



**UNIVERSITÀ
DEGLI STUDI
DI BERGAMO**

Dipartimento di
Ingegneria Gestionale, dell'Informazione e della Produzione

Corso di laurea in
Ingegneria informatica

Classe n. L-8

CoinStatus

Candidato:

*Silviu filote
Pandelea Mihaita*

Matricola n. 1059252

Relatore:

*Chiar.mo/Chiar.ma Prof. (Prof.ssa)
Mattia Cattaneo*

Anno Accademico
2022/2023

Alla mia famiglia e ai miei amici

Indice

1	Introduzione	1
2	Le criptovalute	2
2.1	Cenni storici criptovalute	3
2.2	Cosa sono le transazioni	4
2.3	Come acquistare le criptovalute	5
2.4	Cosa si può fare con le criptovalute	5
2.5	Nuovi sistemi di business	5
2.6	Le initial coin offering (ICO)	7
2.7	Piattaforme di trading	7
3	La blockchain	8
3.1	Le caratteristiche principali	9
3.2	Il funzionamento della blockchain	9
3.3	Mining	10
4	Analisi preliminari	12
4.1	Osservazioni conclusive	14
5	Progetto	15
5.1	Domanda di ricerca	15
5.2	Suddivisione del progetto	16
5.3	Framework utilizzati	19
5.3.1	Javascript: Axios e jQuery	19
5.3.2	Html e css: bootstrap	22
5.3.3	Python: flask e librerie di data science	23
5.4	Funzionamento generale	24
5.5	Ottimizzazioni applicate	27
5.6	Importanza dei strumenti di previsione	27

6	Modelli di previsione	29
6.1	Rete neurale	30
6.2	Le RNN	31
6.3	LSTM	32
6.4	Funzioni di attivazione	33
6.4.1	Sigmoid	34
6.4.2	tanh	35
6.4.3	ReLU	36
6.5	Struttura interna	37
6.6	Il modello	38
6.7	Analisi modello	41
7	Risultati	43
8	Conclusioni	44

Elenco delle figure

2.1	Bitcoin e la sua simbologia	4
3.1	Gestione di una transazione all'interno di una blockchain	10
3.2	Computer con hardware dedicato che esegue il processo di mining . .	11
4.1	Numero totale di criptovalute legittime quotate su CoinMarketCap dal 2013 ad oggi	12
4.2	Numero di utilizzatori per continente	13
4.3	Capitalizzazione massima criptovalute	13
5.1	Esempio richiesta POST di una risorsa	21
5.2	Pagina HTML con inclusione di CSS	23
5.3	Semplice richiesta dati	26
6.1	Rete neurale	30
6.2	Struttura di un neurone	30
6.3	RNN	31
6.4	Struttura interna LSTM	32
6.5	Struttura interna	33
6.6	funzione σ	34
6.7	funzione \tanh	35
6.8	funzione ReLU	36
6.9	Suddivisione LSTM in gates	37
6.10	Esecuzione modello	38
6.11	Neural network nel caso di Ethereum	39
6.12	Metriche di bontà del modello	40
6.13	Modello LSTM	41
7.1	Pagina index.html	43

Capitolo 1

Introduzione

La storia dell'umanità, sin dagli albori, è stata segnata da grandi rivoluzioni tecnologiche che hanno stravolto significativamente la nostra quotidianità.

Con il passare delle ere storiche le tecnologie hanno preso forme sempre più complesse e con impieghi sempre più specifici in determinati ambiti all'interno della società, con l'obiettivo di migliorare processi produttivi esistenti o di portare innovazione.

In età contemporanea questo effetto è molto più tangibile ed evidente. La ricerca ha spostato l'attenzione dalle architetture meccanico-analogiche verso quelle digitali, avviando così un processo di digitalizzazione, che ha portato con sé la creazione di nuove architetture hardware e software in grado di elaborare dati e produrre dei risultati. Con l'era digitale o dell'informazione si intende proprio questo periodo di transizione caratterizzato dalla diffusione di prodotti digitali e del loro impatto all'interno dei vari settori.

L'introduzione di queste tecnologie in un qualsiasi tipo di contesto genera un pacchetto di conoscenze necessarie da apprendere per l'utilizzo e possono riscrivere o dare origine a nuove attività.

Le criptovalute sono state impiegate all'interno del progetto di tesi e sono un esempio di tecnologia digitale che ha rivoluzionato completamente il sistema economico e fiscale, introducendo un mezzo di pagamento virtuale per effettuare qualsiasi tipo di transazione sul web.

L'obiettivo di questa tesi è dunque quello di sviluppare uno strumento che sia in grado di analizzare le criptovalute e di fornire dei metodi di previsione del prezzo delle stesse per combattere l'instabilità del mercato crypto. E però importante comprendere il funzionamento della criptovalute e dell'architettura di base su cui è fondata prima di passare al progetto vero e proprio. Dunque di seguito nei prossimi capitoli si analizzeranno in maniera dettagliata le funzionalità e gli aspetti chiave delle due.

Capitolo 2

Le criptovalute

Le criptovalute sono monete digitali che utilizzano la crittografia per garantire la sicurezza, l'immutabilità e la trasparenza delle transazioni.

Le criptovalute non esistono in forma fisica, si generano e si scambiano esclusivamente per via telematica. Lo scambio inteso come transazione avviene in modalità peer-to-peer, ossia all'interno delle rete i nodi che partecipano non sono gerarchizzati unicamente sotto forma di client o server, ma sono tra loro paritari e possono allo stesso tempo fungere da server e client. La transazione quindi avviene direttamente tra i nodi senza la necessità di un intermediario che faccia da tramite e che si interponga nell'operazione, questa funzionalità prende il nome di decentralizzazione del servizio e viene garantita dall'architettura di base, ovvero dalla blockchain, che verrà spiegata ed approfondita in seguito.

Ad ogni criptovaluta viene quindi associato un valore virtuale il quale poi viene convertito attraverso le classiche monete a corso legale. Durante il tempo, il valore associato ad ogni singola moneta può subire drastiche variazioni, che oltre che dipendere dalla legge della domanda e dell'offerta dipende anche dai fattori esterni, nati a causa della decentralizzazione dell'architettura che hanno maggior impatto sul loro valore:

- Numero totale di criptovalute circolanti, chiamato anche **max supply**. All'interno di questo numero sono comprese:

$$Max\ supply = Cripto\ minate + Cripto\ perse + Cripto\ da\ minare$$

- Il valore di tutte le valute in circolazione e la percezione degli utenti sugli sviluppi futuri del mercato;

⁰non esiste in forma fisica

- Criptovalute come materia prima. Il valore delle criptovalute non è legato ad nessuna performance economica o politica monetaria di alcun paese;
- L'opinione su una criptovaluta e quanta copertura mediatica riceve;
- La capacità di integrazione della relativa criptovaluta all'interno dei portali di pagamento;
- Eventi fondamentali come aggiornamenti normativi, violazioni della sicurezza e rallentamenti economici dovuti alla struttura logica della criptovaluta stessa;

Le criptovalute sono nate per essere un'alternativa alle valute tradizionali, anche se in realtà sono state create con l'intento di funzionare come metodo di pagamento. Purtroppo a causa dell'elevata volatilità le criptovalute sono impiegate come mezzo di investimento.

2.1 Cenni storici criptovalute

Nel 1982 David Chaum, noto ricercatore americano, pubblicò la sua prima ricerca introducendo il concetto di crittografia e nel 1990 applicò questi principi per fondare DigiCash Inc. DigiCash fu un servizio che permise agli utenti di effettuare transazioni usando un software proprietario. L'applicativo consentì di inviare pagamenti strutturando per la prima volta il concetto di chiave privata e chiave pubblica, lo stesso principio di base utilizzato oggi dalle criptovalute.

Dopo il fallimento di DigiCash nel 2002 si succedettero numerose pubblicazioni in ambito tecnico, portando alla luce nuove tecnologie, che contribuirono alla nascita vera e propria della prima criptovaluta. Le pubblicazioni più importanti furono:

- "How to Make a Mint: the Cryptography of Anonymous Electronic Cash"[3] della National Security Agency, in cui si cerca di realizzare un sistema che implementi un modello di pagamento elettronico;
- La tesi di ricerca di Wei Dai, ingegnere informatico laureato alla università di Washington, che definisce da un punto di vista teorico una nuova moneta, i "B-money"[2]. All'interno della tesi vengono ulteriormente definiti altri argomenti importanti come: la decentralizzazione e protezione dell'anonimato.
- "Bitcoin a peer-to peer electronic cash system"[4], pubblicato il 31 ottobre 2008 da Satoshi Nakamoto. La ricerca presenta la blockchain, un'architettura peer-to-peer rivoluzionaria in grado di svolgere pagamenti online in maniera decentralizzata.

A seguito di queste pubblicazioni nacque nel 2009 la prima criptovaluta, il Bitcoin, che rappresenta il risultato evolutivo di tutti i concetti teorici definiti nelle ricerche del tempo.

Dopo il successo di Bitcoin le criptovalute aumentarono esponenzialmente così come le piattaforme di trading e di exchange, che provocarono la rapida crescita dell'ecosistema crypto.



Figura 2.1: Bitcoin e la sua simbologia

2.2 Cosa sono le transazioni

Nel linguaggio tecnico-economico il termine transazione indica un contratto diretto a prevenire o comporre una lite¹ tra due o più parti, giungendo quindi ad un accordo. Nel mondo delle crypto con il termine transazione si indica ogni tipo di operazione che comprende il trasferimento di criptovalute tra due parti. Tutte le transazioni sono registrate in modo decentralizzato, senza quindi l'intermediazione di istituti centralizzati come banche e governi, velocizzando dunque il processo e mantenendo l'anonimato all'interno della rete. Gli attori di una transazione devono possedere i wallet e le chiavi pubblica/privata associate ad ogni wallet. Il trasferimento delle monete avviene quando una delle due parti pone come indirizzo finale la chiave pubblica di un wallet destinatario e per completare l'operazione si utilizza la chiave privata come firma per convalidare il processo. La procedura richiede solamente la chiave pubblica del destinatario per completare l'operazione di trasferimento, quindi non è strettamente necessario che i due attori abbiano un accordo, ma basta semplicemente che il mittente sia consenziente. La maggior parte delle violazioni ai wallet mirano proprio a svuotare il portafoglio virtuale effettuando queste transazioni "non confermate".

¹Controversia giudiziaria

2.3 Come acquistare le criptovalute

Per entrare in possesso di criptovalute è necessario aprire un conto titoli presso broker specializzato in materia, oppure un exchange compilando un modulo di iscrizione online. Inoltre è indispensabile creare un wallet, ossia un portafoglio virtuale che racchiuderà tutte le criptovalute acquistate. Il wallet viene indetificato da una chiave pubblica e da una privata. La chiave pubblica identifica il wallet e tramite quest'ultima possiamo ricevere criptovalute, mentre la chiave privata, che non è assolutamente da condividere, serve a convalidare la transazione. Depositando poi sul nostro conto il capitale di interesse possiamo effettuare un ordine contenente le criptovalute che desideriamo, una volta completato l'ordine il wallet verrà aggiornato di conseguenza.

2.4 Cosa si può fare con le criptovalute

Le valute a corso legale svolgono le funzioni: di unità di conto, di mezzo di pagamento e deposito di valore. Le criptovalute non hanno la stessa applicazione a causa dell'elevata volatilità realizzata dalla blockchain. I prezzi delle criptovalute tendono ad oscillare molto anche durante la stessa giornata rendendo impossibile prezzare beni e servizi in unità di criptovalute, tuttavia esistono negozi fisici e online che offrono comunque la possibilità di acquisto utilizzando le criptovalute. Dunque la maggior parte delle attività che vengono intraprese mirano a fare trading sul mercato per realizzare profitto.

2.5 Nuovi sistemi di business

Le infrastrutture tecnologiche che supportano tali monete virtuali hanno dato origine a nuovi sistemi di business e attività che si sono radicati all'interno della nostra società. Le più importanti sono:

- **Holding:** la strategia dell'holding, detta anche HODL consiste nell'acquistare criptovalute conservandole per un periodo di tempo per poi rivenderle;
- **Trading:** si cerca di trarre il massimo profitto nel minor tempo possibile. Consiste nell'effettuare molteplici transazioni per poi rivenderle, ma su una scala temporale molto ristretta se non immediata;

- **Mining:** è una tecnica che impiega la capacità di elaborazione di un calcolatore per risolvere particolari calcoli complessi. A fine di questo carico computazione si sbloccano delle ricompense che possono variare;
- **Arbitraggio:** questa modalità permette di guadagnare dalla differenza di prezzo che si crea tra due o più Exchange. Esistono due tipi di arbitraggio: quello regolare e quello triangolare.
 Nell'arbitraggio regolare si acquista e si rivende la stessa cryptovaluta su differenti piattaforme di exchange e il profitto si trae dalla differenza di prezzo associata alla stessa moneta. Nell'albitraggio triangolare si sfrutta la differenza di prezzo ottenuta attraverso una valuta intermedia. Acquisto una cryptovaluta in una valuta a corso legale, converto la cryptovaluta in un'altra con valore superiore ed effettuo il cambio;
- **Staking:** attività simile a quella di mining ma impiegando una potenza di calcolo minore;
- **Syndacate:** sono delle promozioni effettuate da varie piattaforme che rivedono criptovalute a un prezzo inferiore;
- **Airdrop:** ti permette di ricevere criptovalute gratuitamente effettuando piccole operazioni, come ad esempio: l'iscrizione ad un canale Telegram oppure ad una newsletter;
- **Fork:** si definisce una proposta di modifica al codice originario (protocollo) il cui fine è il miglioramento in termini prestazionali di una valuta digitale, generando una nuova versione della blockchain. Vengono messe in palio delle ricompense per coloro che emettono una nuova versione del protocollo;
- **Dark Pool:** sono dei mercati che si differenziano dai soliti exchange in quanto consente agli investitori istituzionali di acquistare o vendere grandi quantità di azioni, senza comunicare pubblicamente i dettagli dell'ordine al mercato. Alcuni traders sfruttano questo sistema per avere accesso ad informazioni privilegiate o per guadagnare attraverso l'albitraggio tra dark pool e mercato pubblico.

