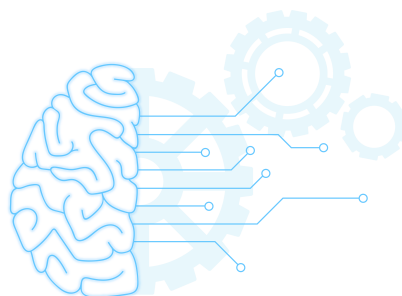


BORSISTA: SARAH ADAMO - TUTOR: PROF. GIOVANNI D'ADDIO



THE ITALIAN
EDUCATION
& RESEARCH
NETWORK

PROGETTAZIONE DI UN SISTEMA DI
INTELLIGENZA ARTIFICIALE DEDICATO
AD APPLICAZIONI DI **TELEMEDICINA** ED
ASSISTENZA DOMICILIARE INTEGRATA



GIORNATA DI INCONTRO
BORSE DI STUDIO GARR
"ORIO CARLINI"
ROMA

Ente ospitante:



IRCCS ICS SPA SB Maugeri –
Telese Terme (BN)



Scopo del Progetto



Progettazione di un
sistema AI



Supporto nel clinical
decision making



Garanzia dei LEA nelle
aree interne





PERCHE' LA TELEMEDICINA?



Insieme delle tecniche
mediche ed informatiche
per la cura a distanza



Utilizzo di sistemi per
acquisizione e/o **trasmissione**
di dati anamnestici



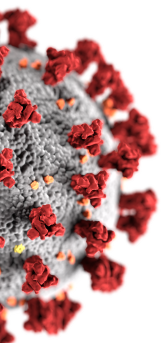
Risorsa per rendere
possibile l'**autogestione**
del paziente



Ma soprattutto...
...Mezzo per collezionare
BIG DATA



**Analisi predittive di Machine
Learning**

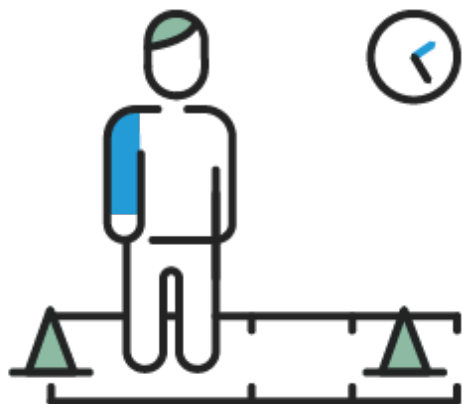


COVID-19 : sintomi medio-lievi a gravi complicazioni.

Dopo la fase acuta, si parla del paradigma: «**long-term-COVID**».



Riabilitazione multidisciplinare ad hoc (ancora non consolidata¹)



6-minute walking test² (6MWT) :

- Efficace e consolidato
- Di facile realizzazione



Dataset

Istanze

- 189 pazienti

Acquisizioni

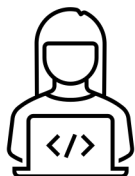
- Analisi di laboratorio
- Test cardiorespiratorio da sforzo
- Spirometria

Features

- 157 attributi
- Parametri **pre e post** riabilitazione
- Parametri anamnestici
- Parametri ematochimici
- Parametri respiratori
- Comorbidità



Normalizzazione $6MWD_{pre}$ e $6MWD_{post}$



FUNZIONE MATLAB



Massimale teorico per ogni paziente

- Sesso, età, BMI;
- «ATS Statement: Guidelines for the Six-Minutes Walk Test»



