



Politecnico di Torino

Corso di laurea in Ingegneria Gestionale classe L-8

Indagine sui collegamenti e profittabilità delle aziende in una rete globale

Relatore
Prof. Averta Giuseppe Bruno

Candidato
Abrate Enrico

Indice

| | |
|---|-----------|
| 1 – Proposta di progetto | 3 |
| 1.1 Studente proponente | 3 |
| 1.2 Titolo della proposta | 3 |
| 1.3 Descrizione del problema proposto | 3 |
| 1.4 Descrizione della rilevanza gestionale del problema | 3 |
| 1.5 Descrizione del data-set utilizzato per l'analisi | 4 |
| 1.6 Descrizione preliminare degli algoritmi coinvolti | 5 |
| 1.7 Descrizione preliminare delle funzionalità previste per l'applicazione software | 6 |
| 2 – Descrizione del problema affrontato | 8 |
| 2.1 Contesto operativo | 8 |
| 2.2 Input output | 8 |
| 2.3 Potenzialità e criticità | 9 |
| 3 – Descrizione del data-set utilizzato per l'analisi | 10 |
| 4 – Descrizione ad alto livello delle strutture dati e degli algoritmi coinvolti | 12 |
| 4.1 Pattern MCV DAO | 12 |
| 4.2 Algoritmi principali | 12 |
| 5 – Diagramma delle classi delle parti principali dell'applicazione | 16 |
| 6 – Videate dell'applicazione e link al video dimostrativo | 17 |
| 6.1 Immagini con risultati | 17 |
| 6.2 Link video presentazione | 20 |
| 7 – Tabelle con risultati sperimentali | 21 |
| 7.1 Creazione del grafo | 21 |
| 7.2 Analisi singola azienda | 21 |
| 7.3 Analisi gruppo di aziende | 22 |
| 7.4 Analisi ricorsiva | 23 |
| 8 – Valutazione risultati ottenuti e conclusioni | 24 |
| 8.1 Valutazione risultati ottenuti | 24 |
| 8.2 Conclusioni | 24 |

1 – Proposta di progetto

1.1 Studente proponente

Abrate Enrico (matricola s299727)

1.2 Titolo della proposta

Indagine sui collegamenti e profittabilità delle aziende in una rete globale

1.3 Descrizione del problema proposto

Tramite questo applicativo sarà possibile effettuare un'analisi dettagliata delle più importanti ed influenti aziende mondiali, suddivise per settore nelle quali esse operano e per paese di appartenenza. All'interno della piattaforma è possibile visualizzare le caratteristiche fondamentali delle imprese, applicare vari vincoli e metterle in relazione tra loro. Questo tipo di analisi risulta utile in due diversi contesti: da un lato, per il personale aziendale, il quale può individuare quali sono le proprietà e prestazioni di un'impresa concorrente nello stesso settore per metterle a confronto; dall'altro, per gli investitori, perché, tramite le informazioni ricavate dall'analisi, riescono a stabilire quali sono le aziende che rispecchiano i propri criteri di investimento.

1.4 Descrizione della rilevanza gestionale del problema

Questo tipo di analisi è molto rilevante in ambito gestionale, perché integra diverse competenze che sono richieste nel settore aziendale, legando aspetti di decisioni aziendali, strategici e di ottimizzazione. I motivi principali sono:

- **Analisi dei dati e decision making:** l'ingegnere gestionale si occupa di analizzare dati complessi per prendere determinate decisioni. Creare un grafo aziendale e identificare le connessioni tra le aziende per paese è un esercizio di analisi strutturata di dati reali, fondamentale per comprendere le dinamiche di mercato e il posizionamento competitivo.
- **Ottimizzazione dei profitti:** un compito fondamentale del personale di un'azienda è quello di riuscire a massimizzare i profitti. Creare un algoritmo che permetta di trovare

un percorso tale per cui i profitti siano massimizzati riflette il processo di ottimizzazione che un ingegnere gestionale deve seguire per migliorare le performance dell'impresa.

- Valutazione della competitività: mettere a confronto le diverse imprese e raggrupparle per i settori di competenza permette di effettuare un'analisi dettagliata di quanto sia influente un'azienda all'interno del settore e permette, inoltre, l'implementazione di nuove strategie o di valutare diversi accordi tra cui alleanze, fusioni o collaborazioni tra aziende.

