Immagine che contiene emblema, cresta, Marchio, simbolo

Il contenuto generato dall'IA potrebbe non essere corretto.

**UNIVERSITÀ DEGLI STUDI DI BARI   
“ALDO MORO”**

Dipartimento di Informatica

Corso di Laurea in Informatica

*Tesi di Laurea in Algoritmi e Strutture Dati*

**Analisi automatizzata del mercato crypto attraverso tecniche di intelligenza artificiale.**

*Relatore*:  
Prof. Gianvito Pio

*Laureando*:  
Gioele Panico

Anno Accademico 2025/2026

**Abstract**

L’evoluzione delle tecnologie digitali e dell’intelligenza artificiale sta modificando profondamente i processi di analisi e previsione nel settore finanziario, con particolare attenzione al mercato delle criptovalute. Queste, a partire dal lancio di Bitcoin nel 2009, hanno visto una crescita esponenziale, caratterizzata da una forte volatilità e da una continua influenza del flusso informativo proveniente dal web e dai social media.

In questo contesto, la capacità di raccogliere, sintetizzare e classificare automaticamente le notizie rappresenta un elemento chiave per comprendere le dinamiche del mercato crypto e supportare decisioni di investimento più consapevoli.

Il lavoro di questa tesi propone un sistema automatizzato per l’analisi del mercato delle criptovalute basato su tecniche di intelligenza artificiale. In particolare, è stato sviluppato un flusso che prevede: la raccolta delle notizie tramite web scraping dalla piattaforma *CryptoPanic*; l’utilizzo di un modello linguistico di grandi dimensioni (LLaMA, versione 3.2) per la generazione di riassunti brevi e lunghi; la classificazione delle notizie tramite un approccio ibrido che combina un clustering non supervisionato per l’individuazione iniziale delle categorie e un classificatore supervisionato (Naïve Bayes) per l’assegnazione automatica delle etichette a nuove news; l’analisi del sentiment e la stima del peso informativo tramite un modello di regressione.

Infine, il sistema integra i risultati ottenuti all’interno di un report settimanale, arricchito dall’analisi del *Fear & Greed Index*, con l’obiettivo di fornire una panoramica sintetica ma completa sull’andamento del mercato.

L’approccio sviluppato dimostra come l’integrazione di tecniche di NLP, machine learning e metriche di sentiment analysis possa rappresentare un valido strumento di supporto per interpretare la complessità e la dinamicità del settore delle criptovalute, confermandone la rilevanza nell’ambito dell’analisi dei mercati finanziari digitali.

**Sommario**

1. **Introduzione1**

1.1 Contesto e motivazioni2

1.2 Obiettivi della tesi2

1.2 Struttura del documento 2

1. **Il mercato delle criptovalute1**

2.1 Cosa è una criptovaluta? 2

2.2 Evoluzione storica e tendenza attuali del mercato crypto2

2.3 Fonti informative del mondo crypto (Twitter, CryptoPanic, blog, ecc.)2

2.4 Limiti e problematiche dell’ analisi tradizionale del mercato2

1. **Tecnologie di riferimento1**

2.1 Intelliegnza Artificiale e Machine Learning: concetti di base 2

2.2 Natural Language Processing (NLP) e tecniche di analisi testuale2

2.3 Large Language Models (LLM): caratteristiche e modelli open-source2

2.4 Tecniche di classificazione supervisionata e non supervisonata 2

2.5 Modelli di regressione per sentiment e scoring 2

1. **Tecnologie di riferimento1**

2.1 Panoramica della pipeline sviluppata 2

2.2 Workflow del sistema2

2.3 Scelte progettuali e criteri di implementazione 2

1. **Raccolta e gestione dei dati (Data Collection)1**

2.1 Obiettivi della fase di scraping 2

2.2 Scelta delle fonti informative (CryptoPanic) 2

2.3 Struttura del database 2

2.2 Problemi comuni e soluzioni (duplicati, encoding, contenuti mancanti)2

1. **Sintesi automatica con LLM1**

2.1 Introduzione ai modelli LLM e scelta di LLaMA 3.2 2

2.2 Prompt engineering per la generazione dei riassunti 2

2.3 Generazione di riassunti brevi e lunghi 2

2.2 Integrazione della sintesi nel flusso di lavoro 2

1. **Classificazione delle notizie 1**

2.1 Classificazione non supervisionata (clustering)2

2.2 Etichettatura manuale delle categorie emerse 2

2.3 Addestramento del classificatore supervisionato (Naïve Bayes)2

2.2 Metriche di valutazione e risultati 2

2.3 Deployment della classificazione per nuove notizie 2

1. **Analisi quantitativa: sentiment e peso informativo 1**

2.1 Definizione dei concetti di “sentiment” e “peso”2

2.2 Linee guida per l’etichettatura manuale iniziale 2

2.3 Preprocessing dei testi per la regressione 2

2.2 Modelli di regressione testati (*Random Forest, XGBoost, regressore neurale*)2

2.3 Selezione del miglior modello e metriche di valutazione 2

2.2 Interpretazione dei risultati 2

1. **Generazione del report settimanale 1**

2.1 Obiettivi del report automatizzato 2

2.2 Struttura del PDF generato 2

2.3 Integrazione del Fear & Greed Index (media ponderata) 2

2.2 Visualizzazione del Fear & Greed Index (media ponderata) 2

2.3 Visualizzazione dinamica (grafici, gauge index) 2

2.2 Automazione del processo di reportistica 2

2.3 Esempio completo di report settimanale generato 2

1. **Implementazione e ambiente di sviluppo 1**

2.1 Linguaggi e strumenti utilizzzati (Python, librerie, GoogleColab) 2

2.2 Organizzazione del codice e struttura del repository GitHub2

2.3 Ambiente di sviluppo (IDE, versionamento, test)2

2.2 Risorse hardware e software utilizzate 2

1. **Discussione e sviluppi futuri 1**

2.1 Criticità riscontrate durante lo sviluppo 2

2.2 Limiti del sistema attuale 2

2.3 Possibili miglioramenti futuri2

* 1. Impatti nel mondo del trading e della finanza 2

**Bibliografia 1**

**Ringraziamenti 1**