UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Informatica

Corso di Laurea Triennale in Informatica

TESI DI LAUREA

Utilizzo di Intelligenza Artificiale nel contesto delle Blockchain

RELATORE

Prof. **Fabio Palomba**

Università degli studi di Salerno

CANDIDATO

Antonio La Marca

Matricola: 0512107486

Anno Accademico 2021-2022

Sommario

L'intelligenza artificiale e la blockchain sono due delle tecnologie più dirompenti e di tendenza. La tecnologia Blockchain ha la capacità di fornire l'accesso a un registro condiviso di dati, transazioni e registri in modo decentralizzato, sicuro e affidabile. La blockchain ha la capacità di governare le interazioni tra i partecipanti della rete senza intermediari o una terza parte "fidata". L'intelligenza artificiale, d'altra parte, offre capacità decisionali per macchine simili a quelle degli esseri umani. Andremo ad analizzare i vantaggi derivanti dall'integrazione di queste due tecnologie e di come potrebbero rivoluzionare le applicazioni usate oggigiorno ed essere di ispirazione per applicazioni future.

Indice

Elenco delle figure

Elenco delle tabelle

CAPITOLO 1

Introduzione

1.1 Contesto applicativo

La Blockchain è una delle innovazioni più acclamate in questi anni che sta guadagnando molta popolarità come tecnologia da adottare ampiamente in vari campi. L'Intelligenza Artificiale è un concetto in continua evoluzione ed attiene allo sviluppo di algoritmi che permettono alle macchine di comportarsi in maniera intelligente, riproducendo sistemi vicini al ragionamento umano.

1.2 Obiettivi e metodologia

Lo scopo del lavoro di ricerca è analizzare in che modo l'Intelligenza Artificiale viene utilizzata in ambito blockchain e quali conseguenze comporta l'integrazione tra le due tecnologie. La metodologia di ricerca scelta è una revisione sistematica della letteratura (SLR) in quanto questo metodo è un modo definito e metodico di identificare, valutare e analizzare la letteratura pubblicata al fine di indagare su una specifica domanda di ricerca. Nello specifico sono state adottate le linee guida proposte da Kitchenham e Charters.

1.3 Struttura della tesi

Il lavoro di tesi è organizzato in 5 capitoli ognuno dei quali è costituito da sottosezioni. Il secondo capitolo descrive la tecnologia blockchain e l'intelligenza artificiale; nel terzo capitolo

viene effettuato il processo di ricerca descrivendo passo passo ogni fase che l'ha costituito; nel quarto capitolo vengono analizzati i risultati mentre nel quinto capitolo vengono discussi gli sviluppi futuri oltre alle conclusioni.

CAPITOLO 2

Background e stato dell'arte

In questo capitolo viene descritta la tecnologia Blockchain e tutte le parti che la compongono oltre che l'Intelligenza Artificiale con un focus su alcune tecniche della stessa.

2.1 Blockchain

Le idee alla base della tecnologie blockchain, nascono già tra la fine degli anni '80 e l'inizio degli anni '90, anche se un'utile e reale implementazione è stata applicata solo nel 2008 da Satoshi Nakamoto (pseudonimo) con Bitcoin [25], creando un sistema monetario che non prevedeva la presenza di un'autorità centrale (come la Banca Centrale Europea per l'euro) affidandosi alla blockchain. Bitcoin ha poi ispirato tutte le altre cryptovalute progettate successivamente. Blockchain, infatti, è una tecnologia molto utile per eliminare la presenza di un'autorità centrale per governare e verificare le interazioni e le transazioni tra i diversi partecipanti della rete. Ogni transazione è firmata crittograficamente e verificata da tutti i partecipanti della rete che la compongono. Difatti, ogni membro della rete possiede una replica di tutte le transazioni. Questo crea un record sicuro, sincronizzato e condiviso che non può essere modificato. Blockchain si propone come mezzo per risolvere tutti i problemi del web 2.0 che hanno portato molte persone a ritenere necessaria una rivoluzione strutturale del web, per riportarlo alla sua idea iniziale di piattaforma decentralizzata, aperta ed universale. I problemi principali riguardano la gestione dei dati. Più nello specifico, non c'è la possibilità di competere con multinazionali come Google, Facebook, Amazon, Microsoft

ecc. che possiedono il monopolio dei dati controllando la totalità del traffico web abbinata ad una potenza di calcolo che possiamo considerare illimitata. Gli utenti non sono in possesso dei loro dati che vengono venduti gratuitamente, ed in maniera poco trasparente, in cambio di un servizio. Grazie alla sua naturale trasparenza, gli utenti potrebbero riprendere il possesso dei propri dati ed aver la facoltà di tracciare facilmente come essi vengano utilizzati.

Il processo di rivoluzione del web 3.0 sta pian piano avvenendo, con la nascita di progetti decentralizzati che però non sono ancora paragonabili in termini di velocità, scalabilità, costi e user experience ai progetti decentralizzati del web 2.0. Questi sono solo alcuni dei fattori che devono essere migliorati prima di poter pensare ad un effettivo passaggio al web 3.0 e dell'adozione delle applicazioni decentralizzate. A partire dal 2020 sono stati sviluppati numerosi progetti come Etherum, EOS, Golem in ambito cloud computing che sono alternative a Google Cloud e Amazon EC2, Filecoin, Storj, IPFS in ambito cloud storage che sono alternative a Dropbox, Google Drive ed Amazon S3 e progetti crypto che si propongono come nuovo metodo di pagamento nel web 3.0 andando a rimpiazzare i tradizionali metodi di pagamento online come Paypal, Visa.

2.1.1 Definizione e funzionamento

É possibile trovare molte definizioni di blockchain, alcune basate sulla struttura altre basate sulle tecnologie alla base di essa, altre sul business e la società[10]. Tutti questi aspetti sono ugualmente importanti e contribuiscono a dare una definizione esaustiva di blockchain [16]. La blockchain è un libro mastro (chiamato "ledger") digitale, decentralizzato e distribuito su un network, strutturato come una catena di registri (blocchi/nodi della rete) responsabili dell'archiviazione dei dati (dalle transazioni di valore a intere applicazioni digitali). È possibile aggiungere nuovi blocchi di informazioni, ma non è invece possibile la modifica o la rimozione di blocchi precedentemente aggiunti alla catena. In questo ecosistema, la crittografia e i protocolli di consenso garantiscono sicurezza e immutabilità. Il risultato è un sistema aperto, neutrale, affidabile e sicuro, dove la nostra capacità di utilizzare e di avere fiducia nel sistema non dipendono dalle intenzioni di nessun individuo o istituzione. La blockchain è molto più di un'infrastruttura di pagamento, di un sistema di monitoraggio della supply chain o di un gestore di identità digitale. È un sistema con le potenzialità per portare un nuovo livello di fiducia nelle applicazioni, introducendo un cambio di paradigma nelle modalità con cui esse vengono realizzate e dandoci l'opportunità di innovare liberamente. Analizziamo ora questi e molti altri concetti chiave della blockchain.

Uno degli scopi principali di una blockchain è quello di salvare informazioni. Come già detto, le informazioni salvate possono essere di qualunque tipo, da una semplice transazione di un bene a interi programmi (smart contract). Al centro della blockchain c'è il concetto di registrazione delle transazioni all'interno del ledger. I ledger sono in uso da ben prima della blockchain, essendo parte integrante dei processi commerciali fin dai tempi antichi. Mentre il concetto di ledger non è cambiato nel tempo, la tecnologia a supporto di esso si è evoluta, passando da registri cartacei ad archivi digitali. Essendo la blockchain completamente digitale, richiede ovviamente che tutto sia salvato in un ledger in forma digitale. Ledger e database potrebbero sembrare molto simili. Alla base di entrambe le tecnologie c'è infatti l'idea di salvare dei dati, ma mentre in un database è possibile inserire, cancellare e modificare i dati, in un ledger è possibile unicamente l'aggiunta di nuove informazioni. Questo è reso possibile da una combinazione di vari fattori, tra i quali la decentralizzazione, la crittografia, la teoria dei giochi ecc ...

A questo punto si potrebbe pensare di implementare un ledger utilizzando un database tradizionale e imponendo dei vincoli alle operazioni disponibili. La blockchain, però, garantisce molte altre proprietà che vanno oltre i semplici database e formano un vero e proprio ecosistema piuttosto che un semplice archivio di informazioni.

In una blockchain, il ledger digitale è strutturato come una catena di blocchi, raffigurata dall'immagine in basso, ognuno dei quali è responsabile della memorizzazione di informazioni.

Figura 2.1: Rappresentazione grafica di una rete blockchain

[1]https://bim.acca.it/cosa-e-la-blockchain/

Per capire come funziona questa tecnologia è importante conoscere i tre elementi chiave contenuti in ciascun blocco: (i) il proprio codice hash che è una rappresentazione alfanumerica dei dati. L'hash di un blocco è univoco e cambia se uno dei dati relativi al blocco viene

modificato; (ii) il codice hash del blocco che lo precede nella catena che serve come riferimento per mantenere i blocchi in ordine cronologico lineare. Se ogni blocco si riferisce all'hash univoco del blocco che lo precede, la catena rimane intatta. Se le informazioni nel blocco cambiano, cambia anche l'hash; (iii) le informazioni da memorizzare.