2.6 Le initial coin offering (ICO)

Le ICO rappresentano un tipo di criptovaluta con una funzionalità ben diversa dalle altre. Questo strumento finanziario è finalizzato alla raccolta fondi necessari a finanziare un progetto imprenditoriale tramite l'emissione di coin o token. Coloro che vogliono sostenere tale progetto e finanziare l'attività acquistano questi token attraverso le valute in corso legale oppure più spesso con le criptovalute.

Solitamente questo tipo di attività vengono intraprese da realtà imprenditoriali piccole o di start up, che per stabilizzarsi sul mercato necessitano di capitali. E' dunque molto importante per il proprio progetto:

- avere un'idea di progetto innovativa e creativa che incoraggi gli investitori a contribuire/acquistare questi token;
- far conoscere la propria idea all'interno di varie comunità di investitori legati al mondo ICO attraverso pubblicità ben studiate;
- riuscire a raccogliere fondi per completare lo sviluppo della propria idea;
- avere un team con solide competenze;

2.7 Piattaforme di trading

Le piattaforme di trading sono dei applicativi disponibili sotto forma di siti web oppure app, che raccolgono al loro interno la maggior parte delle informazioni sull'andamento economico e sulle origini delle criptovalute. Esistono piattaforme che oltre a raccogliere e rendere disponibili informazioni, offrono la possibilità di effettuare operazioni di compravendita attraverso l'accesso in tempo reale ai dati dei mercati finanziari. Inoltre queste piattaforme mettono a disposizione tool finanziari, come per esempio grafici, che permettono di effettuare analisi aggiuntive sull'evoluzione delle criptovalute.

Capitolo 3

La blockchain

Le criptovalute sono la prima applicazione della tecnologia blockchain e di seguito in questo capitolo si darà una spiegazione dettagliata su questa architettura.

La tecnologia blockchain appartiene alla famiglia delle distributed ledger. Con il termine Distributed Ledger Technologies (DLT) si fanno riferimento a quelle tecnologie, protocolli e infrastrutture che sono in grado di salvare le transazioni e tutti i dettagli della transazione all'interno di libri mastri o registri, che sono distribuiti geograficamente su diversi nodi. Dunque si può considerare la blockchain come un registro distribuito all'interno di una rete che va a memorizzare le transazioni effettuate da una determinata criptovaluta in maniera sicura, verificabile e permanente. Il libro mastro possiede varie copie che sono salvate su diversi nodi della rete in maniera ridondante per garantire l'integrità del registro e per questioni di sicurezza. I nodi della rete sono collegati tra di loro in maniera distribuita e ogni nodo viene associato ad un miner che partecipa al sistema.

La blockchain archivia le transazioni andate a buon fine in blocchi e i vari blocchi sono collegati tra loro in maniera cronologica e resi sicuri mediante l'uso della crittografia. I dati che vengono registrati in un blocco sono per loro natura immutabili, quindi non è possibile alterare o modificare la struttura senza l'approvazione da parte della rete (modello cooperativo), il sistema dispone inoltre di meccanismi integrati di sicurezza che impediscono l'inserimento di transazioni non autorizzate. Il processo di validazione delle transazioni, noto anche come mining, garantisce sicurezza e stabilità, ma richiede tempo e grande potenza di calcolo per la risoluzione di complessi calcoli algoritmici.

3.1 Le caratteristiche principali

Le caratteristiche principali della tecnologia blockchain si possono sintetizzare come segue:

- **Decentralizzazione:**

L'archiviazione di una transazione passa attraverso un processo decisionale che non è riservato a una singola entità, bensì è condiviso tra tutti i nodi della rete o una parte di essi. La registrazione della transazione avviene se e solo se la maggioranza viene raggiunta tra i vari partecipanti. Inoltre i partecipanti della rete sono tutti posti allo stesso livello, in questo modo nessun nodo può esercitare la propria autorità o controllare il comportamento dei suoi simili.

- **Immutabilità**

Le transazioni non possono essere manomesse o alterate in alcun modo da operazioni esterne al ecosistema. Una volta aggiunto il blocco alla catena servono alti privilegi per l'accesso al registro. Nel caso di errore durante la validazione dei blocchi viene eseguita una nuova transazione per eliminare quelle precedenti. La tracciabilità dell'operazione annullata così come dell'operazioni che non sono andate a buon fine vengono comunque salvate dal sistema.

- **Consenso**

Il consenso alla validazione di una transazione è condiviso tra i partecipanti scelti della rete. È possibile validare un'operazione soltanto con la maggioranza tra i partecipanti. I partecipanti sono i nodi scelti che fanno parte della rete.

- **Sicurezza**

Le transazioni non possono essere manomesse, perchè esistono nel meccanismo di sicurezza che vanno a controllare e a verificare periodicamente l'integrità del registro mastro con le copie residenti sui nodi scelti.

3.2 Il funzionamento della blockchain

Il processo automatizzato segue le seguenti fasi:

- **Registrazione della transazione**

Ogni volta che avviene una transazione questa viene registrata come un blocco dati. All'interno del blocco si registrano le informazioni più importanti e quest'ultime possono variare.

- **Validazione**

All'interno della rete distribuita della blockchain i nodi scelti devono essere d'accordo sulla validità della transazione registrata.

- **Collegamento blocchi**

Una volta raggiunto il consenso, al blocco viene aggiunto un codice hash crittografico. Il codice hash viene generato a partire da tutti i blocchi precedentemente aggiunti e dal contenuto del blocco in questione, creando attraverso questo processo una catena di blocchi certificata. La modifica interna delle informazioni del blocco corrente provoca un cambiamento del codice hash, in questo modo vengono eliminate ogni forma di manomissione del contenuto del blocco.

- **Rettifica registro mastro**

Completata l'operazione viene aggiornato il registro e le copie sui nodi scelti.

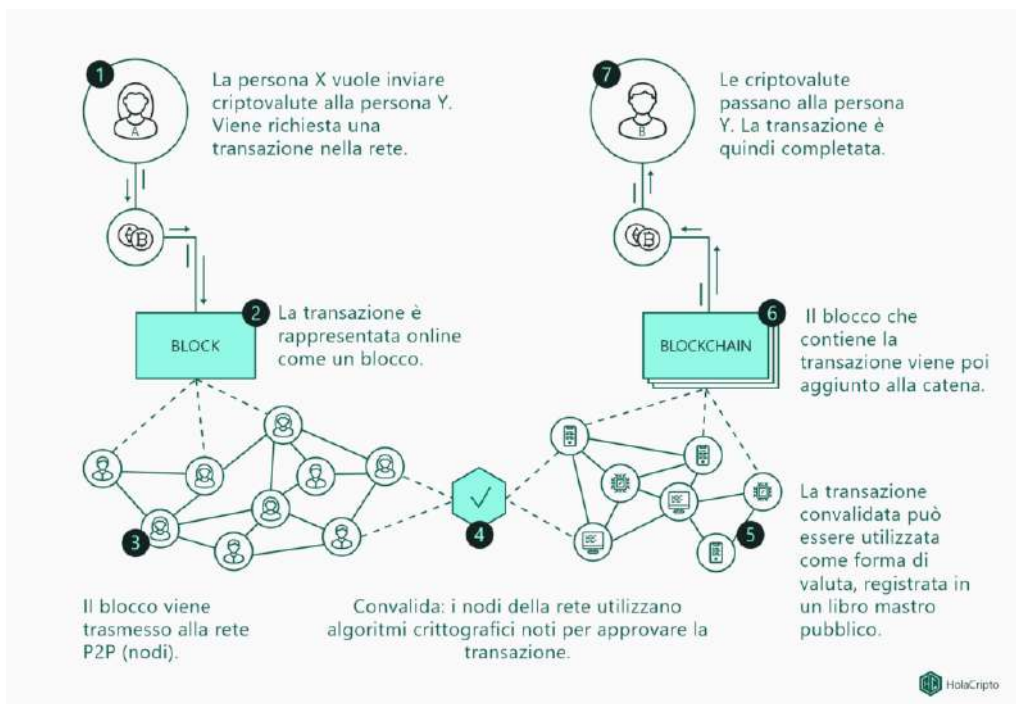


Figura 3.1: Gestione di una transazione all'interno di una blockchain

3.3 Mining

Per mining si intende il processo di verifica, registrazione di una transazione sul libro mastro della blockchain e della creazione di nuove criptovalute all'interno della

rete. Il processo nella sua interezza risulta essere molto oneroso da un punto di vista computazionale, è quindi necessario utilizzare delle macchine con una potenza di calcolo elevata per completare queste operazioni.

Coloro che si occupano di portare a termine il processo di mining sono detti miner e ognuno di essi è un nodo all'interno della rete che si occupa di raccogliere tutte le richieste di esecuzione di una transazione, verificarle e di organizzarle in blocchi. Una volta generato il blocco avviene la fase di validazione che richiede la produzione del codice hash che serve a garantire l'autenticità dei dati presenti all'interno del blocco. Una volta fabbricato il codice quest'ultimo viene controllato da tutti gli altri partecipanti della rete e in caso di esito positivo il blocco contenente le transazioni viene aggiunto sul libro mastro.

Come si è detto precedentemente l'operazione di mining risulta essere molto complessa e onerosa, proprio per questa ragione al completamento del processo viene data una ricompensa al miner che risulta essere proporzionale al numero di task portati a termine.

Visto il profitto che viene ricavato dall'operazione di mining, le strutture hardware di supporto si sono sempre più evolute creando quelle che oggi vengono chiamate mining farm. Le mining farm sono strutture che riproducono la stessa macchina hardware e/o macchine con hardware differenti che sono impiegate nell'esecuzione dell'attività di mining. Questa struttura complessa può essere redditizia ma deve far fronte all'investimento iniziale e alle spese per gli elevati consumi energetici.



Figura 3.2: Computer con hardware dedicato che esegue il processo di mining

Capitolo 4

Analisi preliminari

L'evoluzione finanziaria e tecnologia ha portato alla creazione delle criptovalute. Dalla nascita della prima criptovaluta avvenuta nel 2009 ad oggi si contano circa 10K criptovalute legittimamente quotate.

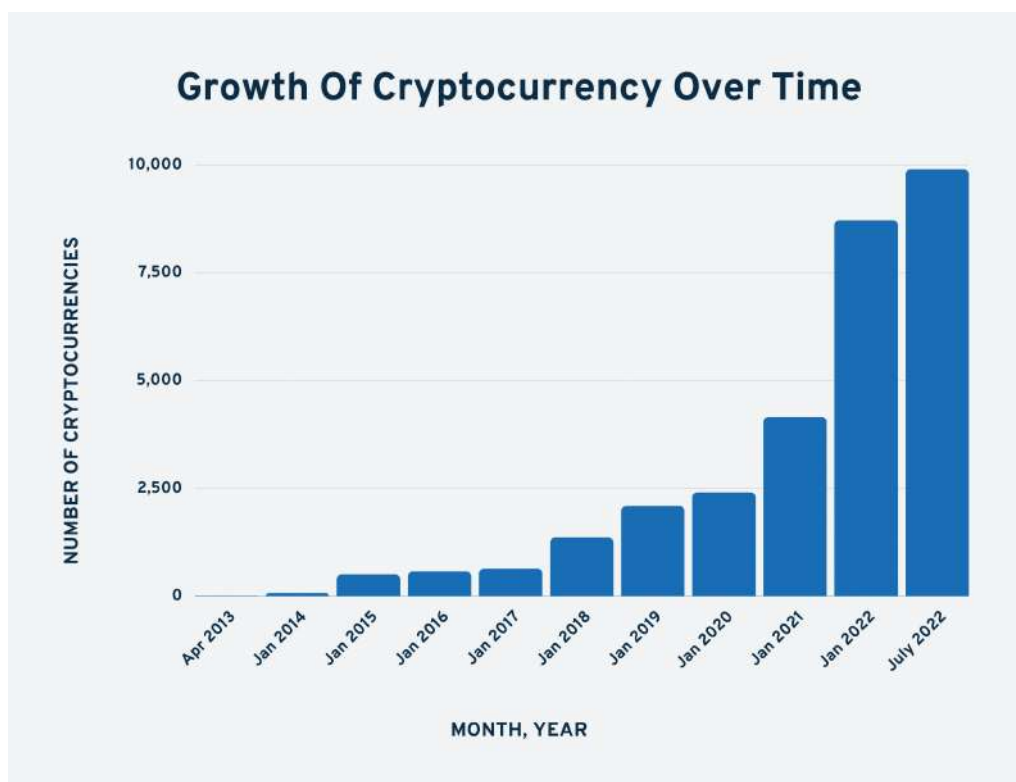


Figura 4.1: Numero totale di criptovalute legittime quotate su CoinMarketCap dal 2013 ad oggi

Questa crescita esponenziale nel tempo è stata incentivata dal positivismo¹ sociale, che ha impresso curiosità nell'animo della gente comune.

