

TRACCE

ESTRATTE



Yll

10.

φ

## GRUPPO MATERIE "B"

### Traccia n. 1

Una azienda deve ampliare la propria infrastruttura di rete in modo da:

- garantire la connessione di rete tra le diverse sedi e permettere ai dipendenti fuori sede il collegamento da remoto ai servizi e alle applicazioni aziendali;
- garantire elevata disponibilità a fronte di possibili guasti degli apparati e dei collegamenti di rete.

Sulla base della precedente descrizione, la candidata/il candidato presenti una soluzione architettonica che garantisca la sicurezza delle comunicazioni suindicate in termini di confidenzialità, di integrità e di elevata disponibilità dei collegamenti.

Con riferimento alle reti e ai sistemi informatici, la candidata/il candidato, inoltre:

- illustri le caratteristiche principali del protocollo IPSec;
- illustri tecniche per il rilevamento dei guasti e tecniche di ridondanza nei sistemi e nelle reti informatiche.

**GRUPPO MATERIE "B"****Traccia n. 2**

In un contesto *smart city*, si intende realizzare un servizio di parcheggi intelligenti. Tramite un sistema di sensori, è possibile rilevare la disponibilità dei posti auto. Le informazioni vengono raccolte da un sistema che elabora la mappa dei posti vacanti. Tramite un *app mobile*, gli utenti possono visualizzare lo stato dei parcheggi e/o prenotare l'utilizzo di un posto auto per un determinato intervallo di tempo e possono effettuare il pagamento sia della prenotazione del posto auto che del parcheggio utilizzato.

La candidata/il candidato descriva una soluzione per realizzare quanto descritto tenendo conto della grande quantità di dati da gestire e della natura distribuita del sistema. La candidata/il candidato descriva, in particolare:

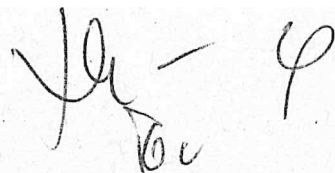
- l'architettura generale dell'applicazione;
- le soluzioni adottate per garantire la consistenza e la tolleranza ai guasti.

Con riferimento ai sistemi e alle infrastrutture distribuite la candidata/il candidato, inoltre:

- illustri il problema del consenso nei sistemi distribuiti e descriva uno o più algoritmi di consenso distribuito;
- illustri i diversi modelli di consistenza e i relativi *trade-offs*;

TRACCE

ESTIMATE



## GRUPPO MATERIE "B"

### Traccia n. 3

Al fine di realizzare un sistema di *file storage* distribuito, si intende realizzare una soluzione basata su tecnologia *blockchain*. La candidata/il candidato presenti una soluzione che distribuisca i file degli utenti su una rete decentralizzata di nodi. Si assuma che gli utenti spendano *tokens* per memorizzare e leggere dati e che i *miners* guadagnino *tokens* per la memorizzazione dei dati.

Con riferimento alle blockchain, la candidata/il candidato, inoltre:

- descriva il ruolo e l'utilizzo del *Merkle Tree*;
- descriva il ruolo dell'algoritmo di consenso.

**GRUPPO MATERIE "A"****Traccia n. 1**

Negli ultimi decenni il Ministero del Turismo ha raccolto i dati relativi ai soggiorni nelle strutture alberghiere italiane. Ciascun soggiorno è relativo ad una persona di cui si conosce il codice identificativo, l'età, la città di nascita (con nazione e continente), la città di residenza (con regione, nazione e continente) ed eventualmente (non sempre) il titolo di studio. Ciascun soggiorno è anche caratterizzato dalla data di inizio del soggiorno (con giorno, mese e anno), dalla durata in numero di notti, dal costo in euro e dalla struttura che ha ospitato la persona. Di ogni struttura alberghiera è noto il comune in cui si trova (con provincia e regione), il numero di posti letto disponibili, il capitale sociale, la categoria a cui appartiene ed il tipo. Sono di interesse tre tipi di strutture: hotel, ostello e familiare. A loro volta i tipi si distinguono in categorie: gli hotel si distinguono in pensioni, alberghi, motels e villaggi; gli ostelli si distinguono in ostelli della gioventù e grandi ostelli e le strutture familiari si distinguono in agriturismo e bed&breakfast. Si richiede di rispondere ai seguenti quesiti.