Questi tre elementi combinati fanno in modo che i blocchi di una blockchain non possano essere modificati e siano quindi immutabili. Se qualcuno tentasse di modificare un blocco, cambierebbe anche l'hash di quel blocco. Di conseguenza, il blocco successivo nella catena non includerebbe più l'hash del blocco che lo precede, rendendo immediatamente evidente che la blockchain è stata alterata. L'hash del blocco non viene memorizzato come riferimento, ma viene calcolato ogni volta che sia necessario. Questa funzione crittografica di hash permette la connessione tra i blocchi la quale crea un collegamento matematico indissolubile tra di essi visto che per ogni nuovo blocco generato, l'hash del blocco precedente viene inserito nell'input per generare l'hash del nuovo blocco.

2.1.2 Caratteristiche della rete e tipi di nodi di una blockchain

Una rete blockchain deve possedere alcune caratteristiche: (i) essere caratterizzato da un'autorità decentralizzata che ne detiene il controllo, (ii) essere logicamente centralizzata, ovvero essere sempre caratterizzata da un singolo stato logico. É necessario che tutti i partecipanti siano d'accordo su quale sia lo stato del sistema; (iii) ovviamente essere una rete distribuita dove ogni nodo possiede una copia della blockchain.Ogni macchina connessa alla rete blockchain è un nodo che può essere di due tipi: full-node o light-node.

Un **full-node** è un nodo che scarica e archivia localmente una copia completa della blockchain e controlla che nessuna transizione viola le regole definite dal sistema. É a tutti gli effetti indipendente, non ha bisogno di avere fiducia (il concetto di fiducia sarà analizzato successivamente) di alcun altro nodo, segue le regole imposte propagando i blocchi e le transizioni valide ignorando quelle che non ritiene valide. Usare un full-node è il modo più sicuro per interagire con una blockchain ma risulta scomodo dal momento che richiede il download dell'intera rete blockchain;

Un **light-node** non memorizza l'intera blockchain ma riceve solo i dati di cui ha bisogno da un nodo fidato (obbligatoriamente un full-node). L'utilizzo di questa tipologia di nodo implica che il nodo stesso debba delegare la fiducia da un full-node in cambio della semplicità di utilizzo. Una delle ragioni che ha contribuito alla nascita della blockchain è stata la ricerca di un sistema che fosse libero dai condizionamenti e dagli errori causati dagli esseri umani [25]. Un full-node segue categoricamente le regole imposte dal sistema a prescindere dalle

decisioni di tutti gli altri nodi e ciò ne consegue che è un sistema intrinsecamente privo di problemi che da sempre affliggono le istituzioni centralizzate, come la corruzione o la mancanza di imparzialità nelle scelte effettuate.

2.1.3 Modelli di blockchain

Le caratteristiche descritte precedentemente evidenziano di come la blockchain può essere utilizzata in un contesto in cui è richiesto uno stato globale logicamente centralizzato ma una struttura del sistema decentralizzata. Uno stato logicamente centralizzato è fondamentale per il funzionamento di ogni blockchain. Tuttavia è possibile avere un'autorità che tende alla centralizzazione. In base al modello di autorità adottato e di come agisce, esistono tre modelli di blockchain: pubblica o permissionless, ad autorizzazione e privata. Se chiunque può pubblicare un nuovo blocco parliamo di blockchain pubblica. In caso contrario, se solo determinati utenti possono pubblicare blocchi, parleremo di blockchain ad autorizzazione. Per comprendere meglio, possiamo paragonare una rete blockchain ad autorizzazione ad una intranet aziendale, mentre una rete blockchain senza autorizzazioni è come la rete internet pubblica.

Il modello di **blockchain pubblico** è quello più utilizzato ed è un sistema ad architettura decentralizzata, autorità decentralizzata e logica centralizzata [32]. Le reti blockchain permissionless sono reti decentralizzate aperte dove chiunque può pubblicare blocchi, senza bisogno dell'autorizzazione di alcuna autorità o nodo "superiore". Visto che ognuno ha il diritto di pubblicare blocchi, implica che chiunque può leggere ed emettere transazioni sulla rete blockchain.Non richiedere autorizzazioni comporta come svantaggio che gli utenti malintenzionati possano pubblicare blocchi in modo da alterare il sistema. Per evitare ciò, le reti blockchain pubbliche utilizzano un accordo tra nodi o un sistema di "consenso" (di cui parleremo successivamente) che richiede agli utenti di spendere o mantenere risorse quando tentano di pubblicare blocchi. Di solito, una blockchain aperta è anche open source, rendendo pubblicamente disponibile e consultabile il codice che ne regola il funzionamento. Ciò consente a tutti di verificarne la correttezza o di suggerire dei miglioramenti. Quando si parla di blockchain, solitamente ci si riferisce alle blockchain pubbliche.

Le **blockchain ad autorizzazione** sono reti decentralizzate aperte dove gli utenti che vogliono pubblicare blocchi, devono essere autorizzati da un' autorità che in questo caso può essere centralizzata o decentralizzata[32]. Anche queste reti utilizzano sistemi di consenso, che sono generalmente più veloci e meno costosi dal punto di vista computazionale rispetto ai modelli di consenso utilizzati dalle reti blockchain pubbliche, in quanto non richiedono

il mantenimento delle risorse per ogni nodo. Per farsi si che un utente si unisca alla rete, esso deve creare la propria identità e ricevere fiducia dagli altri nodi della rete. Questo meccanismo è utile per impedire agli utenti malintenzionati di sovvertire il sistema, infatti in caso di comportamento scorretto, la fiducia viene revocata e di conseguenza viene revocata l'autorizzazione di leggere/scrivere transizioni. Le reti blockchain ad autorizzazione possono essere utilizzate da organizzazioni che necessitano di controllare e proteggere più strettamente la propria blockchain, e le relative transazioni che avvengono su di essa. Questo controllo è di solito affidato ad un insieme di nodi, ed è la loro fiducia che devono ottenere i membri che vogliono unirsi alla rete. Un'altra applicazione è data dall'esigenza di organizzazioni diverse che vogliono lavorare insieme, ma che non si fidano completamente l'una con l'altra. Entrambe possono stabilire una rete blockchain ad autorizzazione e invitare i partner commerciali a registrare le loro transazioni su un registro distribuito condiviso. Queste organizzazioni possono determinare il modello di consenso da utilizzare, e controllare le decisioni aziendali rilevando chi tra le due parti si comporta in maniera scorretta grazie alla caratteristica della blockchain di fornire trasparenza su tutte le informazioni e tracciabilità di tutte le transizioni.

Le **blockchain private** sacrificano una completa decentralizzazione in cambio di un controllo sui permessi di accesso e solitamente delle migliori performance [16]. Come nelle blockchain ad autorizzazione, le blockchain private possiedono un livello di verifica degli accessi controllato da una o più autorità. Il livello di verifica degli accessi ha il compito di decidere chi può leggere e scrivere i dati sulla blockchain e chi può partecipare al processo di verifica delle transazioni. Il sistema viene considerato affidabile solo se gli attori scelti per il processo di verifica sono affidabili. A differenza delle blockchain ad autorizzazione, le blockchain private affidano il controllo in una singola entità. Dal momento che le blockchain completamente private rimuovono totalmente la decentralizzazione e con essa gran parte dei vantaggi peculiari della tecnologia, quelle che riscuotono il maggiore interesse sono le blockchain ad autorizzazione, dal momento che si presentano come una soluzione ibrida tra blockchain pubbliche e private. Per governi, istituzioni o aziende, entrambi i modelli possono risultare più convenienti specialmente quando è necessario un certo grado di controllo sui dati o sui partecipanti nel sistema, quando si vuole instaurare un regime di collaborazione autonomo tra diverse aziende, o per mantenere confidenziali dati sensibili[16].

2.1.4 Algoritmi di consenso

Più il numero di nodi è alto in un network più aumenta l'incertezza del sistema finale in quanto potrebbero esserci milioni di nodi che funzionano in modo indipendente in modo da non poter prevedere come si comporterà ciascuno di essi. Soprattutto in una blockchain permissionless non è possibile fidarsi di nessun partecipante alla rete coinvolto. Andremo quindi a vedere come sia possibile che il network possa decidere se accettare o meno transizioni o semplicemente decidere cosa sia giusto o sbagliato dal momento che non c'è la presenza di un'autorità centralizzata che prenda queste decisioni. Il network deve raggiungere una decisione su cosa è avvenuto all'interno della blockchain ed i nodi devono giungere ad un accordo sullo stato della rete. Difatti, essendo la blockchain una rete distribuita il problema principale è quello di mantenere coerenza tra i nodi che la costituiscono. Questo avviene tramite un processo chiamato consenso. Gli algoritmi di consenso sono un elemento essenziale di una rete blockchain. Hanno l'obiettivo di far sì che un insieme di nodi "concordino" sullo stato complessivo del sistema in ogni istante di tempo, per evitare che i nodi che compongono la rete possano comportarsi in modo malizioso con l'obiettivo di violare la coerenza del sistema. Hanno la responsabilità di proteggere il network, oltre che a verificare e convalidare le transazioni.