¹movimento filosofico e culturale che mira all'esaltazione del progresso scientifico

In questo modo si è diffuso un interesse all'interno del globo che ha condotto a notorietà il mondo crypto.

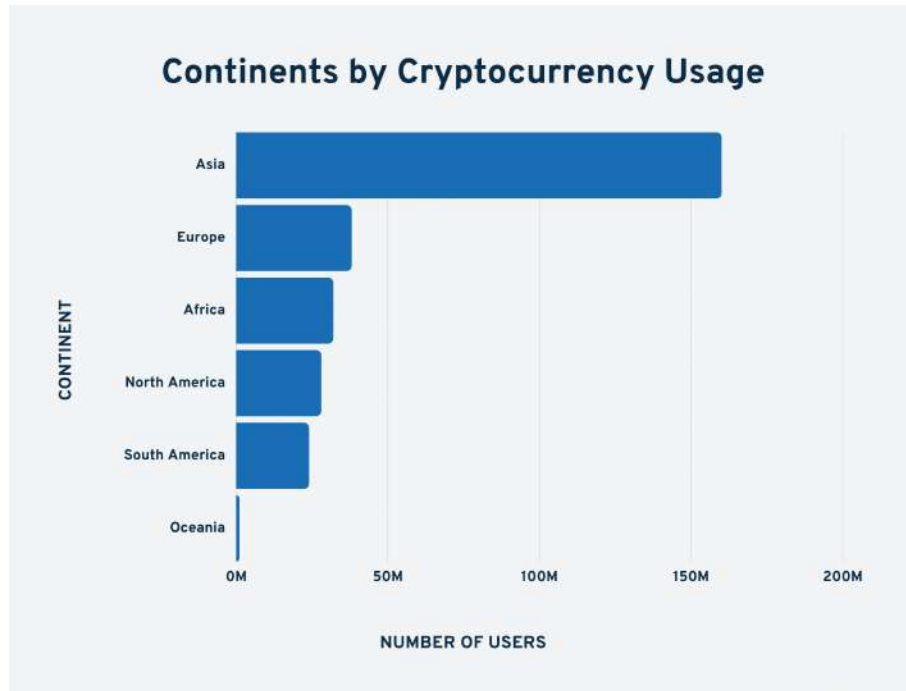


Figura 4.2: Numero di utilizzatori per continente

Oltre alla crescita del numero di criptovalute ed utilizzatori anche la capitalizzazione di mercato complessiva è aumentata durante il tempo.

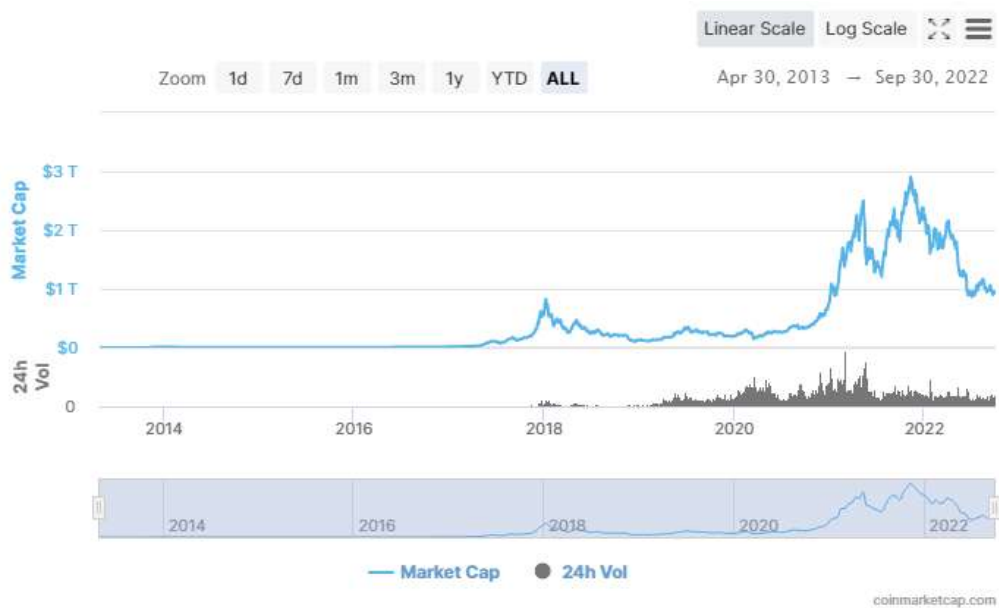


Figura 4.3: Capitalizzazione massima criptovalute

Dall'introduzione delle criptovalute sul mercato fino al 2017 l'innovazione ha conosciuto un periodo di incertezza per poi affermarsi solidamente nel 2018. L'ultimo periodo che inizia nel 2020 e interessa ancora oggi si caratterizza da fluttuazioni molto evidenti e frequenti. Questo comportamento combacia direttamente con l'inizio della pandemia e delle crisi economica odierna.

Si può dunque notare come in un periodo di incertezza economica le criptovalute siano state la base di investimento per molti, perchè il numero di criptovalute su cui si può scommettere è molto ampio e si possono effettuare investimenti variabili di qualsiasi cifra per avere un guadagno.

Tuttavia queste forti oscillazioni così repentine che denotano rialzi e abbassamenti di prezzi in assi temporali così ristretti confermano l'instabilità del mercato crypto e la sua alta volatilità. La principale motivazione vede come protagonista la speculazione stessa. Solitamente in questo scenario, quanto più il prezzo di un asset sale velocemente, tanto più è probabile che ci sia un crollo in futuro. Alcuni di questi crolli sono semplici "correzioni di prezzo" per mantenere in equilibrio il mercato, mentre i restanti sono veri e proprio tracolli.

Recentemente nel "Rapporto sulla Stabilità Finanziaria"[1] del Comitato di Politica Finanziaria della BoE, pubblicato a Luglio 2022, viene trattato il mercato delle criptovalute e le sue peculiarità. I partecipanti sono giunti alla conclusione che l'alta volatilità del prezzo delle criptovalute, che generano squilibri all'interno del mercato, potrebbero causare instabilità finanziaria e vi è la necessità di rafforzare i quadri normativi all'interno del settore.

4.1 Osservazioni conclusive

Il mercato della criptovalute rappresenta una realtà ancora poco stabile ed efficiente, questo è dovuto: alla tecnologia blockchain che è sempre in continua evoluzione e data l'alta complessità della struttura ancora ad oggi può fallire e infine dalla mancanza di un sistema di regolamentazione che gestisca i rischi nel mondo delle criptovalute. Comprendere dunque il funzionamento del mondo crypto e sapere leggere i dati statistici a volte non basta a coltivare investimenti redditizi.

Capitolo 5

Progetto

5.1 Domanda di ricerca

Il mercato delle criptovalute ha una capitalizzazione massima di circa $944 \cdot 10^9$ euro ad oggi e rappresenta una realtà di investimento. Tuttavia questo tipo di mercato è ancora giovane e presenta buchi normativi che provocano l'alta volatilità dei prezzi delle criptovalute. Dunque per generare profitto in un scenario instabile come il medesimo bisogna conoscere l'ambiente, sapersi informare e leggere i dati che si raccolgono. Le piattaforme di trading rappresentano le prime risorse da consultare quando si tratta di reperire informazioni sull'evoluzione delle criptovalute, dove all'interno di queste pagine web sono presenti strumenti finanziari e grafici che riassumono l'andamento delle criptovalute nel tempo. Purtroppo queste piattaforme di trading non sono così intuitive da utilizzare e rendono dunque la navigazione molto pesante. Inoltre le piattaforme di trading offrono la possibilità di consultare diversi grafici per avere un riscontro alternativo, ma alcuni di questi grafici non sono facilmente leggibili e hanno un'interfaccia molto complessa da adoperare.

L'obiettivo del progetto di tesi è quello di sviluppare una dashboard che sia più semplice da utilizzare rispetto alle solite piattaforme di trading. L'intenzione è quella di realizzare un'interfaccia con un design più minimale, in modo tale che si possa reperire le informazioni che si cercano nella maniera più veloce possibile e inoltre aggiungere uno strumento statistico più comprensibile rispetto ai grafici presenti nelle piattaforme di trading, ossia la previsione dell'andamento del prezzo della criptovaluta sfruttando un modello di neural network LSTM.

La funzionalità di previsione è uno strumento che può servire all'utente per condurre ulteriori indagini statistiche e trarre una conclusione più accurata. La finalità dell'introduzione di questa feature da un punto di vista più filosofico è quello di inserire un tool in grado di contrastare l'alta incertezza del mercato crypto. Ovviamente

questo tipo di previsione fornisce un dato qualitativo da prendere con le pinze, visti i svariati fattori di influenza che possono alterare il prezzo delle criptovalute.

5.2 Suddivisione del progetto

Il risultato finale del progetto è una dashboard suddivisa in diverse sezioni, ognuna con la propria funzionalità. Ogni sezione si indentifica in una pagina web il cui scopo è riassumibile all'interno del nome. L'implementazione delle pagine si è condotta in modo tale che vengano visualizzati i dati più importanti mantenendo un design il più minimale e curato possibile, senza caricare inutilmente la GUI¹ e allo stesso tempo accattivante.

Di seguito vengono riportate le pagine web e la loro funzione.

- **index.html**

La pagina di index è la facciata della dashboard. Questa pagina è molto bella da un punto di vista del design e risulta essere molto semplice, ossia non viene abbellita da ulteriori informazioni inutili. Inoltre deve invogliare l'utente ad approfondire il concetto senza dare troppe indicazioni.

- **home.html**

La pagina di homepage raccoglie al suo interno le 100 criptovalute più influenti al momento e le indicizza/inserisce in una tabella dinamica. La dinamicità della pagina è realizzata dall'utente in base ai filtri che vuole applicare sulla tabella e quest'ultima si adatta e restituisce un output diverso ogni qualvolta viene eseguito un filtro.

Ogni riga della tabella contiene una criptovaluta e ne raffigura i parametri caratteristici:

- immagine png che identifica la criptovaluta;
- il nome della criptovaluta;
- il simbolo di indentificazione univoco della criptovaluta, quest'ultimo viene scritto in maiuscolo;
- il prezzo della criptovaluta in euro;
- il volume rappresenta l'attività di acquisto e vendita di criptovalute in termini di liquidità nelle ultime 24h;
- il cambiamento di prezzo percentuale della criptovaluta nell'arco delle 24h;

¹Graphical User Interface

Ogni riga della tabella è cliccabile e si viene reindirizzati alla pagina `check.html` che prende come input la criptovaluta cliccata.

- **statistics.html**

Nella pagina `statics.html` come da nome, si raffigurano le statistiche più importanti sull'evoluzione delle criptovalute. La pagina è suddivisa in 4 riquadri:

- trending, in alto a sinistra, dove si rappresentano le 7 criptovalute che hanno suscitato maggior scalpore in data odierna;
- global info, informazioni sulle prestazioni delle ICOs;
- numero totale di criptovalute attive;
- e infine, nell'ultimo riquadro si mostra la criptovaluta con il prezzo più alto.

- **check.html**

La pagina `check.html` raccoglie al suo interno diverse sezioni di analisi che si possono condurre sulla criptovaluta. Ci sono essenzialmente due modi per giungere in questa sezione della dashboard:

- si può selezionare la criptovaluta di cui si vuole conoscere più dettagli nella pagina `home.html` e tramite la selezione avviene il reindirizzamento in `check.html` precompilando il campo della criptovaluta da cercare in maniera automatica;
- oppure grazie al menù laterale si sceglie direttamente questa sezione e si cerca la criptovaluta di interesse;

Nella parte superiore dopo che si è scelta la criptovaluta si trova una sezione che raffigura al suo interno: indici economici generici, la genesi della criptovaluta e parametri caratteristici della criptovalute come total supply, max supply e circulating supply.

La max supply di una criptovaluta rappresenta il numero totale di coins che esiste sul mercato, si può vederlo come un limite di produzione e una volta raggiunta questa capacità non si produrranno ulteriori coins. Non tutte le criptovalute possiedono un tale limite e nella dicitura max supply si affiancherà il risultato Nan. Il total supply a differenza del max supply conta il numero totale di coin presenti in questo momento senza considerare il numero di coin andati persi (burned coin), sempre facendo riferimento alla criptovaluta che si è selezionata precedentemente. Infine la circulating supply specifica il numero di criptovalute effettivamente minate e scambiate tra investitori.