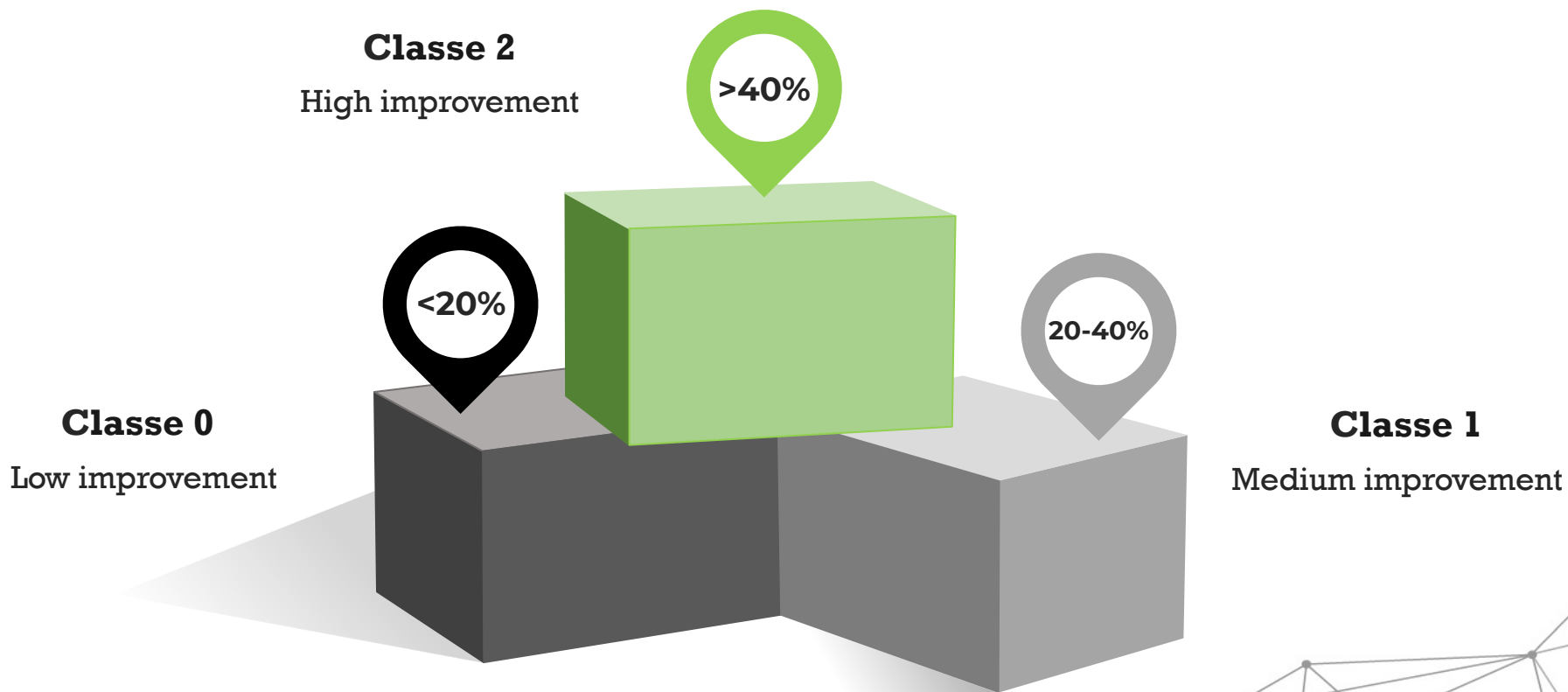
$$6MWD_{MassimalePre\%} = \frac{6MWD_{before}}{6MWD_{maximum}} * 100;$$

$$6MWD_{MassimalePost\%} = \frac{6MWD_{after}}{6MWD_{maximum}} * 100;$$

$$\Delta 6MWD = 6MWD_{MassimalePost\%} - 6MWD_{MassimalePre\%}$$



$\Delta 6\text{MWD}$ post-riabilitazione





Subset 1

30 features:
Anamnestiche
Funzionali
Respiratorie

MISSING VALUES

SMOTE
OVERSAMPLING

19 variabili continue

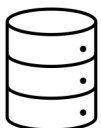
11 variabili nominali
trasformate in binarie

- Mediana per valori numerici
- Valore più frequente per valori categorici

Sovracampionamento della classe minoritaria



HOLD-OUT PARTITIONING



70% TRAINING – 30% TEST



SUPERVISED ML ALGORITHMS

TREE-BASED

Random Forest (RF)
Gradient Boosted (GB)
Adaptive Boosting (AdaBoost)

INSTANCE-BASED

K nearest neighbor (KNN)





CLASSI TARGET

Senza SMOTE

Classe 0: 64 istanze
Classe 1: 95 istanze
Classe 2: 30 istanze



Con SMOTE

Classe 0: 95 istanze
Classe 1: 95 istanze
Classe 2: 95 istanze
TOT: 285





Metriche di validazione per ogni algoritmo

Algorithm	Accuracy	Sensitivity (%)	Specificity (%)	AUCROC (%)
RF	83.7	84.0	91.8	94.5
ADA-B	81.4	71.0	92.7	88.5
GB	79.1	71.0	87.3	84.6
KNN	80.2	74.2	89.1	93.4



- **RF** migliori Acc, Sens, AUCROC
- **ADA-B** migliore Spec

AUROC: area under the receiver operating characteristic curve; RF: random forest; ADA-B: adaptive boosting; GB: gradient boosting; KNN: k-nearest neighbors.