Spiegheremo nelle prossime sezioni i più comuni ed efficaci algoritmi di consenso da adottare per garantire l'integrità della nostra rete blockchain.

La Proof of Work (Pow) è una tecnica algoritmica introdotta nel 1997 da Adam Back [7], ideata al fine di evitare attacchi di tipo Denial of Service (DoS). Richiede una potenza computazionale molto elevata, e, quindi, un grosso dispendio elettrico, per risolvere un problema matematico estremamente complesso, ma molto facile da verificare il cui obiettivo è verificare il lavoro compiuto da un nodo. Lo scopo della PoW è quindi di dimostrare l'impiego di un certo numero di risorse per compiere un dato lavoro. Nonostante sia stata inventata per evitare attacchi Denial of Service, PoW viene utilizzata nelle Blockchain per la risoluzione del problema del consenso. Nella PoW ogni nodo ha un potere decisionale che dipende dalla potenza computazionale che mette a disposizione ed il consenso deve essere raggiunto per decidere e validare i nuovi blocchi dati da aggiungere alla blockchain che si ottiene attraverso la risoluzione di un problema crittografico, la cui soluzione può essere trovata solo tramite bruteforce. Questi problemi sono difficili da risolvere ma data una soluzione è semplice verificarne la correttezza. Maggiore è la potenza di calcolo a disposizione di un nodo più sono alte le probabilità di risolvere il problema prima degli altri e decidere il

prossimo blocco da aggiungere alla blockchain. Questo processo prende il nome di mining. Nel caso in cui durante il processi di mining, due nodi trovino contemporaneamente due soluzioni accettate (ma distinte) al problema crittografico viene creata un fork della rete blockchain. Per un breve periodo esisteranno due versioni valide e parallele della blockchain, e tra le due, grazie all'algoritmo PoW ne verrà scelta una che raggiungerà il consenso. Nello specifico, il consenso tra le due versione ricavate dalla fork verrà raggiunto dalla versione che sarà supportata dalla maggioranza della potenza di calcolo della rete, e che quindi si estenderà più velocemente.

Avendo parlato di fork, e dal momento che nell'ambito blockchain si sente spesso parlare di fork, è opportuno fare una digressione per spiegare meglio questo argomento. Il termine viene utilizzato per indicare una divisione di una blockchain (chain-split) anche se racchiude un insieme di diversi possibili scenari [16] e può avvenire in tre situazioni. Nel primo caso nodi diversi hanno opinioni temporaneamente diverse sulla cronologia delle transazione, dove comunque non è presente nessuna transazione che non rispetti le regole della blockchain. In questo caso parliamo di fork regolare. Nel secondo caso la fork ha creato una situazione in cui le regole della blockchain sono cambiate in maniera retro compatibile e tutti i nodi condividono la stessa cronologia delle transazioni. In questo caso parliamo di soft fork. Nel terzo caso la fork ha creato una situazione in cui le regole della blockchain sono cambiate in maniera non retro compatibile e diversi nodi hanno opinioni diverse sulle regole della blockchain, non condividendo la stessa cronologia delle transizioni. In questo caso parliamo di hard fork con chain split, dal momento che il consenso sulla cronologia delle transizioni è definitivamente perso.

Come descritto dal grafico in basso, una fork regolare non altera le regole del consenso, a differenza di una soft fork e di una hard fork con chain split che implicano una modifica di queste regole. Se le nuove regole sono meno rigide (non retrocompatibili) si ottiene una hard fork, nel caso in cui diventino meno rigide (retro compatibili) si ottiene una soft fork.

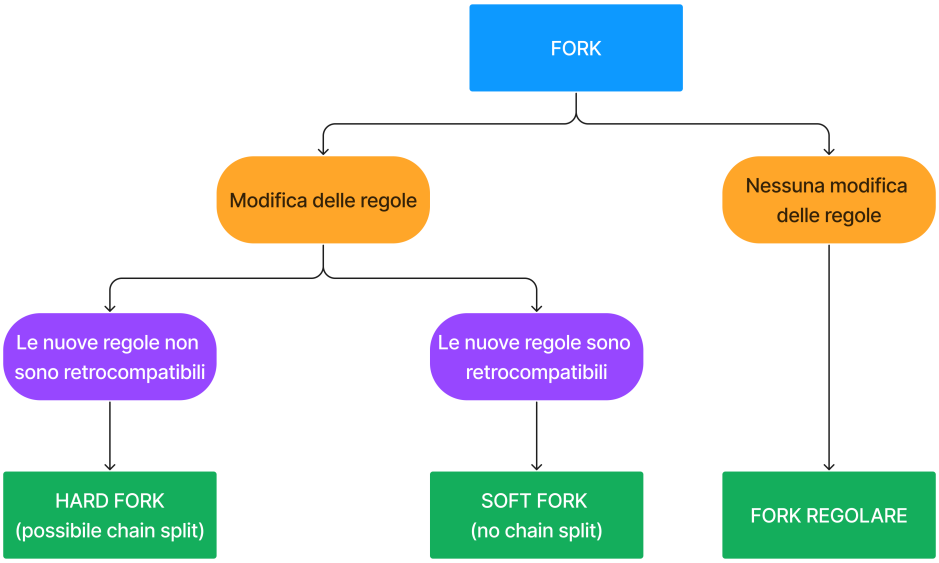

Figura 2.2: Tipologie di fork

La **Proof of Stake (PoS)** è un algoritmo di consenso nato come soluzione alla grande richiesta di potenza computazionale e di consumo energetico della PoW. La validazione dei blocchi non viene effettuata tramite la risoluzione di un problema matematico ma tramite un meccanismo di "staking". I nodi della rete entrano in gioco puntando i propri stake (monete) e il validatore viene scelto in maniera casuale in base alla sua puntata. La quantità minima di stake che vengono investiti dipendono dal nodo e dai requisiti minimi che impone la rete. Naturalmente più monete un nodo è disposto ad impegnare, più aumenta la sua possibilità che venga selezionato come validatore. Facciamo un esempio pratico. Se Antonio decide di puntare 10 stake e un altro partecipante (che chiameremo Matteo) ne punta 30, ovviamente Matteo avrà una probabilità di essere selezionato tre volte superiore a quella di Antonio. Consideriamo il caso in cui Matteo venga selezionato come validatore, e che decida di aggiungere un blocco errato alla rete, per comprometterla. Al prossimo turno di validazione, verrà di nuovo scelto (per la sua maggiore probabilità di essere selezionato) e decida nuovamente di aggiungere un blocco errato, continuando così con l'obiettivo di danneggiare la rete. Prima o poi, però, verrà scelto Antonio, che mediante la trasparenza delle transazioni su blockchain potrà mostrare agli altri nodi della rete i comportamenti scorretti di Matteo. Di conseguenza, a Matteo saranno tolti tutti i suoi stake, che verranno consegnati ad Antonio, come premio. Questo meccanismo previene e scoraggia eventuali attacchi alla rete, in quanto prima o poi si viene scoperti, con la conseguenza di perdere l'intero stake. Inoltre,

è bene ricordare che ogni utente è incoraggiato a far crescere la rete per fare in modo che i suoi stake non perdano di valore.

L'algoritmo **Byzantine Fault Tolerance (BFT)** si basa sul problema dei generali bizantini, un problema logico che spiega come un gruppo di generali bizantini potrebbe avere problemi di comunicazioni nel tentativo di accordarsi sulla mossa da fare[2]. Il dilemma suppone che ciascun generale abbia la propria armata e che ciascun gruppo sia situato in diverse posizioni intorno alla città che intendono attaccare. I generali devono decidere se attaccare o ripiegare. Non importa se attaccano o ripiegano, purché tutti i generali raggiungano il consenso, ovvero che tutti concordino su una decisione comune per eseguirla in modo coordinato. In particolare devono concordare secondo le seguenti condizioni: (i) ciascun generale deve decidere se attaccare o ripiegare, (ii) una volta presa la decisione, essa non può essere cambiata, (iii) tutti i generali devono concordare sulla stessa decisione ed eseguirla in modo sincronizzato. Un generale è in grado di comunicare con un altro soltanto tramite messaggi e la sfida principale del Problema dei Generali Bizantini è che questi messaggi possono arrivare in ritardo, essere distrutti o smarriti. Inoltre, anche se un messaggio viene consegnato, uno o più generali potrebbero decidere (per qualsiasi ragione) di agire in modo disonesto e inviare un messaggio falso per confondere gli altri generali, portando a un totale fallimento. Se riportiamo il problema dei generali bizantini nel contesto delle blockchain, ogni generale rappresenta un nodo del network e i nodi devono raggiungere il consenso sullo stato attuale del sistema. In altre parole, la maggioranza dei partecipanti di un network distribuito devono concordare ed eseguire la stessa azione al fine di evitare un completo fallimento. L'unico modo per raggiungere il consenso in questi tipi di sistemi distribuiti è avere almeno il 51% di nodi affidabili e onesti. Questo significa che se la maggioranza del network decide di agire in modo disonesto, il sistema è suscettibile a errori e attacchi. In sintesi, la Byzantine Fault Tolerance (BFT) possiamo definirlo come l'algoritmo di consenso a votazione dove viene effettuata la scelta di maggioranza votata dai nodi del network[4].