Scendendo nella visualizzazione della pagina si notano due grafici. Il primo grafico raffigura l'andamento del prezzo in termini di giorni fa, mentre il secondo grafico in base al numero di giorni che si sceglie costruisce la time series del prezzo della criptovaluta selezionata e aggiunge la previsione del prezzo per i successivi 30 giorni. Le previsioni si identificano nel grafico tramite il tratteggio e sono il risultato dell'applicazione di un modello di rete neurale denominato LSTM, che verrà poi approfondito nel capitolo successivo.

E' possibile cambiare la scala, ossia il numero di giorni che i grafici elaborano attraverso l'opzione a inizio pagina. Il numero massimo di giorni che si può selezionare è 1000 e si è scelto come limite questo numero per questioni di efficienza di elaborazione.

A fondo pagina è presente lo schema riassuntivo dei indici statistici dei due grafici calcolati. Per il primo grafico si indicano:

- la media;
- il prezzo più alto registrato;
- il prezzo più basso registrato;
- e il numero di giorni che si è selezionato;

Tutte le statistiche sopra elencate si calcolano sulla base del numero di giorni che l'utente seleziona. Per il secondo grafico si sono inseriti indici statistici di bontà del modello e ulteriori informazioni riassuntive di come si è effettivamente calcolato il modello LSTM. Nel capitolo dedicato alla previsione verranno ampiamente trattati questi argomenti.

- **compare.html**

In questa pagina è possibile confrontare contemporaneamente 3 criptovalute. Il modello di confronto prevede: una prima parte dove vengono mostrati gli indici economici delle monete mettendo in parallelo i risultati e una seconda parte costituita da un grafico riassuntivo con tanto di previsione.

- **glossary.html**

Questa pagina svolge la funzione di glossario. Tutte le risposte alle domande e i dubbi che si possono sollevare nella lettura del sito sono presenti all'interno di questa pagina.

In aggiunta si definiscono concetti teorici e metriche di lettura a dati e indici statistici che si sono impiegati nella descrizione delle criptovalute.

5.3 Framework utilizzati

Lo sviluppo della dashboard vede raccogliere al suo interno diversi tipi di framework. Per framework si intendono un insieme di librerie o per meglio dire di classi astratte che cooperando tra di loro realizzano un'architettura logica di supporto sulla quale un software può essere progettato e realizzato. Per i programmatori questi framework agevola lo sviluppo e permette di risparmiare tempo offrendo degli strumenti di programmazione già pronti da utilizzare.

Di seguito si analizzeranno in dettaglio i framework che si sono impegnati per la creazione della dashboard.

5.3.1 Javascript: Axios e jQuery

JavaScript è un linguaggio di programmazione a eventi utilizzato nella programmazione web lato client e esteso successivamente lato server per lo sviluppo di applicazioni web dinamiche.

Javascript viene utilizzato nell'ambito della programmazione come linguaggio di scripting. Per linguaggio di scripting si intendono tutti quei linguaggi che presentano le seguenti caratteristiche:

- sono linguaggi interpretati. L'interpretazione avviene tramite un software che prende il nome di interprete e si occupa di prendere il codice sorgente così come è stato scritto, si dice ad alto livello ed eseguirlo istruzione per istruzione senza doverlo compilare. La compilazione invece traduce il codice sorgente ad alto livello in un'altro linguaggio a basso livello, chiamato anche linguaggio macchina perchè direttamente comprensibile dalla macchina stessa. Il risultato della compilazione genera un file direttamente eseguibile dal processore;
- si tratta di un linguaggio destinato ad automatizzare procedure;
- tali linguaggi possono essere integrati all'interno di un altro codice;

Javascript può essere eseguito sia lato client sia lato server. Lato client, quando viene visitata una pagina web contenente codice javascript, il codice viene interpretato direttamente all'interno di un qualunque browser, come: firefox, chrome, edge. L'interpretazione diretta del browser è possibile perchè integra direttamente un interprete javascript, chiamato "JavaScript Virtual Machine".

Lato server invece viene creata una infrastruttura javascript in grado di rispondere alle chiamate HTTP effettuate dai client.

Javascript è un linguaggio orientato agli oggetti e per oggetti si intendono strutture logiche che realizzano una determinata funzione ed è possibile utilizzarli nella programmazione. Esistono due tipologie di oggetti utilizzabili in javascript:

- oggetti predefiniti, impiegabili sempre durante la programmazione, come: array, Boolean, Date, ..
- e infinite altri oggetti chiamati come "ospiti", definiti non dal linguaggio ma dall'ambiente di esecuzione. Javascript può impiegare gli oggetti presenti nella pagina web visualizzata dal browser. Per poter utilizzare questi oggetti è necessario accedere al DOM della pagina web.

DOM (Document Object Model) è la rappresentazione di un documento in formato a oggetti, questa rappresentazione è stata standardizzata a livello globale e viene adoperata per modificare o accedere al oggetto di interesse.

JavaScript come si è detto precedentemente è un linguaggio di programmazione ad eventi, cioè il codice scritto non viene eseguito in maniera sequenziale bensì si ramifica in procedure le quali l'esecuzione è determinata dal verificarsi di un determinato evento specifico esterno. Ogni oggetto del DOM che può esercitare un'azione presenta un insieme di eventi predefiniti che sono richiamabili all'interno delle procedure, per esempio:

un bottone in una pagina web viene visto tramite DOM come un oggetto di tipo Button, quest'ultimo preseta i seguenti eventi:

- onclick = "*faiQualcosa()*": quando clicco richiamo la procedura/funzione *faiqualcosa()*;
- onmouseover = "*faiQualcosa()*": quando sono sopra il bottone con il mouse viene richiamata la funzione *faiqualcosa()*;

Nel progetto oltre ad utilizzare javascript si sono impiegate le seguenti librerie:

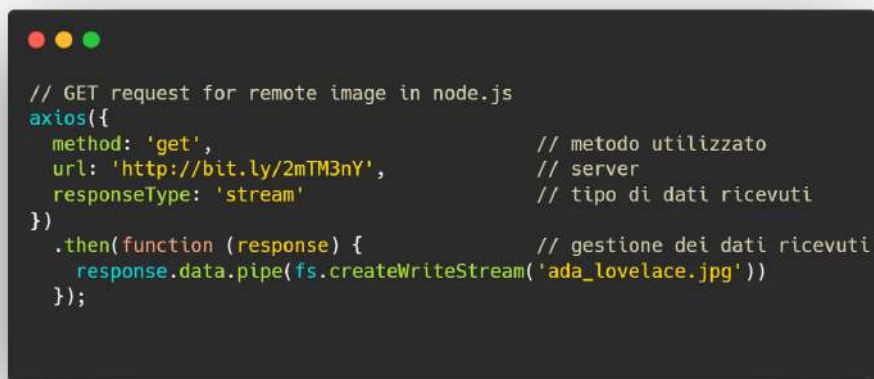
- **Axios.js**

Axios è una libreria javascript che permette di effettuare richieste HTTP ed è in grado di gestire la risposta a questa richiesta.

HTTP è un protocollo che definisce dei standard di comunicazione per instaurare una chiamata all'interno della rete tra client e server. Con il protocollo HTTP, una volta stabilita la comunicazione con il server (web server), è possibile effettuare delle operazioni sulle risorse residenti sul server o inviare dati attraverso 4 metodi:

- GET richiede al server web una specifica informazione o risorsa mediante l'URL. L'URL (Uniform Resource Locator) identifica univocamente l'indirizzo di una risorsa sulla rete attraverso un insieme di caratteri, solitamente accessibili dai client;
- POST si utilizza per inviare grandi quantità di dati a un server, questo perchè il metodo POST, a differenza del GET che allega i dati da inviare nel URL, inserisce il tutto all'interno dell'header HTTP;
- PUT ha una funzionalità simile a POST, tuttavia si occupa di apportare modifica a risorse già esistenti;
- HEAD si ricevono solo le informazioni di intestazione della pagina che si vuole;

Axios gestisce la comunicazione HTTP utilizzando codice javascript. Esempio:

A screenshot of a code editor with a dark background and light-colored text. The code is a JavaScript snippet using the Axios library to make a GET request. It includes comments in Italian explaining each part of the configuration. The code is as follows:

```
// GET request for remote image in node.js
axios({
  method: 'get',           // metodo utilizzato
  url: 'http://bit.ly/2mTM3nY', // server
  responseType: 'stream'   // tipo di dati ricevuti
})
.then(function (response) { // gestione dei dati ricevuti
  response.data.pipe(fs.createWriteStream('ada_lovelace.jpg'))
});
```

Figura 5.1: Esempio richiesta POST di una risorsa

- **jQuery.js**

jQuery è una libreria di funzioni javascript che si propone come obiettivo quello di semplificare la programmazione lato client delle pagine HTML. Con jQuery si può svolgere tutte le funzionalità che javascript è in grado di effettuare, ma in maniera più efficiente e rapida.

5.3.2 Html e css: bootstrap

HTML è un linguaggio di markup nato inizialmente per la formattazione e impaginazione di documenti ipertestuali, per poi successivamente impierlo nella creazione di pagine web. Un linguaggio di markup definisce un insieme di regole strutturali e semantiche per la rappresentazione di codice. Il codice si scrive facendo uso dei marcatori, ossia blocchi elementari che realizzano una funzione logica ben specifica e combinando questi blocchi si possono costruire sistemi più complessi.


CSS (Cascading Style Sheets) è un linguaggio usato per definire la formattazione dei documenti HTML. CSS è in grado di specificare:

- lo stile dei blocchi, ossia la loro apparenza grafica e attraverso complesse istruzioni si possono aggiungere effetti grafici per creare dinamicità all'interno della pagina web;
- la posizione all'interno della pagina web dei blocchi;
- e infine la dimensione;

Il CSS si può scrivere in un file separato e successivamente aggiungere il riferimento nella pagina HTML che si vuole modificare, oppure direttamente all'interno della pagina HTML stessa inglobando il codice all'interno del tag `<style></style>`. La combinazione di questi due linguaggi creano quelle che vengono definite pagine statiche, html costruisce lo scheletro della pagina web e CSS ne regola il design. Le pagine statiche una volta determinato il loro contenuto mediante codice presenteranno sempre lo stesso aspetto e le stesse informazioni a meno che non vengano effettuate modifiche.

- **Bootstrap**

Bootstrap è una libreria, framework, open source per la programmazione front-end di pagine web responsive. Una libreria open source è accessibile a tutti e completamente gratuita, scaricando le dipendenze dalla pagina proprietaria oppure inglobando i riferimenti si possono utilizzare le strutture implementate al loro interno. Una libreria front-end definisce una serie di strumenti logici già pronti per la realizzazione di una pagina statica.



```

<!DOCTYPE html>
<html lang="en-us">

<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1">
  <title>Page Title</title>

  <!-- External css file -->
  <link rel="stylesheet" href="styles.css">

  <!-- Internal css -->
  <style>
    h1 {
      color: blue;
      font-family: verdana;
      font-size: 300%;
    }

    p {
      color: red;
      font-family: courier;
      font-size: 160%;
    }
  </style>
</head>

<body>
  <!-- Page structure -->
  <h1> My First Heading </h1>
  <p> My first paragraph </p>
</body>

</html>

```

Figura 5.2: Pagina HTML con inclusione di CSS

5.3.3 Python: flask e librerie di data science

Python è un linguaggio di programmazione ad alto livello, interpretato, object oriented, che si utilizza nello sviluppo di applicazioni: web, software e nella data science. Python data la sintassi semplice e pulita, molto simile al linguaggio naturale, possiede un'alta curva di apprendimento, ossia le quantità di informazioni corrette sul funzionamento del linguaggio tendendo a salire molto velocemente con il passare del tempo, inoltre risulta essere uno dei linguaggio più utilizzati al giorno d'oggi e data la sua semplicità gode di un'abbondante quantità di librerie e framework per qualsiasi ambito.

- **Flask**

Flask è un microframework che possiede una struttura in grado di realizzare una web application. Un microframework è una libreria, una struttura logica, che solitamente non possiede alcuna dipendenza con altre librerie esterne e possiede un nucleo interno molto semplice ma estensibile. Flask più in specifico svolge due principali funzioni:

- url routing, ossia interpreta l'URL esterno di una richiesta e lo trasforma in un URI interno per capire quale modulo/azione chiamare;
- template engine, cioè è in grado di creare una pagina di template base sulla quale poi costruire nuove funzionalità aggiungive senza ogni volta dover partire dal principio nello sviluppo di una pagina web;

- **Librerie di data science**

Come si è detto precedentemente esistono librerie di python per qualsiasi cosa, specialmente nel campo della data science. Brevemente nel progetto si sono impiegate:

- **matplotlib**: libreria per tracciare dati grafici di alta qualità bi e tridimensionali (2D e 3D).
- **pandas**: libreria che fornisce strutture di dati ottimizzate per manipolare i dati in maniera efficiente;
- **tensorflow**: libreria che consente di creare modelli di machine learning per applicazioni;
- **plotly**: libreria per tracciare grafici interattivi;
- **sklearn**: libreria che contiene algoritmi di machine learning;

5.4 Funzionamento generale

Le informazioni effettive che si utilizzano all'interno della dashboard si ottengono mediante richieste API a una web application di trading online che prende il nome di CoinGecko.