Feature importance: le 10 features più rilevanti per la classificazione

Feature	IG
6MWD, meters	10.62%
DLCO%, % predicted	6.25%
FVC, L	5.85%
DLCO, ml/min/mmHg	5.09%
FEV ₁ , L	4.68%
PaO ₂ , mmHg	4.67%
TLC, L	4.59%
CAT	4.57%
Age, years	4.53%
FVC%, % predicted	4.41%



Confronto tra le 3 classi di miglioramento delle 10 features più importanti

Features	Group 0 (n=64)	Group 1 (n=95)	Group 2 (n=30)	P-value
6MWD, meters	193.13 ± 131.77	171.20 ± 90.16	31.10 ± 56.69	<0.001
DLCO%, % predicted	55.70 ± 15.63	56.27 ± 12.82	49.97 ± 11.58	0.230
FVC, L	2.97 ± 0.62	2.92 ± 0.85	2.42 ± 0.65	0.001
DLCO, ml/min/mmHg	11.73 ± 6.77	10.97 ± 5.19	10.47 ± 4.11	0.682
FEV ₁ , L	2.44 ± 0.52	2.34 ± 0.67	2.03 ± 0.58	0.003
PaO ₂ , mmHg	75.19 ± 13.12	72.89 ± 12.33	65.03 ± 13.32	0.005
TLC, L	4.51 ± 0.85	4.71 ± 1.02	4.53 ± 1.11	0.736
CAT	26.92 ± 2.53	26.61 ± 1.36	27.00 ± 3.53	0.163
Age, years	62.56 ± 12.84	62.88 ± 8.62	63.97 ± 9.18	0.972
FVC%, % predicted	76.81 ± 15.56	76.57 ± 14.83	64.33 ± 14.22	<0.001



DLCO, FEV1, FVC E TLC come principali indicatori per l'outcome riabilitativo



DLCO: indicatore di anomalie vascolari polmonari o interstiziali



FVC, FEV1 e TLC: stretta correlazione tra residual restrictive pattern e manifestazioni più gravi di sintomi dopo la fase acuta



E' possibile dunque monitorare tali indicatori durante il corso della riabilitazione polmonare per predire e agevolare i miglioramenti del paziente

Adamo, S.; Ambrosino, P.; Ricciardi, C.; Accardo, M.; Mosella, M.; Cesarelli, M.; d'Addio, G.; Maniscalco, M. **A Machine Learning Approach to Predict the Rehabilitation Outcome in Convalescent COVID-19 Patients.** J. Pers. Med. 2022, 12, 328. <https://doi.org/10.3390/jpm12030328>





Subset 2

78 istanze

22 features:

