

# Software Visuali per Analisi Avanzate

Corso sostitutivo di Tirocinio



UNIVERSITÀ DEGLI STUDI  
DEL SANNIO Benevento  
DEMM  
DIPARTIMENTO DI DIRITTO, ECONOMIA,  
MANAGEMENT E METODI QUANTITATIVI

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

1

## Calendario del corso



UNIVERSITÀ DEGLI STUDI  
DEL SANNIO Benevento  
DEMM  
DIPARTIMENTO DI DIRITTO, ECONOMIA,  
MANAGEMENT E METODI QUANTITATIVI

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 Percezione e rappresentazione funzionale delle informazioni
Martedì 7/4 14.00 – 18.00	4 ore	Il Quadrante di Gartner per l'Advanced Analytics Ripasso concetti di Machine Learning Sas: Introduzione, ambiente operativo, Visual Analytic
Venerdì 17/4	4 ore	SaS – Ospite SaS Italia
Sabato 18/4	4 ore	SaS Conclusione del Modulo 1

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

2

## Sommario

Presentazione Modulo 2

Riepilogo delle risorse del corso

Elaborati del corso

Elaborato Finale

Bozza  
Calendario  
Modulo II



Giorno	Durata	Argomenti
TBD	2 ore	Tipologie di dati e loro trattamento: <ul style="list-style-type: none"> <li>• Strutturati</li> <li>• Semistutturati</li> <li>• Destrutturati</li> </ul>
TBD	4 ore	Database relazionali Linguaggio SQL MS Access
TBD	2 ore	Linguaggio SQL Select, Where, Join
TBD	4 ore	Linguaggio SQL Query di aggregazione Indici
TBD	4 ore	Formati dati semi-strutturati: Json, XML Database NoSQL MongoDB
TBD	4 ore	Analitica e Machine Learning in ambito Big Data Strumenti Cloud Databricks, Zeppelin, SQL, Grafici, R
TBD	4 ore	Analisi di dati e Machine Learning con Spark, Databricks ed R

## Attestazione di partecipazione di SaS

- Sulla chat:

«Digitando nome, cognome ed indirizzo email autorizzo il trattamento dei miei dati personali al fine di predisporre un attestato di partecipazione per la lezione tenuta da SaS il 17/04 nell'ambito del corso sostitutivo di tirocinio Software Visuali per Analisi Avanzate presso il dipartimento DEMM dell'Università del Sannio»

- Se accettate, digitate nome, cognome ed indirizzo email istituzionale

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

5

## Risorse del coso



<https://github.com/valeriovv/Software-Visuali-per-Analiti-Avanzate>



Risorse Tableau



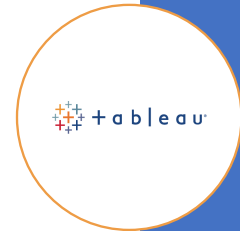
Risorse SaS

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

6

# Risorse Tableau Learning

- Corso e-learning gratuito per 90 gg.
  - <https://www.tableau.com/learn/training/elearning>
- Versione Educational di Tableau (1 anno gratuito)
  - <https://www.tableau.com/academic/students>
- Versione Trial Creator + Server per usi aziendali

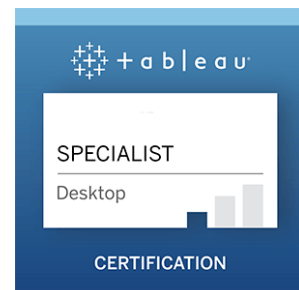


7

## 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/DesktopSpecialist\\_ExamGuide.pdf](https://mkt.tableau.com/files/DesktopSpecialist_ExamGuide.pdf)



8

## Link utili per SaS Viya

- Istruzioni per iscriversi al portale: Lezione-3-Iscrizione-PortaleSAS\_slide.pdf
- Home del Virtual Learning Environment per SCYP Machine Learning with SAS: <https://vle.sas.com/course/view.php?id=3436>
- Cliccare su: LEARN SAS per accedere all'e-learning (link diretto: <https://vle.sas.com/course/view.php?id=3376>)
- Cliccare su Access SaS Software per accedere a Sas Viya (link diretto: <https://vle.sas.com/course/view.php?id=3436&section=1>)