1. Definire lo schema concettuale del dominio descritto, espresso secondo un opportuno formalismo a scelta, del quale occorre fornire una breve descrizione.
2. Definire il *Dimensional Fact Model* (DFM) per un Data Warehouse ministeriale che consenta opportune analisi dei dati sul dominio descritto.
3. Definire lo star schema oppure lo snowflake schema, oppure ancora un qualunque schema in un altro formalismo, a scelta del quale occorre fornire una breve descrizione.
4. Definire opportune query (in SQL o in un linguaggio adeguato al formalismo scelto al punto 3) per condurre attività di analisi con i seguenti obiettivi: a) calcolare gli introiti delle strutture ottenuti dai soggiorni nei vari mesi del 2021 (per assegnare un soggiorno ad un mese vale la data di inizio del soggiorno stesso), nelle varie regioni in cui le strutture si trovano e per le varie categorie a cui le strutture appartengono; b) calcolare il numero di soggiorni negli ostelli e nelle strutture familiari, distinti per categoria dell'ostello, e per età dell'ospite; c) per ogni nazione di nascita e per ogni nazione di residenza, calcolare l'età media delle persone con quella nazione di nascita, quella nazione di residenza e quella età che hanno soggiornato in hotel nel decennio 2011-2020.

Per tutto ciò che non è specificato nella descrizione fornita qui sopra, il candidato può formulare e giustificare opportune assunzioni e svolgere la prova sulla base di esse.

**GRUPPO MATERIE "A"****Traccia n. 2**

Un agricoltore è interessato ad acquistare da un venditore campi coltivati in una certa area geografica. I campi in vendita sono  $N$ , ciascuno caratterizzato dalla dimensione (in metri quadrati), dal costo al metro quadrato e dal profitto (un numero intero positivo) che l'agricoltore pensa di ottenere acquistando il campo. Ciascun campo appartiene ad una categoria tra 10 possibili, a seconda del tipo di coltivazione ideale per quel campo, e sussiste la regola che nessun acquirente può acquistare campi che appartengono a più di 2 categorie. Inoltre, l'agricoltore ha un preciso budget che può utilizzare e mira ad ottenere il massimo profitto dall'acquisto, ma rispettando il budget stesso. Infine, il venditore è disponibile a vendere anche una frazione di ciascun campo coltivato, ovvero un certo numero di metri quadrati tra quelli disponibili. Si richiede di rispondere ai seguenti quesiti.

1. Definire formalmente il problema computazionale corrispondente allo scenario descritto.
2. Descrivere lo spazio delle potenziali soluzioni (detto anche spazio degli stati) del problema computazionale.
3. Dare la definizione di algoritmo di tipo enumerativo (*brute-force*), illustrare brevemente le sue proprietà e specificare se sia possibile definire un algoritmo di tipo enumerativo per risolvere ogni istanza del problema suddetto. In caso di risposta positiva, descrivere tale algoritmo in pseudocodice nel modo più dettagliato possibile, dimostrandone la correttezza e caratterizzandone la complessità computazionale rispetto al tempo di esecuzione; in caso contrario, motivare in dettaglio il ragionamento seguito per giungere alla risposta negativa.
4. Indicare quale sia l'algoritmo ritenuto dal candidato il più efficiente per affrontare il problema, senza escludere la possibilità di confermare quello scelto per il punto 3, specificando la tecnica algoritmica alla base della sua formulazione, motivando la scelta di tale tecnica algoritmica in contrapposizione ad altre e descrivendo l'algoritmo e le strutture di dati nel modo più dettagliato possibile, utilizzando un qualunque formalismo ritenuto adatto.

**GRUPPO MATERIE "A"****Traccia n. 3**

Si consideri un dominio  $D$  relativo ai beni culturali, descritto come segue. I libri sono opere letterarie, non figurative. I quadri sono opere figurative, non letterarie. Entrambi i suddetti tipi di opere sono prodotti dell'intelletto. I libri per l'infanzia sono libri particolari. Sia i libri per l'infanzia sia i libri d'arte sono classificati come opere figurative. Anche le prestazioni sportive sono opere, ma opere dell'attività fisica e quindi di svago, piuttosto che dell'intelletto. Si richiede di rispondere ai seguenti quesiti.