Le API (Application Programming Interface) consistono in un insieme di procedure e strutture dati che svolgono un specifico compito e sono richiamabili esternamente dal programmatore. Il programmatore non vede effettivamente come tali procedure

siano state implementate a livello di codice, bensì sono visibili le interfacce di richiamo e il tipo di dato resituito dalla procedura chiamata, proprio per questa ragione le API vengono anche definite programmazione "black box".

La dashboard tratta i dati richiesti tramite API e svolge delle operazioni interne per rendere i dati utili al caso che si studia. Di seguito si illustrano i possibili scenari di come effettivamente vengono richiamati e trattati i dati necessari all'interno della dashboard:

(i) **Richiesta dati**

La semplice richiesta dati avviene tramite una chiamata $HTTP \rightarrow GET$ alle API di CoinGecko, impiegando Axios.js. La risposta di CoinGecko avviene restituendo, in caso l'esito sia positivo, un file JSON contenente i dati richiesti. Il formato JSON è un formato per lo scambio di dati, di facile lettura, direttamente interpretabile da javascript. Il file JSON è organizzato in un insieme di oggetti, indentificati all'interno delle parentesi graffe, strutturati da più campi, che sono accessibili via codice. Una volta ottenuti i dati, quest'ultimi vengono formattati e inseriti all'interno della dashboard.

(ii) **Creazione del modello di neural network**

In questo caso con Axios.js si invia una richiesta $HTTP \rightarrow GET$ ad un URL interno, al quale flask risponde inviando il risultato dell'esecuzione della funzione di creazione della neural network. La funzione neural network reperisce i dati richiamando tramite codice python le API di CoinGecko e utilizza i dati per la creazione del modello. Una volta creato il modello, si utilizza il modello stesso per fare previsione sul costo della criptovaluta di riferimento per i prossimi 30 giorni a partire da oggi stesso.

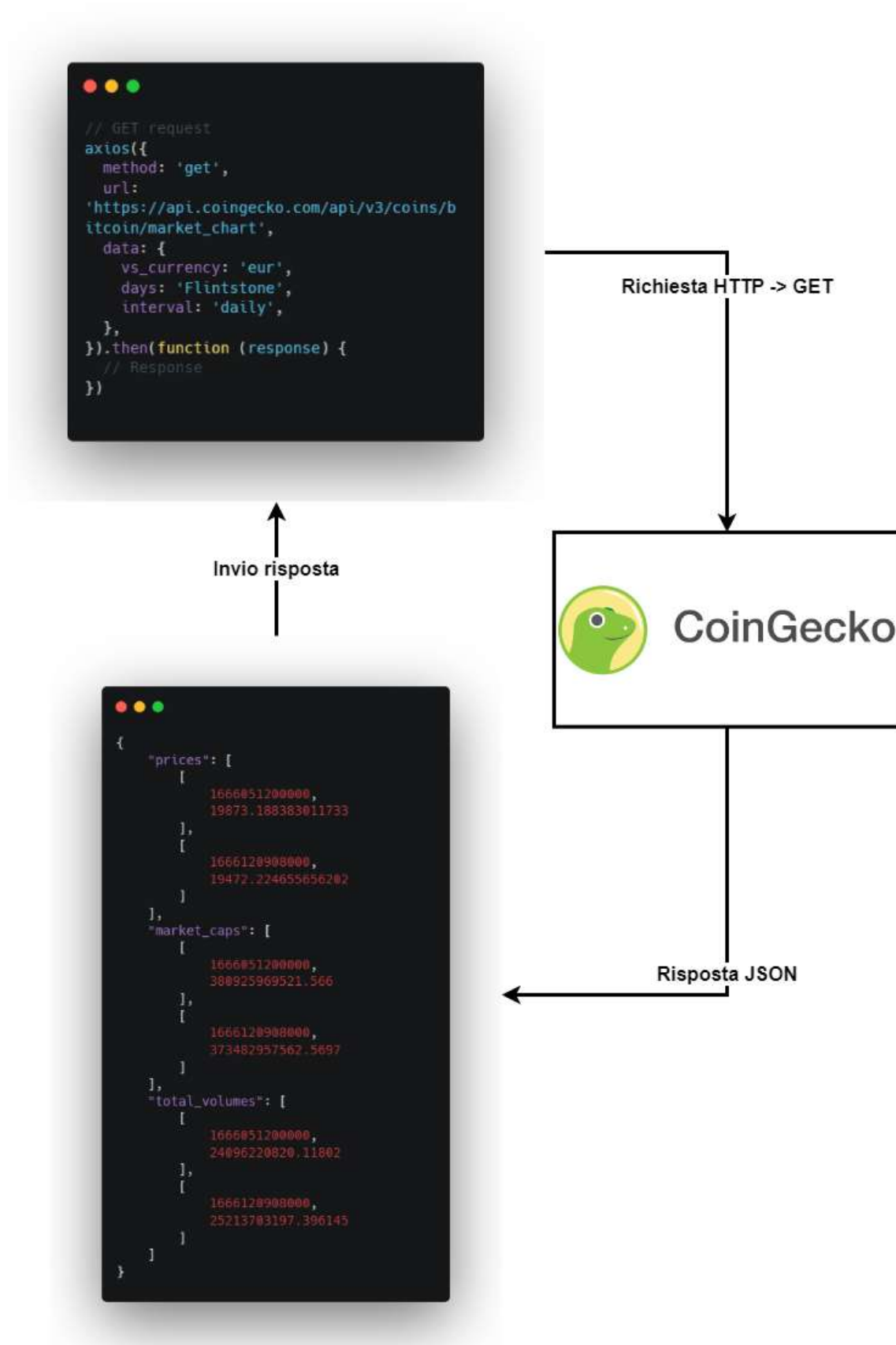


Figura 5.3: Semplice richiesta dati

5.5 Ottimizzazioni applicate

Le ottimizzazioni applicate da un punto di vista strutturale per velocizzare l'esecuzione operativa sono:

- al caricamento della criptovaluta che si seleziona tramite ricerca, si carica la struttura dati relativa alla time series del prezzo di quest'ultima andando a prendere il limite massimo di giorni possibile, ossia 1000. Indipendentemente dal range di giorni fa che l'utente sceglie, in memoria si ha la struttura completa, quindi il processamento risulta essere immediato e più efficace. Questo modo di operare evita ogni volta di effettuare una richiesta API ogni qualvolta vi è la necessità di inserire un range di giorni diverso.
- il processo per la creazione del modello di neural network e la relativa previsione richiede un tempo non indifferente di processamento. In particolar modo nella sezione *compare.html* dove si confrontano 3 criptovalute vi è la necessità di eseguire l'algoritmo 3 volte di fila. Per diminuire il tempo di attesa e rendere il tutto più efficiente si sono utilizzati i thread.

Un processo di esecuzione molto complesso che realizza diverse diverse funzionalità si suddivide in sottoprocessi chiamati thread. Ogni thread svolge un compito ben specifico e lavora parallelamente rispetto agli altri thread dello stesso processo. L'esecuzione concorrente realizzata da una applicazione multithread è possibile se il processore della macchina su cui si sta eseguendo il codice è multiprocessore/multithreading, ovvero a livello hardware possiede la struttura fisica per poter gestire tali operazioni. Per risolvere il problema di attesa si sono impiegati 3 thread che lavorano parallelamente per la risoluzione del modello di neural network.

5.6 Importanza dei strumenti di previsione

All'interno di un qualsiasi tipo di mercato domanda e offerta mutano durante il tempo a causa di eventi, bisogni o novità che possono cambiare il trend di mercato per instaurarne uno nuovo.

Un modo per comprendere lo stato di un mercato è attraverso lo studio dei dati che da esso ne deriva. I dati che si raccolgono possono essere di varia natura e forma così come possono descrivere diversi tipi di contesti, per questo è molto importante sapere cosa si vuole analizzare e che tipo di informazione si vuole ricavare.

Per poter estrapolare informazioni dai dati è necessario condurre una serie di studi matematici per evincere un modello statistico in grado di spiegare il comportamento

dei dati. Un buon modello oltre a spiegare i dati effettivamente osservabili si può applicare per fare previsione e ricavare quello che con grande probabilità succederà in futuro. La previsione gioca un ruolo fondamentale in questi tipi di scenari economici, perchè permette di avere un riscontro sul lungo o breve termine di una o più variabili finanziarie di cui si vuole conoscere l'evoluzione e prendere dunque dei provvedimenti in merito.

Come si è discusso precedentemente le criptovalute fanno parte di un mercato ancora troppo volatile e la sola conoscenza della struttura economica su cui si ergono le criptovalute non basta a realizzare investimenti remunerativi. Il progetto di tesi si pone dunque come obiettivo quello di creare un modello statistico in grado di spiegare i dati del prezzo delle criptovalute e fare delle previsioni su quello che sarà il loro prezzo per desumere critiche e analisi più accurate in merito.

Capitolo 6

Modelli di previsione

L'era della digitalizzazione ha portato alla creazione di nuove tecnologie digitali capaci di elaborare dati e trasmettere informazioni. Con il passare del tempo le tecnologie hanno trovato applicazioni in qualsiasi tipo di ambito e si sono evolute diventando sempre più complesse ed efficienti.

Di pari passo a questa rivoluzione informatica che ha quindi introdotto nuovi sistemi informatici, la quantità effettiva di dati in circolazione è aumentata considerevolmente, basta pensare che la quantità di dati emessa quotidianamente al giorno d'oggi ammonta a circa $2.5 \cdot 10^{18}$ bytes.

I dati che vengono emessi rappresentano un bene prezioso per la società perchè riflettono la realtà in cui viviamo. L'analisi e il trattamento dei dati denudano la complessità del mondo reale e offre la possibilità di migliorare il futuro, perchè ci consente di individuare e correggere le incertezze scovate durante le indagini statistiche effettuate. Lo studio specifico e la metodologia di interpretazioni delle informazioni avviene attraverso metodi di statistical learning, ovvero attraverso la creazione di modelli statico/matematici che descrivono il fenomeno studiato e sono in grado di spiegare il presente e di prevedere il futuro.

Per ogni caso di studio esistono diverse modalità di approccio ai dati e di conseguenza l'utilizzo di un modello piuttosto che un'altro. All'interno del progetto di tesi si è scelto di utilizzare un modello di deep learning denominato LTSM, un'evoluzione delle RNN, ossia una rete neurale artificiale capace di prendere in input un'intera sequenza di dati, come la time series di una criptovaluta per fare previsione. Di seguito in questo capitolo si darà la definizione del modello impiegato cercando di ricoprire tutti i concetti teorici per comprendere il meccanismo che sta alla base del modello.

6.1 Rete neurale

Una rete neurale è un'insieme di algoritmi che cerca di apprendere la dipendenza dei dati osservati imitando la struttura del cervello umano. Le reti neurali sono composte da un insieme di nodi collegati tra loro formando così una rete e tale rete è costituita da diversi livelli:

- un livello di input, il punto di ingresso dei dati;
- uno o più livelli nascosti, dove l'informazione di input viene elaborata;
- e un livello di output, dove il nodo decide come procedere in base al dato elaborato;

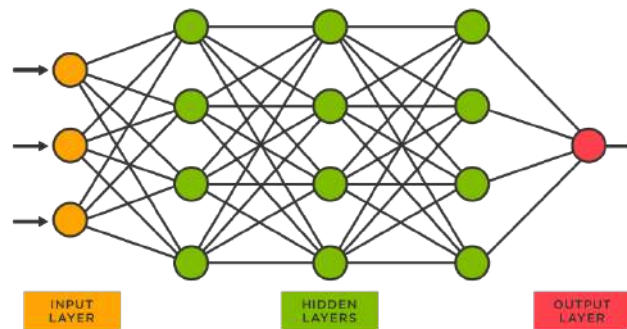


Figura 6.1: Rete neurale

Ogni nodo, chiamato anche neurone, ha un peso e una soglia associati. Se l'output generato da un nodo è al di sopra del valore di soglia specificato, tale nodo si attiva e invia i dati al successivo a cui è collegato. Nel caso invece il valore di output prodotto dal nodo è inferiore alla soglia il dato non passa al nodo successivo.

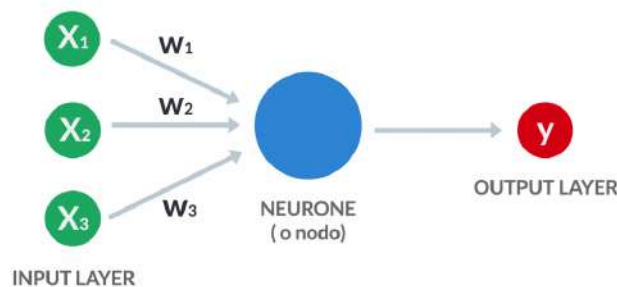


Figura 6.2: Struttura di un neurone

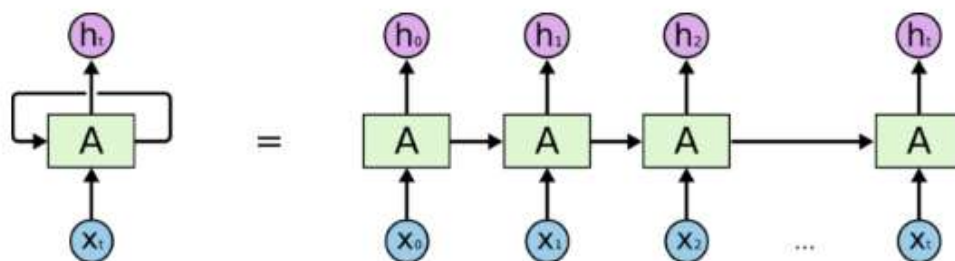
x_1, x_2, x_3 dati di input
 w_1, w_2, w_3 pesi associati ai dati di input
 y è l'output del neurone

Le reti neurali in base ai dati che vengono dati in input al sistema allenano il modello imparando la dipendenza tra i dati e la durata del processo di learning varia molto in base alla complessità della rete neurale che si crea. Durante la fase di apprendimento i pesi associati ai neuroni possono subire cambiamenti e regolarsi di conseguenza al fenomeno che si sta studiando.