1.5 Descrizione del data-set utilizzato per l'analisi

Il dataset utilizzato per l'analisi dei dati proviene da Kaggle (link: <https://www.kaggle.com/datasets/rakkesharv/forbes-2000-global-companies>) ed è basato sui dati del Forbs Global 2000. Questo database elenca le duemila migliori aziende al mondo per l'anno 2022 classificate in base ai criteri finanziari di reddito, profitto, asset e valore di mercato. Successivamente ho ritoccato il dataset per adattarlo alle esigenze di analisi. In particolare, ho modificato il nome della colonna "2022 Ranking" in "ID" inserendo in seguito i valori numerici univoci. Inoltre, ho aggiunto nelle colonne "YearFounded" e "TotalEmployees" i valori mancanti al fine di ottenere un'analisi più dettagliata. Questi indicatori aiutano a comprendere il successo e la stabilità delle imprese in diversi settori e regioni geografiche, fornendo un'ampia visione delle migliori aziende su scala globale.

Il database è costituito da undici colonne contenenti:

- ID: numero identificativo dell'azienda, rappresenta il posizionamento nella classifica;
- OrganizationName: nome dell'azienda;
- Industry: settore in cui opera l'impresa;
- Country: stato in cui è stabilita;
- YearFuonded: anno di fondazione;
- CEO: capo dell'organizzazione;
- Revenue: Reddito dell'impresa nell'anno corrente, in miliardi;
- Profits: Profitto generato nell'anno corrente, in miliardi;
- Assets: Valore totale dei beni appartenenti all'impresa nell'anno corrente, in miliardi;

- Market Value: Valore di mercato nell'anno corrente, in miliardi;
- TotalEmployees: Numero totale di impiegati.

1.6 Descrizione preliminare degli algoritmi coinvolti

All'interno dell'applicazione vengono utilizzate query SQL per la ricerca e l'estrazione di record specifici dal database. Inoltre, sono stati implementati in Python algoritmi ricorsivi e di ottimizzazione per analizzare i dati e ottenere risultati più complessi e approfonditi.

I problemi che l'applicazione web dovrà affrontare sono quelli che seguono.

In primo luogo, dovrà creare un grafo pesato e non orientato seguendo tali criteri: i nodi sono tutte le aziende che sono state fondate dopo un determinato anno (l'anno è scelto dalla DropDown di anni presenti nel database). Gli archi si formano tra aziende se e solo se queste risiedono nello stesso stato, se il valore assoluto della differenza degli anni di fondazione è minore di dieci e se hanno un numero minimo di impiegati n (scritto all'interno della TextBox). Il peso di tale arco è la media dei profitti delle due imprese.

Creato il grafo bisogna popolare le DropDown relative ai nomi dell'impresa, ai settori e agli stati con gli elementi estratti dai nodi.

Successivamente viene separata la pagina tramite tre tab che permettono di navigare all'interno dell'applicazione tramite le tre sezioni. In particolare, la prima contiene l'analisi dettagliata di una singola azienda, la seconda consente di effettuare l'analisi su un gruppo mirato di aziende, mentre il terzo è utilizzato per la ricerca ricorsiva.

Contenuto del primo tab:

Alla pressione del bottone "Cerca Simili" cerca i vicini che operano nello stesso settore dell'azienda selezionata dalla DropDown e stampa nome e peso dell'arco che li connette, in ordine di valore decrescente, all'interno della ListView dedicata alla sezione.

Tramite il pulsante "Calcola ROI" viene calcolato e visualizzato il ROI dell'impresa selezionata dal precedente DropDown, tramite il calcolo $(\text{Profitti}/\text{Assets}) * 100$. Vengono calcolati successivamente gli indici ROI delle altre aziende del grafo che operano nello stesso settore dell'impresa selezionata e dopodiché stampati per ordine decrescente di

valore. Creare, inoltre, un grafico a barre orizzontali contenente le dieci aziende con ROI più elevato, mettendo l'azienda selezionata con un colore differente per visualizzarla meglio.

Contenuto del secondo tab:

Alla pressione del bottone "Volume Affari" viene calcolato il totale dei profitti generati dalle imprese. Viene fornito un primo totale per il settore e un secondo totale per la nazione di appartenenza, selezionati entrambi dai rispettivi DropDown. Per entrambi i gruppi il sistema visualizza tutte imprese che contribuiscono alla crescita del totale all'interno del settore e del paese.