1. Definire sinteticamente il formalismo della logica classica proposizionale per la rappresentazione della conoscenza e per il ragionamento automatico in intelligenza artificiale.
2. Descrivere la conoscenza che emerge dalla suddetta descrizione del dominio  $D$  in termini della logica classica proposizionale.
3. Considerare le seguenti affermazioni sul dominio  $D$ :
  - a) le opere letterarie non sono opere figurative;
  - b) ogni prodotto dell'intelletto o è un'opera letteraria o è un'opera figurativa;
  - d) le prestazioni sportive non sono opere dell'intelletto;
  - e) è possibile che un'opera sia contemporaneamente figurativa e letteraria;e, basandosi sulla conoscenza specificata secondo quanto stabilito al punto 2, dire se l'affermazione è una conclusione valida oppure no, motivando la risposta in dettaglio con l'ausilio di opportune nozioni del formalismo della logica proposizionale.
4. Supponiamo che vengano aggiunte le seguenti considerazioni sul dominio  $D$ .
  - a) Ogni opera ha un certo numero (almeno uno) di autori, che sono persone.
  - b) Gli autori di libri sono scrittori e gli autori di prestazioni sportive sono sportivi.
  - c) Esiste almeno un libro con più di un autore, mentre la stessa proprietà non sussiste per i quadri.

Scegliere un formalismo per la rappresentazione della conoscenza in intelligenza artificiale e, usando tale formalismo, esprimere la conoscenza complessiva sul dominio  $D$  che emerge da quanto detto nei punti 1 e 4. Infine, indicare le caratteristiche che deve avere uno strumento di intelligenza artificiale per eseguire efficacemente deduzioni significative sulla base di conoscenza relativa al dominio  $D$ .

**GRUPPO MATERIE "A"****Traccia n. 1**

L'Associazione Italiana per la Musica (AIM) vuole realizzare un sistema informativo che raccolga i dati, attuali e passati, delle orchestre che operano nel territorio nazionale. I dati che si vogliono trattare riguardano appunto le orchestre. I cosiddetti dati anagrafici delle orchestre sono: codice identificativo, anno di fondazione, nome, provincia di appartenenza, genere principale ed eventuale anno di scioglimento. Oltre ai dati anagrafici, delle orchestre interessano anche: il numero di concerti effettuati nei vari mesi, le collaborazioni artistiche tra orchestre (con periodo di collaborazione, che è un intervallo di date, scopo della collaborazione e valutazione numerica della collaborazione stessa) e la composizione delle orchestre nel tempo, composizione che può variare nei diversi mesi. Si noti che, istante per istante, per ogni orchestra ci si aspetta di censire la collaborazione di al massimo un'orchestra. Riguardo alla composizione, mese per mese un'orchestra è composta da un insieme di musicisti, ciascuno con il suo ruolo (trombettista, violinista, ecc.) e non è possibile che un musicista nello stesso mese faccia parte di più di un'orchestra. Di ogni musicista interessano l'eventuale data di primo diploma al conservatorio ed i seguenti dati anagrafici: il codice identificativo, il nome, il cognome e la data di nascita. L'AIM possiede la lista completa, aggiornata giorno per giorno, delle orchestre, con i relativi dati anagrafici e dei musicisti, anch'essi con i relativi dati anagrafici. Gli altri dati rilevanti per il sistema informativo provengono da diverse fonti: quelli relativi al numero di concerti delle orchestre dai sistemi informativi delle province delle orchestre stesse (ogni provincia manda le informazioni sulle orchestre della propria provincia), quelli relativi alla composizione delle orchestre dai sistemi informativi delle associazioni musicali regionali (ogni regione manda le informazioni sulle orchestre di quella regione) e quelli relativi ai diplomi dai sistemi informativi dei conservatori. Si richiede di rispondere ai seguenti quesiti.

