Software Visuali per Analisi Avanzate

Corso sostitutivo di Tirocinio

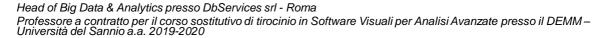


Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

1



VALERIO MORFINO



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@unisannio.it

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

Sommario

Presentazione del corso

Big Data

Analitica Avanzata e Machine learning

Visual analytics

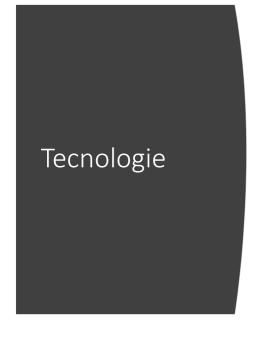
Gli strumenti del corso: SaS e Tableau



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

3

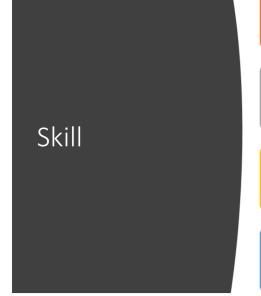








5



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

Risolvere i problemi in modo efficace ed <u>efficiente</u>

16

Comunicare i risultati ottenuti in modo corretto, comprensibile e interessante

N

Trovare e comprendere le informazioni rilevanti con cui interagire

Calendario del corso





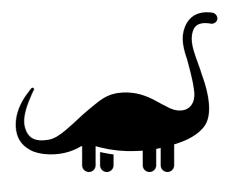


Giorno	Durata	Argomenti
Venerdì 27/3 16.00 – 18.00	2 ore	Presentazione del Corso Il contesto dei Big Data Analitica Avanzata Visual Analytics
Sabato 28/3 9.30 - 13.30	4 ore	Visual Analytic, Il Quadrante Gartner per la Visual Analytics Tableau
Martedì 31/3 16.00 – 18.00	2 ore	Tableau
Sabato 4/4 9.30 - 13.30	4 ore	Tableau - Conclusione Il Quadrante di Gartner per l'Advanced Analytics Ripasso concetti di Machine Learning Sas: Introduzione, ambiente operativo
Martedì 7/4	4 ore	SaS [Ospite SaS] *da confermare
Martedì 14/4	4 ore	SaS [Ospite SaS] *da confermare
Sabato 18/4	4 ore	SaS Conclusione del corso

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

7

Big Data, cosa sono?





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

Intuitivamente, parliamo di volumi enormi di dati.

Ma...

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

UNIVERSITÀ DEGLI STUDI DEL SANNIO BENEVENIO DEMM

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

9



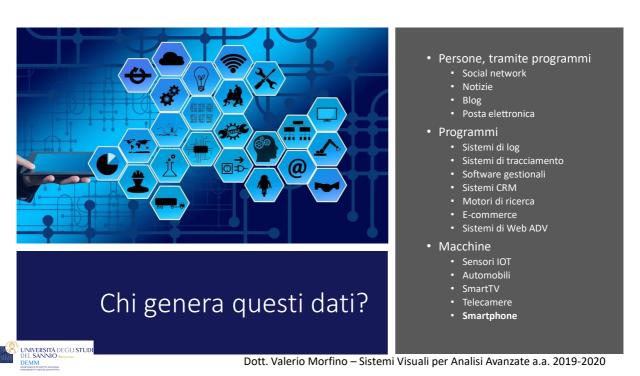


Capacità di memorizzazione

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

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





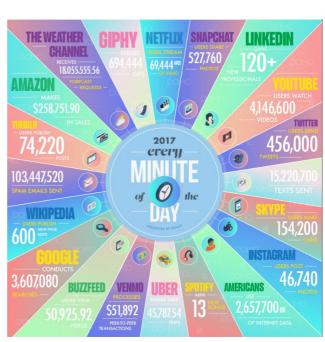


Il fenomeno della convergenza

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

13

Quanti dati vengono generati ogni minuto?



