prime ore di degenza	0 : No 1: Una volta 2: 2 Volte 3: 3volte 4: 4 volte (missings: 10)	
NOT_NA_2_n	Uso di NSAID in terapia intensiva nel secondo giorno di degenza	0 : No 1: Una volta 2: 2 Volte 3: 3volte 4: 4 volte (missings: 110)	
NOT_NA_3_n	Uso di NSAID in terapia intensiva nel terzo giorno di degenza	0 : No 1: Una volta 2: 2 Volte (missings: 131)	

Problemi Ricontrati

- La variabile binaria **NOT_NA_KB** presenta un osservazione dal valore anomalo 2
- Molte variabili presentano valori mancanti

Possibili Soluzioni

....

Ci sono quattro possibili momenti temporali per la previsione della complicazione: sulla base delle informazioni note:

Feature di input misurata al momento dell'ammissione in ospedale

Feature di input misurata prima della fine del primo giorno (24 ore dopo il ricovero in ospedale)

R_AB_1_n, NA_R_1_n, NOT_NA_1_n

Feature di input misurata prima della fine del secondo giorno (48 ore dopo il ricovero in ospedale)

R_AB_2_n, NA_R_2_n, NOT_NA_2_n

Feature di input misurata prima della fine del terzo giorno (72 ore dopo il ricovero in ospedale)

R_AB_3_n, NA_R_3_n, NOT_NA_3_n

Feature di input misurata al momento dell'ammissione in ospedale

Tutte le altre feature di input.

Data Cleaning