



SELEZIONE PUBBLICA, PER TITOLI ED ESAMI, PER IL RECLUTAMENTO DI N. 1 UNITÀ DI TECNOLOGO DI PRIMO LIVELLO CON RAPPORTO DI LAVORO SUBORDINATO A TEMPO DETERMINATO DELLA DURATA DI 36 MESI, PRESSO L'UNIVERSITÀ DEGLI STUDI MILANO - DIPARTIMENTO DI ECONOMIA, MANAGEMENT E METODI QUANTITATIVI, PER L'ATTUAZIONE DEL PROGRAMMA DI RICERCA E INNOVAZIONE DAL TITOLO "MUSA - MULTILAYERED URBAN SUSTAINABILITY ACTION" (CODICE IDENTIFICATIVO ECS00000037) NELL'AMBITO DEL PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) - CODICE 22206

La Commissione giudicatrice della selezione, nominata con Determina Direttoriale n. 18621 del 6.12.2022, composta da:

Prof. Andrea Maurino	Presidente
Dott. Andrea Geraci	Componente
Dott.ssa Federica Nicolussi	Componente
Dott.ssa Alessandra Maria Magenga	Segretaria

comunica i quesiti relativi alla prova orale:

TRACCIA 1

1. Quali tecniche di Machine learning utilizzeresti per un algoritmo di clustering; parla di un esempio di implementazione in R o in Python
2. Si parli della funzione logistica
3. Progettazione di una base di dati realizzata dal candidato
4. Illustrare un esempio di tecniche di Web Scraping
5. Come progettare una dashboard interattiva sul dominio universitario

Si legga e traduca il seguente brano:

"Random forests or random decision forests is an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned.[1][2] Random decision forests correct for decision trees' habit of overfitting to their training set.[3]:587-588 Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees.[citation needed] However, data characteristics can affect their performance."

TRACCIA 2

1. Quali tecniche di Machine learning utilizzeresti per un algoritmo di classificazione; parla di un esempio di implementazione in R o in Python
2. Quando conviene utilizzare la funzione logistica
3. Modellazione dati in ambito universitario
4. Acquisizione dati tramite Web Scraping
5. Come progettare una piattaforma interattiva per la visualizzazione dati

Si legga e traduca il seguente brano:

"The first algorithm for random decision forests was created in 1995 by Tin Kam Ho[1] using the random subspace method,[2] which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg.[6][7][8]

An extension of the algorithm was developed by Leo Breiman[9] and Adele Cutler, who registered "Random Forests" as a trademark in 2006 (as of 2019, owned by Minitab, Inc.).[12] The extension combines Breiman's



"bagging" idea and random selection of features, introduced first by Ho[1] and later independently by Amit and Geman[13] in order to construct a collection of decision trees with controlled variance."

Milano, 11 gennaio 2023

La Commissione

Prof. Andrea Maurino - Presidente

Dott. Andrea Geraci - Componente

Dott.ssa Federica Nicolussi - Componente

Dott.ssa Alessandra Maria Magenga - Segretaria