

Software Visuali per Analisi Avanzate

Corso sostitutivo di Tirocinio



Dott. Ing. Valerio Morfino – Software Visuali per Analisi Avanzate aa. 2020-2021

1

Docente del
corso

VALERIO MORFINO

Solution Principal Big Data & Analytics @DXC

Professore a contratto per il corso sostitutivo di tirocinio in Software Visuali per Analisi Avanzate presso il DEMM – Università del Salerno a.a. 2020-2021

Valerio Morfino si occupa di informatica e di Internet dal 2000. Laureato in Ingegneria Informatica, nel corso della propria carriera ha lavorato in società di consulenza, università, grandi e medie aziende occupandosi di consulenza, formazione, ricerca, direzione di progetti. Autore di articoli scientifici, relatore in conferenze e meet-up su temi relativi a web, e-commerce, machine learning, bioinformatica, cyber security, analitica avanzata.

E-mail: vmorfino@unisanno.it



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

2

Sommario

Presentazione del corso

Big Data

Analitica Avanzata e Machine learning

Visual analytics

Gli strumenti del corso: SaS e Tableau



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

3

Obiettivi del corso



Pontetra l'Università e le Aziende



Fornire una conoscenza del contesto dell'analitica avanzata



Fornire competenze su tecnologie e prodotti utili alle aziende



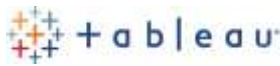
Sensibilizzare e stimolare skill utili per lavorare in azienda



Avviare percorsi di certificazioni di prodotto

4

Tecnologie



5

Skill



Comprendere Requisiti e Necessità del cliente (Interno o Esterno all'azienda)



Risolvere i problemi in modo efficace ed efficiente



Comunicare i risultati ottenuti in modo corretto, comprensibile e interessante



Trovare e comprendere le informazioni rilevanti: con cui interagire

6

Calendario del corso
Parte 1



Giorno	Orario	Argomenti
Venerdì 12/3 16.00 - 18.00	2 ore	Presentazione del Corso Il contesto dei Big Data Analitica Avanzata Visual Analytics
Giovedì 17/3 16.00 - 18.00	4 ore	Visual Analytics, il Quadrante Gartner per la Visual Analytics
Venerdì 18/3 9.30 - 13.30	2 ore	Tableau
Venerdì 19/3 16.00 - 18.00	2 ore	Tableau
Giovedì 25/3 9.30 - 13.30	4 ore	Design visuale - Ospite Tableau - Esercitazione finale
Venerdì 26/3	2 ore	il Quadrante di Gartner per Advanced Analytics Ripasso concetti di Machine Learning Sas: introduzione, ambiente operativo
Giovedì 8/4	4 ore	SaS - Ospite SaS
Venerdì 9/4	2 ore	SaS
Giovedì 15/4	4 ore	SaS - Esercitazione finale

Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

7

Calendario del corso
Parte 1
Alternativa



Giorno	Orario	Argomenti
Venerdì 12/3 16.00 - 18.00	2 ore	Presentazione del Corso Il contesto dei Big Data Analitica Avanzata Visual Analytics
Venerdì 19/3 16.00 - 18.00	2 ore	Visual Analytics, il Quadrante Gartner per la Visual Analytics
Sabato 20/3 9.00 - 13.00	4 ore	Tableau
Venerdì 26/3 16.00 - 18.00	2 ore	Design visuale - Ospite
Sabato 27/3 9.00 - 13.00	4 ore	Tableau Tableau - Esercitazione Finale
Venerdì 26/3 16.00 - 18.00	2 ore	il Quadrante di Gartner per Advanced Analytics Ripasso concetti di Machine Learning Sas: introduzione, ambiente operativo
Giovedì 8/4 14.00 - 18.00	4 ore	SaS - Ospite SaS
Sabato 10/4 9.00 - 13.00	4 ore	SaS SaS - Esercitazione finale

Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

8

Materiale del
corso su
GitHub

<https://github.com/valerio75/Software-Visuali-per-Analisi-Avanzate-2020-2021>

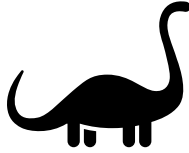
Per questa lezione:
• Lezione_1_Slide.pdf



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2019-2020

9

Big Data, cosa sono?



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

10

Intuitivamente, parliamo di volumi enormi di dati.

Ma...

- Perché memorizziamo questi dati?
- Chi genera questi dati?
- Quanti sono questi dati?



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

11



Capacità di memorizzazione

1956, IBM 350 Disk File (1.000 Kg) 5 MB
2020, Usb DataTraveler (pochi grammi) 2 TB



12

I dati hanno un grande valore

- Offerta personalizzata di prodotti
- Programmi TV personalizzati in base ai gusti
- Previsione della domanda del mercato
- Previsione dei trend globali
- Previsione del traffico
- Ricerca scientifica
- Medicina personalizzata
- Prezzi dinamici in base alla domanda
- Il resto lo scopriremo con il tempo

Università Ca' Foscari di Venezia
Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

13

Chi genera questi dati?