Quando viene premuto il bottone "Trova Cluster di Aziende" vengono identificate le componenti connesse nel grafo, analizzando i gruppi di aziende interconnesse. Per ogni componente calcola e visualizza il numero totale di aziende, il settore maggiormente presente all'interno del gruppo, il paese di riferimento e i redditi totali. Viene stampata inoltre una classifica dei cluster per numero di aziende coinvolte.

Contenuto del terzo tab:

Alla pressione del bottone "Massimizza Profitti" si genera un algoritmo ricorsivo che identifica un percorso semplice per massimizzare i profitti totali delle imprese utilizzando il peso degli archi. Per tale algoritmo si devono considerare esclusivamente le aziende che hanno un valore di mercato sopra una soglia prestabilita e con peso degli archi maggiore di dieci.

Ogni sezione è dotata di una propria ListView per la visualizzazione dei risultati ottenuti.

1.7 Descrizione preliminare delle funzionalità previste per l'applicazione software

L'applicazione è progettata per aiutare il fruitore ad esplorare ed analizzare una rete di aziende internazionali, sfruttando un'interfaccia interattiva che offre diverse funzioni di visualizzazione e analisi. Gli utenti possono esplorare le connessioni tra le aziende in base a criteri come settore, paese e numero di dipendenti, nonché ottenere insight sui profitti e distribuzioni delle aziende all'interno di cluster.

Nella prima riga viene visualizzata una DropDown contenente gli anni di fondazione e una TextBox per indicare il numero minimo dei dipendenti. Selezionati questi parametri,

premendo il pulsante “Crea Grafo”, l'utente genera un grafo pesato e non orientato che gli permetterà di visualizzare nella riga successiva il numero di nodi e archi presenti.

Dopo la creazione del grafo l'utente visualizza i tab contenenti le tre divisioni. Nella parte relativa alla singola azienda egli seleziona un'azienda da un menù a discesa e, con la pressione del pulsante “Cerca Simili”, visualizzerà tutte le aziende vicine a quella selezionata, ossia le imprese connesse direttamente che operano nello stesso settore. Per ciascuna organizzazione vengono mostrati il nome e il peso dell'arco che li collega, ordinati in ordine decrescente di valore. Premendo sul pulsante “Calcola ROI” viene calcolato il Roi dell'azienda selezionata e quello delle imprese che lavorano nello stesso settore, visualizzati in ordine decrescente di valore insieme al grafico.

Nella seconda sezione dell'interfaccia, alla prima riga, sono presenti due DropDown, uno per selezionare il settore e l'altro per lo stato, seguiti dal bottone “Volume Affari”. Premendo il pulsante l'utente può calcolare il totale dei profitti generati dalle aziende in due modalità: per settore o per paese. Verranno poi visualizzati i due totali e le classifiche delle aziende che contribuiscono a ciascuno di essi.

Nella riga successiva, premendo il pulsante “Trova Cluster di Aziende”, verranno visualizzati i gruppi di imprese fortemente connesse tra loro. L'applicazione mostrerà una classifica dei primi tre più grandi cluster, per numero di aziende coinvolte, sui quali verranno calcolati il numero totale di aziende, il settore maggiormente presente e i redditi totali.

Nell'ultima sezione, dopo aver scelto la soglia del valore di mercato, tramite il pulsante “Massimizza Profitti” viene individuato un percorso per ottimizzare i guadagni delle imprese, utilizzando un algoritmo ricorsivo. Vengono quindi visualizzati i nomi delle organizzazioni, i settori in cui esse operano e i profitti generati (peso dell'arco).

2 – Descrizione del problema affrontato

2.1 Contesto operativo

Il progetto nasce dall'esigenza di analizzare le dinamiche competitive e i collegamenti tra le 2000 aziende più influenti a livello globale dell'anno 2022, basandosi sui dati del *Forbes Global 2000*. Queste imprese rappresentano un campione significativo dell'economia mondiale, coprendo una vasta gamma di settori e paesi. L'obiettivo principale è quello di fornire uno strumento analitico in grado di supportare decisioni strategiche sia per il personale aziendale sia per gli investitori. Da un lato, le aziende possono confrontare le proprie performance con quelle di concorrenti diretti; dall'altro, gli investitori possono identificare opportunità di investimento coerenti con i propri criteri di selezione.

L'analisi effettuata utilizza specifici vincoli che permettono di mettere in relazione tra loro le diverse imprese, per poter studiare le singole performance, e i collegamenti che si creano tra di esse. L'obiettivo è quello di comprendere come caratteristiche comuni possano influenzare la competitività e la profittabilità, identificando pattern e cluster di aziende con dinamiche simili.