L'origine del **Proof of Elapsed Time (PoET)** risale al 2016 quando i ricercatori di Intel hanno iniziato a partecipare al progetto Hyperledger, guidato da Linux Foundation[9]. Questo è un algoritmo di consenso progettato da zero per essere altamente scalabile e mirato a blockchain private (o ad autorizzazione). PoET è potenzialmente utile per le aziende che necessitano di sistemi di controllo che garantiscano l'immutabilità delle informazioni ad esempio catene di assemblaggio altamente tecniche e automatizzate, laboratori chimici e di medicina ... PoET crea una cerchia di fiducia in cui un gruppo di partecipanti è coordinato da un controllore[28]. Il controllore ha il compito di prendere il lavoro delle persone nella cerchia

di fiducia e verificare che sia corretto. Per fare ciò, il controllore condivide un cronometro e una serie di test crittografici che consentono ai partecipanti in modo casuale di produrre blocchi all'interno della blockchain. Una volta scelto un partecipante, esso genera un blocco, sviluppa la prova crittografica e la invia al controllore che quando la riceve, la verifica e, se corretta, accetta il blocco altrimenti lo scarta. Dopo questa decisione fa scattare nuovamente il timer in modo che il processo di selezione ricominci e il test sia generato da un altro partecipante. Il processo segue questo ciclo di ripetizioni consentendo di mantenere il funzionamento della rete a tempo indeterminato.

2.2 Intelligenza Artificiale

Abbiamo fornito una panoramica sulla tecnologia Blockchain. Adesso ci focalizzeremo sull'Intelligenza Artificiale, per poter ricercare quali vantaggi possano derivare dall'integrazione di tecniche di intelligenza artificiale in ambito Blockchain. L'intelligenza artificiale può essere considerata come un'entità che porta nella sua mente un modello in scala reale del mondo esterno e delle proprie possibili azioni, che sarà in grado di trovare alternative e decidere quale sia la migliore, utilizzando l'esperienza delle azioni passate per agire in modo più opportuno e affidabile nel presente e nel futuro[30]. Paragonando un'Intelligenza Artificiale al funzionamento del cervello umano, essa può essere in grado di: agire umanamente, in modo analogo ad un essere umano; pensare umanamente, risolvendo problemi tramite le proprie funzioni cognitive; pensare razionalmente, sfruttando la logica come un essere umano; agire razionalmente, ottenendo il miglior risultato atteso tramite un processo in cui si utilizzano le informazioni a disposizione.

In base a come si comporta l'IA possiamo classificarla in due principali tipologie: Intelligenza Artificiale debole ed Intelligenza Artificiale forte. L'Intelligenza Artificiale debole agisce e pensa come se avesse capacità di pensiero, ma senza averla [15]. Lo scopo dell' IA debole non è quello di andare a creare un modelli che emulino il nostro pensiero o il nostro modo di risolvere determinati problemi, ma di creare modelli che siano in grado di agire in situazioni complesse solite di un essere umano, ad esempio la traduzione di un testo. Questo tipo di intelligenza nasce nel momento esatto in cui la complessità e capacità di calcolo è tale che nessuna persona umana è capace di fronteggiare tali calcoli e di avere una memoria così stabile anche quando si parla di un infinità di dati. Quindi l'AI debole va a risolvere tutti quei problemi che ad oggi hanno già una soluzione conosciuta, ma che per problemi di calcolo o memoria le persone non sono in grado di svolgere in modo autonomo.

L'Intelligenza Artificiale forte ha sia capacità cognitive sia conoscenza delle proprie capacità e dei propri limiti [15]. Alla base della AI forte abbiamo una serie di sistemi o algoritmi "esperti e ben allenati" i quali sono in grado di andare a riprodurre, con successo, le conoscenze e prestazioni degli uomini in determinati settori. É bene specificare che nello sviluppo della loro intelligenza non emuleranno processi di pensiero e capacità cognitive analogamente all'uomo. Ciò in cui si lavora attualmente è uno sviluppo dell'Intelligenza Artificiale forte all'interno del campo chiamato Machine Learning.

2.2.1 Machine Learning

Il Machine Learning è un sottoinsieme dell'Intelligenza Artificiale che si occupa di creare algoritmi che apprendono dai dati migliorano le performance in base ai dati che utilizzano. Sono due le caratteristiche che hanno portato alla popolarità del Machine Learning, ovvero la disponibilità dei dati e la disponibilità di strumenti computazionali adeguati. Distinguiamo quattro tipi di apprendimento: apprendimento supervisionato, apprendimento non supervisionato, apprendimento semi supervisionato, apprendimento per rinforzo. Nell'apprendimento supervisionato l'agente apprende usando dei dati etichettati, dove le etichette determinano la variabile dipendente, ovvero quello che l'agente dovrà apprendere. Oltre ai dati di input è noto anche il valore della variabile dipendente: l'apprendimento si dice supervisionato proprio perché il progettista deve fornire all'agente delle etichette che abiliteranno l'apprendimento.

Nell'apprendimento per rinforzo, l'agente compirà azioni in maniera sequenziale e, al termine di ogni sequenza, gli verrà assegnata una "ricompensa" che ha lo scopo di incoraggiare comportamenti corretti. Il sistema impara attraverso l'errore e le sequenze di decisioni che prende, e non da un insieme di dati di esempio.

Nell'apprendimento semi-supervisionato alcuni dei dati sono etichettati ed altri no. L'agente intelligente sarà tenuto ad apprendere quali sono le etichette mancanti.

L'apprendimento non supervisionato è utilizzato per problemi più complessi di quelli supervisionati. L'agente sarà quindi in grado di imparare senza conoscere il valore reale della variabile dipendente. L'apprendimento si dice non supervisionato perché il progettista lascerà all'agente il compito di apprendere sulla base dei dati a disposizione.

Per quanto riguarda l'apprendimento supervisionato descriveremo nelle prossime sezioni tecniche di Deep Learning e di Regressione. Per quanto riguarda l'apprendimento non supervisionato descriveremo tecniche di clustering.

2.2.2 Deep Learning

Il Deep Learning è una sotto categoria del Machine Learning e indica quella branca dell'intelligenza artificiale che fa riferimento agli algoritmi ispirati alla struttura e alla funzione del cervello. Si tratta di un metodo ad hoc di Machine Learning che utilizza reti neurali artificiali in strati successivi per apprendere dai dati in maniera iterativa. Una rete neurale si presenta come un sistema "adattivo" in grado di modificare la sua struttura (i nodi e le

interconnessioni) basandosi sia su dati esterni sia su informazioni interne che si connettono e passano attraverso la rete neurale durante la fase di apprendimento.

2.2.3 Regressione

Un problema di regressione porta alla costruzione di un modello che fa uso di un algoritmo di apprendimento, chiamato regressore, per predire i nuovi elementi sulla base del training set. I regressori sono essenzialmente delle funzioni matematiche che descrivono i dati. Diversi regressori si distinguono tra di loro per via delle assunzioni fatte sui dati così come delle specifiche proprietà che portano alla regressione, ma anche del numero di variabili indipendenti (predittori) di cui disponiamo. Sono disponibili numerosi approcci di analisi di regressione per fare previsioni. Inoltre, la scelta della tecnica è determinata da vari parametri, tra cui il numero di variabili indipendenti, la forma della retta di regressione e il tipo di variabile dipendente. Alcuni tipi di regressori sono la regressione singola e la regressione multipla. La differenza tra modelli singoli e multipli dipende dal numero di predittori che abbiamo a disposizione.

2.2.4 Clustering

Un problema di clustering ha l'obiettivo di classificare i dati, senza assegnare loro un'etichetta. In questo caso non abbiamo classi predefinite, ma ogni cluster può essere interpretato come una classe di oggetti simili, cioè aventi caratteristiche simili. Gli algoritmi di clustering si basano su un concetto fondamentale, che è quello della similarità. La qualità di un algoritmo di clustering dipenderà dalla misura di similarità utilizzata e dall'algoritmo stesso. Il problema frequente è quello dell'identificazione del numero ideale di cluster che un algoritmo dovrà generare. D'altro canto, essendo un apprendimento non supervisionato, gli algoritmi di clustering non hanno a disposizione delle informazioni su quante classi dovranno produrre. Altre proprietà che rendono un algoritmo di clustering buono possono essere la scalabilità, robustezza agli outlier, o interpretabilità dei risultati. Gli algoritmi di clustering si suddividono in varie tipologie:

- Esclusivi e non esclusivi. Un algoritmo di clustering è esclusivo se ogni pattern appartiene solo ad un cluster. Al contrario, se ogni pattern può essere assegnato a più di un cluster, allora parleremo di algoritmo non esclusivo.