Anamnestiche

Parametri

ematochimici



Variabili continue

MISSING VALUES



Mediana per valori numerici

NORMALIZATION



Normalizzazione delle variabili numeriche



UNSUPERVISED ML ALGORITHM

- K-means clustering

EVALUATION

- Silhouette coefficient

VISUALIZATION

- Parallel coords

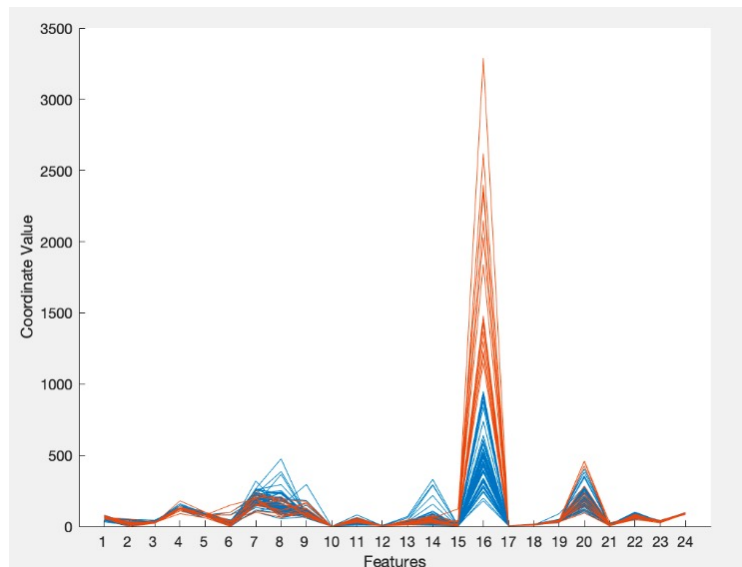


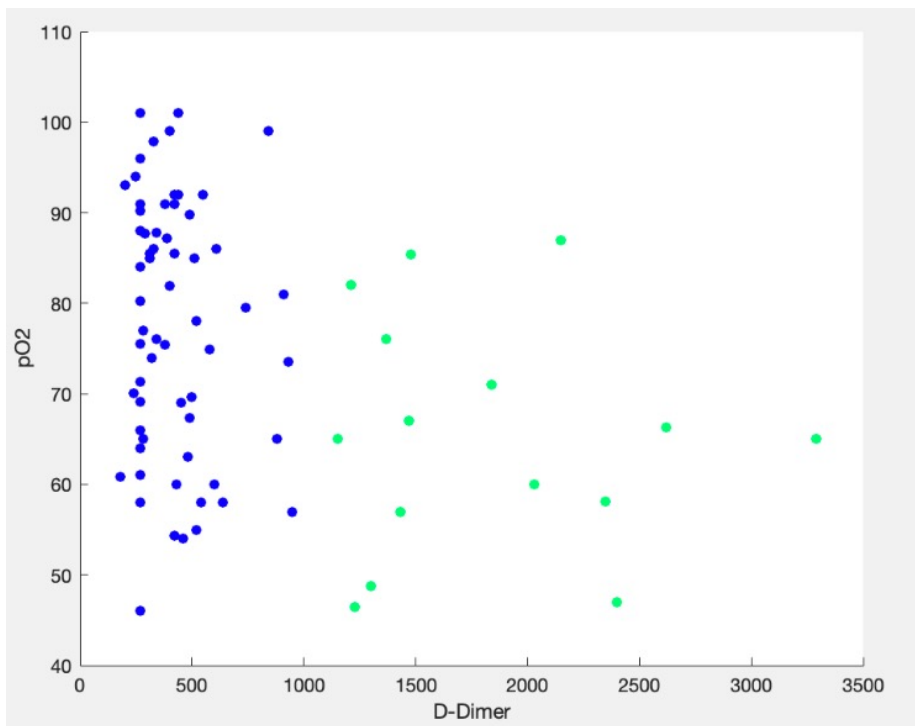


Silhouette coefficients per il clustering al variare del numero di clusters, con k da 2 a 4.

	$k=2$	$k=3$	$k=4$
Silhouette Coefficient	0.85	0.83	0.61

Parallel coords plot, dove i colori indicano i due gruppi. La feature 16 rappresenta il **D-Dimero**.





La figura mostra i dati di PaO_2 clusterizzati in funzione del D-Dimero, dove la PaO_2 rappresenta un indice di ossigenazione del sangue.



