



fondo  
sociale europeo



Unità Formativa (UF): AI e Machine Learning - Python

Docente: Fabio Giuseppe Antonio Gagliardi

Titolo argomento: esame finale

in collaborazione con:



per una crescita intelligente,  
sostenibile ed inclusiva  
[www.regione.piemonte.it/europa2020](http://www.regione.piemonte.it/europa2020)  
INIZIATIVA CO-FINANZIATA CON FSE

## ▼ REGOLE ESAME

- Commentare i passaggi importanti. Nello specifico:
  - a cosa serve un blocco di codice, una classe o una funzione
  - motivare le scelte effettuate nel codice.
  - L'ASSENZA DI COMMENTI SARÀ VALUTATA NEGATIVAMENTE
  - QUALORA DUE O PIÙ PROVE DOVESSERO RIPORTARE GLI STESSI COMMENTI E LO STESSO CODICE, SARANNO TUTT VALUTATE CON IL VOTO MINIMO.

es.

```
# Ho creato una classe Foo per questo motivo ...  
# Funzione per calcolare ...  
# Ho optato per questa soluzione piuttosto che ....
```

- I dati devono obbligatoriamente essere caricati dall' url fornito su questo documento. Il caricamento dal drive locale dello studente verrà valutato negativamente.
- È richiesta la stesura di una legenda in quei grafici in cui non è chiara l'associazione della rappresentazione scelta con il dato.
- Si è liberi di utilizzare qualsiasi strumento per rispondere alle richieste.
- Non è consentito il copia-incolla da chatgpt. Qualora i test di controllo dovessero confermarne l'uso, l'eame verrà valutato direttamente con voto mininmo.

---

## ▼ DATI

Sono inclusi due set di dati, relativi a campioni di vino rosso e bianco provenienti dal nord del Portogallo.

L'obiettivo è creare un modello in grado di predire la qualità del vino sulla base dei test fisico-chimici.

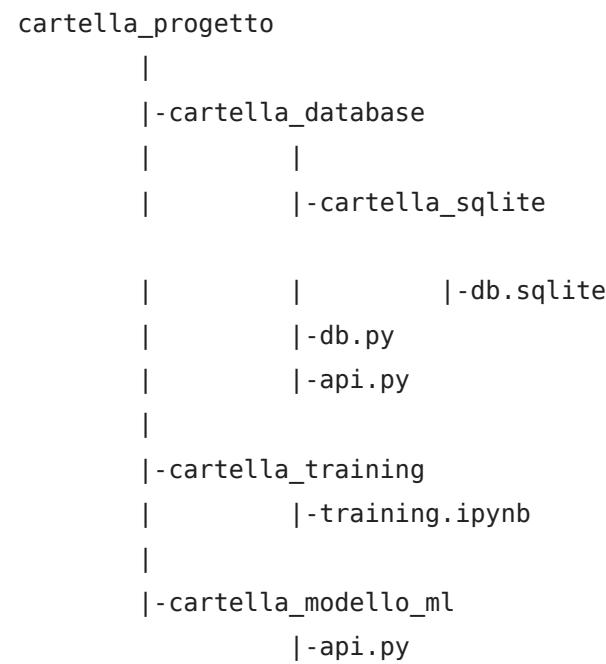
[https://github.com/FabioGagliardi/datasets/tree/main/wine\\_quality](https://github.com/FabioGagliardi/datasets/tree/main/wine_quality)

---

## ▼ PROGETTO

Il progetto è diviso in tre moduli e deve essere consegnato seguendo secondo lo schema delle cartelle indicato di seguito.

### Schema



***N.B Nomi di file e cartelle utilizzati nell'esempio sono solo a scopo descrittivo. La scelta della nomenclatura è a discrezione dello studente, purché sia pertinente al progetto e rispetti lo schema.***

### Moduli

- Il primo e il terzo modulo devono essere sviluppati come servizi da esporre tramite API REST.
- Il secondo modulo riguarda lo sviluppo del modello di machine learning. Si richiede l'uso di colab e l'accesso ai dati tramite le API precedentemente esposte. Inoltre, per questo modulo è richiesta anche la parte di dataviz.

#### 1. Modulo Database

**Cosa fare:**

- impostare un database a partire dai dati forniti.
- implementare le operazioni CRUD (Create, Read, Update, Delete)
- creazione di API REST

**Sviluppo:**

- python + sqlite.

**Cosa verrà valutato:**

- l'organizzazione e stesura del codice
- la progettazione del database
- il corretto funzionamento di tutti gli endpoint da swagger

#### 2. Modulo di training (Machine Learning)

**Cosa fare:**

- richiesta dei dati dal db usando la libreria reuquest. (<https://www.geeksforgeeks.org/python-requests-tutorial/>)
- analisi espplorativa

- selezione delle feature
- training
- salvataggio del modello addestrato

**Sviluppo:**

- Questo modulo va realizzato in colab e non necessita dell'esposizione di API REST.

**Cosa verrà valutato:**

- l'organizzazione e stesura del codice
- l'analisi dei dati con relativa visualizzazione
- la creazione della pipeline di ml e di tutte le fasi di verifica intermedie.
- l'impiego diretto dei file csv verrà valutato negativamente

### 3. Deploy del modello (Machine Learning)

**Cosa fare:**

- Un servizio (API REST) in grado di prendere dati e restituire una predizione utilizzando il modello addestrato.

**Sviluppo:**

- python

**Cosa verrà valutato:**

- l'organizzazione e stesura del codice
- la modalità di passaggio dei dati via API
- il corretto funzionamento di tutti gli endpoint da swagger