- Gerarchico e partizionale. Un algoritmo di clustering è detto gerarchico se mira a costruire delle gerarchie di cluster, anche dette sequenze innestate di partizioni. Al contrario, un algoritmo partizionale effettua solo una partizioni dei pattern.

- agglomerativi e divisivi. Un algoritmo di clustering è agglomerativo se parte da cluster atomici che punta ad unire iterativamente in cluster più grandi. Un algoritmo divisivo parte invece da ampi cluster per dividerli poi in cluster più piccoli.

- Seriali e simultanei. Un algoritmo di clustering è seriale se elabora i pattern uno alla volta. Un algoritmo è simultaneo se invece elabora i pattern insieme.

- Graph-theoretic e algebrici. Un algoritmo di clustering è graph-theoretic se elabora i pattern sulla base della loro collegabilità. Un algoritmo è algebrico se invece elabora i pattern sulla base di criteri di errore.

CAPITOLO 3

Metodologia di ricerca

La metodologia di ricerca scelta è una revisione sistematica della letteratura (SLR) in quanto questo metodo è un modo definito e metodico di identificare, valutare e analizzare la letteratura pubblicata al fine di indagare su una specifica domanda di ricerca. Sono state adottate le linee guida proposte da Kitchenham e Charters [21] composte da diverse fasi che vedremo in seguito. Come fonte per effettuare le ricerche si è deciso di utilizzare IEEE Xplore. IEEE Xplore è un database che permette l'accesso ad articoli di riviste, atti di conferenze, standard tecnici e materiali correlati su informatica, ingegneria elettrica ed elettronica e campi affini. Contiene materiale pubblicato principalmente dall'Institute of Electrical and Electronics Engineers (IEEE) e da altri editori partner.

3.1 Obiettivi della ricerca

L'obiettivo della SLR è quello di individuare in che modo l'Intelligenza Artificiale sia attualmente utilizzata a supporto della tecnologia Blockchain per migliorarne le caratteristiche quali l'efficienza energetica, la scalabilità, l'adattabilità o come possa essere affrontato in maniera "intelligente" il problema del consenso. La fase di ricerca è iniziata con l'individuazione di due domande che guideranno tutto il processo di ricerca, che sono:

- *Quali tecniche di intelligenza artificiale sono utilizzate a supporto di sviluppo di reti blockchain?*
- *Quali vantaggi ha comportato utilizzare tecniche di intelligenza artificiale a supporto di reti blockchain?*

Vedremo nella prossima sezione come l'individuazione di queste domande sia particolarmente utile ai fini di individuare le giuste keywords per formulare le queries di ricerca.

3.2 Formulazione della query di ricerca

Per quanto riguarda la prima fase, abbiamo individuato la population, l'intervention e l'outcome sulla seconda domanda individuata. Analizzando la domanda "*Quali **vantaggi** [OUTCOME] sono stati riscontrati nello sviluppo di **reti blockchain** [POPULATION] utilizzando tecniche di **intelligenza artificiale** [INTERVENTION]?*" notiamo subito che la population è rappresentata dallo sviluppo di reti blockchain, le tecniche di intelligenza artificiale rappresentano le intervention mentre l'outcome sono rappresentati dai vantaggi riscontrati. Sono stati considerati i termini "reti blockchain", "intelligenza artificiale" e "vantaggi" in lingua inglese, considerando tutti i loro sinonimi ed altre parole che possono aiutare nella ricerca, con le relative abbreviazioni, al fine di poter ricavare un maggior numero di risultati. Per il termine "Intelligenza Artificiale" sono state considerate le parole: *(i)* AI, *(ii)* Artificial Intelligence, *(iii)* Machine Learning, *(iv)* ML, *(v)* Deep Learning. Per il termine "Blockchain" sono state considerate le parole: *(i)* Blockchain, *(ii)* Network Blockchain. Infine, per quanto riguarda i "vantaggi" abbiamo individuato delle aree più specifiche che sono: *(i)* scalability; *(ii)* security; *(iii)* consensus algorithm; *(iv)* energy efficiency; *(vi)* energy consumption;

La query di ricerca è stata formulata nel seguente modo:

- tutti i termini che rappresentavano lo stesso concetto sono stati accorpati tra parentesi tonde ed uniti attraverso l'operatore logico OR. Ad esempio avremo ('Blockchain' OR 'Network Blockchain') ed ('AI' OR 'Artificial Intelligence' OR 'Machine Learning' OR 'ML' OR 'Deep Learning');
- le frasi ottenute tra parentesi sono state unite dall'operatore logico AND.

In questo modo abbiamo ricavato la seguente query di ricerca, che ci permetterà di estrarre i dati dal database IEEE Explore:
('Blockchain' OR 'Network Blockchain') AND ('AI' OR 'Artificial Intelligence' OR 'Machine Learning' OR 'ML' OR 'Deep Learning') AND ('Scalability' OR 'Security' OR 'Consensus Algorithm' OR 'Energy Efficiency' OR 'Energy Consumption')

3.3 Estrazione e pulizia dei dati

La query nel database IEEE Explore ha prodotto 1578 risultati. Al fine di eliminare gli articoli non aderenti allo scopo della ricerca sono stati applicati dei criteri di esclusione/inclusione. I criteri di esclusione e inclusione consentono di filtrare le risorse che rispondono alle domande di ricerca di una revisione sistematica della letteratura al fine di evitare di ottenere risultati non aderenti allo scopo della ricerca.

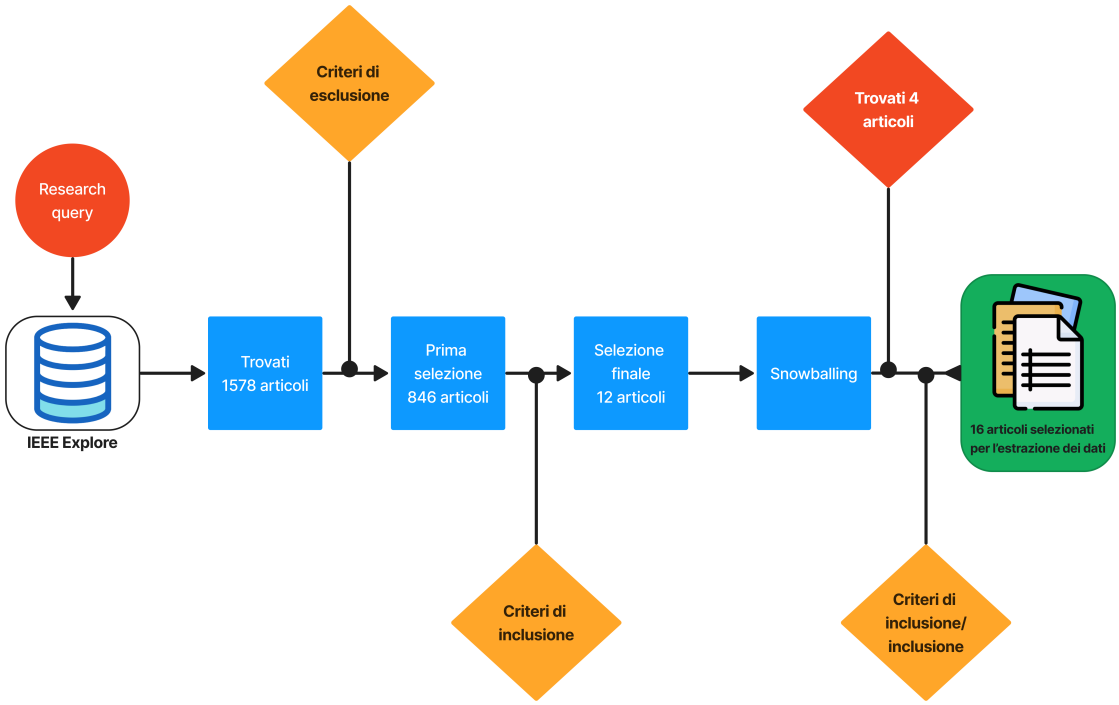

Figura 3.1: Overview del processo di estrazione dei documenti

Per quanto riguarda i criteri di esclusione, nel nostro studio abbiamo scartato gli articoli che rispettassero i seguenti vincoli:

- articoli non in inglese;
- materiale di tesi;
- materiale di sondaggi;
- articoli non completamente accessibili;
- articoli prodotti prima del 2017;
- documenti di conferenze;