1. Descrivere il risultato della modellazione dei dati rilevanti per il sistema informativo, utilizzando un qualunque formalismo (o un qualunque insieme di formalismi, se il metodo scelto dal candidato ne richiede più d'uno), illustrando sinteticamente le caratteristiche essenziali di tale formalismo (o tali formalismi).
2. Illustrare sinteticamente le principali dimensioni note in letteratura per valutare la qualità dei dati e fornire una lista di controlli di qualità ragionevoli per il sistema informativo dell'AIM, illustrando per ognuno la dimensione di qualità pertinente ed il metodo da utilizzare per realizzarlo.
3. Individuare i meta-dati rilevanti per il sistema informativo e descrivere il risultato della loro modellazione, tenendo presente che i meta-dati da modellare dovranno consentire anche di registrare le valutazioni sui controlli di qualità delle fonti informative descritti al punto 3.
4. Fornire una descrizione dei possibili utilizzi dei meta-dati modellati nelle varie fasi e nelle varie attività proprie del ciclo di vita del sistema informativo dell'AIM.

Per tutto ciò che non è specificato nella descrizione fornita qui sopra, il candidato può formulare e giustificare opportune assunzioni e svolgere la prova sulla base di esse.

TRACCE NON ESTIMATE  
Z  
P  
101

## GRUPPO MATERIE "A"

### Traccia n. 2

Un'associazione scientifica organizza sedute di "brainstorming", che possono durare anche diversi giorni. In ogni seduta diversi temi vengono affrontati da gruppi di partecipanti in sessioni di un'ora. Ogni tema deve essere trattato in una sessione e ad ogni sessione viene assegnato il tema da trattare (esattamente uno), lo slot di tempo in cui si deve svolgere e la sala da utilizzare per la discussione dei partecipanti. Si noti che non ci sono problemi sul numero di sale e quindi non ci sono limiti sul numero di sessioni allocate nello stesso slot di un'ora. All'inizio di ogni seduta sono noti sia quali sono i temi da affrontare, sia quali sono i partecipanti che devono discutere tali temi e che quindi devono partecipare alle corrispondenti sessioni e sia anche quali sono gli slot di tempo utilizzabili (ogni slot denotato dalla corrispondente ora del giorno).

L'organizzazione del congresso ha deciso di acquisire un software che la supporti nella decisione di come organizzare ciascuna di queste sedute di brainstorming. Per ogni seduta, lo scopo è quello di decidere, per ognuno dei temi previsti, in quale sessione discuterlo ed in quale slot di tempo svolgere la sessione, rispettando la condizione che tutti i partecipanti che devono poter partecipare alle sessioni in cui si discutono i temi di suo interesse. È ovvio che assegnare uno slot diverso ad ogni sessione risolverebbe questa esigenza, ma l'organizzazione mira a minimizzare il tempo in cui si svolge la seduta di brainstorming e quindi in sostanza a minimizzare il numero di slot previsti. Si richiede di rispondere ai seguenti quesiti.

1. Definire formalmente il problema computazionale corrispondente allo scenario descritto.
2. Descrivere lo spazio delle potenziali soluzioni (detto anche spazio degli stati) del problema computazionale.
3. Dare la definizione di algoritmo di tipo "backtracking", illustrare brevemente le sue proprietà e specificare se sia possibile definire un algoritmo di tipo "backtracking" per risolvere ogni istanza del problema suddetto. In caso di risposta positiva, descrivere tale algoritmo in pseudo-codice, dimostrandone la correttezza e caratterizzandone la complessità computazionale rispetto al tempo di esecuzione; in caso contrario, motivare in dettaglio il ragionamento seguito per giungere alla risposta negativa.
4. Indicare quale sia l'algoritmo ritenuto dal candidato il più idoneo per affrontare il problema, senza escludere la possibilità di confermare quello scelto per il punto 3, specificando la tecnica algoritmica alla base della sua formulazione, motivando la scelta di tale tecnica algoritmica in contrapposizione ad altre ed illustrando l'algoritmo in pseudo-codice. Nel caso in cui il candidato scelga un algoritmo diverso da quello indicato per il punto 3, si richiede di illustrare le proprietà dell'algoritmo, inclusa la complessità computazionale rispetto al tempo di esecuzione.



