





**Le nuove tecnologie a supporto della PA**  
Angela Peduto, 15/09/2022

Università degli Studi di Salerno  
[anpeduto@unisa.it](mailto:anpeduto@unisa.it)

1

**Indice degli argomenti**

- **dove siamo arrivati**
  - Piattaforme per la PA
  - Interoperabilità
  - App
- **dove stiamo andando**
  - Cloud
  - Banda larga
  - IoT
  - Intelligenza Artificiale
- **casi di studio**
  - Prometea
  - TREE4NB



Angela Peduto - [anpeduto@unisa.it](mailto:anpeduto@unisa.it)

2



## Tecnologie a supporto della PA *dove siamo arrivati*

Angela Peduto - anpeduto@unisa.it

3

## Strategia per l'innovazione tecnologica e la digitalizzazione del Paese 2025

- prima sfida - **una Società digitale**: promozione di modelli virtuosi e di creazione di nuovi servizi digitali
- seconda sfida - **un paese innovativo**: attraverso la collaborazione con le diverse realtà locali, regionali, nazionali e internazionali
- terza sfida - **Sviluppo inclusivo e sostenibile**: attraverso la semplificazione dei servizi e il conseguente rafforzamento delle capacità digitali dei cittadini

Angela Peduto - anpeduto@unisa.it

4

## Le piattaforme per la PA

- Le Piattaforme sono soluzioni che offrono funzionalità fondamentali, trasversali e riusabili nella digitalizzazione dei procedimenti amministrativi delle PA, uniformandone le modalità di erogazione.

Angela Peduto - anpeduto@unisa.it



5

## Le piattaforme per la PA

In una visione funzionale alle scelte delle PA, le Piattaforme sono suddivisibili nelle seguenti 3 categorie:

- *process service*
- *task service*
- *data service*

Angela Peduto - anpeduto@unisa.it

6

## Le piattaforme per la PA (1/3)

**process service**, l'insieme delle Piattaforme che realizzano digitalmente un processo completo di cui le PA diventano utilizzatori

Esempi:

- Public e-procurement
- NoiPA
- SGPA

Angela Peduto - anpeduto@unisa.it

7

## Le piattaforme per la PA (2/3)

**task service**, che implementano singole funzionalità trasversali alla realizzazione di procedimenti amministrativi digitali che le PA integrano all'interno dei loro sistemi

- PagoPA
- SPID
- CIE

Angela Peduto - anpeduto@unisa.it

8

## Le piattaforme per la PA (3/3)

**data service**, che assicurano l'accesso a fonti di dati validati di cui le PA necessitano per dare seguito alle proprie funzioni istituzionali

- ANPR
- SIOPE

Angela Peduto - anpeduto@unisa.it

9

*Rappresentazione sintetica del modello di implementazione e dello stato delle piattaforme*

- <https://www.agid.gov.it/it/piattaforme>

Categoria	Piattaforma	Modello di implementazione	Stato
Process service	Public e-procurement	multiple instances	<i>implementation</i>
	NoiPA	single instance	<i>evolution</i>
	SGPA - Sistema di Gestione Procedimenti Amministrativi nazionali	single instance	<i>design</i>
Task service	pagoPA	single instance	<i>available</i>
	SPID	single instance	<i>available</i>
	CIE	single instance	<i>available</i>
	Poli di conservazione	multiple instances	<i>implementation</i>
Data service	ANPR	single instance	<i>available</i>

Angela Peduto - anpeduto@unisa.it

10

## Obiettivo di garantire l'uniformità e l'interoperabilità

SGPA garantisce un modello unico di dematerializzazione



Data: 18/05/2021

Publicato il nuovo testo dell'Allegato 5 e dell'Allegato 6 alle Linee guida su formazione, gestione e conservazione dei documenti informatici e posticipato al 1° gennaio 2022 l'obbligo di attuazione

Angela Peduto - anpeduto@unisa.it

11

**Modello di interoperabilità e standard tecnologici**

- Tutte le amministrazioni devono aderire agli standard tecnologici ed utilizzare pattern e profili del nuovo Modello di interoperabilità, che consenta di definire ed esporre **Application Programming Interface (API)** conformi agli standard consolidati anche in ambito EU.

Angela Peduto - anpeduto@unisa.it

12

## Le API

- le **API** sono degli strumenti di programmazione e possono assumere diverse forme:
  - possono essere delle librerie di funzioni che permettono al programmatore di interagire con un programma o una piattaforma software
  - Possono essere "chiamate" a parti di un programma che uno sviluppatore può utilizzare per abbreviare il suo lavoro
- Rappresentano l'interfaccia aperta attraverso la quale interagire con programmi (o parti di essi) altrimenti inaccessibili.



Angela Peduto - anpeduto@unisa.it

13

## Tutte le amministrazioni devono aderire agli standard tecnologici

Garantendo:

- tracciabilità delle diverse versioni
- documentazione coordinata (*documentation*);
- limitazioni di utilizzo collegate alle caratteristiche e agli utilizzatori (*throttling*);
- tracciabilità delle richieste ricevute e del loro esito (*logging* e *accounting*);
- un adeguato livello di servizio (SLA);
- configurazione scalabile delle risorse.

Angela Peduto - anpeduto@unisa.it

14

## Servizi della PA fruibili con smartphone

- 28 febbraio 2021
- **I servizi pubblici fruibili attraverso lo smartphone**
- Ad aprile 2022 risultavano **28.6 milioni di download**.



Angela Peduto - anpeduto@unisa.it

15

## App IO, punto di accesso per i servizi digitali attraverso smartphone – i vantaggi

- **18 aprile 2020** l'app "IO" della pubblica amministrazione italiana è disponibile
- **luglio 2020** risultavano già un milione di download
- **Giugno 2021** App IO "**bloccata**" dal Garante privacy causa tracker
- A **novembre 2021**, il Garante privacy ha dato il proprio via libera alle linee guida Agid sull'accesso ai servizi della PA anche tramite App IO



Angela Peduto - anpeduto@unisa.it

16

## SPID e CIE, l'identità digitale unica

- **28 febbraio 2021**
- l'identità digitale SPID e CIE le sole credenziali per accedere ai servizi digitali della pubblica amministrazione.
- **fino al 31 dicembre 2022** – per l'accesso, esclusivamente tramite SPID e CIE, ai servizi consolari per i connazionali residenti all'estero.



Angela Peduto - anpeduto@unis.it

17

## SPID, l'identità digitale unica – I vantaggi

Le credenziali uniche semplificano la vita

- Vantaggi per il cittadino
- Vantaggi per le amministrazioni



18

## PA: la strada è ancora lunga



### PROCESSI

I processi amministrativi correlati alle interazioni tra cittadini/aziende e PA sono complessi e di difficile applicazione



### CONDIVISIONE

Le Amministrazioni non condividono le informazioni ed è compito del cittadino raccoglierle affinché il servizio sia erogato. Tali informazioni sono spesso in forma cartacea



### COSTI

I costi in tempo consumato da cittadini e PA e gli importi spesi per la produzione delle certificazioni sono sproporzionati al servizio