6.2 Le RNN

Una recurrent neural network (RNN) è una classe di reti neurali artificiali in cui le connessioni tra nodi possono creare un ciclo, consentendo all'output di alcuni nodi di influenzare l'input dei successivi.

La propagazione di informazioni tra nodi successivi genera una dipendenza a lungo termine, ossia l'informazione iniziale va a condizionare i risultati di tutti i nodi successivi rimanendo in memoria. Tuttavia può essere difficile addestrare RNN per risolvere problemi che richiedono l'apprendimento delle dipendenze temporali a lungo termine, questo succede perchè le informazioni processate tendono a svanire dalla memoria esponenzialmente nel tempo e questo fenomeno prende il nome di vanishing gradient problem.



An unrolled recurrent neural network.

Figura 6.3: RNN

6.3 LSTM

Le LSTM sono un tipo speciale di RNN in grado di apprendere dipendenze a lungo termine utilizzando una struttura interna molto più complessa per eliminare il vanishing gradient problem. Il processo di trasferimento dei dati è lo stesso delle RNN, tuttavia l'operazione di propagazione delle informazioni è diversa, perchè le informazioni che vengono processate subiscono un filtro. Il filtro ha il compito di selezionare le informazioni rilevanti conservandole e dimenticando quelle irrilevanti al processo di learning dei dati, ossia della comprensione del dataset da parte del modello.

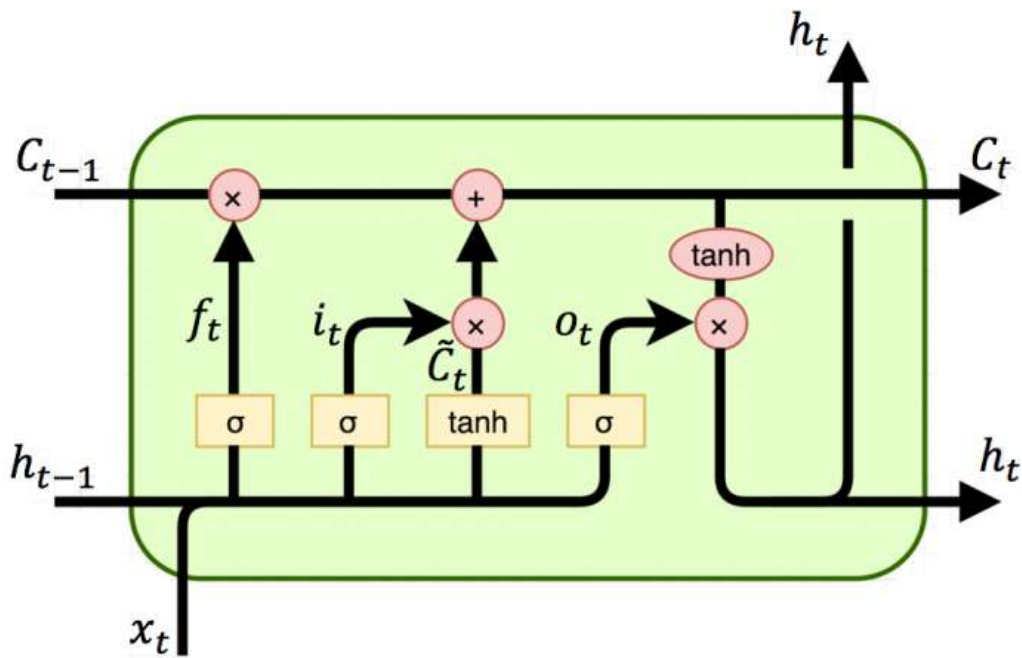


Figura 6.4: Struttura interna LSTM

h_t, C_t : layer nascosti
 x_t : dati di input
 o_t : vettore di output
 b_f, b_i, b_c, b_o : vettori di bias
 W_f, W_i, W_c, W_o : parametri della matrice
 σ : funzione di attivazione Sigmoid
 $tanh$: funzione di attivazione tanh

6.4 Funzioni di attivazione

La funzione di attivazione è una funzione matematica che dato un determinato input di dati produce in output il risultato. Il risultato che deriva dall'applicazione della specifica funzione di attivazione genera l'attivazione o meno del prossimo neurone della neural network.

Si può dunque considerare la funzione di attivazione un filtro applicato ai dati di ingresso di un nodo e la specificità dell'azione che il filtro applica sui dati dipende dalla particolare funzione scelta.

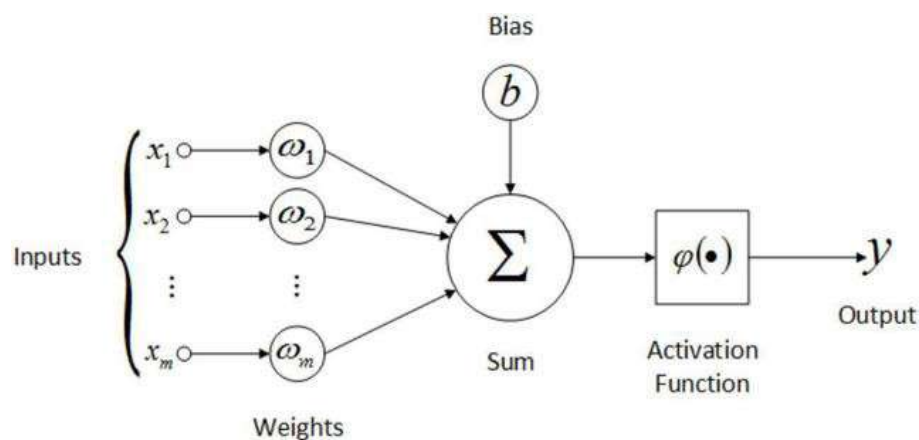


Figura 6.5: Struttura interna

Esistono diverse funzioni di attivazione:

- Sigmoid or Logistic Activation Function;
- Tanh or hyperbolic tangent Activation Function;
- ReLU (Rectified Linear Unit) Activation Function;
- Leaky ReLU;

Di seguito si analizzeranno le prime 3 funzioni di attivazione.

6.4.1 Sigmoid

I gates interni contengono una funzione di attivazione denominata sigmoid o anche detta logistic function. La sigmoid function è una funzione matematica utilizzata nel campo delle reti neurali e serve a convertire un numero $x \in \mathcal{R}$ in una probabilità, in questo modo si riesce a interpretare in maniera più semplice i dati.

$$\sigma = \frac{1}{1 + e^{-x}}$$

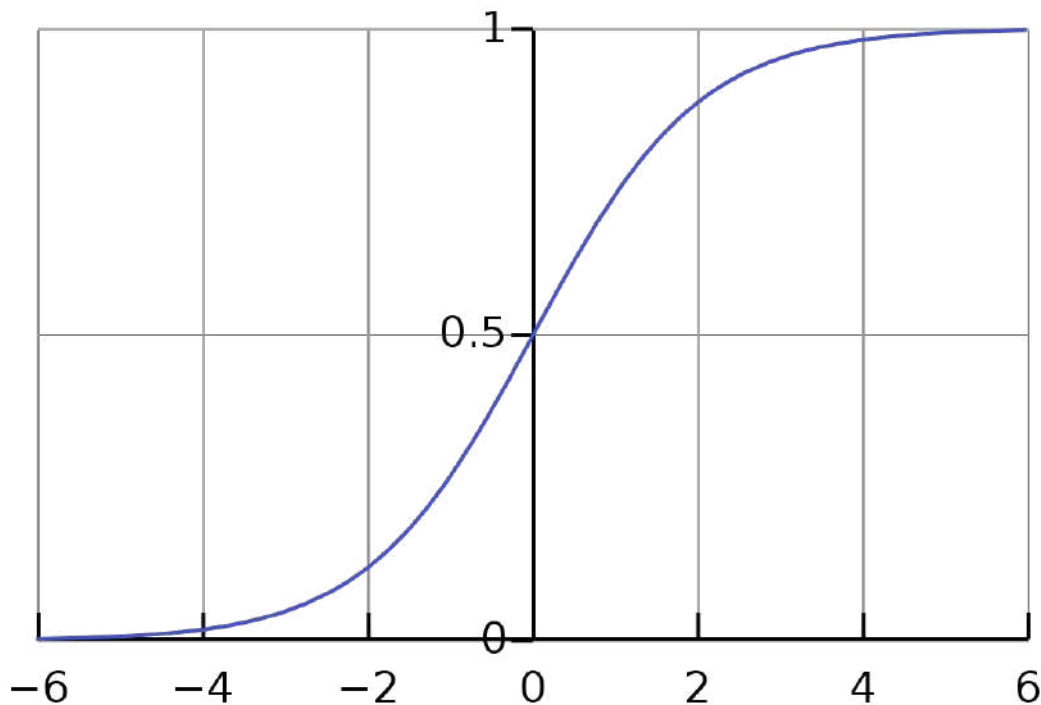


Figura 6.6: funzione σ

Ogni gate contiene una funzione sigmoid, la quale prende come input un dato e lo filtra. Il filtro serve a far dimenticare o a mantenere le informazioni in memoria attraverso l'eplicitazione di un valore soglia. Dato che la funzione assume valori x compresi tra $0 \leq x \leq 1$, se l'informazione ricevuta come input è particolarmente importante per la compresione del fenomeno osservato tale informazione avrà una probabilità superiore alla soglia e gli verrà attribuita la probabilità 1, in caso contrario se l'informazione non è preziosa la probabilità finale sarà 0 e dunque sarà scartata.

6.4.2 tahm

La funzione tahm è una funzione di attivazione così come la funzione sigmoid. La funzione risulta essere limitata tra -1 e 1 , in questo modo permette una migliore mappatura della precedente.

$$tahm = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

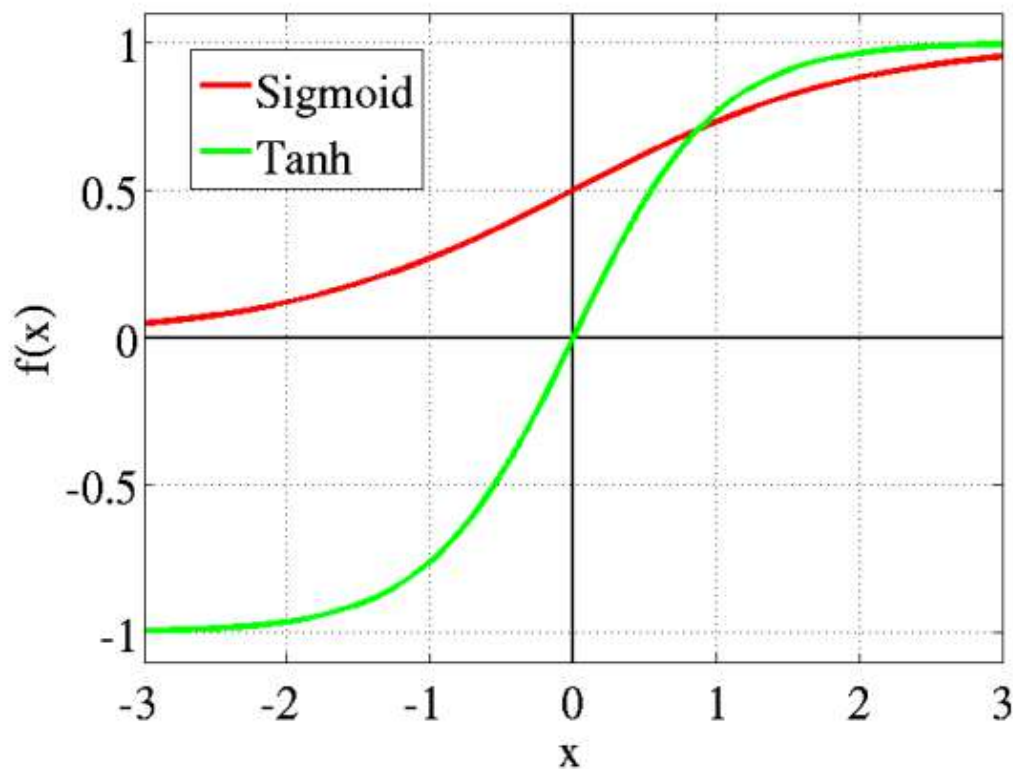


Figura 6.7: funzione tahm

Dato un confronto più approfondito tra la funzione tahm e la sigmoid emerge che applicando la funzione tahm la neural network apprende molte più informazioni sul dataset di ingresso, ma indipendentemente da questo le due funzioni soffrono di vanishing gradient e dunque durante l'esecuzione dell'algoritmo i pesi dei neuroni tendono a non aggiornarsi correttamente rendendo in questo modo il processo di learning molto lento fino ad arrestarsi completamente nei peggiori dei casi.