TRACCE NON ESTIMATE

*[Handwritten signature]*

## GRUPPO MATERIE "A"

### Traccia n. 3

In una amministrazione pubblica Z si è scelto di utilizzare il concetto di Knowledge Graph per rappresentare le informazioni su una porzione del dominio dell'anagrafe (nel quale si assume che si possano tralasciare problematiche di privacy delle informazioni stesse). Le caratteristiche generali del dominio sono descritte come segue. Sono di interesse le persone, con codice identificativo, età, città di residenza, reddito e categoria alla quale appartengono, una tra lavoratore (a sua volta suddivisa in operaio, insegnante, dirigente), studente e pensionato. Di ogni persona interessa anche l'eventuale persona che ne è coniuge e le eventuali persone che ne sono figli. Delle città interessa il numero di abitanti, la nazione in cui si trova e se è la capitale o meno della nazione in cui si trova. Le caratteristiche particolari del dominio sono descritte attraverso un insieme di fatti, che sono, come di consueto, di due tipi: i fatti che indicano che un individuo è istanza di una classe (ad esempio, Roma è una città) ed i fatti che indicano link tra individui, ossia che un individuo è legato ad un altro mediante una relazione (ad esempio, la persona con codice 120 ha la persona con codice 140 come coniuge). Si noti che le caratteristiche particolari non sono considerate complete. In altre parole, siamo nel contesto di open-world assumption, ovvero non possiamo assumere che tutti i fatti che valgono del dominio siano noti. Si richiede di rispondere ai seguenti quesiti.

1. Fornire la definizione di knowledge graph ed illustrare nel modo più dettagliato possibile la forma del knowledge graph che si reputa adatto alla rappresentazione del dominio descritto, considerando che tale knowledge graph deve modellare sia le caratteristiche generali sia le seguenti caratteristiche particolari: Roma è una città, (la persona con codice) 120 è un operaio, 120 ha 140 come coniuge, 120 ha 200, 230 e 250 come figli, 120 ha un reddito di 30.000 e 120 risiede a Firenze.
2. Definire un algoritmo che, dato un knowledge graph della forma stabilita al punto 1 e data la specifica di un fatto della forma "X è istanza di Y" verifichi se tale fatto è derivabile mediante inferenza logica dal knowledge graph.
3. Discutere la correttezza dell'algoritmo e la complessità computazionale rispetto al tempo di esecuzione.
4. Il problema del "link prediction" ha l'obiettivo di valutare come plausibili alcuni link che sono assenti nel knowledge graph ma che si possono indurre dalla conoscenza rappresentata. Si consideri la domanda: si può formulare il problema del link prediction come un problema di apprendimento automatico (machine learning)? Se la risposta è negativa, motivare la risposta. Se la risposta è positiva, illustrare sinteticamente un possibile metodo per realizzare il task di link prediction.

TRACCE NON ESTIMATE

24 - 4  
700

## GRUPPO MATERIE "A"

### Traccia n. 1

La Protezione Civile vuole realizzare un sistema informativo relativo al volontariato nelle RSA (Residenze Sanitarie Assistenziali). I dati rilevanti per questo sistema sono distribuiti in diversi sistemi esterni. In particolare, i dati sui volontari (con codice fiscale, data di nascita, comune di nascita, comune di residenza, titolo di studio e reddito annuo) registrati per il servizio sono presenti in basi di dati relazionali regionali (una per ogni regione), i dati sulle RSA (con codice identificativo, anno di fondazione, direttore sanitario, numero di posti letto e numero di dipendenti) si trovano in file XML prodotti dalle macroregioni (nord-ovest, nord-est, centro-nord, centro-sud, sud, isole) e i dati sulle offerte mensili di servizio dei volontari si trovano in file JSON prodotti mensilmente da un ufficio centrale che opera per le associazioni di volontari nel territorio nazionale. Ciascuna offerta è descritta mediante il mese a cui si riferisce, il codice fiscale del volontario che presenta l'offerta e i giorni di disponibilità, ciascuno con la RSA in cui vale la disponibilità, e altre informazioni che possono variare da offerta a offerta. Si tenga presente che ogni offerta di servizio presentato da un volontario per un certo giorno può riguardare al massimo una RSA. I dati sulle offerte mensili, una volta prodotti, non vengono modificati, mentre i dati relativi ai volontari e alle RSA sono soggetti ad aggiornamenti, ad esempio del reddito di un volontario, del direttore di una RSA, o altro. Il sistema informativo deve supportare diverse funzioni, ma qui interessano due servizi fondamentali: il processo di assegnazione mensile di servizi ai volontari ed il servizio di data analytics. Il processo di assegnazione mensile ha lo scopo di analizzare le offerte presentate dai volontari e decidere quali offerte vengono accettate. Il servizio di data analytics riguarda le analisi descrittive e predittive sui dati del sistema al fine di prendere decisioni informate; per supportare tali analisi il sistema deve mantenere i dati storici, che includono anche i dati sulle accettazioni delle offerte nei vari mesi. Si richiede di rispondere ai seguenti quesiti.