Angela Peduto - anpeduto@unisa.it

19



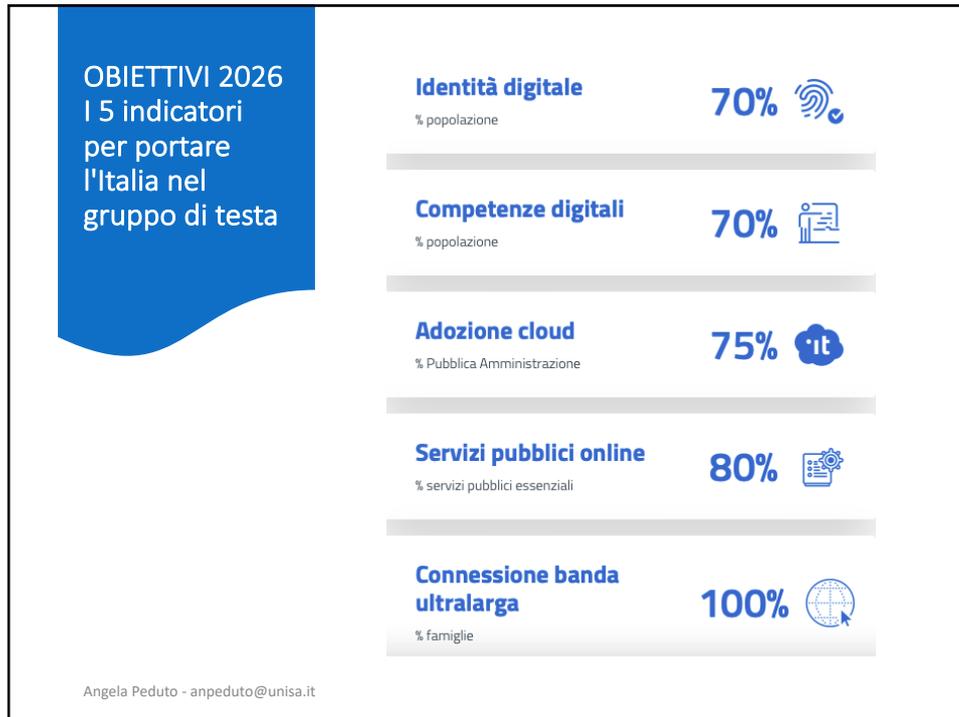
## Tecnologie a supporto della PA

*Dove stiamo andando*

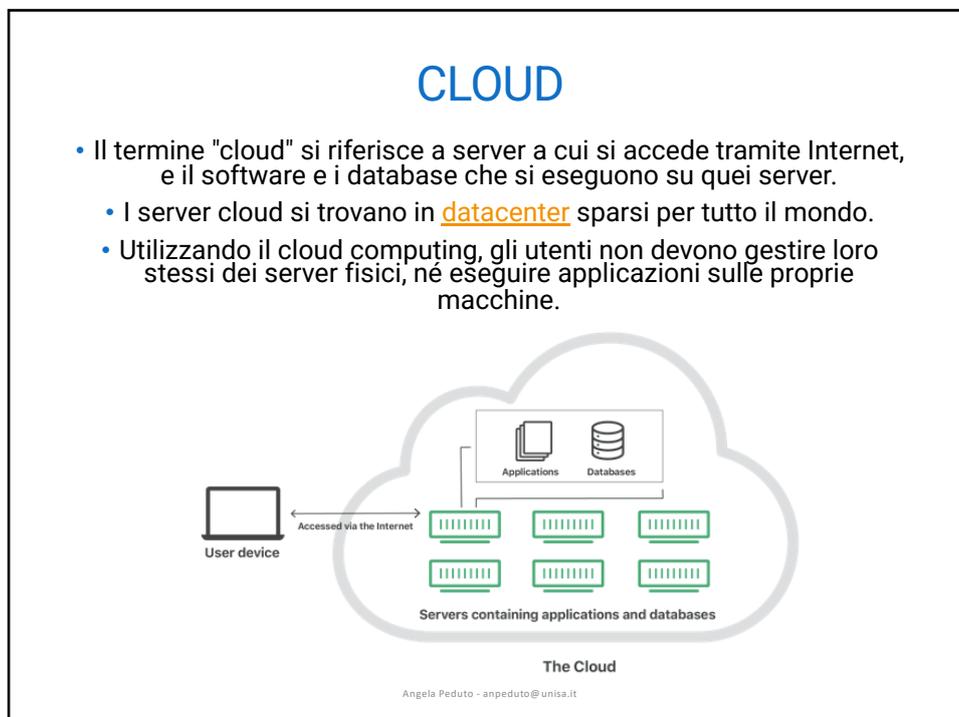
*Accelerare il processo di Semplificazione Amministrativa attraverso l'uso delle tecnologie emergenti*

Angela Peduto - anpeduto@unisa.it

20



21



22

## Quali sono i diversi tipi di implementazioni cloud?

- **Cloud privato:** è un server, un datacenter o una rete distribuita interamente dedicati a una sola organizzazione.
- **Cloud pubblico:** è un servizio eseguito da un fornitore esterno che può includere server in uno o più datacenter.
- **Cloud ibrido:** le implementazioni di **cloud ibridi** combinano cloud pubblici e privati e possono includere anche server legacy on-premise.

Angela Peduto - anpeduto@unisa.it

23

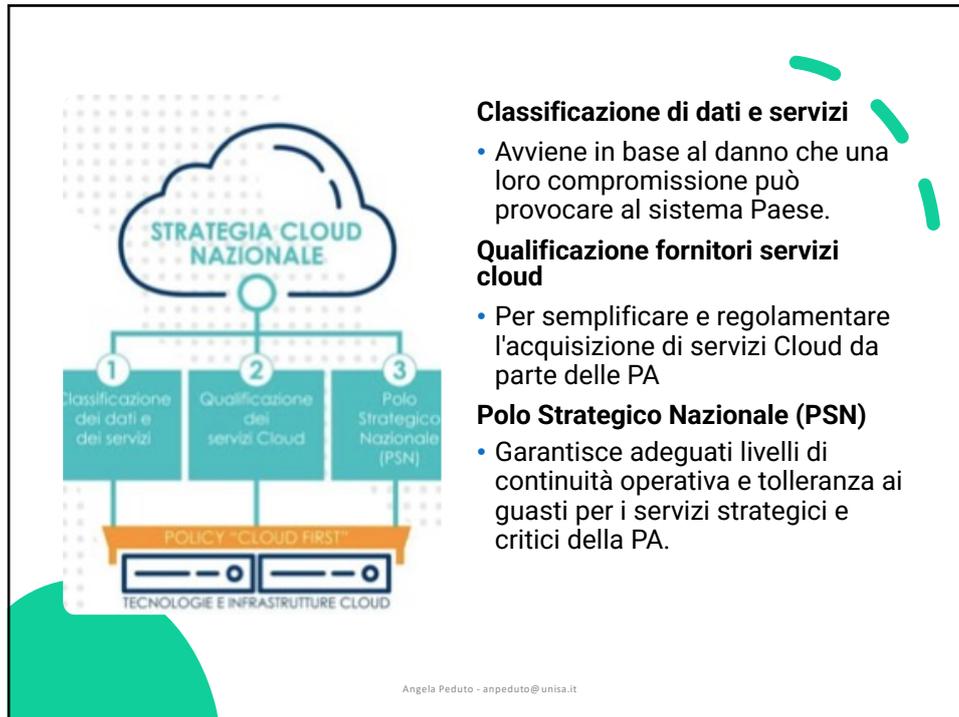
## Strategia Cloud Italia



- infrastrutture digitali sicure, efficienti ed affidabili;
- tutela della privacy e;
- autonomia tecnologica del Paese, di sicurezza e controllo nazionale sui dati.

Angela Peduto - anpeduto@unisa.it

24



25



26

## La migrazione della Pubblica Amministrazione sul Cloud

- Le amministrazioni dovranno:
  - Classificare dati e servizi sulla base delle indicazioni dell'Agenzia per la Cybersicurezza Nazionale, con il supporto del Dipartimento per la trasformazione digitale.
  - Definire i piani di migrazione che saranno validati dal Dipartimento.
- La migrazione dovrà concludersi entro il 2025, attraverso un processo uniforme per tutte le amministrazioni

Angela Peduto - anpeduto@unisa.it

27

2021

Pubblicazione del bando di gara per la realizzazione del Polo Strategico Nazionale entro la fine dell'anno

2022

- Aggiudicazione e realizzazione del PSN
- Classificazione di dati e servizi pubblici
- Qualificazione dei servizi cloud

2023

A partire dalla fine del 2022 le amministrazioni avvieranno la migrazione verso il cloud qualificato che dovrà concludersi entro la fine del 2025 (75% della PA in cloud, PNRR)