6.4.3 ReLU

La funzione di attivazione in questione restituisce come output il valore massimo tra 0 e il dato di input x .

$$f(x) = \max(0, x)$$

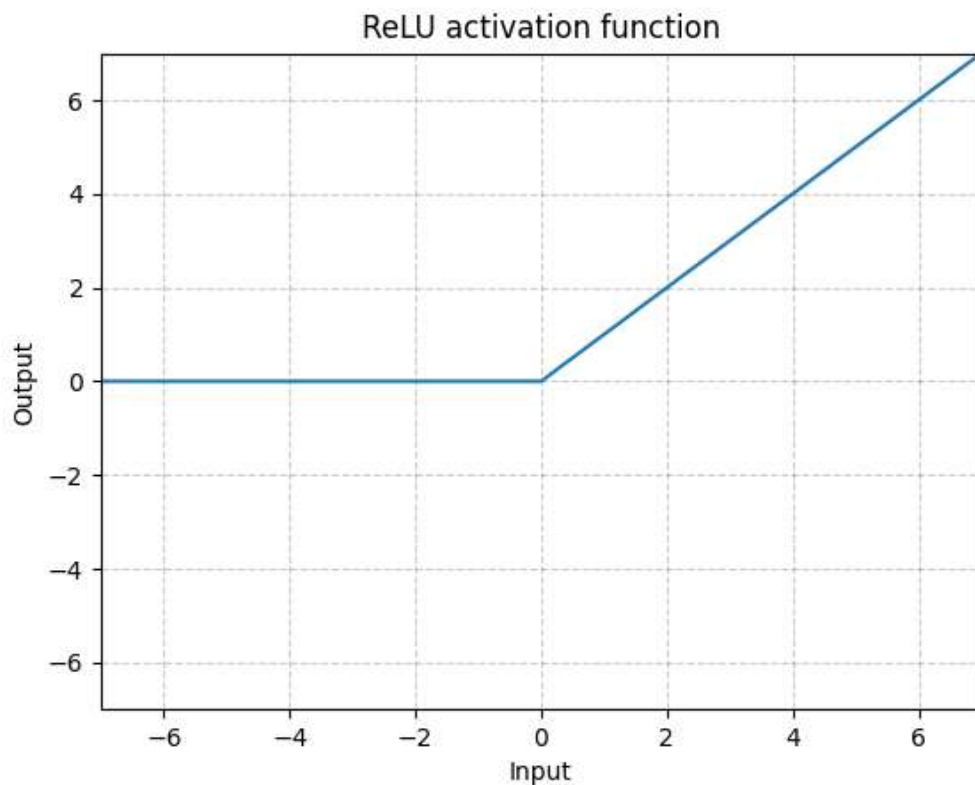


Figura 6.8: funzione ReLU

A differenza delle funzioni di attivazioni precedenti la Relu function non subisce il fenomeno del vanishing problem e consente di preservare le proprietà del fenomeno osservato semplificando l'ottimizzazione dell'algoritmo di neural network. Inoltre la ReLu function risulta essere molto sensibile nell'attribuire i pesi ai vari neuroni durante l'apprendimento, evitando dunque di saturarli.

Putroppo l'utilizzo della ReLU genera:

- **Exploding Gradient** è il problema inverso del Vanishing Gradient, ossia i pesi dei neuroni tendono a diventare molto importanti crescendo esponenzialmente nel tempo e causando instabilità nel processo di apprendimento;

- **Dying ReLU** questo si verifica quando i neuroni si bloccano nel negativo dell'output producendo costantemente in questo modo 0, senza più riprendersi.

6.5 Struttura interna

Le LSTM sono suddivisibili internamente in 3 cancelli ognuno dei quali svolge un ruolo ben specifico:

- **Input gate**
I dati di input insieme ai dati elaborati precedentemente vengono passati alla funzione di sigmoid e alla funzione tanh.
- **Forget gate**
Le informazioni che passano attraverso questo gate subiscono un processo di filtraggio il quale si sceglie quali informazioni elaborare e quali dimenticare.
- **Output gate**
Le informazioni prodotte che sono passate attraverso le funzioni di attivazione e che sono state quindi elaborate vengono indirizzate al prossimo nodo della rete neurale.

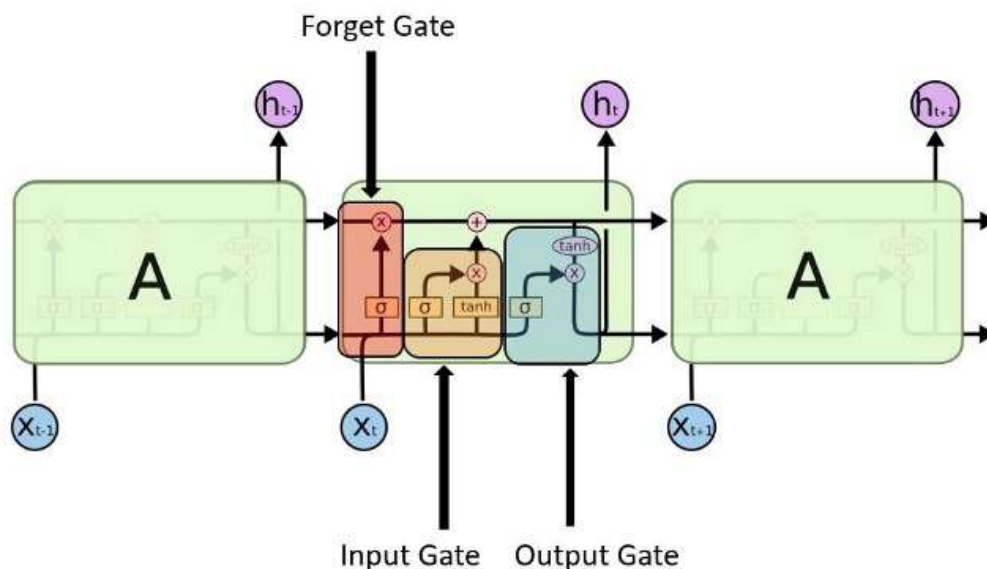


Figura 6.9: Suddivisione LSTM in gates

6.6 Il modello

Come si spiega nella sezione 5.4 la creazione e il processo di learning del modello di neural network LSTM viene effettuata attraverso una richiesta get con axios per l'accesso alla risorsa `/forecast.html` o nel caso di confronto di più criptovalute `/neural.html`.

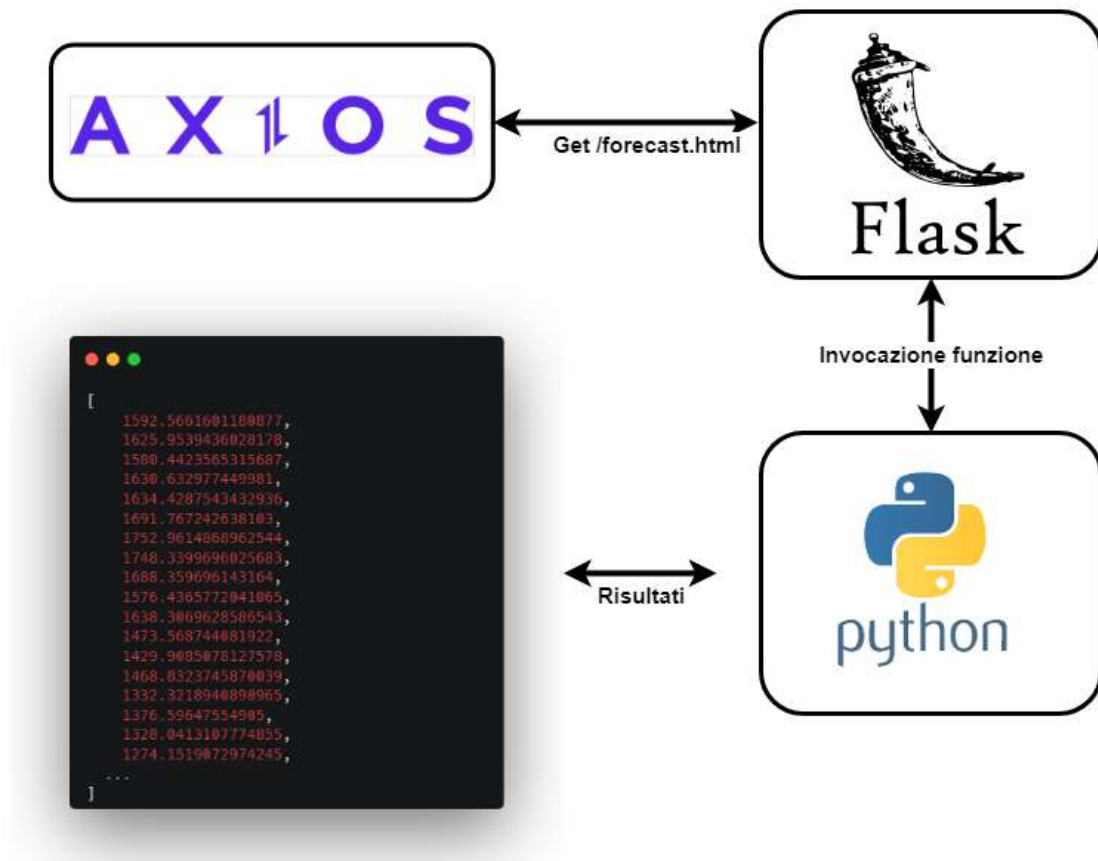


Figura 6.10: Esecuzione modello

Flask associata alla risorsa `/forecast.html` o `/neural.html` una funzione python da eseguire che alla fine del processo restituisce come risultato finale la previsione del prezzo della criptovaluta per i prossimi 30 giorni. Infine il risultato viene ridato ad axios che si occupa di generare il grafico con i relativi dati ottenuti.

Il modello di neural network che viene allenato impiega come dataset di input la time series della criptovaluta selezionata e la lunghezza di tale dataset dipende da quanti giorni si sceglie di passare al modello. La selezione di quanti gioni passare alla funzione viene indicata dall'utente completando il campo `daysAgo` della dashboard, ovviamente il processo di learning e l'accuratezza del modello aumetano per dataset più grandi, ma è richiesto un tempo più lungo di processamento.

Il grafico ritrae l'andamento della criptovaluta nel tempo e aggiunge in fondo alla time series i 30 giorni di previsione restituiti dal modello di neural network. La previsione viene indicata nel grafico con una linea tratteggiata ed è possibile visualizzare la previsione avvicinando il cursore nel punto del grafico interessato.

C'è la possibilità di inserire il tema chiaro o scuro per favorire la lettura di indici e grafici ed è abilitabile nel pannello del menù laterale, tale impostazione si salverà in memoria tramite cookie.

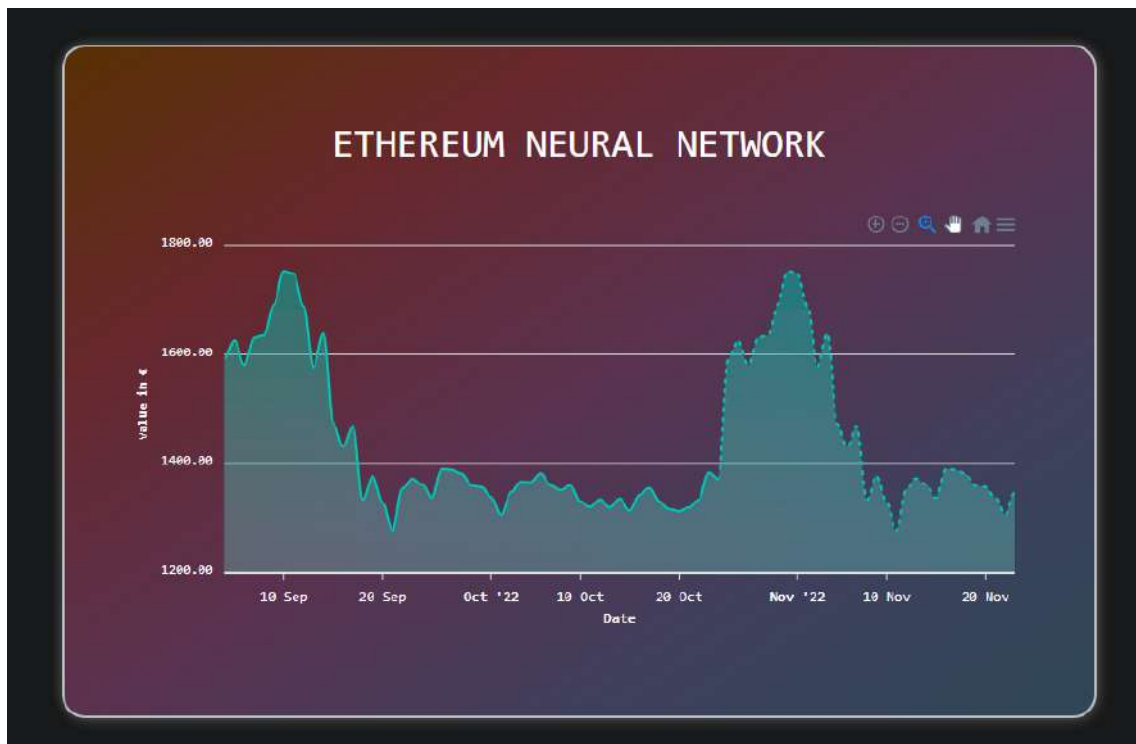


Figura 6.11: Neural network nel caso di Ethereum

E' buona pratica dividere l'intero set di dati input in due parti, training set e test set. Il training set viene utilizzato per addestrare il modello (fase di learning), mentre la parte di test set per testare la bontà del modello.