GIORNATA DI INCONTRO BORSE DI STUDIO GARR “ORIO CARLINI” BORSISTI DAY 2022



EVALUATION: STATISTICAL ANALYSIS - subset 1

CASE 1: POST-COVID-19

Features (n)	Cluster 0 (n=15)	Cluster 1 (n=63)	p
Age, years	67.47 ± 7.00	59.21 ± 10.90	0.007
BMI, m ² /kg	26.52 ± 3.82	29.22 ± 5.57	Ns
Hospitalization, days	19.13 ± 14.15	16.29 ± 14.75	Ns
Cigarettes, packs per year	35.80 ± 43.03	10.56 ± 19.37	0,003
TC ^b , mg/dl	174.53 ± 41.68	188.14 ± 41.52	Ns
TGs ^c , mg/dL	132.07 ± 52.30	164.17 ± 75.16	Ns
Glycemia, mg/dL	108.60 ± 36.75	91.97 ± 33.51	Ns
Creatinine, mg/dL	0.77 ± 0.18	0.81 ± 0.14	Ns
Urea, mg/dL	42.60 ± 11.51	37.94 ± 11.97	Ns
Uricemia, mg/dL	4.03 ± 1.54	5.23 ± 1.59	0,010
AST ^d , UI/L	23.07 ± 11.70	21.70 ± 10.90	Ns
ALT ^e , UI/L	44.13 ± 19.94	58.89 ± 65.34	Ns
PCR ^f , mg/dL	18.68 ± 31.16	6.18 ± 7.60	0.026
D-Dimer, ng/mL	1821.33 ± 632.67	419.84 ± 186.14	<0.001
Red blood cells, 10 ⁶ /mL	4.11 ± 0.60	4.57 ± 0.53	0.005
Hemoglobin, g/dL	12.04 ± 2.02	13.03 ± 1.54	0.039
Hematocrit, %	36.64 ± 5.44	40.48 ± 7.60	0.026
Platelet, 10 ³ /mL	242.20 ± 119.318	204.03 ± 66.56	Ns
Leukocyte, 10 ³ /mL	9.16 ± 5.33	7.84 ± 2.56	Ns
PaO ₂ ^g , mmHg	65.47 ± 13.13	77.62 ± 14.09	0.006
PaCO ₂ ^h , mmHg	34.84 ± 3.97	36.32 ± 3.34	Ns
SpO ₂ ⁱ , %	92.53 ± 4.17	95.11 ± 3.07	0.011



- ✓ Identificazione di due fenotipi sulla base dei valori ematochimici in ingresso alla riabilitazione
- ✓ D-Dimero come indice principale per la discriminazione: marker di infiammazione e causa di trombosi venose in pazienti COVID-19 e post³
- ✓ PaO₂ e SpO₂ più bassi nel fenotipo con condizioni peggiori, coerentemente con i valori di emoglobina e globuli rossi.
- ✓ E' possibile dunque monitorare tali indicatori per predire eventi avversi e per identificare le condizioni di salute prossime del paziente

Adamo S., Ricciardi C., Ambrosino P., Maniscalco M., Biancardi A., Cesarelli G., D'Addio G. Unsupervised machine learning to identify convalescent COVID-19 patients. *IEEE MeMeA 2022, submitted.*





CASE 2: TELESORVEGLIANZA CHF



Un modello innovativo di servizio per la gestione domiciliare di pazienti con scompenso cardiaco cronico medio/grave, che si affianca - integrandola - all'attività del medico di medicina generale

- Arruolamento di pazienti di pazienti affetti da **CHF** con classe **NHYA ≥ 2** ;
- Riarruolamento di pazienti già seguiti in telesorveglianza;
- Introduzione del percorso a **bassa intensità**;
- Presenza di un **infermiere tutor**.



Riduzione dei ricoveri, degli accessi al PS e
del ricorso a prestazioni ambulatoriali