- Persone, tramite programmi
 - Social network
 - Notizie
 - Blog
 - Posta elettronica
- Programmi
 - Sistemi di log
 - Sistemi di tracciamento
 - Software gestionali
 - Sistemi CRM
 - Motori di ricerca
 - E-commerce
 - Sistemi di Web ADV
- Macchine
 - Sensori IOT
 - Automobili
 - SmartTV
 - Telecamere
 - Smartphone

Università Ca' Foscari di Venezia
Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

14

Il fenomeno della convergenza

Università Ca' Foscari di Venezia
Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

15

Quanti dati
vengono
generati ogni
minuto?



Fonte: https://web-assets.domo.com/blog/wp-content/uploads/2017/07/17_domo_data-per-secondo-5-01.png

16

Tirando le
somme

2,5 Quintilioni di byte
generati ogni giorno

Il 90% dei dati sono
stati generati negli
ultimi due anni



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

17

Ecco a voi i **Big Data**



Volume



Varietà



Velocità



Valore

<https://www.linkedin.com/pulse/caratteristiche-del-big-data-6-valerio-morfino/>

18

Ecco a voi i **Big Data**... continua



Veridicità



Vulnerabilità



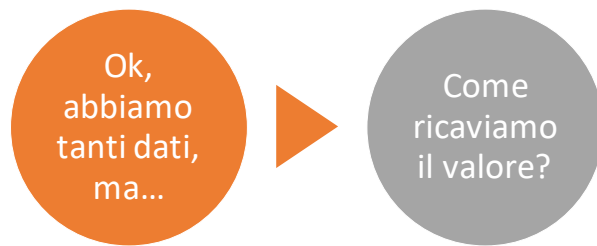
Volatility



Visualization

<https://www.linkedin.com/pulse/caratteristiche-dei-big-data-le-5-veridic-morfino/>

19




20

Tipologie di analisi

- **Analisi descrittive**, strumenti che permettono di descrivere la situazione attuale e passata di fenomeni (es. processi aziendali). Permettono di accedere ai dati in maniera interattiva (inserendo ad esempio filtri o effettuando operazioni di drill-down) e di visualizzare in modo sintetico e grafico indicatori di prestazione
- **Analisi predittive**, strumenti avanzati che effettuano l'analisi dei dati per rispondere a domande relative a cosa potrebbe accadere nel futuro (sono caratterizzati da tecniche matematiche quali regressione, *forecasting*, modelli predittivi, ecc.)
- **Analisi prescrittive**, strumenti avanzati che, insieme all'analisi dei dati, sono capaci di proporre soluzioni operative/strategiche sulla base delle analisi svolte

21

Un esempio con Google Maps



- Abbiamo appuntamento alle 10.00 presso un cliente
- Innanzitutto **vogliamo sapere che percorso fare**. Inseriamo l'indirizzo di destinazione. Viene mostrato il percorso. Stiamo usufruendo di un'analisi **descrittiva**.
- Impostiamo l'orario dell'appuntamento, in modo da **sapere a che ora partire per arrivare puntuali**. L'app ci comunica che per arrivare alle 10,00 dobbiamo partire alle 9,15. Ecco che abbiamo usufruito di un'analisi **predittiva**.
- Mentre siamo in auto e stiamo seguendo le indicazioni, viene segnalato che sul percorso c'è un ingorgo e **viene suggerita una deviazione**. Ed ecco un'analisi **prescrittiva**.

<https://www.galileo.com/nuke/big-data-analisi-descrittive-predittive-e-prescrittive-sono-velocita-marfina/>

22



MACHINE LEARNING

VISUAL ANALYTICS

Analitica avanzata

23



Machine Learning

24

Facciamo un esempio

Velocità	Azione
50	Frena
40	Accelera
30	Accelera
100	Frena

Velocità	Azione
10	?



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

25

Facciamo un altro esempio

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	Azione
12	44	A	1	Y	0	1	0.2	Y	100	0	C	11	Accelera
22	43	A	1	Y	1	1	0.33	N	200	0	C	11	Accelera
11	23	B	1	Y	1	2	0.3	N	150	0	C	12	Frena
12	43	A	1	N	1	2	0.4	N	100	0	C	13	Accelera
3	33	C	1	Y	1	1	0.1	N	150	0	C	13	Frena
2	54	B	2	Y	0	3	0.11	N	300	0	C	13	Frena
32	54	C	2	N	0	3	0.12	N	300	0	C	11	Accelera
12	33	B	3	N	0	3	0.12	N	100	1	A	12	Frena
3	45	C	1	Y	1	2	0.11	N	150	1	A	13	Accelera
2	56	A	2	Y	1	2	0.3	Y	150	0	A	13	Frena

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	Azione
2	33	B	2	Y	0	1	0.1	Y	100	0	A	12	?