Angela Peduto - anpeduto@unisa.it

28

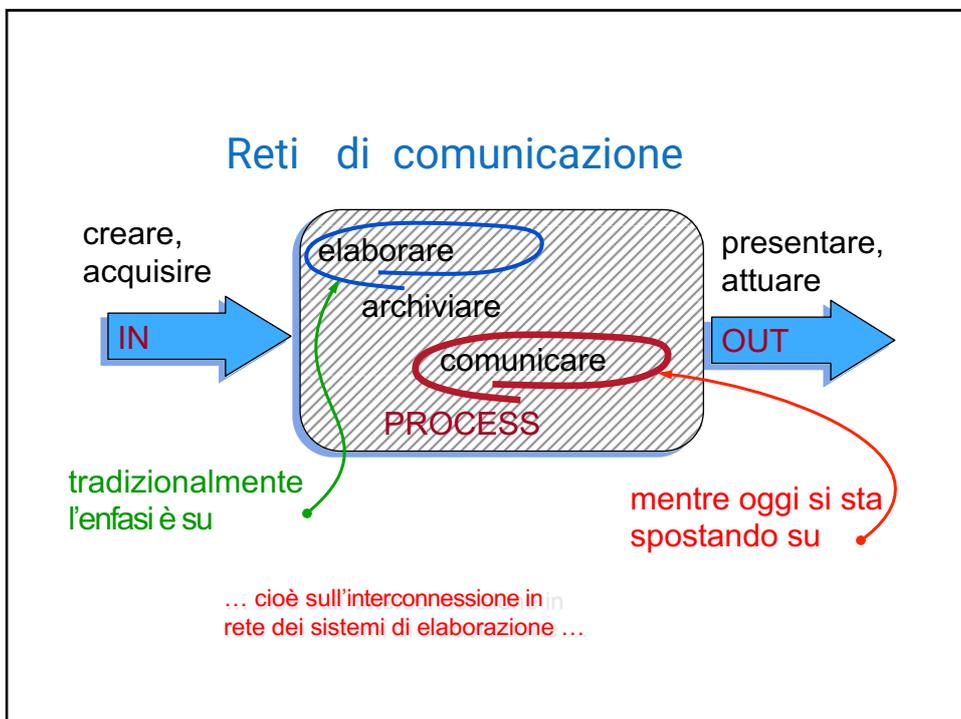


Banda larga e ultralarga

Verso l'infinito e oltre...

Angela Peduto - anpeduto@unisa.it

29



30

## Architetture di rete: caratteristiche

- I mezzi di trasmissione sono costituiti da:
  - apposti cablaggi;
  - rete telefonica dati;
  - satelliti;
  - sistemi di comunicazione wireless.



31

## Componenti di una rete

- Componenti di una rete sono:
  - **nodi**: un nodo è un qualsiasi dispositivo hardware del sistema, in grado di comunicare con gli altri dispositivi che fanno parte della rete
  - **collegamenti** (links) tra i nodi

32

## Scala delle reti

- Un criterio per classificare le reti è legato alla loro scala in base alla distanza tra nodi.

10 m	stanza	Rete locale LAN
100 m	edificio	LAN
1 km	università	LAN
10 km	città	Rete metropolitana MAN
100 km	nazione	Rete geografica WAN
1000 km	continente	Internet
10000 km	pianeta	Internet

33

33

## Prestazioni della rete

- La trasmissione avviene in base alla **banda della rete** (larghezza di banda-utilizzo di più canali)
- La velocità misurata in bit al secondo:
  - Kilobit/s (Kb/s) 1.024 bit per secondo.
  - Megabit/s (Mb/s) circa 10 milioni di bit al secondo

34

## La larghezza di banda

Una banda stretta è come una strada a due corsie

Una banda larga è come un'autostrada

*La larghezza di banda*

35

## Perché la banda larga?

- Banda larga significa:
  - più servizi (TV, telefonia tradizionale, “on demand”, etc.) sulla stessa linea e senza cambiare il supporto fisico;
  - contenuti più “pesanti” (e interessanti).
- Ecco quanti dati si devono trasmettere per un video di qualità media:

Elemento	Quantità
Pixel dello schermo (640 × 480)	307 200
Bit per pixel	8
Totale bit dello schermo	2 457 600
Fotogrammi per secondo	30
Totale bit per secondo	73 728 000

*Le cifre della trasmissione video*

36

## LARGA E ULTRALARGA

- **BANDA LARGA:** Il termine banda larga (in lingua inglese *broadband*), nel campo delle telecomunicazioni, indica generalmente la trasmissione e ricezione di dati informativi ad una velocità di connessione superiore a **2Mbit/s**.
- **BANDA ULTRALARGA:** Si parla di banda ultralarga (ultra broadband) in riferimento ad una velocità di connessione effettiva in download di almeno **30 Mbit/s (fino a 1Gb/s)**  
Le reti che possono fornire tali velocità vengono indicate come reti NGA (Next Generation Access).

Angela Peduto - anpeduto@unisa.it

37

## Banda Ultra Larga = bene primario

- La Banda Ultra Larga è ormai considerabile come un bene primario.
- L'esigenza diffusa di connettività veloce, sia fissa sia mobile, richiede infrastrutture strategiche in grado di fornire prestazioni elevate e di risultare "a prova di futuro", cioè adeguatamente flessibili e scalabili per veicolare servizi innovativi come, ad esempio, quelli in ambito smart cities, Internet of Things, le auto a guida autonoma, le case intelligenti.

Angela Peduto - anpeduto@unisa.it

38

## Investimenti PNRR

# 27%

delle risorse del PNRR destinate alla transizione digitale

# 6,7

miliardi di € dedicati ai progetti per la Banda Ultra Larga

Tutte le misure:

<https://padigitale2026.gov.it/misure>

Angela Peduto - anpeduto@unisa.it

39

- Il 24 maggio 2022 è stato assegnato il bando **Italia a 1Giga** per portare Internet veloce a circa 7 milioni di indirizzi civici distribuiti su tutto il territorio italiano. Il bando partiva da una base d'asta di circa 3,6 miliardi di euro per le 14 aree geografiche aggiudicate (cosiddetti lotti): ne sono stati assegnati 3,4 miliardi.
- **Italia 5G** è il piano di intervento pubblico nazionale per incentivare la realizzazione delle infrastrutture di rete mobile per lo sviluppo e la diffusione del 5G. L'Italia è stata il primo Paese nell'Unione ad aver assegnato diritti d'uso dello spettro radio in tutte le tre bande di frequenze

Il 28 gennaio 2022 sono stati pubblicati i **bandi** (tutto il territorio italiano suddiviso in otto aree geografiche):

- **Scuole connesse** per connettere quasi 10 mila sedi scolastiche. Il bando è stato assegnato il 6 giugno 2022, per circa 166 milioni.
- **Sanità connessa**, che interessa oltre 12 mila strutture del servizio sanitario pubblico, il 6 giugno 2022 è stato aggiudicato per circa 314 milioni

Angela Peduto - anpeduto@unisa.it

40



Internet of Things (IoT)

Da un computer in ogni casa ad un computer in ogni cosa

Angela Peduto - anpeduto@unisa.it

41

## Il progresso diventa esponenziale

**I visionari:**  
[1975]

**Ieri:**

**Oggi:**

“Nel futuro vedo un computer su ogni scrivania e uno in ogni casa”. Bill Gates

- ➔ Un pc per ogni persona, anzi, più di uno.
- ➔ I notebook: un pc in ogni borsa
- ➔ Gli smartphone: un pc in ogni taschino

L'Internet of Things: un pc in ogni cosa



1995  
"La casa digitale"  
Negroponte  
"La casa digitale"  
Bill Gates



42

# IoT

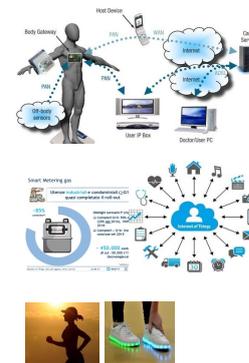
- è nato negli ambienti di ricerca del MIT di Boston .
- Un concetto semplice: tutti gli oggetti possono acquisire un ruolo attivo grazie al collegamento alla rete.
- Gli oggetti si rendono riconoscibili e acquisiscono intelligenza grazie al fatto di poter comunicare dati su se stessi e accedere ad informazioni aggregate da parte di altri.