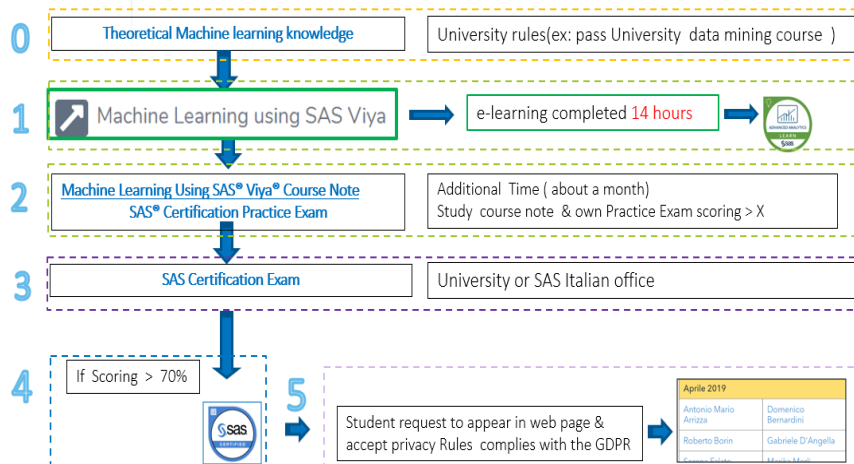


9

## SaS Viya - Certificazione



### SAS® Certified Specialist: Machine Learning Using SAS Viya



10

## Superamento del corso

- Elaborato finale COVID-19 entro fine mese
- Elaborato SaS, sviluppato nella lezione del 18/04
- L'elenco degli studenti che hanno superato il corso verrà pubblicato con le consuete modalità



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

11



## Churn Prediction

Elaborato Finale SaS

12

# Churn Prediction

## Introduzione

- Si richiede di sviluppare un modello di machine learning che affronti il problema della Churn Prediction.
- Utilizzare:
  - SASVisualAnalytics
  - Data Mining e Machine Learning (Creazione dei Modelli)
- L'elaborato dovrà essere composto da 4 pagine:
  - Pagina 1: Copertina con Nome, Cognome, Matricola, Titolo del Lavoro, breve abstract (5 righe max)
  - Pagina 2: Analisi esplorativa (1 pagina)
  - Pagina 3: Selezione del modello da utilizzare (1 pagina)
  - Pagina 4: Tuning del modello (1 pagina)

13

# Churn Prediction

## Analisi esplorativa

- Descrizione del Dataset (**COMMSDATA da 555 Mbyte**)
  - Numero e tipo di predittori, classe da predire e tipo
- Analisi delle principali caratteristiche del dataset
  - Analisi della distribuzione dei valori della classe da predire
  - Correlogramma delle variabili
  - Boxplot
  - Istogramma di alcune variabili (almeno 1)
- Eventuale testo libero descrittivo

14

## Elaborato SaS Selezione Algoritmo di ML

- Splittare in Dataset in Training (almeno 70%) e Validation
- Eliminare features non utili e impostare variabile target
- Inserire un task di esplorazione
- Se necessario, utilizzare la gestione dei Missing Value
- Selezionare 3 algoritmi a piacere
- Selezionare una metrica di performance
- Selezionare l'algoritmo più performante
- Inserire alcuni grafici da Esplorazione, Pipeline e diagrammi nel documento ed eventuale testo libero descrittivo



15

## Elaborato SaS Ottimizzazione del modello

- Creare una nuova pipeline ed effettuare il tuning del modello che ha avuto prestazioni migliori
- E' possibile utilizzare il tuning automatico o manuale degli iperparametri
- Riportare nel documento Word:
  - Matrice di confusione (TP, TN, FP, FN)
  - Accuratezza
  - Error Rate
  - I grafici di performance che si ritengono più significativi (es. modello esplicito dell'albero, ROC, F1 score, ecc.)
  - Codice sorgente del classificatore
- Eventuale testo libero descrittivo



16



# Grazie per l'attenzione



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



[vmorfino@unisannio.it](mailto:vmorfino@unisannio.it)