26

Machine Learning supervisionato

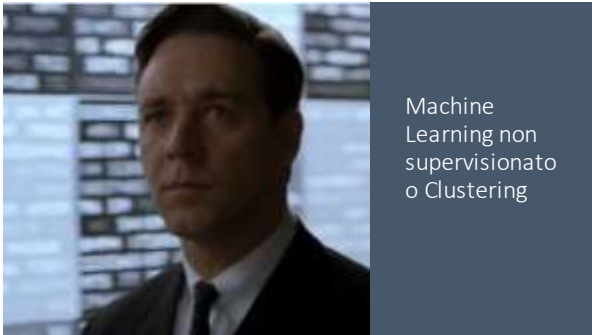
Training set	Velocità	Azione
	50	Frena
	40	Accelera
	30	Accelera
	100	Frena

Trained Model


If Velocità <= 40	=>	Accelera
Else	=>	Frena

New Data	Velocità	Azione
	10	?

27




28



Cosa abbiamo visto nel video

1. Schemi sulla parete: tanti dati incomprensibili all'occhio umano
2. Una persona, ma nel caso del Machine Learning è un algoritmo, che analizza i dati
3. Dall'analisi dei dati emerge uno schema: i numeri
4. Quel piccolo «cluster» di dati identificato è stato poi interpretato dall'analista come la coordinata sulla mappa.

In modo analogo, applicando il Machine Learning ai nostri dati possiamo ottenere conoscenza nascosta nei dati.



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

29



Clustering

<https://www.guru99.com/unsupervised-machine-learning.html>



30

Domanda

Quale è la principale differenza tra il primo ed il secondo tipo di Machine Learning?


31

Risposta


-  Nel Machine Learning supervisionato abbiamo un training set, ossia dei dati da cui un algoritmo impara una formula per poi applicarla a dati nuovi.
-  Nel Machine learning non supervisionato, viene inferita nuova conoscenza senza aver addestrato preventivamente un modello. Viene estratta nuova conoscenza esplorando i dati.

32


Machine Learning - Tipologie




Machine Learning supervisionato
Algoritmi di Intelligenza Artificiale che imparano dai dati (da esempi) cosa devono fare




Machine Learning non supervisionato
Algoritmi che scoprono informazioni, relazioni o schemi nei dati. In genere, necessitano di un analista per interpretare il risultato.



Alcuni algoritmi sono ispirati al modello del cervello (Reti Neurali)
In questa categoria rientra anche il Deep Learning

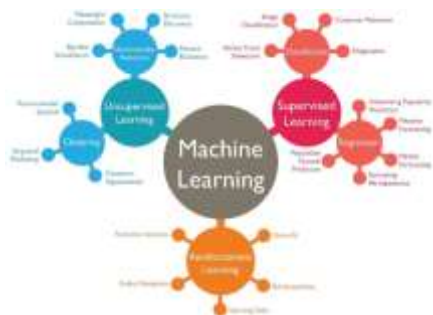


Permettono di analizzare grandi quantità di dati in modo automatico ed «ottenere» nuova conoscenza



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2020-2021

33



<https://www.kynhix.it/qualidiversi-tipi-machine-learning/215006/>

34

Applicazioni del Machine Learning

Machine Learning non supervisionato

- Clustering di informazioni (clienti, offerte, pazienti, casi, ecc.)
- Analisi delle frodi e della anomalie

Machine Learning supervisionato

- Motori di raccomandazione
- Sentiment ed Emotions Analysis
- Diagnosi di malattie
- Riconoscimento facciale
- Prevenzioni atti terroristici
- Riconoscimento scrittura e voce

Reinforcement learning

- Guida autonoma

Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

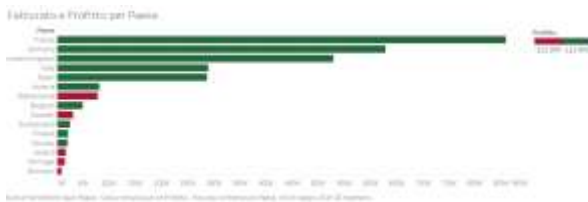
35

Esempio di API per face recognition



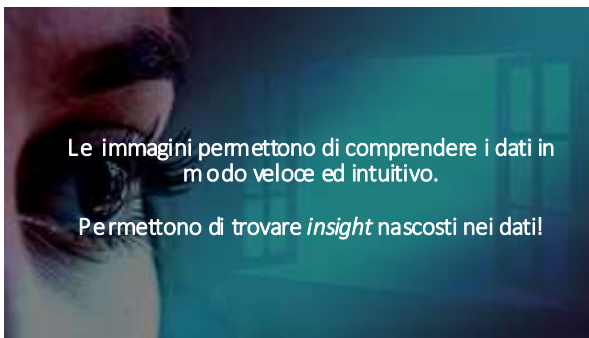
<https://api.facedetection.com/docs/#/api/cognitive-services/face/>

36



In quale paese abbiamo venduto di più?
In quale paese abbiamo profitti? E in quali perdite?