Aumento della qualità della vita



Dataset

Istanze

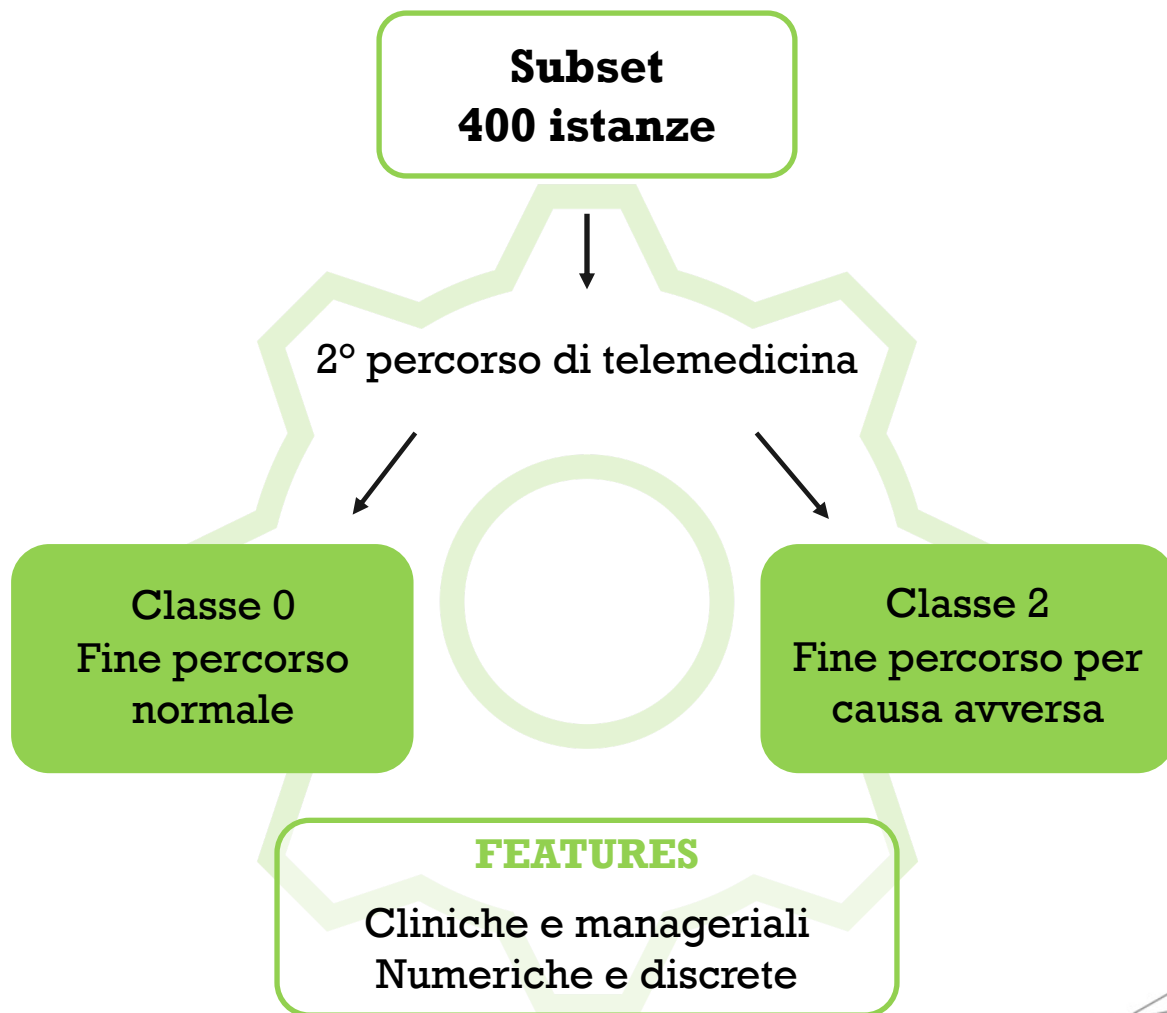
- 926 pazienti

Acquisizioni

- ECG
- ECO
- Dispositivi e-Health

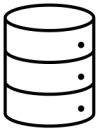
Features

- Parametri anamnestici
- Parametri clinici
- Parametri gestionali
- Comorbidità





HOLD-OUT PARTITIONING



70% TRAINING – 30% TEST



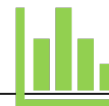
SUPERVISED ML ALGORITHMS

- Decision tree

FEATURE SELECTION

- Wrapper



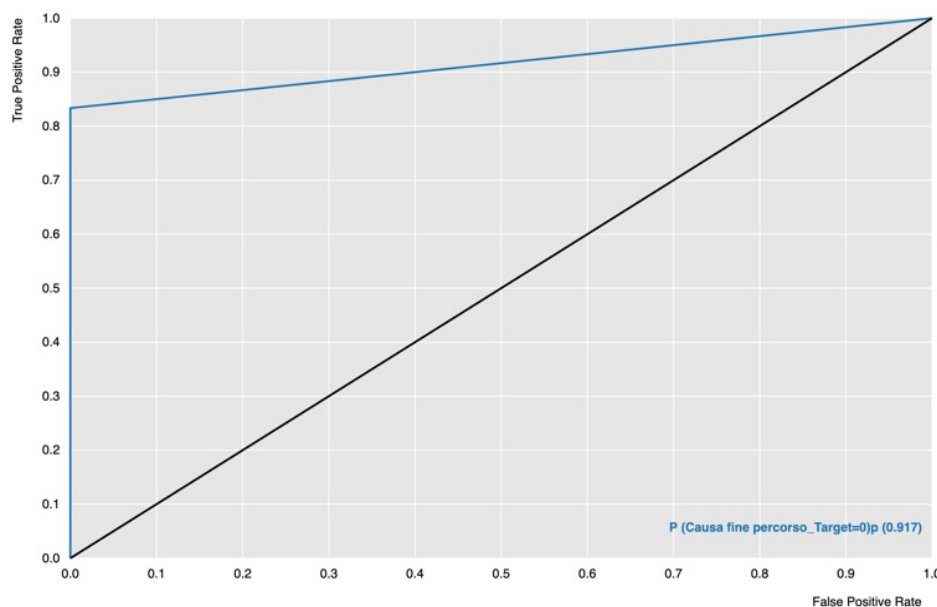


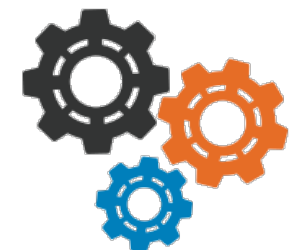
Features selezionate con il wrapper method