Angela Peduto - anpeduto@unisa.it

43

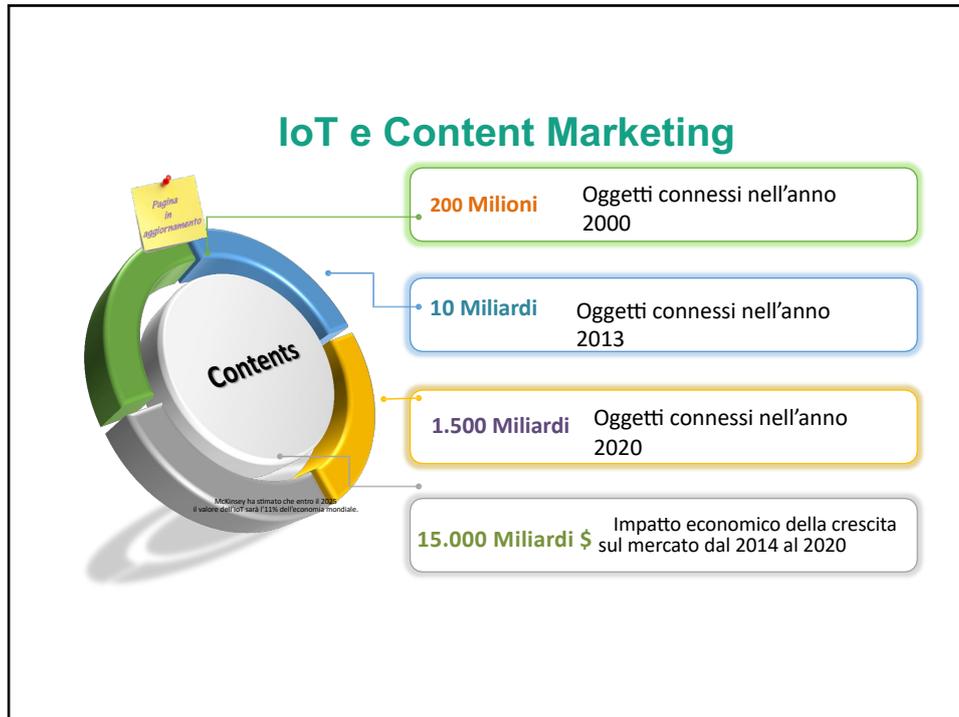
## Esempi di utilizzo IoT

- Le sveglie suonano prima in caso di traffico,
- le piante comunicano all'innaffiatoio quando è il momento di essere innaffiate,
- le scarpe da ginnastica trasmettono tempi, velocità e distanza per gareggiare in tempo reale con persone dall'altra parte del globo,
- i vasetti delle medicine avvisano i familiari se si dimentica di prendere il farmaco.



Angela Peduto - anpeduto@unisa.it

44



45

## SMART CITY

- Uno smart building consuma meno ed è più sicuro, è un edificio che pulsa come il cuore di un organismo in risposta alle esigenze interne ed agli stimoli dell'ambiente esterno
- Una smart city è una proiezione astratta di comunità del futuro, definita attraverso bisogni che trovano risposte in tecnologie, servizi e applicazioni, a partire dagli edifici, capaci di scambiare dati ed integrarsi con il sistema città.

Angela Peduto - anpeduto@unisa.it

46

## Big data

- L'immensa mole di dati generati quotidianamente da dispositivi elettronici di ogni genere (v. IoT) connessi alla Rete, genera l'esigenza di gestire e raccogliere un'enorme quantità di informazioni, analizzarla per individuare correlazioni e pattern interpretativi al suo interno e infine riutilizzarla dopo averne tratte le dovute conclusioni.
- Si tratta di trasformare dati grezzi in informazioni utili e in molti casi per creare modelli predittivi, per risolvere problemi in questo o quel campo di attività. (Sicurezza, meteo, business intelligence, salute, smart city, ecc.)

Angela Peduto - anpeduto@unisa.it

47



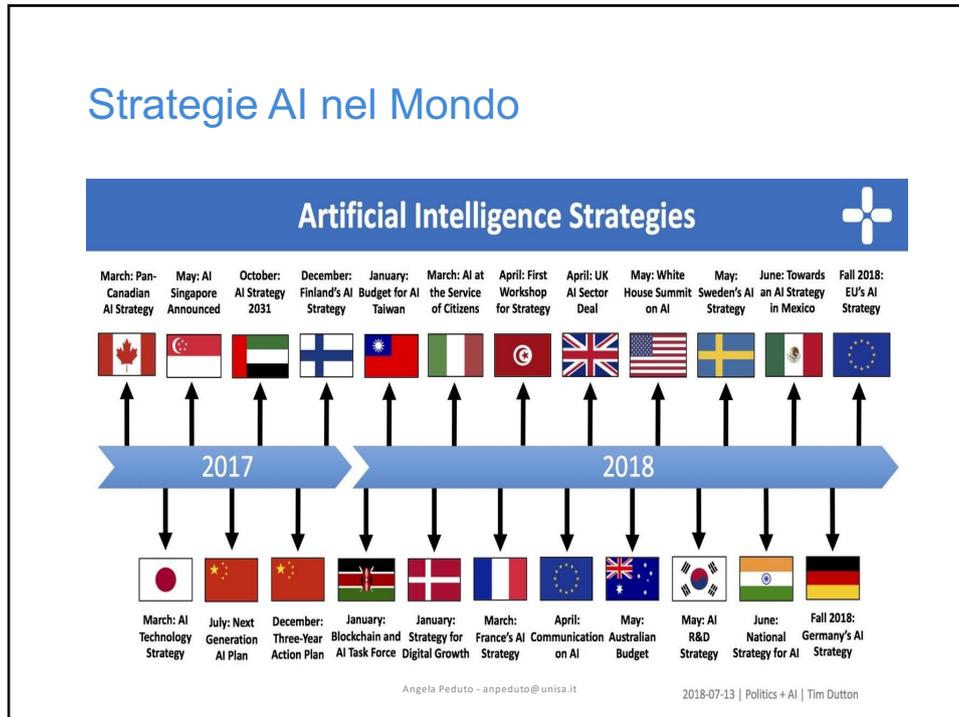
IA

Focus su machine learning

Angela Peduto - anpeduto@unisa.it

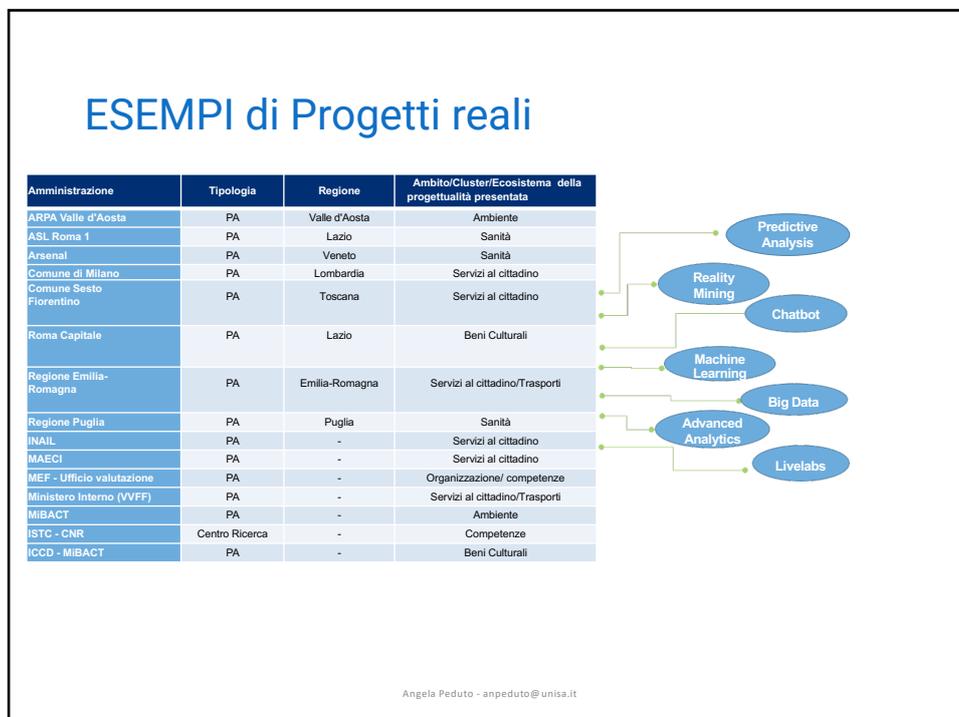
48

## Strategie AI nel Mondo



49

## ESEMPI di Progetti reali



50

## Machine Learning - DEFINIZIONE

- Il Machine Learning è quella branca dell'informatica che dà ai computer la possibilità di imparare qualcosa senza che questo gli venga esplicitamente insegnato.
- Il "fare cose che prima non potevamo fare" o "automatizzare i processi" non sono elementi caratteristici del Machine Learning.