In fondo alla pagina di *check.html*, che analizza la singola criptovaluta, si riassumono gli indici statistici di bontà del modello e ulteriori informazioni generali. Ogni indice viene definito e analizzato nella sezione *glossary.html*.

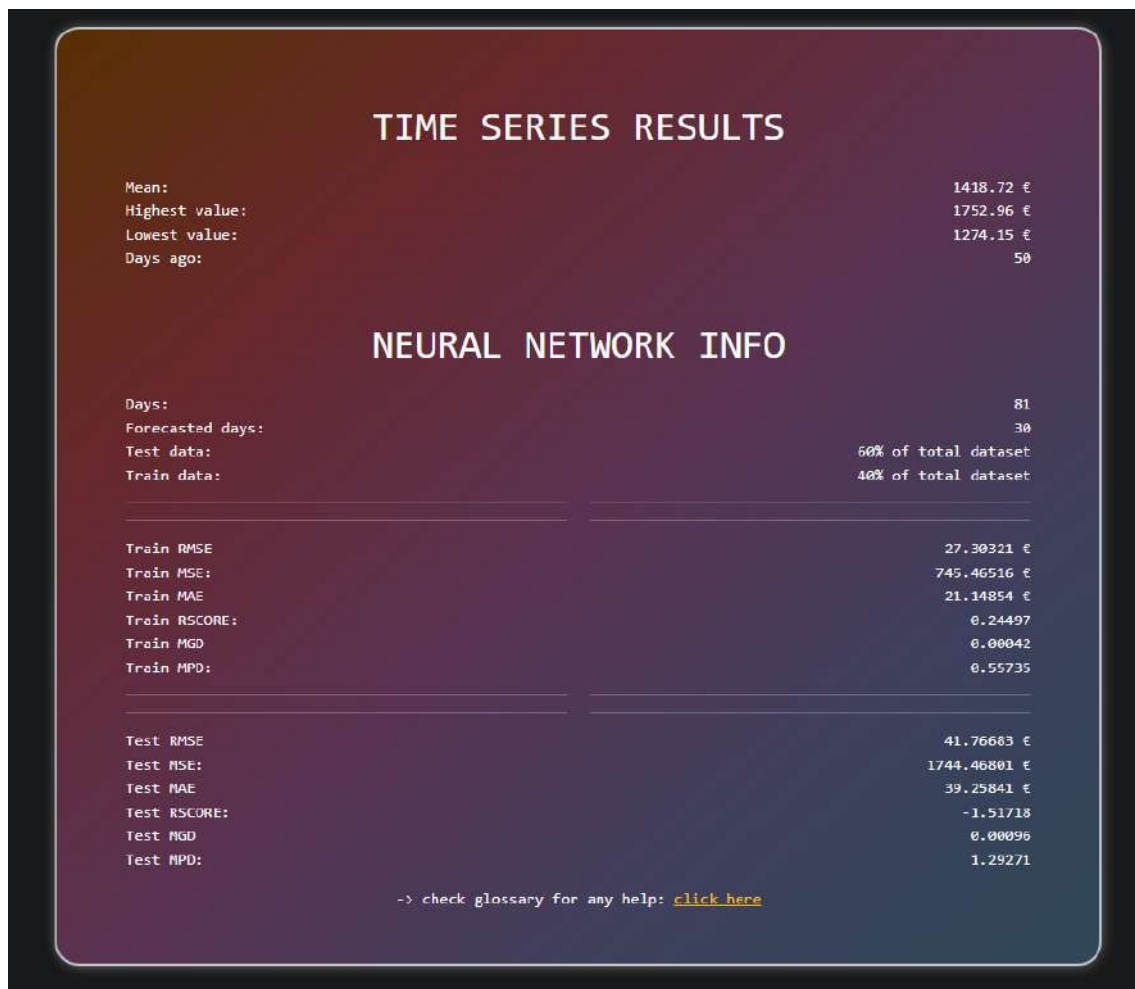
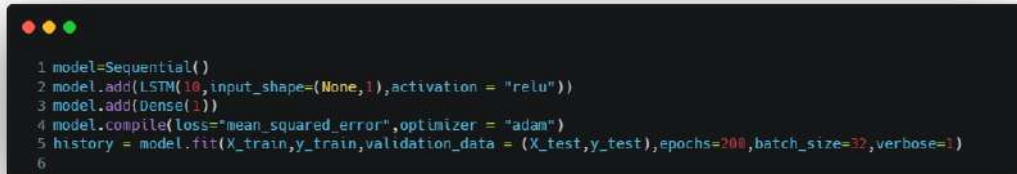


Figura 6.12: Metriche di bontà del modello

6.7 Analisi modello

In seguito si analizza il modello da un punto di vista semantico, ossia si cerca di attribuire un significato ad ogni istruzione che contribuisce alla sua creazione.



```
1 model=Sequential()  
2 model.add(LSTM(10,input_shape=(None,1),activation = "relu"))  
3 model.add(Dense(1))  
4 model.compile(loss="mean_squared_error",optimizer = "adam")  
5 history = model.fit(X_train,y_train,validation_data = (X_test,y_test),epochs=200,batch_size=32,verbose=1)  
6
```

Figura 6.13: Modello LSTM

In seguito si esaminano le istruzioni che vengono effettuate in python per riga:

1. Un modello sequenziale è appropriato per una semplice pila di livelli in cui ogni livello ha esattamente un input e un output. In questo caso lo stiamo inizializzando e la struttura di partenza risulta vuota;

model = keras.Sequential()

2. Con il metodo *.add()* andiamo ad aggiungere in maniera incrementale i vari livelli al modello. In questo caso stiamo aggiungendo un **hidden layer** di 10 neuroni/blocchi LSTM in grado di ricevere in ingresso una matrice $(x, 1)$, ossia il training data.

La *activation* é una funzione che mappa gli input di un nodo sull'output corrispondente. Relu sta per "*rectified linear unit*" e dato un input x restituisce il massimo tra x e 0. L'idea alla base di questa funzione di attivazione è che più il neurone è positivo, più è attivato.

model.add(LSTM(10,input_shape = (None, 1),activation = "relu"))

3. Il dense layer è un normale strato completamente connesso in una rete neurale. Un singolo neurone collega i 10 blocchi LSTM e restituisce l'output.

model.add(Dense(1))

4. Il *.compile()* compila il modulo. É necessario specificare la **loss function**, che viene utilizzata per trovare errori o deviazioni nel processo di apprendimento e

la **optimization** che ottimizza i pesi di input confrontando la previsione con la loss.

```
model.compile(loss = "mean_squared_error", optimizer = "adam")
```

5. Lo scopo principale della funzione *.fit()* é di allenare il modello utilizzando il dataset di training e validandolo su quelli di test. Il modello viene valutato durante l'addestramento **epochs** volte in modo tale che la rete neurale risca ad apprendere i dati di input. La **batch size** è il numero di campioni che vengono passati alla rete contemporaneamente durante una epoch.

$$batches\ in\ epoch = \frac{training\ size}{batch_size}$$

```
history = model.fit(X, Y, epochs, batch_size)
```

Dopo l'allenamento del modello LSTM é stato poi impiegato per la previsione del prezzo delle criptovalute per i successivi 30 giorni. A tali previsioni vengono affiancate dei indicatori statistici di misurazione della bontà del modello, che mostrano le performance del modello stesso.

La lettura delle previsioni e del quadro economico generale dello stato della criptovaluta fornisce agli utenti materiale aggiuntivo da esaminare. La lettura complessiva di questi dati può quindi portare ad una decisione differente rispetto alla sola analisi degli indici economici.

Capitolo 7

Risultati

La dashboard CoinStatus è scaricabile al seguente indirizzo:

https://github.com/silviufilete/Progetto_tesi

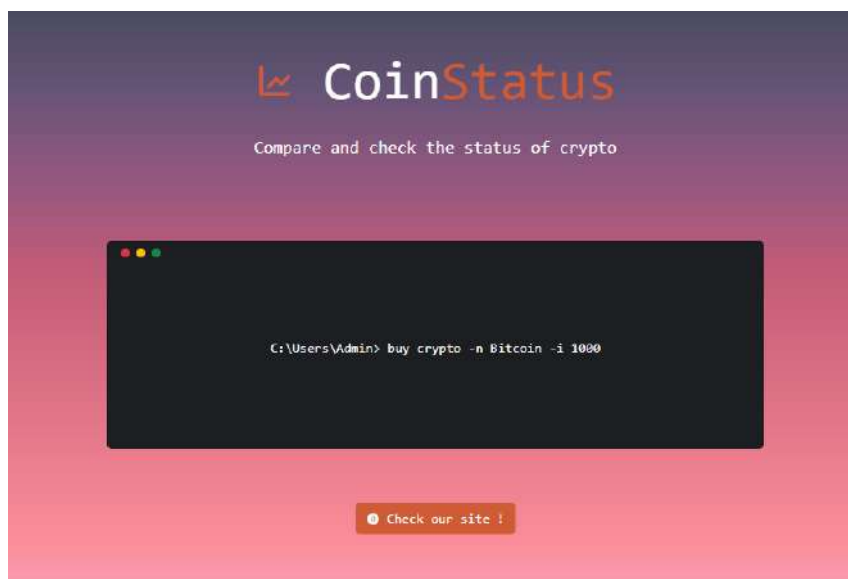


Figura 7.1: Pagina index.html

Sono però presenti dei requisiti necessari per l'avvio della dashboard sulla macchina su cui si esegue il codice:

- Python 3.7–3.10 ;
- Microsoft Visual C++ redistribuibile per Visual Studio 2015, 2017 e 2019;
- Installazione delle librerie/dipendenze nel file *requirements.txt*

Tutte le procedure sono indicate sulla pagine di github.

Capitolo 8

Conclusioni

La dashboard raccoglie al suo interno le 100 criptovalute più influenti al momento e cattura gli indici statistici ed economici più importanti, con l'aggiunta di previsione tramite un modello di neural network LSTM sul possibile valore del prezzo.

La previsione rappresenta uno strumento statistico molto importante che si può applicare in qualunque tipo di ambito laddove esistano dati su cui realizzare un modello statistico, a maggior ragione nell'ambito delle criptovalute in cui sapere come si comporterà la variabile di studio in un tempo futuro aiuta a prendere decisioni sul possibile investimento o sulla possibile vendita, riducendo in maniera progressiva l'incertezza su queste operazioni.

Dunque l'applicazione di questi modelli aiuta a comprendere più profondamente un set di dati che a prima vista potrebbe sembrare del tutto incorrelato e mettere in risalto informazioni che non sono leggibili direttamente dai semplici dati raccolti. Come si è spiegato precedentemente il mercato delle criptovalute rappresenta un'incertezza ed è importante avere degli strumenti statistici che siano in grado di misurare questa incertezza e oltretutto prevederla. Proprio per il motivo precedentemente riassunto l'impiego di tale modello all'interno della dashboard risulta essere una delle parti applicative più importanti.

La dashboard CoinStatus da un punto di vista funzionale risulta essere un prodotto utilizzabile da qualsiasi tipo di utente, perchè offre un glossario generale dove si può approfondire tutti i concetti teorici e ampliare le conoscenze statistiche.

L'interfaccia della dashboard risulta essere minimale e curata nell'aspetto, in questo modo gli utenti finali posso navigare nelle pagine e trovare le informazioni che cercano nel minor tempo possibile, senza dover interpretare la funzionalità di ogni pagina.

Le performance generali della dashboard risultano essere nella media. Si è prestata particolare attenzione al calcolo del modello di neural network su più criptovalute

utilizzando i thread in modo tale che le operazioni di apprendimento dei modelli fossero eseguite parallelamente per velocizzare il tutto. La dashboard è un applicativo software che riassume al suo interno diverse tecnologie che si sono spiegate in maniera dettagliata nei capitoli precedenti e inoltre riesce a rispondere a tutti gli obiettivi funzionali e logici che si sono posti. L'applicazione CoinStatus costituisce uno strumento invidiabile alle solite piattaforme di trading e ingloba al suo interno nuovi strumenti statistici più pratici e più comprensibili per l'utente.

Bibliografia

- [1] Jon Cunliffe. speech, vicegovernatore della Banca centrale.
- [2] W. Dai. "b-money, 1998.
- [3] Jerry Solinas Laurie Law, Susan Sabett. How to make a mint: The cryptography of anonymous electronic cash, 1996-1997.
- [4] Satoshi Nakamoto. Bitcoin a peer-to peer electronic cash system, 2008-2009.