L'applicazione dei criteri di esclusione ci ha permesso di escludere 732 articoli. Dopodiché i risultati sono stati esportati su un foglio Excel tramite la funzione messa a disposizione da IEEE Explore. Ogni entry della tabella risultante è formata da molteplici attributi tra cui titolo, autori, data di pubblicazione, data di aggiunta a IEEE Xplore, ISSN, ISBNs, IEEE Terms, citazioni. In un primo momento, per rendere il processo di ricerca più fluido abbiamo considerato per ogni entry solo titolo che ci ha permesso di identificare immediatamente il contesto dello studio analizzato per poter applicare i criteri di inclusione. Per i criteri di inclusioni abbiamo considerato gli articoli che applicavano metodi di intelligenza artificiale a soluzioni blockchain nel contesto dell'ingegneria del software. Analizzando i risultati, buona parte degli articoli erano incentrati su altri contesti come IoT, electric vehicular network, healtcare, 5G Network, E-voting. Tutti gli studi effettuati su questi contesti ed altri che non rispettassero i criteri di inclusione, sono stati scartati. Successivamente, per la selezione finale degli articoli, abbiamo considerato oltre al titolo, anche l'abstract. Dopo questa selezione manuale sono stati ritenuti idonei i seguenti 12 articoli:

Improving the performance of the Proof-of-Work Consensus Protocol Using Machine learning[27]
Deep Reinforcement Learning Empowered Adaptivity for Future Blockchain Networks[26]
Methods and Techniques for Privacy Preserving in Blockchain[6]
Energy-recycling Blockchain with Proof-of-Deep-Learning[12]
A Novel Delegated Proof of Work Consensus Protocol[18]
The Research of An Improved Blockchain Consensus Mechanism[33]
Performance Analyses for Applying Machine Learning on Bitcoin Miners[14]
Efficient Privacy-Preserving Machine Learning for Blockchain Network[20]
Improvement and optimization of consensus algorithm based on PBFT[35]
Improved Raft Algorithm exploiting Federated Learning for Private Blockchain performance enhancement[19]
A Hierarchy of Deep Reinforcement Learning Agents for Decision Making in Blockchain Nodes[23]
Predictive Proof of Metrics – a New Blockchain Consensus Protocol[8]

Tabella 3.1: Risultato della selezione manuale

L'individuazione dei seguenti articoli ci ha permesso di poter passare alla fase di snowballing che affronteremo nella prossima sezione.

3.4 Fase di snowballing

Lo snowballing è un metodo di ricerca per individuare articoli pertinenti all'argomento di interesse. La tecnica dello snowballing consiste nell'utilizzare l'elenco dei riferimenti

o le citazioni di un articolo per identificare altri articoli ed escludendo i lavori che non soddisfano i criteri di esclusione ed inclusione precedentemente stabiliti come la lingua, l'anno di pubblicazione e il tipo di pubblicazione. Durante la fase di snowballing sono state condotte due iterazioni. La prima di esse ha portato all'identificazione dei seguenti articoli:

When machine learning meets blockchain: A decentralized, privacy-preserving and secure design [11]
DRBFT: Delegated randomization Byzantine fault tolerance consensus protocol for blockchains [34]
Consensus-based distributed optimization: Practical issues and applications in large-scale machine learning [31]

Tabella 3.2: Risultati della prima iterazione della fase di snowballing

Successivamente è stata effettuata una seconda iterazione che ha permesso l'individuazione di un ulteriore articolo. La lista completa di articoli ricavati dalla fase di snowballing è la seguente:

When machine learning meets blockchain: A decentralized, privacy-preserving and secure design [11]
DRBFT: Delegated randomization Byzantine fault tolerance consensus protocol for blockchains [34]
Consensus-based distributed optimization: Practical issues and applications in large-scale machine learning [31]
Securing majority-attack in blockchain using machine learning and algorithmic game theory: a proof of work. [29]

Tabella 3.3: Risultati della seconda iterazione della fase di snowballing

3.5 Analisi dei dati

Prima di estrarre ed analizzare i dati per rispondere alle nostre domande di ricerca abbiamo valutato la qualità e la completezza dei documenti che sono stati selezionati nelle fasi precedenti. In particolare abbiamo formulato le seguenti domande:

Q1. Le tecniche di intelligenza artificiale utilizzati sono facilmente identificabili ?

Q2. I vantaggi ottenuti dall'implementare tecniche di intelligenza artificiale a sistemi blockchain sono chiaramente descritti ?

Queste domande saranno utili per associare un valore numerico per valutare al meglio la qualità e la completezza di ogni fonte. Ogni domanda ha tre possibili risposte esclusive che sono: "Si", "Parzialmente", "No". Al "Si" è stato associato il valore 1, al "Parzialmente" è stato assegnato il valore 0.5, mentre al "No" è stato associato il valore 0. Il punteggio di qualità complessivo è stato calcolato sommando il punteggio delle risposte alle due domande

e sono stati accettati gli articoli con punteggio di qualità almeno pari a 1. Abbiamo così valutato ogni documento in base alla loro chiarezza e alla disponibilità di informazioni utili a rispondere alle nostre domande di ricerca. L'individuazione delle informazioni per rispondere alla prima domanda sono state immediate e facili da raccogliere. Per quanto riguarda la seconda domanda, l'individuazione delle informazioni è risultata più impegnativa e ci siamo concentrati nelle sezioni degli articoli in cui venivano trattati i risultati e le limitazioni della soluzione proposta. Dal momento che i seguenti articoli avevano un punteggio di 0.5, sono stati esclusi.

Methods and Techniques for Privacy Preserving in Blockchain[6]
The Research of An Improved Blockchain Consensus Mechanism[33]

Tabella 3.4: Articoli con il punteggio di qualità minore di 1

Dopo aver individuato il set definitivo di articoli da considerare, abbiamo estratto le informazioni rilevanti per rispondere alle nostre domande di ricerca. Per estrarre i dati abbiamo definito un modulo di estrazione dei dati. Il modulo di estrazione dei dati è così descritto:

Attributo	Descrizione
Limitazione	Quale limitazione viene riscontrata nell'uso delle rete blockchain?
Blockchain	Che categoria di reti blockchain è stata utilizzata ?
Intelligenza Artificiale	Che tipo di tecniche di Intelligenza Artificale sono state utilizzate ?
Topic	Come l'Intelligenza Artificiale ha limitato/risolto il problema?

Tabella 3.5: Modulo di estrazione dei dati

CAPITOLO 4

Analisi dei risultati

Abbiamo analizzato le informazioni ottenute grazie al modulo di estrazione dei dati definito nel capitolo precedente. Riportiamo nella figura sottostante in che ambito si sono concentrati gli studi analizzati.

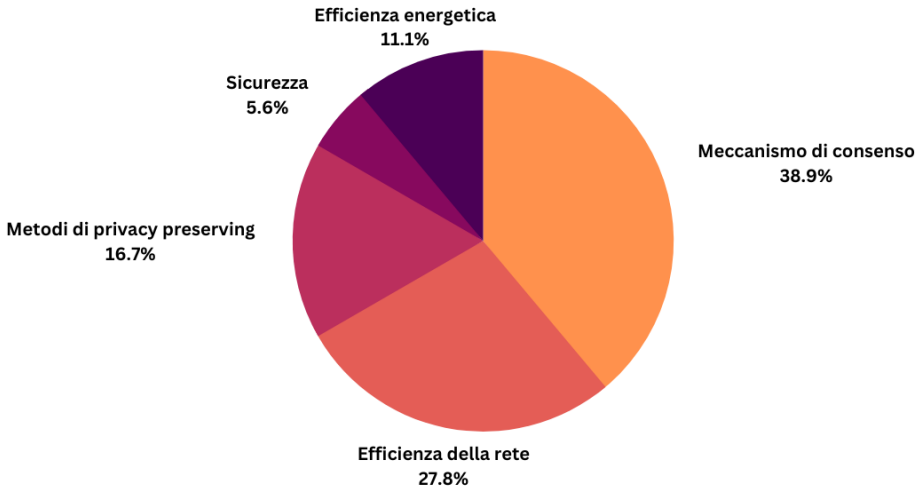

Figura 4.1: Statistica degli ambiti considerati negli studi analizzati

Possiamo osservare che la maggior parte degli studi si sono concentrati sull'integrazione dell'intelligenza artificiale in ambito blockchain per cercare di migliorare l'**efficienza della rete**, e il **meccanismo del consenso**. Dal momento che cercando di ottimizzare il meccanismo del consenso

e l'efficienza della rete è inevitabile andare ad influenzare le caratteristiche degli altri ambiti considerati, cominceremo ad analizzare i modelli proposti a partire dagli ambiti più coinvolti nella ricerca. Il 38.9% degli studi si è concentrato esplicitamente sul **meccanismo del consenso**.