Angela Peduto - anpeduto@unisa.it

51

## PAROLA CHIAVE 1

1 **imparare** qualcosa

"imparare" significa che viene acquisita una conoscenza e che questa può essere messa in pratica anche in futuro.

Angela Peduto - anpeduto@unisa.it

52

## PAROLA CHIAVE 2

2 senza che gli venga **esplicitamente** insegnato  
l'intervento umano è necessario affinché il  
computer possa apprendere, ma tale intervento  
non consiste nel dire esplicitamente al computer  
come comportarsi.

Angela Peduto - anpeduto@unisa.it

53

## DEFINIZIONE 2

- Il Machine Learning esplora lo studio e la costruzione di **algoritmi** che possono **imparare dai dati** e sulla base di questi fare delle previsioni.
- **data-driven predictions**

Angela Peduto - anpeduto@unisa.it

54

## STORIA

Sebbene le applicazioni di Machine Learning stiano dilagando solo in questi anni, il termine è stato introdotto nel 1959 da Arthur Samuel quando lavorava presso IBM.

Alcuni degli algoritmi risalgono ancora a prima del 1959.

Il termine "regressione", ad esempio, è stato coniato da Galton nel 1889, ed il primo algoritmo di regressione lineare fu realizzato addirittura agli inizi del 1800 da Legendre e Gauss.

Angela Peduto - anpeduto@unisa.it

55

## CASI D'USO

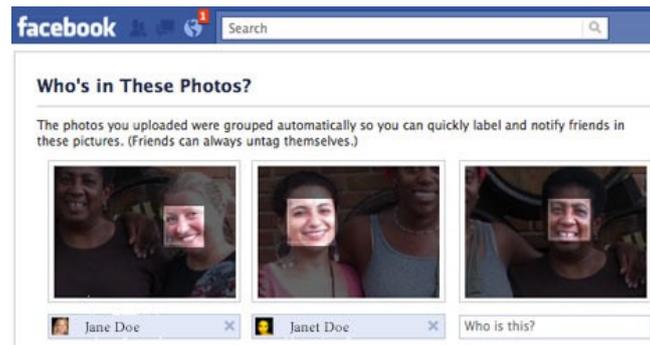
Il Machine Learning viene utilizzato in tantissimi settori, ... vediamo insieme alcuni esempi.

Angela Peduto - anpeduto@unisa.it

56

## CASI D'USO - COMPUTER VISION

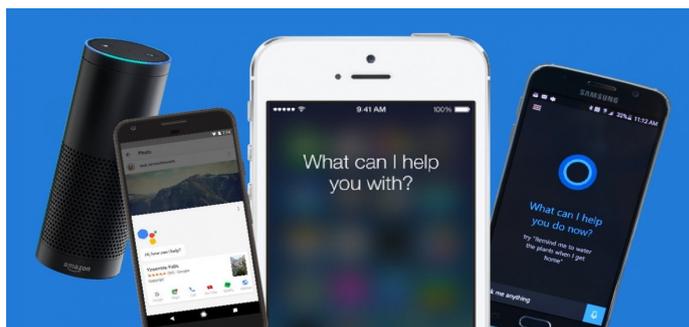
Riconoscimento volti su Facebook, identificazione di persone, oggetti, animali, paesaggi, ... in un'immagine



Angela Peduto - anpeduto@unisa.it

57

## CASI D'USO - RICONOSCIMENTO VOCALE



Siri, Ok Google, Cortana, Alexa.  
Riconoscimento del parlato e sua interpretazione per capire cosa la persona ha chiesto e cosa si aspetta come risposta.

Angela Peduto - anpeduto@unisa.it

58

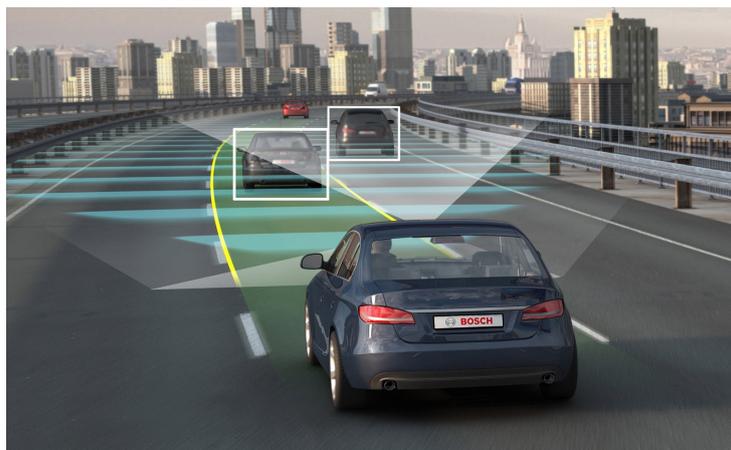
Per suggerire cosa guardare, comprare, leggere, ...



59

## CASI D'USO - SELF-DRIVING CARS

Il Machine Learning ha persino permesso le auto che si guidano da sole!



Angela Peduto - anpeduto@unisa.it

60

## CASI D'USO - IDENTIFICAZIONE DELLO SPAM

Sistema di classificazione per individuare le email di spam



61

## CASI D'USO - ATTIVITÀ FRAUDOLENTE

Sistema di classificazione che identifica quando una transazione online (e.g. pagamento con carta di credito) è legittima e quando invece è fraudolenta

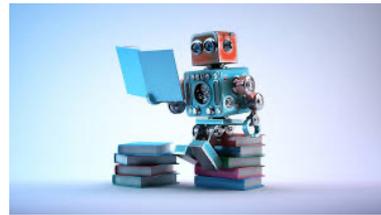


62



## TIPO DI APPRENDIMENTO

- L'apprendimento del computer nel Machine Learning viene di solito suddiviso in due grandi macro-categorie:
  - apprendimento supervisionato
  - apprendimento non-supervisionato



65

## APPRENDIMENTO SUPERVISIONATO - DEFINIZIONE

L'apprendimento supervisionato è quello in cui il computer apprende usando dei dati "etichettati".

Le etichette riportano il valore della grandezza che si vuole che il computer impari a predire, spesso indicata semplicemente come "output".

Quindi, per ogni osservazione, oltre ai valori di input, è noto anche il valore dell'output su cui il computer impara.

Angela Peduto - anpeduto@unisa.it

66

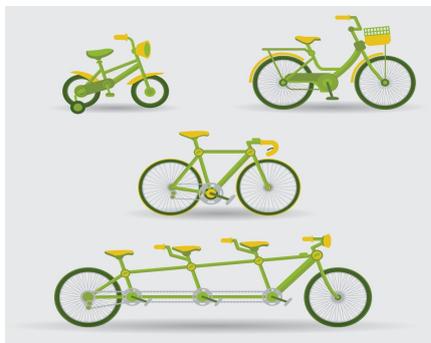
## APP. SUPERVISIONATO - ESEMPIO

Supponiamo di voler insegnare ad un bambino a riconoscere una bicicletta da una macchina, senza che lui ne abbia mai viste in vita sua.



67

Mostriamo al bambino le immagini con le biciclette e gli diciamo che quelle sono biciclette, lasciando che lui capisca da solo quali sono le caratteristiche delle biciclette.