Fonte: https://web-assets.domo.com/blog/wp-content/uploads/2017/07/17_domo_data-never-sleeps-5-01.png

Tirando le somme

2,5 Quintilioni di byte generati ogni giorno

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



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

15

Ecco a voi i Big Data







Varietà



Velocità



Valore

https://www.linkedin.com/pulse/caratteristiche-dei-big-data-le-6-v-valerio-morfino/

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









Veridicità

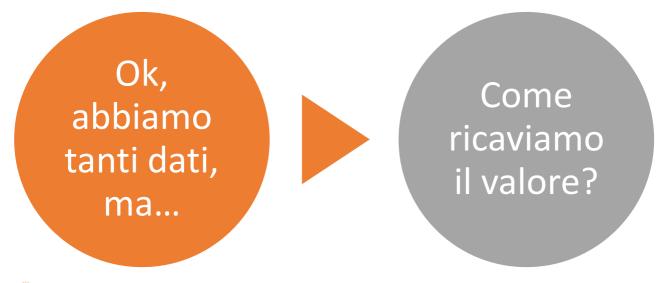
Vulnerabilità

Volatility

Visualization

https://www.linkedin.com/pulse/caratteristiche-dei-big-data-le-6-v-valerio-morfino/

17





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

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

19

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.linkedin.com/pulse/big-data-analisi-descrittive-predittive-e-cosa-sono-valerio-morfino/





MACHINE LEARNING

VISUAL ANALYTICS

Analitica avanzata

21





Machine Learning

Facciamo un esempio

Velocità	Azione
50	Frena
40	Accelera
30	Accelera
100	Frena

Velocità	Azione
10	?



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

100

23

Facciamo un altro esempio

							ı						
X1	X2	Х3	X4	X5	Х6	X7	Х8	Х9	X10	X11	X12	X13	Azione
12	44	Α	1	Υ	0	1	0.2	Υ	100	0	С	11	Accelera
22	43	Α	1	Υ	1	1	0.33	N	200	0	С	11	Accelera
11	23	В	1	Υ	1	2	0.3	N	150	0	С	12	Frena
12	43	Α	1	N	1	2	0.4	N	100	0	С	13	Accelera
3	33	С	1	Υ	1	1	0.1	N	150	0	С	13	Frena
2	54	В	2	Υ	0	3	0.11	N	300	0	С	13	Frena
32	54	С	2	N	0	3	0.12	N	300	0	С	11	Accelera
12	33	В	3	N	0	3	0.12	N	100	1	Α	12	Frena
3	45	С	1	Υ	1	2	0.11	N	150	1	Α	13	Accelera
2	56	Α	2	Υ	1	2	0.3	Υ	150	0	Α	13	Frena
X1	X2	Х3	X4	Х5	Х6	Х7	Х8	Х9	X10	X11	X12	X13	Azione

24

Machine Learning supervisionato

Training set

New Data

	Velocità	Azione
ָ ע	50	Frena
٥	40	Accelera
	30	Accelera
-	100	Frena

Trained Model

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

Velocità	Azione
10	Ş

25



Machine Learning non supervisionato o Clustering



- 1. Schemi sulla parete: tanti dati incomprensibili all'occhio umano
- 2. Una persone, 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. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020



Clustering

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

Domanda

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

29

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.

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

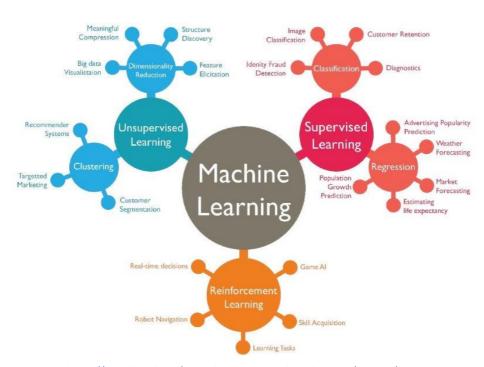


Permettono di analizzare grandi quantità di dati in modo automatico ed «ottenere» nuova conoscenza