1. Descrivere il risultato della modellazione dei dati rilevanti per il sistema informativo, utilizzando un qualunque formalismo (o un qualunque insieme di formalismi, se il metodo scelto dal candidato ne richiede più d'uno), illustrando sinteticamente le caratteristiche essenziali di tale formalismo (o tali formalismi).
2. Descrivere l'architettura che si ritiene più adeguata rispetto all'obiettivo di realizzare tutte le funzionalità del sistema informativo ed indicare le tecnologie che si ritengono adatte a tale scopo.
3. Illustrare il metodo che si ritiene adatto per gestire gli aggiornamenti dei dati ed il metodo per realizzare il processo di assegnazione mensile sopra menzionato.
4. Scegliere un qualunque task di query o di data analytics ragionevole nell'ambito del sistema informativo, descriverlo in linguaggio naturale, darne la specifica in un formalismo scelto e discutere le tecnologie che si reputano adatte per realizzarlo.

Per tutto ciò che non è specificato nella descrizione fornita qui sopra, il candidato può formulare e giustificare opportune assunzioni e svolgere la prova sulla base di esse

TRACCE NON ESTRATTE

201  
201

## GRUPPO MATERIE "A"

### Traccia n. 2

Una società di consulenza deve decidere, all'inizio di ogni semestre, su quali progetti lavorare tra quelli ad essa proposti dalle  $N$  aziende partner. In particolare, ogni azienda partner presenta  $M$  progetti alla società di consulenza, ognuno caratterizzato sia dal numero di mesi uomo necessari per portare a termine il progetto stesso, sia dal compenso che verrà corrisposto alla società di consulenza al momento della sua conclusione. L'obiettivo della società di consulenza è ovviamente massimizzare l'introito derivante dai compensi, rispettando due condizioni: *i)* non superare un numero prefissato  $U$  di mesi uomo, *ii)* lavorare, in ogni semestre, sui progetti proposti al massimo da due aziende partner. Si richiede di rispondere ai seguenti quesiti.

1. Definire formalmente il problema computazionale corrispondente allo scenario descritto e descrivere lo spazio delle potenziali soluzioni (detto anche spazio degli stati).
2. Descrivere in pseudo-codice l'algoritmo ritenuto dal candidato il più idoneo per risolvere ogni istanza del problema, specificando la tecnica algoritmica alla base della sua formulazione, motivando la scelta di tale tecnica algoritmica in contrapposizione ad altre, dimostrando la correttezza dell'algoritmo e caratterizzandone la complessità computazionale rispetto al tempo di esecuzione.
3. Illustrare un metodo per rappresentare i dati dello scenario descritto in termini di una base di dati espressa in un modello di dati scelto da candidato e descrivere l'architettura di un sistema software che includa la base di dati suddetta ed il modulo che realizza l'algoritmo scelto per il punto 2.
4. Fornire il codice che realizza l'algoritmo scelto al punto 2 in termini di un linguaggio di programmazione a scelta, motivato opportunamente, che acceda alla base di dati per analizzare l'istanza del problema da risolvere ed esegua le opportune operazioni previste per il calcolo della sua soluzione.