2.2 Input output

Per poter effettuare l'analisi vengono messi a disposizione i dati contenuti nel data-set pubblicato dal *Forbes Global 2000*. All'interno sono contenute informazioni quali settore in cui operano, stato di appartenenza e dati numerici come ad esempio il reddito, il valore di mercato e degli asset, ricavi totali e numero di dipendenti. L'utente può inserire specifici parametri, come l'anno minimo di fondazione, la soglia per il numero di dipendenti o il valore di mercato, per affinare la ricerca in base agli aspetti di suo interesse. I risultati che si ottengono comprendono un grafo non orientato pesato contenente le aziende connesse tra di loro in base ai vincoli selezionati, le classifiche delle imprese più redditizie suddivise per settore e paese, l'identificazione di cluster di aziende, i cammini minimi ottimi tra i nodi del grafo e un grafico contenente i ROI delle imprese.

2.3 Potenzialità e criticità

I punti di forza di questo progetto sono notevoli. Esso offre la possibilità di inserire vincoli per la ricerca e renderla molto flessibile in base ai criteri desiderati dall'utente. Il programma offre supporto alle decisioni che l'utilizzatore deve effettuare grazie alla possibilità di comparare le aziende e vedere come si collocano in base alle diverse classifiche presenti, utile sia per persone interne che esterne ad essa, come ad esempio dei potenziali investitori.

Nonostante i punti di forza, il programma presenta alcune criticità. A causa dell'elevata mole di dati che il software deve gestire, possono esserci dei rallentamenti nella visualizzazione dei risultati ed è richiesta una buona efficienza degli algoritmi. Ciò si nota soprattutto nella creazione del grafo e nella parte di ricerca del cammino ottimo tra le aziende, in quanto utilizza un algoritmo di tipo ricorsivo complesso.

3 – Descrizione del data-set utilizzato per l'analisi

Il database utilizzato per effettuare l'analisi si basa sui dati forniti dal Forbes Global 2000, una delle classifiche più riconosciute a livello mondiale per la valutazione delle imprese in tutto il mondo. Forbes è una rinomata rivista finanziaria che, ogni anno, pubblica la classifica delle 2000 maggiori aziende pubbliche a livello mondiale, offrendo una dettagliata panoramica dello stato dell'economia del mondo.

La classifica si basa su quattro metriche fondamentali:

- Vendite (Revenue): il fatturato generato dall'azienda nell'anno corrente;
- Profitti (Profits): Il guadagno al netto delle imposte e dei costi sostenuti a fine anno;
- Asset: il valore dei beni e delle risorse ottenute dall'impresa, non solo al fine della produzione, ma anche come investimento;
- Valore di mercato (Market Value): valore totale dell'azienda calcolato anche in base al prezzo delle azioni in borsa.

Queste metriche offrono un quadro completo delle prestazioni finanziarie delle aziende, aiutando a comprendere il loro impatto nei rispettivi settori e a livello globale. La classifica Global 2000 non solo identifica le imprese più potenti e influenti, ma permette anche di analizzare le tendenze economiche internazionali. La lista aggiornata può essere consultata sul sito ufficiale di Forbes (<https://www.forbes.com/lists/global2000>).

Il data-set proviene dal sito Kaggle (link del database: <https://www.kaggle.com/datasets/rakkesharv/forbes-2000-global-companies>) dove è scaricabile gratuitamente. Oltre ai dati relativi alla ricchezza delle imprese sopra citati, sono presenti ulteriori informazioni:

- ID: numero identificativo dell'azienda, rappresenta il posizionamento nella classifica;
- OrganizationName: nome dell'azienda;
- Industry: settore in cui opera l'impresa;
- Country: stato in cui è stabilita;
- YearFounded: anno di fondazione;

- CEO: capo dell'organizzazione;
- TotalEmployees: Numero totale di impiegati.

Il database originariamente presentava alcuni valori nulli che potevano compromettere la completezza dell'analisi. Di conseguenza, per rendere l'analisi più dettagliata, sono stati sostituiti i valori nulli con dati verosimili. In particolare, sono stati completati i valori che non si conoscevano nelle colonne di MarketValue, TotalEmployees e YearFounded, mentre nelle altre colonne non vi erano valori mancanti. I dati mancanti sono stati stimati facendo una media dei valori relativi alla colonna a cui facevano riferimento, in modo tale da non turbarla eccessivamente.