Algorithm	Features selected
DT	1. Intensity program 2. Echocardiogram done (n) 3. Cardiology consultations (n) 4. Diabetes

Evaluation metrics per Decision Tree

Algorithm	ACC [%]	SENS [%]	SPEC [%]	AUCROC [%]
DT	91.7	83.3	100.0	91.7





IN PROGRESS...

- Analisi dei fattori di rischio in soggetti asmatici in collaborazione con AstraZeneca;
- 260 istanze, 20 features;
- Algoritmi di ML unsupervised e supervised learning;
- Clustering e KNN;
- Eosinofili come principale indice discriminante.





UTILIZZO RETE GARR

IRCCS ICS Maugeri



**Dipartimento DIETI – UNINA
« Federico II » – Napoli**



Servizi Essenziali

- Rete e accesso
- Sicurezza
- Identità e mobilità

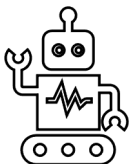


Cloud & Applicazioni

- Trasferimento dati
- Storage



Conclusioni: risultati ottenuti



Sono stati implementati algoritmi in grado di fornire **predizioni accurate** in termini di **Accuracy, Sensitivity, AUCROC** a partire dai dati collezionati in pazienti con patologie cronico-disabilitanti



Sono stati ricavati indici **predittivi** relativi all'outcome riabilitativo dei pazienti



Sono stati individuati **indicatori precoci** del peggioramento delle condizioni patologiche che suggeriscono una ri-ospedalizzazione del paziente



Donisi, Ricciardi, C., Cesarelli, G., Coccia, A., Amitrano, F., Adamo, S., & D'Addio, G. (2022). **Bidimensional and Tridimensional Poincaré Maps in Cardiology: A Multiclass Machine Learning Study.** *Electronics*, 11(3), 448. <https://doi.org/10.3390/electronics11030448>

Adamo, S.; Ambrosino, P.; Ricciardi, C.; Accardo, M.; Mosella, M.; Cesarelli, M.; d'Addio, G.; Maniscalco, M. **A Machine Learning Approach to Predict the Rehabilitation Outcome in Convalescent COVID-19 Patients.** *J. Pers. Med.* 2022, 12, 328. <https://doi.org/10.3390/jpm12030328>

Adamo S., Ricciardi C., Ambrosino P., Maniscalco M., Biancardi A., Cesarelli G., D'Addio G. **Unsupervised machine learning to identify convalescent COVID-19 patients.** *IEEE MeMeA 2022, submitted.*

Adamo S. et al, D'Addio G., **Characterization of Patients Admitted to Emergency Department for Asthma Attack by a Machine Learning Approach,** 2022, *in submission.*



Sviluppi futuri (1/2)

01. ASPETTI Qualitativi

- Estensione dello studio di telesorveglianza alla Broncopneumopatia Cronica Ostruttiva (BPCO);
- Estensione del set di features per CHF, introducendo nuove features cliniche, gestionali e farmacologiche





Sviluppi futuri (2/2)



02. ASPETTI Quantitativi

- Analisi dataset CHF con ampliamento da 400 a 926 istanze;
- Analisi dataset post-COVID-19 con ampliamento da 189 a 300 c.a. istanze.



Ricadute Previste



Applicabilità del sistema nelle attività di telemonitoraggio durante i trattamenti riabilitativi



Riprogettazione di alcune pratiche secondo indici predittivi individuati



Equità di accesso alle prestazioni sanitarie anche nelle aree geografiche interne di pazienti con patologie cronico/disabilitanti





Realizzazione di un **cruscotto di indicatori**

Monitoraggio in real-time

Integrazione con
architettura preesistente



Strumento concreto a supporto del
clinical decision making





Grazie per l'attenzione