TRACCE NON ESTRATTE

100 / 100 P

## GRUPPO MATERIE "A"

### Traccia n. 3

Si supponga di avere raccolto i dati relativi al gradimento di film da parte di una persona ed il nostro scopo è imparare il concetto di "Gradito" usando la tecnica degli alberi di decisione (Decision Tree). I dati raccolti sono descritti nel seguente dataset D:

Film	AppenaUscito	ConAlmenoUnaStar	ValutazioneSuMyMovie	Gradito
1	sì	no	3	sì
2	no	sì	3	no
3	no	no	4	sì
4	no	no	3	no
5	no	sì	4	sì

Si richiede di rispondere ai seguenti quesiti.

1. Formalizzare il problema di apprendimento automatico fornendo sia una descrizione formale della funzione da imparare e del dataset D, sia una descrizione generale del concetto di albero di decisione.
2. Definire formalmente il concetto di coerenza (anche detta consistenza) di un albero di decisione rispetto ad un dataset. Mostrare un qualsiasi albero di decisione coerente (anche detto consistente) con il dataset D.
3. Descrivere un algoritmo generale per costruire un albero di decisione a partire da un dataset di training, discutere le sue proprietà ed applicare tale algoritmo al caso in esame, mostrando l'albero di decisione T costruito.
4. Descrivere una possibile rappresentazione dell'albero mostrato al punto 3 o quello mostrato al punto 2 in termini di un formalismo a regole, illustrando come, mediante tale formalismo, sia possibile determinare se l'albero è coerente con il seguente insieme di campioni (sample):  $S = \{ s_1 = \langle \text{sì}, \text{no}, 3, \text{sì} \rangle, s_2 = \langle \text{no}, \text{no}, 4, \text{sì} \rangle, s_3 = \langle \text{sì}, \text{sì}, 4, \text{no} \rangle, s_4 = \langle \text{no}, \text{no}, 4, \text{no} \rangle \}$ .
5. Illustrare un algoritmo che, avendo come input un albero di decisione rappresentato in termini del formalismo a regole scelto al punto 4 ed un campione (sample), stabilisce se l'albero è coerente rispetto al campione. Discutere la correttezza e la complessità computazionale dell'algoritmo.

TRACCE NON ESTRATTE

✓  
76

## GRUPPO MATERIE "B"

### Traccia n. 3

Al fine di realizzare un sistema di voto elettronico distribuito, si intende sviluppare una soluzione basata su tecnologia *blockchain*. La candidata/il candidato presenti una soluzione che soddisfi i seguenti requisiti: ogni utente abilitato deve poter esprimere il suo voto in forma anonima e poter votare al più una volta; il sistema deve poter permettere di esprimere il voto durante un intervallo di tempo predefinito; al termine delle operazioni di voto, l'applicazione rende disponibili i risultati della votazione.

Con riferimento ad una *blockchain* pubblica, come ad esempio Bitcoin:

- si descriva il ruolo dei *miner*;
- si descriva il problema del *double spending*.

TRACCE NON ESTIMATE

✓  
100  
φ

## GRUPPO MATERIE "B"

### Traccia n. 2

Si supponga di dover realizzare un'applicazione distribuita per lo *storage* di file in grado di servire una vasta popolazione di utenti distribuita su scala geografica. Oltre alla registrazione e alla autenticazione di un utente, l'applicazione deve consentire le usuali operazioni CRUD (*Create, Read, Update, Delete*) su file e la condivisione di file tra utenti. La candidata/il candidato presenti una possibile architettura ad alto livello del sistema caratterizzata da bassi tempi di latenza e da elevata affidabilità, illustrandone vantaggi e possibili svantaggi. Si evidenzi come il teorema CAP è stato applicato all'architettura proposta e con quale impatto sulle funzionalità offerte dal sistema.

Con riferimento ai sistemi e alle infrastrutture distribuite, la candidata/il candidato, inoltre:

- illustri i diversi modelli di consistenza e i relativi *trade-offs*;
- illustri un protocollo di consistenza che implementi uno dei modelli sopra descritti.

TRACCE NON ESTIMATE  
W

T&U

Ver  
4

## GRUPPO MATERIE "B"

### Traccia n. 1

Un negozio intende offrire ai propri clienti la possibilità di acquistare i propri prodotti *online*. A tal fine si rende necessaria la realizzazione di una applicazione *web-based* di commercio elettronico che permetta: la visualizzazione, l'aggiunta/rimozione di prodotti su un carrello e il loro acquisto.