Nel complesso sono presenti duemila righe all'interno del data-set (una per azienda), consentendo una approfondita analisi su scala globale. Questo permette di suddividere il problema in base alle diverse categorie presenti, quali il settore o il paese di appartenenza, il numero di dipendenti dell'impresa e tanto altro.

4 – Descrizione ad alto livello delle strutture dati e degli algoritmi coinvolti

4.1 Pattern MCV DAO

Per lo sviluppo dell'applicazione è stata utilizzata la combinazione delle due architetture di tipo MCV e DAO in linguaggio python.

Il modello MCV (Model Controller View) è un'architettura che permette di organizzare la logica dell'applicazione suddividendola in tre parti con scopi diversi.

All'interno della *View* viene gestita la parte grafica e l'interfaccia con cui l'utente interagisce e visualizza l'output, fornito dal Controller, senza modificarlo.

Il *Controller* ha lo scopo di gestire la logica di interazione tra il View e il Model. Egli prende le informazioni fornite in input dalla View, quali eventi dei pulsanti o testi, per poi mandarli al model per permettere la manipolazione. Quando il model ha effettuato gli algoritmi i dati vengono mandati al controller e successivamente alla View per renderli visibili all'utente.

Il *Model* gestisce la logica applicativa del programma. Tramite i dati che gli arrivano dal Controller e dal DAO, effettua gli algoritmi manipolativi e trasferisce i risultati al Controller per poterli visualizzare.

Il *DAO (Data Access Object)* è un'architettura utilizzata per l'estrazione dei dati contenuti all'interno del database. In questa sezione sono inserite le query SQL utilizzate per estrapolare i dati, per poi essere mandati al Model per la manipolazione.

4.2 Algoritmi principali

Gli algoritmi utilizzati per lo sviluppo dell'applicazione sfruttano i pattern sopra descritti.

In primo luogo, il programma deve creare un grafo pesato e non orientato utilizzando le informazioni raccolte nella view inserite dall'utente. In particolare, deve creare un grafo in cui i nodi sono tutte le aziende che sono state fondate dopo un determinato anno, gli archi si formano tra aziende se e solo se queste risiedono nello stesso stato, se il valore assoluto della differenza degli anni di fondazione è minore di dieci e se hanno un numero minimo di

impiegati n e il peso di tale arco è la media dei profitti delle due imprese. Quindi, prendendo due aziende A e B il peso è calcolato come segue:

$$Peso(A, B) = \frac{Profitti(A) + Profitti(B)}{2}$$

L'algoritmo è implementato da due Query eseguite nel DAO: una per trovare il grafo, una per individuare gli archi e per prendere il valore del peso degli archi.

Una volta creato il grafo viene visualizzata la parte sottostante in cui vi è la divisione della pagina in tre Tab differenti. Nella prima sezione sono contenuti gli algoritmi inerenti all'analisi sulla singola azienda selezionata (ricerca simili e calcolo ROI), nella seconda risiedono i metodi per la valutazione dei dati su un gruppo ristretto di imprese (classifiche profitti totali e cluster aziende) e nella terza è presente l'algoritmo ricorsivo (massimizzazione profitti).

Nella prima scheda, dopo aver selezionato un'azienda presente nel DropDown, premendo sul pulsante "Cerca Simili" algoritmo cerca tutte le imprese che operano nello stesso settore. Per ogni azienda nello stesso settore, l'algoritmo recupera il peso dell'arco che la connette all'azienda selezionata. I risultati vengono ordinati in ordine decrescente di peso, mostrando il nome delle aziende e il valore del peso associato all'arco. Questo algoritmo consente all'utente di esplorare le aziende simili a quella selezionata in base alla connessione nel grafo.

Tenendo una imprese selezionata è possibile calcolare il ROI (Return on Investment) di tale compagnia e confrontarlo con quello delle altre aziende che operano nel medesimo settore. Il calcolo è effettuato come segue:

$$ROI = \frac{Profitti}{Assets} \times 100$$

Questo indice misura la redditività di un investimento, calcolato come il rapporto tra il guadagno ottenuto e il costo dell'investimento. I risultati vengono ordinati in ordine decrescente di ROI. Inoltre, l'algoritmo crea un grafico a barre orizzontali che visualizza le dieci aziende con i ROI più elevati, evidenziando l'azienda selezionata con un colore diverso per facilitarne l'identificazione. Nel caso in cui essa faccia parte delle dieci migliori viene visualizzata una sola volta nella posizione che la identifica.