Lo studio [27] ha discusso le problematiche relative alle prestazioni del protocollo di consenso Proof-of-Work proponendo un approccio per migliorarne le prestazioni senza senza rinunciare alla sicurezza o alla decentralizzazione del sistema. Grazie agli esperimenti condotti, lo studio è stato in grado di identificare il tempo necessario ai nodi per trovare il valore *nonce* durante il processo di mining. Questo tempo è stato definito come il tempo che richiede più effort all'interno del processo di consenso, indipendentemente dal traffico della rete, dalle dimensioni della stessa e dal numero di nodi connessi. Si è scelto di ottimizzare il calcolo del valore *nonce* invece che abolirlo in quanto è un pilastro nella decentralizzazione del sistema, caratteristica principale delle reti blockchain. Lo studio definisce le tecniche di Machine Learning come la tecnica adatta per ottimizzare questo processo. Lo studio [12] ha proposto come meccanismo di consenso il Proof of Deep Learning (PoDL). Tramite il Deep Learning i minatori vengono addestrati per essere più abili nel processo di mining. Questa scelta influisce anche sugli ambiti di **efficienza della rete** e **efficienza energetica**. Dagli esperimenti effettuati risulta che la media su 1.000 transazioni convalidate è di 1,96 minuti, che è un valore soddisfacente paragonato alla media di alcune blockchain popolari. Ad esempio su Bitcoin e Bitcoin cash la media è di 10 minuti mentre per Litecoin 2,5 minuti. Allo stesso tempo, PoDL non è ancora paragonabile alle prestazioni che ottiene la rete Ethereum dove la media su 1.000 transazioni si attesta tra i 10 e i 19 secondi. I paper [35] [34] hanno integrato all'algoritmo di consenso PBFT algoritmi di intelligenza artificiale. Il primo studio [35] definendo una funzione Modified Random Select (MRS) esegue uno screening preliminare dei nodi di rete e successivamente forma dei cluster di nodi. In questo modo, il processo di mining non avviene tra i nodi ma tra i cluster. Lo screening preliminare dei nodi avviene attraverso il parametro *Reputazione* di un nodo. Più la reputazione di un nodo è alta, meno è la probabilità che ha intenzioni malevole. In questo modo viene favorita la comunicazione tra nodi che rispettano le regole imposte dalla rete blockchain comportando vantaggi in ambito **sicurezza**. La funzione MRS proposta risulta avere prestazioni complessive più elevate rispetto ad altri metodi di ottimizzazione preliminari già utilizzati in letteratura. Il secondo studio [34] sulla base del PBFT definisce l'algoritmo di consenso Delegated Randomization Byzantine Fault Tolerance (DRBFT). Il DRBFT ha in comune con il meccanismo proposto nel primo studio la formazione dei cluster. A differenza del meccanismo proposto nel primo studio, la schermatura iniziale dei nodi viene fatta su un parametro *RS* reso non prevedibile per motivi di **sicurezza**. La prima selezione si basa su 3 caratteristiche fondamentali:

- Imprevedibilità: dato un insieme di nodi, i risultati della selezione sono imprevedibili prima dell'esecuzione della selezione.
- Distribuzione uniforme: per due insieme identici di nodi, l'insieme dei risultati della selezione è uniformemente distribuita e irreversibile sotto diversi parametri dinamici. Questa proprietà è il cuore dell'algoritmo RS. La distribuzione uniforme dell'insieme dei risultati significa che ogni nodo viene selezionato con la stessa probabilità rendendo la selezione equa.
- Imparzialità: Una volta dato il parametro dinamico, il risultato dell'esecuzione dell'algoritmo RS da parte di qualunque nodo è lo stesso.

L'efficacia del protocollo di consenso proposto è stata confermata dagli esperimenti effettuati. Nello specifico è stato confrontato con i protocolli di consenso tradizionali: PoW, DPoS, PBFT e con i protocolli di consenso basati su PBFT, DBFT e DPoA+PBFT. In PoW e DPoA+PBFT, c'è uno spreco di risorse in quanto i nodi spendono risorse extra per competere per i diritti di consenso. Il DRBFT risulta essere più efficiente del PoW e del PBFT ed efficiente quanto il DPoS, DBFT, DPoA+PBFT. Inoltre, risulta essere tollerante agli errori bizantini. L'integrazione del Machine Learning agli algoritmi di consenso ha permesso la nascita di Un ulteriore algoritmo di consenso chiamato Predictive Proof of Metrics (PPoM) [8]. Il PPoM è basato su metriche predittive per selezionare un nodo "fornitore" di servizi per creare una transazione sulla rete blockchain quando un nodo "cliente" richiede un servizio.

In ambito **sicurezza**, algoritmi di Machine Learning e algoritmi basati sulla teoria dei giochi sono stati utilizzati per fermare attacchi di tipo 51% sulle reti blockchain che utilizzavano Proof-of-Work come algoritmo di consenso [29]. Per proteggere la rete rete blockchain da attacchi di tipo 51% ci si concentra su una funzione di payoff derivante dalla teoria dei giochi. La funzione di payoff è alimentata da algoritmi di Machine Learning per predire se un attacco è probabile o meno in base al valore della informazioni contenute sulla rete. Il sistema proposto implementa una serie di regole che si attiveranno quando l'attacco è probabile per evitare che un nodo attaccante possa confermare blocchi sulla rete relativi ad una transazione malevola.

L'implementazione proposta dallo studio [26] tratta gli aspetti relativi alla **scalabilità**, **larghezza di banda**, **limitate risorse di calcolo** come un problema congiunto. La soluzione va in ambito **algoritmi di consenso** ed utilizza algoritmi di deep learning. Nella soluzione proposta la rete blockchain si comporta in modo differente in base alle esigenze che devono essere soddisfatte al momento. Il modello proposto implementa 4 algoritmi di consenso. La scelta dell'algoritmo di consenso da utilizzare viene affidata ad un modello di deep learning che permetterà la gestione adattiva nella blockchain. I protocolli di consenso candidati, i nodi e i collegamenti disponibili nel livello delle risorse della blockchain adattiva sono gestiti dal coordinatore e dai servizi di supporto per le applicazioni blockchain. Gestire in maniera adattiva la rete implica che il coordinatore debba gestire un gran numero di stati del sistema, il che risulta difficile con le soluzioni tradizionali. Nel modello proposto, il coordinatore raccoglie gli stati del sistema dal livello delle risorse e delle applicazioni, i requisiti QoS richiesti degli utenti, le capacità di calcolo dei nodi e le risorse disponibili e inserisce lo stato del sistema nel modello di deep learning. Il modello di deep learning supporterà il coordinatore nella gestione della rete essendo in grado di apprendere il vantaggio relativo ad un'azione a differenza degli approcci tradizionali.

Nel paper [23] viene proposto un proposto un framework per ottimizzare il processo decisionale tra nodi distribuiti. Il framework introduce l'uso di tecniche di deep learning. Nello specifico, la rete verrà organizzata in layer, ed ogni layer verrà diviso in settori. Ogni settore è formato dal nodo della rete blockchain e dall'agente intelligente che supporterà il nodo nel processo decisionale. In maniera simile, un'ulteriore modello [14] proposto implementa ed analizza la possibilità di integrare algoritmi di Machine Learning alle operazioni di mining dei nodi della rete Bitcoin per valutare come gli algoritmi di ML impattano sulle operazioni di mining al crescere della dimensione della rete. Gli esperimenti hanno dimostrato che a velocità per il calcolo dell'hash di un blocco diminuisce di circa $0,2x10^6 h/s$ quando si utilizzano gli algoritmi di Machine Learning.

Complessivamente, gli algoritmi semi-supervisionati hanno prestazioni di costo peggiori rispetto a quelli supervisionati.

I paper [6] [20] [11] affrontano la tematica della privacy, indicando come le tecniche di Machine Learning siano utili per preservare la privacy e la sicurezza dei dati memorizzati nei nodi della blockchain. Nello specifico si sono occupati di preservare la privacy delle informazioni per qualsiasi sistema di Machine Learning decentralizzato. Inoltre [11] propone l'architettura LearningChain che non ha bisogno di alcun requisito sul tipo di blockchain e quindi implementabile su qualsiasi Blockchain. L'architettura proposta favorisce l'utilizzo di dati privati da fornire agli algoritmi di Machine Learning, la cui privatezza sarà assicurata dall'architettura stessa. L'approccio proposto dallo studio [31] verte in ambito scalabilità utilizzando algoritmi di Machine Learning. I vantaggi dell'approccio riguardano sia la scalabilità che una maggiore resistenza ai guasti dei nodi. Gli esperimenti mostrano come gli algoritmi di Machine Learning integrati alla rete blockchain funzionano correttamente anche in circostanze in cui l'overhead di comunicazione dei nodi non è equamente distribuito o dove i nodi hanno carichi di lavoro sbilanciati.

CAPITOLO 5

Conclusioni e sviluppi futuri

L'obiettivo della tesi è di analizzare le tecniche di Intelligenza Artificiale utilizzate in ambito Blockchain. Possiamo affermare che in letteratura è presente un numero ragionevole di architetture di blockchain che integrano tecniche di Intelligenza Artificiale. Il grafico a torta riporta le tecniche di intelligenza artificiale utilizzate.