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

31



https://www.key4biz.it/quali-diversi-tipi-machine-learning/215006/

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

Esempio di API per face recognition





https://azure.microsoft.com/it-it/services/cognitive-services/face/

33





Visual Analytics

35

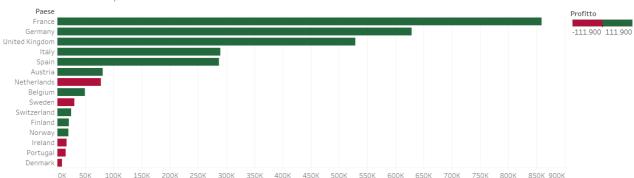


ata spedizione	Modalità spedi	ID cliente	Nome cliente	Segmento	Città	Stato	Paese	Categoria	Sottocate	Nome prodotto	Vendite	Quantità	Profitto
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United Kind	Forniture p	Conservazi	Fellowes Folders, Blue	79,2	3	39,
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United Kind	Tecnologia	Accessori	SanDisk Numeric Keypad, Blu	388,92	7	
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United Kind	Forniture p	Etichette	Avery Legal Exhibit Labels, 50	35,19	3	16,1
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United Kind	Forniture p	Conservazi	Fellowes Folders, Wire Frame	50,94	2	13,
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United Kind	Tecnologia	Accessori	Memorex Memory Card, USB	307,44	3	73,7
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United Kind	Forniture p		Rogers Shelving, Wire Frame	122,4	2	37,9
11/02/2019	Standard	AS-10045	Aaron Smayling	Azienda	Leeds	England	United King	Tecnologia	Telefoni	Apple Signal Booster, Full Siz	413,82	3	20,6
07/08/2017	Due/tre giorni	EB-13840	Ellis Ballard	Azienda	West Bromwich	England	United King	Tecnologia	Fotocopiat	Canon Personal Copier, Laser	428,22	3	192,6
07/08/2017	Due/tre giorni	EB-13840	Ellis Ballard	Azienda	West Bromwich	England	United King	Forniture p	Elettrodom	Hoover Stove, Red	3979,29	7	1989,5
07/08/2017	Due/tre giorni	EB-13840	Ellis Ballard	Azienda	West Bromwich	England	United King	Forniture p	Strumenti p	Accos Push Pins, Bulk Pack	43,56	3	12,
20/06/2017	Stesso giorno	LH-17155	Logan Haushalter	Consumate	Le Bouscat	Aquitaine-Limousin-Poitou-C	France	Forniture p	Etichette	Hon Color Coded Labels, Alph	25,26	2	0.4
20/06/2017	Stesso giorno	LH-17155	Logan Haushalter	Consumate	Le Bouscat	Aquitaine-Limousin-Poitou-C	France	Forniture p	Elettrodom	Hamilton Beach Stove, Silver	2443,905	5	760,30
20/06/2017	Stesso giorno	LH-17155	Logan Haushalter	Consumate	Le Bouscat	Aquitaine-Limousin-Poitou-C	France	Forniture p	Arte	Binney & Smith Pens, Easy-Ei	12,21	1	
20/06/2017	Stesso giorno	LH-17155	Logan Haushalter	Consumate	Le Bouscat	Aquitaine-Limousin-Poitou-C	France	Tecnologia	Telefoni	Samsung Smart Phone, Full S	2167,296	4	790,41
20/06/2017	Stesso giorno	LH-17155	Logan Haushalter	Consumate	Le Bouscat	Aquitaine-Limousin-Poitou-C	France	Forniture p	Conservazi	Rogers Folders, Blue	138,105	5	-12,34
20/02/2018	Standard	KC-16540	Kelly Collister	Consumate	Strasbourg	Alsace-Champagne-Ardenne	France	Forniture p	Conservazi	Smead Trays, Wire Frame	128,385	3	4,27
20/02/2018	Standard	KC-16540	Kelly Collister	Consumate	Strasbourg	Alsace-Champagne-Ardenne	France	Tecnologia	Accessori	Belkin Memory Card, USB	690,12	6	
20/02/2018	Standard	KC-16540	Kelly Collister	Consumate	Strasbourg	Alsace-Champagne-Ardenne	France	Forniture p	Etichette	Hon Removable Labels, Adjus	8,16	1	1,1
18/09/2017	Stesso giorno	JC-15340	Jasper Cacioppo	Consumate	Aachen	North Rhine-Westphalia	Germany	Forniture p	Buste	GlobeWeis Peel and Seal, Se	128,34	6	
16/03/2018	Standard	JE-15715	Joe Elijah	Consumate	Aachen	North Rhine-Westphalia	Germany	Forniture p	Raccoglitor	Wilson Jones Index Tab, Clea.	11,82	2	1,0
16/08/2019	Veloce	BF-11275	Beth Fritzler	Azienda	Thiais	Ile-de-France	France	Forniture p	Raccoglitor	Acco Hole Reinforcements, R	14,04	2	7,0
24/06/2019	Standard	DR-12880	Dan Reichenbach	Azienda	Kassel	Hesse	Germany	Mobili	Librerie	Safco Stackable Bookrack, M	268,164	2	5,90
05/05/2019	Due/tre giorni	LB-16795	Laurel Beltran	Studio	Milan	Lombardy	Italy	Forniture p	Conservazi	Rogers Shelving, Wire Frame	36,72	1	-5,5
05/05/2019	Due/tre giorni	LB-16795	Laurel Beltran	Studio	Milan	Lombardy	Italy	Tecnologia	Accessori	Memorex Mouse, Bluetooth	28,17	1	1,1
16/08/2018	Standard	RO-19780	Rose O'Brian	Consumate	Brandenburg	Brandenburg	Germany	Forniture p	Conservazi	Rogers Shelving, Single Width	279,72	5	-6,3
16/08/2018	Standard	RO-19780	Rose O'Brian	Consumate	Brandenburg	Brandenburg	Germany	Forniture p	Arte	BIC Pens, Fluorescent	108,78	7	35,
04/03/2019	Standard	SP-20920	Susan Pistek	Consumate	Heilbronn	Baden-Württemberg	Germany	Tecnologia	Macchine	Panasonic Calculator, Red	332,01	7	112,7
04/03/2019	Standard	SP-20920	Susan Pistek	Consumate	Heilbronn	Baden-Württemberg	Germany	Forniture p	Conservazi	Rogers Box, Industrial	110,025	5	-12,22
04/03/2019	Standard	SP-20920	Susan Pistek	Consumate	Heilbronn	Baden-Württemberg	Germany	Forniture p	Strumenti p	Stockwell Rubber Bands, Ass	41,94	3	7,9
28/11/2018			Roy Collins	Consumate	Aachen	North Rhine-Westphalia	Germany	Mobili	Librerie	Sauder Floating Shelf Set, Tra	1050,732	6	315,07
28/11/2018	Standard	RC-19825	Roy Collins	Consumate	Aachen	North Rhine-Westphalia	Germany	Forniture p	Raccoglitor	Ibico Index Tab, Clear	58,59	7	21,6
06/06/2016	Standard	DB-13270	Deborah Brumfield	Studio	Barcelona	Catalonia	Spain	Forniture p	Raccoglito	Cardinal Binding Machine, Du	102.54	2	35.8