Angela Peduto - anpeduto@unisa.it

68

Facciamo poi lo stesso con le automobili.



69

Infine, facciamo vedere al bambino una nuova immagine, mai usata prima, e gli chiediamo se si tratta di un'automobile o di una bicicletta.



Angela Peduto - anpeduto@unisa.it

70

## APPRENDIMENTO NON SUPERVISIONATO - DEFINIZIONE

L'apprendimento non-supervisionato, invece, è quello in cui il computer apprende usando dati che **non sono etichettati**.

Il computer deve dunque imparare senza conoscere l'output, ma avendo a disposizione solo i dati di input.

Si tratta di una tipologia di problemi normalmente più complessa di quella supervisionata.

Angela Peduto - anpeduto@unisa.it

71

## APP. NON SUPERVISIONATO - ESEMPIO

Ripetiamo l'esempio precedente in modalità non supervisionata.

Possiamo usare sempre lo stesso dataset, ma questa volta quando facciamo vedere le foto al bambino, non gli diciamo quali sono le biciclette e quali le automobili. Invece, glielie facciamo vedere tutte insieme.

Il bambino deve capire da solo che le immagini appartengono a due categorie diverse, perchè hanno delle caratteristiche diverse.

Angela Peduto - anpeduto@unisa.it

72



73

## APP. NON SUPERVISIONATO - NOTE

Ci sono due note importanti da fare:

- il bambino potrebbe trovare più di due categorie
- le categorie trovate potrebbero essere molto diverse da quanto ci aspettiamo

Ad esempio, il bambino potrebbe decidere di raggruppare le immagini in base al colore, o alla dimensione, e avrebbe tutto il diritto di farlo!

In genere il risultato ha una forte dipendenza dalle immagini presenti nel dataset.

Angela Peduto - anpeduto@unisa.it

74

## RIEPILOGO

Esistono altri tipi di apprendimento, come quello Semi Supervisionato o Rinforzato, che sono in qualche modo delle evoluzioni o combinazioni dei precedenti.

Nome	Caratteristica
Supervisionato	i dati sono etichettati
Non Supervisionato	i dati non sono etichettati
Semi Supervisionato Rinforzato	alcuni dati sono etichettati con un tipo particolare di etichette

Angela Peduto - anpeduto@unisa.it

75

## SUDDIVISIONE PER OUTPUT

Un altro modo di suddividere i problemi di machine learning si basa sul tipo di output che si intende ottenere.

Ci sono tre casi:

- Regressione
- Classificazione
- Clustering

Angela Peduto - anpeduto@unisa.it

76

## REGRESSIONE

Se l'output è un valore **continuo**, si parla di "Regressione".

Esempi:

- stimare lo stipendio di una persona in base al titolo di studio, genere ed età;
- stimare il peso di una persona in base ad età, genere ed altezza.

Di solito sono casi supervisionati.

Angela Peduto - anpeduto@unisa.it

77

## CLASSIFICAZIONE

Se l'output è un valore **qualitativo (discreto)**, si parla di "Classificazione".

Esempi:

- valutare se un'email è spam o no in base al suo oggetto;
- valutare se un testo tratta di qualcosa o di qualcuno in maniera positiva o negativa (*sentiment analysis*).

Di solito sono casi supervisionati.

Angela Peduto - anpeduto@unisa.it

78

## CLUSTERING

Se l'output è la suddivisione dei dati di input in **gruppi**, si parla di "Clustering".

Esempi:

- identificare se esistono gruppi di utenti con comportamento simile su un e-commerce;
- scoprire se esistono relazioni nascoste tra i pazienti di un ospedale.

Di solito sono casi non-supervisionati.

Angela Peduto - anpeduto@unisa.it

79



80

## L'ALGORITMO MIGLIORE per un dato problema

Ogni algoritmo è indicato solo per alcuni tipi di problemi e non per altri.

Quindi il primo passo da fare nella scelta dell'algoritmo per il proprio problema di Machine Learning è capire se ci si muove in uno scenario di apprendimento supervisionato o non-supervisionato, e se è di classificazione, regressione o clustering.

Angela Peduto - anpeduto@unisa.it

81

## Machine Learning Algorithms *(sample)*

	<u>Unsupervised</u>	<u>Supervised</u>
<u>Continuous</u>	<ul style="list-style-type: none"> <li>• Clustering &amp; Dimensionality Reduction               <ul style="list-style-type: none"> <li>○ SVD</li> <li>○ PCA</li> <li>○ K-means</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Regression               <ul style="list-style-type: none"> <li>○ Linear</li> <li>○ Polynomial</li> </ul> </li> <li>• Decision Trees</li> <li>• Random Forests</li> </ul>
<u>Categorical</u>	<ul style="list-style-type: none"> <li>• Association Analysis               <ul style="list-style-type: none"> <li>○ Apriori</li> <li>○ FP-Growth</li> </ul> </li> <li>• Hidden Markov Model</li> </ul>	<ul style="list-style-type: none"> <li>• Classification               <ul style="list-style-type: none"> <li>○ KNN</li> <li>○ Trees</li> <li>○ Logistic Regression</li> <li>○ Naive-Bayes</li> <li>○ SVM</li> </ul> </li> </ul>

82

## COME FARE LA SCELTA FINALE?

- La teoria ci aiuta a fare la prima selezione.
- Poi però dobbiamo vedere i risultati sul campo.
- Procediamo dunque con la **misura dell'errore**.
  - Si costruiscono diversi modelli, si calcola l'errore per ciascuno e si sceglie quello con i risultati migliori.

Angela Peduto - anpeduto@unisa.it

83

## L'errore

La misura dell'errore ci aiuta sia a confrontare modelli che usano algoritmi diversi, sia che usano lo stesso algoritmo ma con impostazioni diverse.



Angela Peduto - anpeduto@unisa.it

84

## DEFINIZIONE DI ERRORE

Per semplificare l'esposizione dei concetti, consideriamo il caso dell'apprendimento supervisionato di regressione, ad esempio un modello che stima l'altezza di una persona in base a peso, genere ed età.

L'**errore** (o residuo) del nostro modello è la **differenza** che c'è **tra l'altezza stimata** dal modello e **l'altezza osservata**.

Angela Peduto - anpeduto@unisa.it

86

## NOTAZIONE - OSSERVAZIONI

Rappresentiamo i dati di input per ciascuna persona del nostro dataset con un vettore:

$$x_i \equiv (x_1, x_2, \dots, x_p)_i$$

Nel nostro caso  $p = 3$  e la terna  $(x_1, x_2, x_3)_i$  rappresenta **genere**, **peso** ed **età** della persona  $i$ -esima.

Chiamiamo poi  $y_i$  l'altezza misurata della persona  $i$ -esima.

Angela Peduto - anpeduto@unisa.it

87

## NOTAZIONE - PREVISIONI

Costruire un modello di machine learning significa trovare quella funzione  $f$  per cui:

$$y_i = f(x_i)$$

Due esempi:

$$f(m, 70 \text{ kg}, 35 \text{ yr}) = 170 \text{ cm}$$

$$f(w, 60 \text{ kg}, 35 \text{ yr}) = 165 \text{ cm}$$

Angela Peduto - anpeduto@unisa.it

88

## VARIABILITÀ INTRINSECA

In realtà sappiamo bene che persone con lo stesso peso, genere ed età possono avere altezze diverse.

Ad esempio, nel nostro dataset potremmo avere le seguenti 3 osservazioni:

$$x_3 = (m, 70 \text{ kg}, 35 \text{ yr}) \text{ e } y_3 = 170 \text{ cm}$$

$$x_4 = (m, 70 \text{ kg}, 35 \text{ yr}) \text{ e } y_4 = 180 \text{ cm}$$

$$x_5 = (m, 70 \text{ kg}, 35 \text{ yr}) \text{ e } y_5 = 175 \text{ cm}$$

Angela Peduto - anpeduto@unisa.it

89

## VARIABILITÀ INTRINSECA

- La nostra funzione  $f$  invece, a parità di input, restituisce sempre lo stesso valore di output.
- Questo implica che il fenomeno ha una variabilità intrinseca che rende impossibile la creazione di un modello perfetto.



