

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

1

Calendario del corso Parte 1



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

Giorno	Durata	Argomenti
Venerdì 12/3 16.00 -18.00	2 ore	Presentazione del Corso Il contesto dei Big Data Analitica Avanzata Visual Analytics
Giovedì 18/3 14.00 – 18.00	4 ore	Visual Analytics, Il Quadrante Gartner per la Visual Analytics
Venerdì 19/3 16.00 - 18.00	2 ore	Tableau
Giovedì 8/4 14.00 - 18.00	4 ore	Tableau Design visuale - Ospite Tableau – Ospite
Venerdì 9/4 16.00 - 18.00	2 ore	Tableau Esercitazione finale
Giovedì 15/4 14.00 - 18.00	4 ore	Il Quadrante di Gartner per l'Advanced Analytics Ripasso concetti di Machine Learning Sas: Introduzione, ambiente operativo
Giovedì 22/4 14.00 - 18.00	4 ore	SaS – Lezione a cura di SAS Italia
Venerdì 23/4 14.00 - 18.00	2 ore	SaS – Elaborato Finale

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

2

Sommario

Presentazione Modulo 2

Riepilogo delle risorse del modulo 1

Elaborati del corso

Elaborato Finale

Churn Prediction

Riepilogo ambiente SAS

Calendario Modulo 2



Giorno	Durata	Argomenti
Giovedì 29/04	4	Presentazione del Corso Dati Strutturati, semi-strutturati e non strutturati Software da installare per il corso I sistemi informativi aziendali - Ospite
Venerdì 30/04	2	Database relazionali MySQL Database, Tabelle, Righe, Colonne Linguaggio SQL: Introduzione, Select, Where,
Giovedì 6/5	4	Linguaggio SQL: Inner Join, Left Join, Right Join, Order By Linguaggio SQL: Query di aggregazione
Venerdì 7/5	2	Cenni su Indici Cenni sulle DDL Esercitazione finale SQL
Giovedì 13/5	4	Programmazione con Sas: Introduzione a SAS Studio Struttura di un programma SAS
Venerdì 14/5	2	Programmazione SAS
Giovedì 20/5	4	Programmazione con Sas SQL con SAS Esercitazione finale SAS
Venerdì 21/5	2	Ospite Ospite

Risorse del coso



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



Risorse Tableau



Risorse SaS

Risorse Tableau Learning

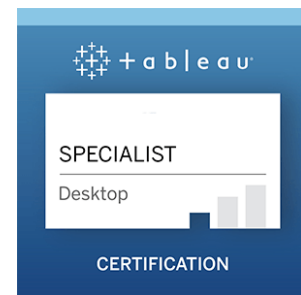
- Corso e-learning gratuito per 90 gg.
 - <https://www.tableau.com/learn/training/elearning>
- Versione Educational di Tableau (1 anno gratuito)
 - <https://www.tableau.com/academic/students>
- Versione Trial Creator + Server per usi aziendali



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



Link utili per SaS Viya

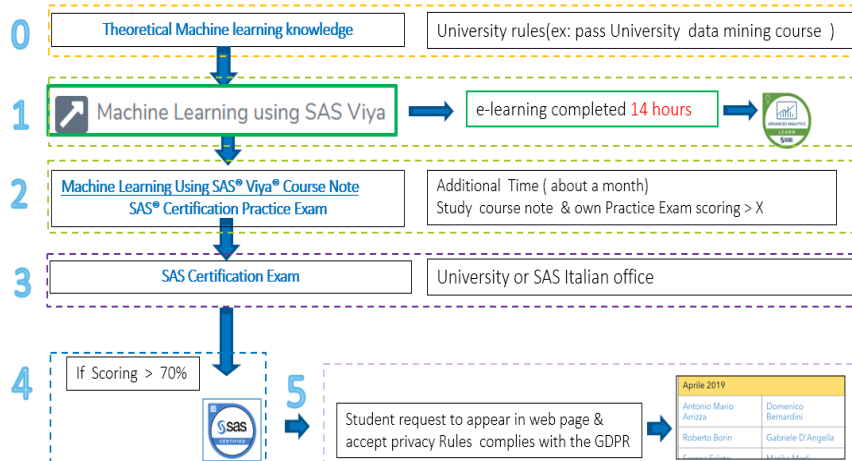
- Istruzioni per iscriversi al portale: Lezione-6_Iscrizione-PortaleSAS.pdf
- Home del Virtual Learning Environment per SCYP Machine Learning with SAS:
<https://vle.sas.com/course/view.php?id=3436>
- Cliccare su: LEARN SAS per accedere all'e-learning (link diretto:
<https://vle.sas.com/course/view.php?id=7768>)
- Cliccare su Access SaS Software per accedere a Sas Viya (link diretto:
<https://vle.sas.com/course/view.php?id=3436§ion=1>)



SaS Viya - Certificazione



SAS® Certified Specialist: Machine Learning Using SAS Viya



9

Acquisizione
dei crediti del
corso

- Elaborato Tableau entro 15 Maggio
- Elaborato SaS entro 15 Maggio
- Frequenza del corso
- L'elenco degli studenti che hanno superato il corso verrà pubblicato con le consuete modalità



Elaborato Finale SaS

Churn Prediction

11

Churn Prediction Introduzione

- Si richiede di sviluppare un modello di machine learning supervisionato che affronti il problema della Churn Prediction – Previsione del tasso di abbandono
- Utilizzare:
 - SASVisualAnalytics
 - SAS 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) - opzionale



12

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



13

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



14

Elaborato SaS Ottimizzazione del modello

- Creare una nuova pipeline ed effettuare i tuning del modello che ha avuto prestazioni migliori
- E' possibile utilizzare il tuning automatico o manuale degli iperparametri
- Riportare nel document Word:
 - Metrica di prestazione selezionata
 - 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



15

Grazie per l'attenzione



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



vmorfino@unisannio.it

16