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

37





 $Sum of Vendite for each Paese. \ Colour shows sum of Profitto. The view is filtered on Paese, which keeps 15 of 15 members.$

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









INFOGRAFICA: IMMAGINI PER MOSTRARE RISULTATI



REPORT E BUSINESS INTELLIGENCE: IMMAGINI PER VISUALIZZARE I DATI



VISUAL ANALYTICS: IMMAGINI PER ANALIZZARE I DATI

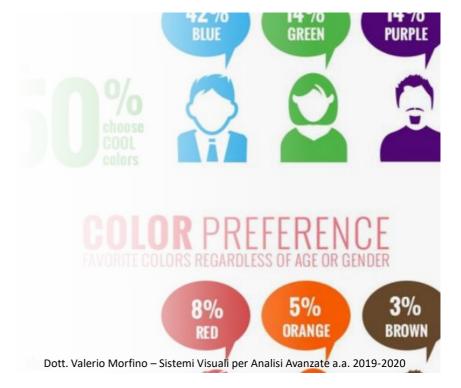


41

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

Infografica

Una rappresentazione artistica per mostrare degli insight





Reporting

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



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



43

Visual Analytics

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





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

Visual Analytics – Funzionalità chiave

Esplorare i dati

Integrare diverse sorgenti Geolocalizzare le informazioni

Navigare le informazioni temporalmente

Condividere le informazioni

Arrivare dal grafico al singolo dato

Forecasting

Clustering



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

45

Gli strumenti del corso





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

47

Formazione e certificazioni Tableau



- https://www.tableau.com/it-it/learn
- https://www.tableau.com/learn/certification
- https://www.tableau.com/learn/certification/d esktop-specialist
- https://mkt.tableau.com/files/DesktopSpecialis t_ExamGuide.pdf



