

## Homeworks Machine Learning A.A. 2022/23

Sul dataset fornito ("Online News Popularity Data Set") dovranno essere implementati (non usando le librerie di scikit-learn) i seguenti modelli:

- 1. Decision Trees
- 2. Regressione lineare (senza regolarizzazione, regolarizzazione L1, L2, e elastic net)
- 3. Classificazione con regressione lineare (usando la funzione di threshold vista a lezione)
- 4. Classificazione con regressione logistica
- 5. Regressione con algoritmo di k-Nearest Neighbors
- 6. Classificazione con algoritmo di k-Nearest Neighbors
- 7. Regressione con Reti Neurali
- 8. Classificazione con Reti Neurali
- 9. Tabella riassuntiva finale

Si ricorda che lo homework finale dovrà contenere nella parte finale (tipicamente le ultime due celle) il riassunto dei risultati degli algoritmi implementati. Una cella per gli algoritmi di regressione e una cella per gli algoritmi di classificazione.

Riporto di seguito le regole stabilite anche a lezione:

0. Scikit-learn può essere usata *solo* per controllare che i modelli che voi implementati si comportino nel modo corretto. Non saranno valutati completati modelli implementati *solo* usando scikit-learn.

1. La consegna deve essere fatta *tassativamente* prima della fine del semestre (entro il 31 maggio 2023)
2. Il codice deve funzionare su Colab. La prima operazione che farò sui notebook sarà farli girare. Tutti i notebook che non funzioneranno saranno considerati invalidi.
3. Il codice deve essere pulito e leggibile. Ogni notebook che sarà difficilmente leggibile sarà penalizzato. In particolare, non dovranno essere consegnati notebook con porzioni di codice commentati, tutto il codice non utilizzato va cancellato. Ovviamente non mi riferisco ai commenti *al* codice che invece dovranno essere presenti e esplicativi.
4. E' possibile usare le funzioni di scikit-learn per controllare il funzionamento del proprio codice
5. Il punteggio massimo che è possibile ottenere allo homework è 24 punti.