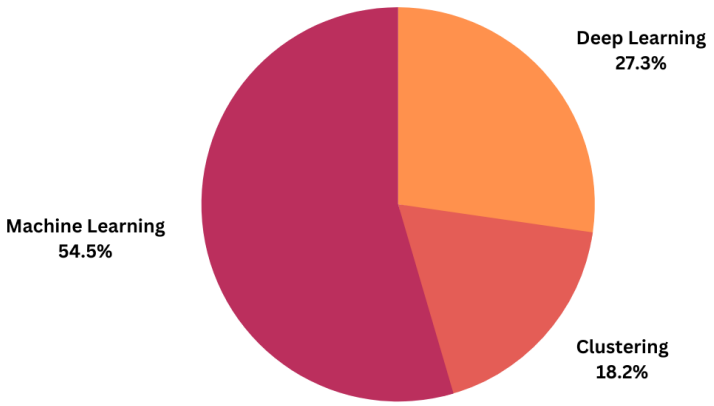

Figura 5.1: Statistica delle tecniche di intelligenza artificiale utilizzate

L'utilizzo di tecniche di Intelligenza Artificiale in ambito blockchain si è rilevata particolarmente utile nell'ambito dei meccanismi di consenso. Nello specifico ha permesso la nascita di ulteriori algoritmi di consenso che risultano aumentare l'efficienza e diminuire il consumo energetico della rete. Gli algoritmi di Machine Learning sono risultati utili nel preservare la privacy delle informazioni contenute nei nodi di una rete blockchain oltre che aumentarne la sicurezza mentre

gli algoritmi di Deep Learning sono risultati utili per supportare i nodi di una rete blockchain nel processo decisionale durante il mining di informazioni.

In futuro, si può estendere il processo di ricerca ad altre fonti, come Scopus. Inoltre si potrà continuare, nei prossimi anni, a tenere d'occhio i paper pubblicati che rispettano i criteri utilizzati per questa ricerca in modo da seguire nel tempo le implementazioni proposte che integrano algoritmi di Intelligenza Artificiale nel contesto delle blockchain.

Bibliografia

[1] Binance Academy. Proof of stake. 2018.

[2] Binance Academy. La byzantine fault tolerance spiegata. 2021. (Citato a pagina 12)

[3] Binance Academy. Cos'è la proof of work (pow)? 2022.

[4] Elli Androulaki, Artem Barger, Vita Bortnikov, Christian Cachin, Konstantinos Christidis, Angelo De Caro, David Enyeart, Christopher Ferris, Gennady Laventman, Yacov Manevich, Srinivasan Muralidharan, Chet Murthy, Binh Nguyen, Manish Sethi, Gari Singh, Keith Smith, Alessandro Sorniotti, Chrysoula Stathakopoulou, Marko Vukolic, Sharon Weed Cocco, and Jason Yellick. Hyperledger fabric: A distributed operating system for permissioned blockchains. *CoRR*, abs/1801.10228, 2018. (Citato a pagina 12)

[5] Antonopoulos. *The internet of Money*. Createspace Indipendent Pub, 2016.

[6] Sowmiya B and Poovammal E. Methods and techniques for privacy preserving in blockchain. In *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, pages 1346–1351, 2020. (Citato alle pagine 21, 23 e 27)

[7] Adam Back. Hashcash: Proof-of-work system. 2010. (Citato a pagina 9)

[8] Venkata Siva Vijayendra Bhamidipati, Michael Chan, Arpit Jain, Ashok Srinivasa Murthy, Derek Chamorro, and Aniruddh Kamalapuram Muralidhar. Predictive proof of metrics – a new blockchain consensus protocol. In *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, pages 498–505, 2019. (Citato alle pagine 21 e 26)

[9] bit2me. Che cos'è proof of elapsed time (poet)? (Citato a pagina 12)

[10] Philip Boucher, S Nascimiento, and M Kritikos. Come la tecnologia blockchain può cambiarci la vita. *How blockchain technology can change our life.) Brussels: Parlamento Europeo*, 2017. (Citato a pagina 4)

[11] Xuhui Chen, Jinlong Ji, Changqing Luo, Weixian Liao, and Pan Li. When machine learning meets blockchain: A decentralized, privacy-preserving and secure design. *2018 IEEE International Conference on Big Data (Big Data)*, pages 1178–1187, 2018. (Citato alle pagine 22 e 27)

[12] Changhao Chenli, Boyang Li, Yiyu Shi, and Taeho Jung. Energy-recycling blockchain with proof-of-deep-learning. In *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pages 19–23, 2019. (Citato alle pagine 21 e 25)

[13] Thang N. Dinh and My T. Thai. Ai and blockchain: A disruptive integration. *Computer*, 51(9):48–53, 2018.

[14] Wenjun Fan, Jinoh Kim, Ikkyun Kim, Xiaobo Zhou, and Sang-Yoon Chang. Performance analyses for applying machine learning on bitcoin miners. pages 1–4, 2021. (Citato alle pagine 21 e 26)

[15] Johnathan Charles Flowers. Strong and weak ai: Deweyan considerations. In *AAAI Spring Symposium: Towards Conscious AI Systems*, volume 22877, 2019. (Citato a pagina 14)

[16] Raffaele Bianchi Gianluca Chiap, Jacopo Ranalli. *Blockchain Tecnologia e applicazioni per il business*. Hoepli. (Citato alle pagine 4, 8 e 10)

[17] Shihab Shahriar Hazari and Qusay H Mahmoud. A parallel proof of work to improve transaction speed and scalability in blockchain systems. In *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)*, pages 0916–0921. IEEE, 2019.

[18] Mostefa Kara, Abdelkader Laouid, Ahcene Bounceur, Farid Lalem, Muath AlShaikh, Romaissa Kebache, and Zaoui Sayah. A novel delegated proof of work consensus protocol. In *2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP)*, pages 1–7, 2021. (Citato a pagina 21)

[19] Donghee Kim, Inshil Doh, and Kijoon Chae. Improved raft algorithm exploiting federated learning for private blockchain performance enhancement. 2021. (Citato a pagina 21)

[20] Hyunil Kim, Seung-Hyun Kim, Jung Yeon Hwang, and Changho Seo. Efficient privacy-preserving machine learning for blockchain network. *IEEE Access*, 7:136481–136495, 2019. (Citato alle pagine 21 e 27)

[21] Barbara Kitchenham, O. Pearl Brereton, David Budgen, Mark Turner, John Bailey, and Stephen Linkman. Systematic literature reviews in software engineering – a systematic literature review. *Information and Software Technology*, 51(1):7–15, 2009. Special Section - Most Cited Articles in 2002 and Regular Research Papers. (Citato a pagina 18)

[22] Batta Mahesh. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, 9:381–386, 2020.

[23] Arafat Abu Mallouh, Omar Abuzaghleh, and Zakariya Qawaqneh. A hierarchy of deep reinforcement learning agents for decision making in blockchain nodes. In *IEEE EUROCON 2021 - 19th International Conference on Smart Technologies*, pages 197–202, 2021. (Citato alle pagine 21 e 26)

[24] David Malone and K.J. O'Dwyer. Bitcoin mining and its energy footprint. pages 280–285, 01 2014.

[25] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2009. (Citato alle pagine 3 e 6)

[26] Chao Qiu, Xiaoxu Ren, Yifan Cao, and Tianle Mai. Deep reinforcement learning empowered adaptivity for future blockchain networks. *IEEE Open Journal of the Computer Society*, 2:99–105, 2021. (Citato alle pagine 21 e 26)

[27] Mujistapha Ahmed Safana, Yasmine Arafa, and Jixin Ma. Improving the performance of the proof-of-work consensus protocol using machine learning. In *2020 Second International Conference on Blockchain Computing and Applications (BCCA)*, pages 16–21, 2020. (Citato alle pagine 21 e 25)

[28] Khaled Salah, M Habib Ur Rehman, Nishara Nizamuddin, and Ala Al-Fuqaha. Blockchain for ai: Review and open research challenges. *IEEE Access*, 7:10127–10149, 2019. (Citato a pagina 12)

[29] School of Computer Science Somdip Dey and Electronic Engineering University of Essex. Securing majority-attack in blockchain using machine learning and algorithmic game theory: A proof of work. (Citato alle pagine 22 e 26)

[30] Peter Norvig Stuart Russell. Artificial intelligence, a modern approach. (Citato a pagina 14)

[31] Konstantinos I. Tsianos, Sean F. Lawlor, and Michael G. Rabbat. Consensus-based distributed optimization: Practical issues and applications in large-scale machine learning. *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1543–1550, 2012. (Citato alle pagine 22 e 27)

[32] Dylan Yaga, Peter Mell, Nik Roby, and Karen Scarfone. Blockchain technology overview. *arXiv preprint arXiv:1906.11078*, 2019. (Citato a pagina 7)

[33] YaJuan Yao, Fang Tian, and Cheng Zhang. The research of an improved blockchain consensus mechanism. In *2020 2nd International Conference on Applied Machine Learning (ICAML)*, pages 305–310, 2020. (Citato alle pagine 21 e 23)

[34] Yu Zhan, Baocang Wang, Rongxing Lu, and Yong Yu. Drbft: Delegated randomization byzantine fault tolerance consensus protocol for blockchains. *Information Sciences*, 559:8–21, 2021. (Citato alle pagine 22 e 25)

[35] Liang Zhao, Bin Li, QingLei Zhou, and XiaoJie Chen. Improvement and optimization of consensus algorithm based on pbft. 2022. (Citato alle pagine 21 e 25)