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





Tableau Public – Il mio profilo

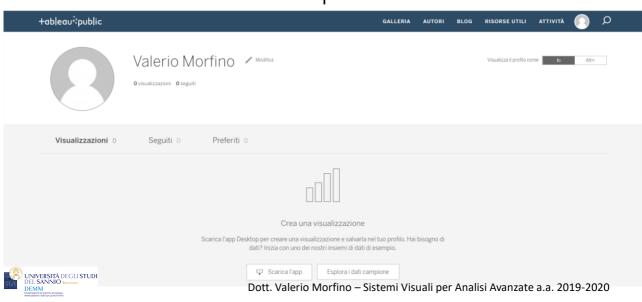


Tableau Public – Il mio profilo

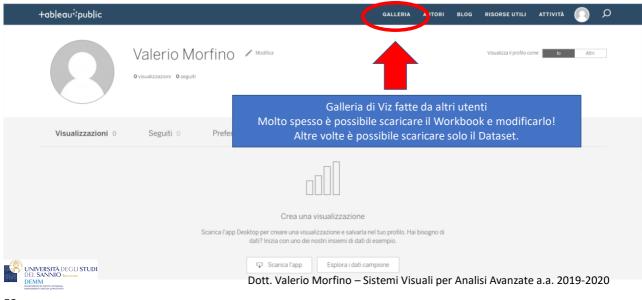


Tableau Public – Il mio profilo

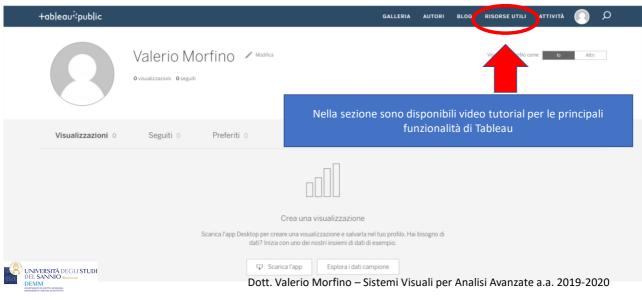
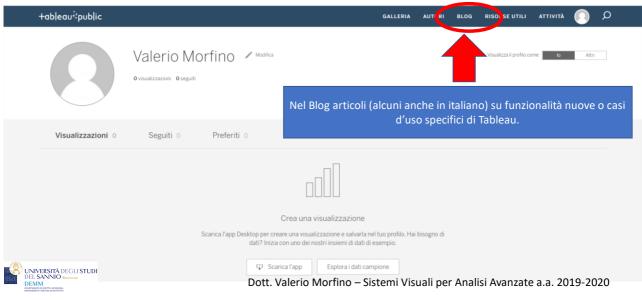
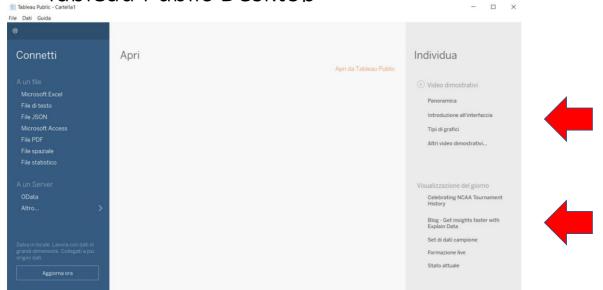


Tableau Public – Il mio profilo



55

Tableau Public Desktop



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 a.a. 2019-2020

57

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

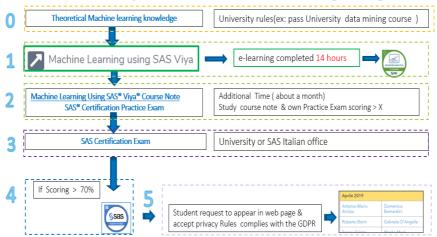


Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

SaS Viya - Certificazione



SAS® Certified Specialist: Machine Learning Using SAS Viya



59

Grazie per l'attenzione



https://it.linkedin.com/in/valerio-morfino



vmorfino@unisannio.it