Sulla base della precedente descrizione, la candidata/il candidato:

- proponga una possibile architettura dell'applicazione usando le moderne tecnologie;
- proponga una soluzione per la gestione delle variazioni di carico che tenga conto dei modelli di costo offerti dalle piattaforme *cloud*.

Con riferimento alle architetture e ai sistemi *cloud* la candidata/il candidato:

- illustri come viene realizzata l'orchestrazione di risorse con particolare attenzione alla gestione dello *auto-scaling*;
- descriva le caratteristiche principali e i relativi vantaggi e svantaggi delle tecniche di virtualizzazione basate su container.

TRACCE NON ESTIMATE

TOL  
CP

## GRUPPO MATERIE "B"

### Traccia n. 3

Per la realizzazione di una applicazione distribuita, si rende necessario sviluppare un sottosistema che implementi un protocollo per il consenso distribuito tra processi che sia robusto rispetto ai guasti bizantini. La candidata/il candidato presenti una possibile soluzione, descrivendo in modo dettagliato, il protocollo di consenso. Si richiede la valutazione del costo computazionale della soluzione presentata.

Con riferimento alle criptovalute, la candidata/il candidato, inoltre:

- motivi e illustri il ruolo degli algoritmi di consenso distribuito nell'ambito delle criptovalute, fornendo adeguati esempi;
- descriva il problema del *double spending* e come questo possa essere risolto.



TRACCE NON ESTIMATE

TOU  
P

## GRUPPO MATERIE "B"

### Traccia n. 2

In un contesto *smart city* si intende realizzare un servizio di monitoraggio della qualità dell'aria. Il servizio di monitoraggio deve fornire, a partire dai dati rilevati in tempo reale da una moltitudine di sensori, opportune statistiche a diversi servizi municipali. La candidata/il candidato descriva una soluzione per realizzare il servizio tenendo conto della grande quantità di dati da gestire, della natura distribuita della rete di sensori, nonché della necessità di fornire dati e statistiche, in tempo reale, su molteplici piattaforme.

La candidata/il candidato illustri, in particolare:

- l'architettura generale dell'applicazione e della relativa infrastruttura di elaborazione e di rete;
- le soluzioni adottate al fine di garantire la scalabilità, l'accesso in tempo reale ai dati e alle relative statistiche e la robustezza rispetto ai guasti.

Con riferimento alle infrastrutture e ai sistemi distribuiti la candidata/il candidato, inoltre:

- illustri le caratteristiche principali dei sistemi *publish-subscribe*;
- illustri una o più soluzioni per la comunicazione "uno a tutti" in una applicazione distribuita che consiste di N agenti geograficamente distribuiti ( $N > 1$ ), valutando l'efficacia delle diverse proposte in termini di semplicità della soluzione, traffico di rete per messaggio inviato e tolleranza ai guasti.

TRACCO NON ESTRATTO



100  
φ  
Yll

## GRUPPO MATERIE "B"

### Traccia n. 1

Un'azienda vende i propri prodotti tramite un *online store*. L'applicazione fornisce, fra l'altro, la possibilità di ricercare i prodotti tramite parole chiave e di visualizzarne le relative informazioni. I prodotti possono essere inseriti/rimossi in un carrello. Per gli acquisti, l'applicazione richiede che gli utenti siano registrati, fornendo i dati personali, l'indirizzo di spedizione e i dati relativi al pagamento.

Sulla base della precedente descrizione, la candidata/il candidato:

- proponga una possibile architettura dell'applicazione basata su un modello a microservizi, evidenziando i relativi vantaggi rispetto ad una soluzione monolitica;
- nel contesto della soluzione proposta, descriva gli accorgimenti necessari a garantire la scalabilità e la disponibilità dell'applicazione.

Con riferimento alle architetture a microservizi, la candidata/il candidato:

- descriva l'utilizzo dei sistemi a code di messaggi per supportare il disaccoppiamento nella comunicazione tra microservizi;
- nel contesto dell'applicazione sviluppata, valuti l'opportunità di usare un servizio di persistenza fornendo adeguate motivazioni.