90

## ERRORE IRRIDUCIBILE

Per questo motivo si introduce il concetto di **errore irriducibile**, chiamato  $S_{irr}$ , che rappresenta una misura della variabilità intrinseca del fenomeno in esame:

$$y_i = f(x_i) + S_{irr}$$

91

## ERRORE DEL MODELLO

In aggiunta all'errore irriducibile, c'è poi l'errore dovuto al modello di machine learning, indicato con  $s$ .

Questo errore è dovuto al fatto che **non sappiamo qual è la vera forma di  $f$**  e dobbiamo invece fare una stima anche per questa.

Se chiamiamo  $\hat{f}$  questa stima, possiamo scrivere:

$$f(x_i) = \hat{f}(x_i) + s$$

Oppure, indicando con  $\hat{y}_i$  l'altezza stimata dal modello:

$$f(x_i) = \hat{y}_i + s$$

Angela Peduto - anpeduto@unisa.it

92

## ERRORE TOTALE

A questo punto possiamo mettere insieme i due pezzi e scrivere l'altezza osservata di una persona come:

$$y_i = \hat{y}_i + s + s_{irr}$$

Dove:

- $\hat{y}_i$  è l'altezza stimata dal modello
- $s$  è l'errore del modello (può essere ridotto)
- $s_{irr}$  è la variabilità del fenomeno (non può essere ridotto)

Angela Peduto - anpeduto@unisa.it

93

## ERRORE DEL MODELLO

Tutti gli sforzi relativi alla ricerca del miglior modello di machine learning puntano a rendere minimo l'errore  $s$ .

E' possibile dimostrare che tale errore è la somma di due parti:

$$s = \text{Bias} + \text{Varianza}$$

Angela Peduto - anpeduto@unisa.it

94

## RELAZIONE TRA BIAS E VARIANZA

Purtroppo si dimostra che **bias e varianza sono inversamente correlati**, cioè tanto più si diminuisce il bias tanto più aumenta la varianza e viceversa.

Per rendere minimo l'errore totale di un modello quindi bisogna puntare a rendere minima **la somma** di bias e varianza, e non uno solo dei due.

Angela Peduto - anpeduto@unisa.it

95

## DEFINIZIONE DI BIAS

Un algoritmo di machine learning ha un certo bias se, quando viene addestrato su diversi dataset (tutti ugualmente buoni), l'output che restituisce è **sistematicamente sbagliato**.

L'errore è "sistematico" nel senso che le sue caratteristiche non variano al variare dei dataset di training.

Angela Peduto - anpeduto@unisa.it

96

## DEFINIZIONE DI VARIANZA

Un algoritmo di machine learning ha un'alta varianza se, quando viene addestrato su diversi dataset (tutti ugualmente buoni), l'errore che si commette è **significativamente diverso** per alcuni dei dataset utilizzati nell'addestramento.

Questa situazione viene descritta anche con il termine "**over-fitting**".

Angela Peduto - anpeduto@unisa.it

97



99

## Il sistema di Intelligenza Artificiale Prometea

- grazie al lavoro svolto congiuntamente da un team di funzionari e magistrati della Procura della città di Buenos Aires e di esperti nel campo dell'Intelligenza Artificiale è stato sviluppato l'innovativo sistema di Intelligenza Artificiale Prometea
- Questo sistema di Intelligenza Artificiale consente un'accelerazione esponenziale dei tempi di preparazione di documenti legali e amministrativi con un impatto che appare decisivo, nel sistema argentino, rispetto a tematiche cruciali quali l'effettività dei diritti, il principio di uguaglianza e il principio di certezza del diritto

Angela Peduto - anpeduto@unisa.it

100

## Il sistema Prometea si basa su cinque livelli di innovazione principali:

1. la creazione di una interfaccia utente intuitiva, che può essere attivata con comando vocale (conversazione) o come chat;
2. la creazione di una interfaccia dello schermo unica e integrata, per ridurre i clic ed eliminare l'apertura di finestre digitali;
3. una migliore gestione dei dati e delle informazioni, automatizzando le attività finalizzate a creare documenti che fungono da base per l'adozione di decisioni maggiormente prevedibili;
4. lo sviluppo di molteplici funzioni di assistenza digitale, come la ricerca avanzata di leggi, documenti, reporting e statistiche, etc.;
5. la possibilità di fare previsioni sulla base dell'utilizzo di un sistema di apprendimento automatico supervisionato, con un tasso medio di successo del 96,34 %.

Angela Peduto - anpeduto@unisa.it

101

## Il passato

- il lavoro richiesto per avviare una gara di appalto di beni e servizi completamente digitalizzati nella città di Buenos Aires, come l'acquisto di "accessori per computer", implicava una media di:
  - 29 giorni lavorativi,
  - 670 clic,
  - l'apertura di circa 60 finestre
  - il "copia e incolla" di numerosi dati.

Angela Peduto - anpeduto@unisa.it

102

## Presente e futuro

- Prometea è stata addestrata a generare un elenco di condizioni generali, un elenco di condizioni specifiche nell'acquisizione di merci (ad esempio, computer), e per completare un progetto di atto amministrativo di approvazione delle specifiche e di bando per la selezione del co-contraente. Il tutto:
  - nello spazio **di quattro minuti**,
  - con l'utilizzo di una sola schermata
  - attraverso un sistema di domande e risposte tramite comando vocale (chat di conversazione)

Angela Peduto - anpeduto@unisa.it

103




## TREE4NB – Training, Reorganizing, Evaluating, Enabling For Natural Birth

- soluzione innovativa per il supporto decisionale clinico e di governance e il potenziamento del *patient empowerment* attraverso valutazioni previsionali di *risk assessment* derivate da un processo di correlazione di informazioni provenienti da dataset clinici ed extra-clinici e basate su algoritmi di **Machine Learning e Artificial Intelligence**.

Angela Peduto - anpeduto@unisa.it

104

## Modello di classificazione

- Il mio contributo, nei tre mesi in cui ho collaborato alla fase di sperimentazione e validazione della soluzione tecnologica complessiva ottenuta come *outcome* delle attività di ricerca del progetto TREE4NB:
  - Valutazione del modello di classificazione (Machine Learning)
  - Usabilità della Mobile App.

Angela Peduto - anpeduto@unisa.it

105

## Le fasi del machine learning

Pre-Processing

Apprendimento

Valutazione.

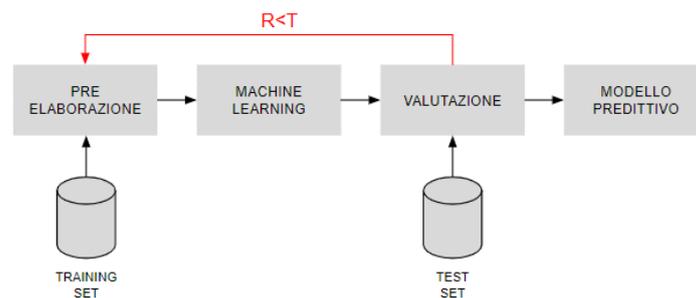
Predizione.

Angela Peduto - anpeduto@unisa.it

106

## Valutazione

- Per valutare la qualità predittiva fisso una percentuale T di risposte esatte come parametro di riferimento.
- Provo il modello predittivo con un dataset di test ( test set ), diverso dal dataset di apprendimento, e analizzo i risultati.
  - $R \geq T$  **promosso**
  - $R < T$  **bocciato**
- I dati di training e di test saranno quelli forniti dal AOU San Giovanni di Dio e Ruggi d'Aragona.