40



41



42

Come usare le immagini con i dati



INFOGRAFICA: IMMAGINI PER MOSTRARE RISULTATI

REPORTE BUSINESS INTELLIGENCE: IMMAGINI PER VISUALIZZARE I DATI

VISUAL ANALYTICS: IMMAGINI PER ANALIZZARE I DATI

Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

43

Infografica

Una rappresentazione artistica per mostrare degli insight



50% choose color colors

COLOR PREFERENCE

8% RED

5% ORANGE

3% BROWN

44

Reporting

Report, Grafici, Tabella su Excell o su sistemi di Business Intelligence tradizionale



45

Visual Analytics

- Analisi interattiva e visuale dei dati
- Può essere usata in modalità self-service da esperti di dominio
- Visualization ed analitica avanzata insieme



nted analytics

Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

46

Visual Analytics – Funzionalità chiave



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

47

Gli strumenti del corso



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021

48

Tableau

- E' una soluzione di Visual Analytics, ossia un software che permette di raccogliere, aggregare ed analizzare i dati in modo visuale
- Rappresenta un'evoluzione dei software che rientrano nella categoria della Business Intelligence
- Grazie ad un'ampia disponibilità di connettori verso numerosi sorgenti dati, permette di esplorare sorgenti dati sia tradizionali che Big Data
- Tableau è Leader nel Magic Quadrant 2020 di Gartner per le piattaforme di Modern Analytics and Business Intelligence Platforms.



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021



49

Formazione e certificazioni Tableau



- <https://www.tableau.com/it-it/learn>
- <https://www.tableau.com/learn/certification>
- <https://www.tableau.com/learn/certification/desktop-specialist>
- <https://mkt.tableau.com/files/DesktopSpecialistExamGuide.pdf>



50

Tableau – Configurazione dell'ambiente

- E' disponibile in versione commerciale ed in versione public
- Nel corso utilizzeremo la versione Public
- <https://public.tableau.com/it-it/s>
- Scaricare Tableau Public Desktop per Windows o Mac
- Creare un account su Tableau Public



Dott. Ing. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2020-2021



51

Tableau Public Home page



52

Tableau Public Home page



53

Tableau Public – Il mio profilo



54

Tableau Public – Il mio profilo



55

Tableau Public – Il mio profilo



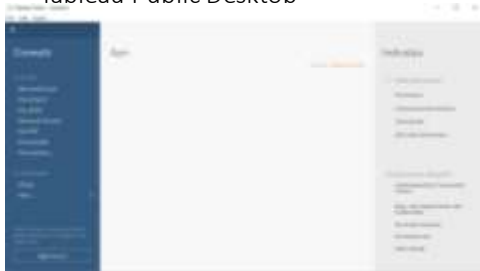
56

Tableau Public – Il mio profilo



57

Tableau Public Desktop



58

SaS Viya



- SAS Viya è un'architettura unificata con un environment analitico centralizzato per consentire la gestione end-to-end del dato, dall'esplorazione al risultato di business
- SAS è Leader nel Magic Quadrant 2020 di Gartner per le piattaforme di Data Science e Machine Learning.
- SAS Visual Data Mining and Machine Learning fornisce un'interfaccia intuitiva per velocizzare la costruzione di modelli e la generazione di codici.
- Ha funzioni di ricerca dati, feature engineering, riduzione della dimensione, analisi esplorativa, modellizzazione, training, tuning e implementazione dei modelli nei processi produttivi.



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2019-2020

59

SaS Viya



- Durante il corso utilizzeremo la piattaforma per esplorare le funzionalità di advanced analytics, con particolare riferimento al machine learning supervisionato
- Piattaforma e-learning di SaS per il corso "Machine Learning Using SAS® Viya"
- Badge pubblicabile su LinkedIn dopo 14 h di lezione e-learning
- Un contributo direttamente da SaS Italia
- Possibilità di certificarsi SaS Certified Specialist: Machine Learning Using SAS® Viya



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate aa. 2019-2020

60

SaS Viya - Certificazione



61

Grazie per l'attenzione


<https://it.linkedin.com/in/valerio-morino>


vmorino@unisannio.it

62