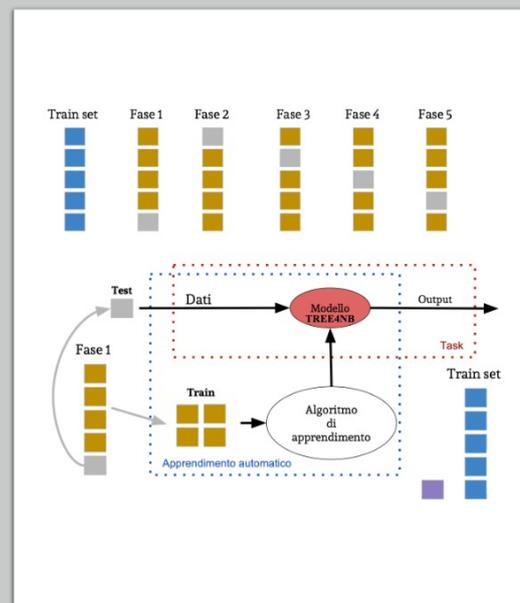


Angela Peduto - anpeduto@unisa.it

107

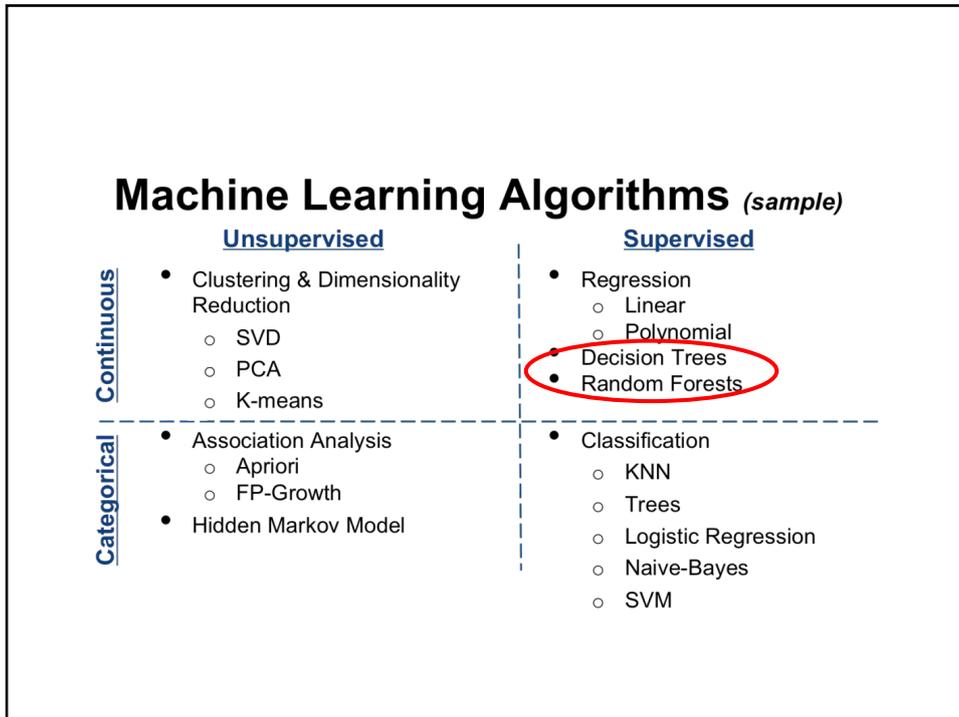
## cross validation

- suddividendo il training set in k parti di uguale dimensione.
- Poi selezionando una parte 1/k per utilizzarla come *validation set*.
- Le restanti parti k-1/k invece continueranno a comporre il training dataset.
- Si dovrà ripetere la stessa procedura per k volte

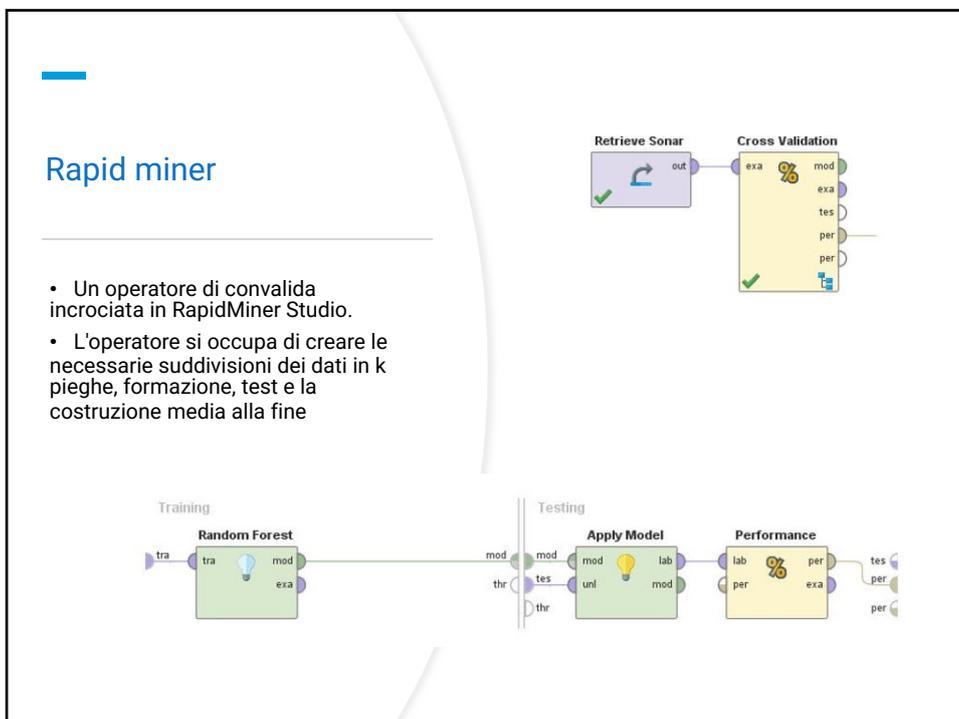


Angela Peduto - anpeduto@unisa.it

108



109



110

## classificazione di Robson

- classificazione di Robson che divide le pazienti sottoposte a TC in 10 classi mutuamente esclusive definite in base alla parità, numero di feti, presentazione fetale, età gestazionale e decorso del travaglio e del parto)

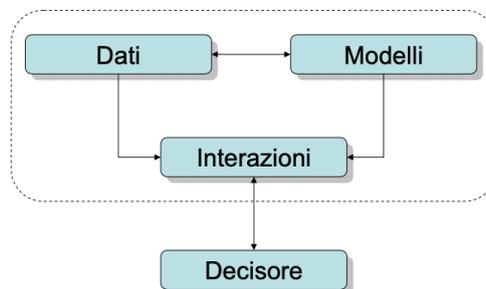
CLASSE	
I	Nullipare, Grav. singola, pres. Cefalica, > = 37 sett. Travaglio spontaneo
II	Nullipare, Grav. Singola , pres. Cefalica, >= 37 sett. Travaglio indotto o TC prima del travaglio
III	Multipare (esc.preg. TC) Grav. Singola, pres. Cefalica, > = 37 sett. Travaglio spontaneo
IV	Multipare (esc. preg. TC) grav. Singola, pres. Cefalica, > = 37 sett. Travaglio indotto o TC prima del travaglio
V	Pregresso TC, Grav. Singole in pres. Cefalica, >= 37 sett. Preg. TC indotte
VI	Nullipare Presentazione podalica. Indipendentemente dall'età Gest
VII	Pluripare Presentazione podalica. Indipendentemente dall'età Gest
VIII	Grav. Multiple anche con preg. TC
IX	Posizioni anomale anche con preg. TC
X	Tutte la grav. Singole, cefaliche, inclusi i preg. TC Età Gest. < 36 sett.

Angela Peduto - anpeduto@unisa.it

111

## Definizione di DSS

- Un sistema informatico interattivo che aiuta il decision maker a utilizzare dati e modelli per risolvere problemi semi-strutturati e non strutturati



112

## Mobile app

- La mobile app è una sorta di interfaccia del DSS.

113

## Conclusioni

- I tassi di taglio cesareo (TC) sono importanti indicatori globali per misurare l'accesso ai servizi ostetrici.
- Avere un sistema di supporto decisionale basato sulla Robsons Classification è sufficiente per ridurre i tassi di CS?
- L'analisi mensile dei gruppi combinati da 1 a 4 dalla Classificazione di Robsons consente lo sviluppo di obiettivi per migliorare l'assistenza e anche prevenire i TC nei gruppi di partorienti ove prevenibile.

Angela Peduto - anpeduto@unisa.it

114