Nella seconda sezione sono presenti due DropDown, uno per selezionare lo stato di appartenenza e l'altro per individuare il settore in cui opera. Scelti i valori di partenza, tramite la pressione del pulsante "Calcola Volume Affari" vengono generati due totali. Il primo è relativo al totale dei profitti sviluppati dalle imprese presenti all'interno del grafo secondo il criterio di stato di appartenenza, mentre il secondo è relativo alla somma dei profitti prodotti dalle aziende in base al settore in cui lavorano. Oltre ai totali sono presenti due liste, una per ogni rispettivo totale, in cui vengono visualizzati i nomi delle imprese con relativo valore dei profitti, scritti in ordine decrescente.

L'algoritmo per la ricerca dei cluster di aziende individua le componenti connesse nel grafo. Utilizzando una ricerca in profondità (DFS) o in ampiezza (BFS), l'algoritmo identifica tutti i gruppi di aziende che sono interconnesse tra loro. Per ogni cluster identificato, l'algoritmo calcola il numero totale di aziende nel cluster, il settore prevalente, lo stato di appartenenza e il totale dei profitti generati dalle aziende del cluster. I cluster vengono poi ordinati in base al numero di aziende coinvolte, mostrando i cluster per dimensione decrescente.

Nell'ultima scheda risiede l'algoritmo ricorsivo. Questo algoritmo utilizza un approccio, il quale permette di individuare il percorso che massimizza i profitti totali lungo le aziende del grafo. Il percorso è determinato utilizzando il peso degli archi, rappresentato dalla media dei profitti delle aziende collegate. Il sistema considera solo le aziende che hanno un valore di mercato superiore a una soglia prestabilita e con peso dell'arco maggiore di dieci, filtrando così le aziende che non soddisfano tali criteri. L'algoritmo esplora ricorsivamente le connessioni nel grafo e seleziona il percorso che massimizza la somma dei profitti delle aziende lungo il cammino. Alla fine, il percorso ottimale viene visualizzato, mostrando le aziende selezionate e il totale dei profitti. L'algoritmo è scritto come segue:

```
def getPercorso(self, valMercato):
    self._bestPath = []
    self._bestVal = 0
    self._memo = {}
    minPeso = 10 # Peso minimo dell'arco per considerarlo

    for n in self._graph.nodes:
        if n.MarketValue > valMercato:
            parziale = [n]
            self.ricorsione(parziale, valMercato, 0, minPeso)

    return self._bestPath, self._bestVal

def ricorsione(self, parziale, valMercato, profittoAttuale, minPeso):
    chiaveMemo = (parziale[-1], frozenset(parziale))
    if chiaveMemo in self._memo and self._memo[chaveMemo] >= profittoAttuale:
```

```

        return
    self._memo[chiaveMemo] = profittoAttuale

    # Aggiorna il miglior percorso trovato
    if profittoAttuale > self._bestVal:
        self._bestVal = profittoAttuale
        self._bestPath = copy.deepcopy(parziale)

    for node in self._graph.neighbors(parziale[-1]):
        if node not in parziale and node.MarketValue > valMercato:
            peso = self._graph[parziale[-1]][node]["weight"]

            if peso < minPeso:
                continue

            max_profitto = profittoAttuale + peso + sum(
                [self._graph[parziale[-1]][n]["weight"] for n in
self._graph.neighbors(parziale[-1])]
            )
            if max_profitto <= self._bestVal:
                continue

            parziale.append(node)
            self.ricorsione(parziale, valMercato, profittoAttuale+peso, minPeso)
            parziale.pop()

```

In sintesi, gli algoritmi implementati nell'applicazione permettono di costruire un grafo dinamico di aziende basato su vari criteri, di analizzare i dati attraverso vari strumenti (ROI, volume affari, cluster, ecc.) e di ottimizzare i profitti lungo percorsi selezionati. Ogni algoritmo è progettato per rispondere a specifiche richieste dell'utente, permettendo un'esperienza di interazione completa con il sistema di analisi.

5 – Diagramma delle classi delle parti principali dell'applicazione

L'applicazione è strutturata utilizzando il pattern MCV DAO in modo tale da poter garantire una chiara suddivisione del codice. L'architettura del programma è suddivisa come segue:

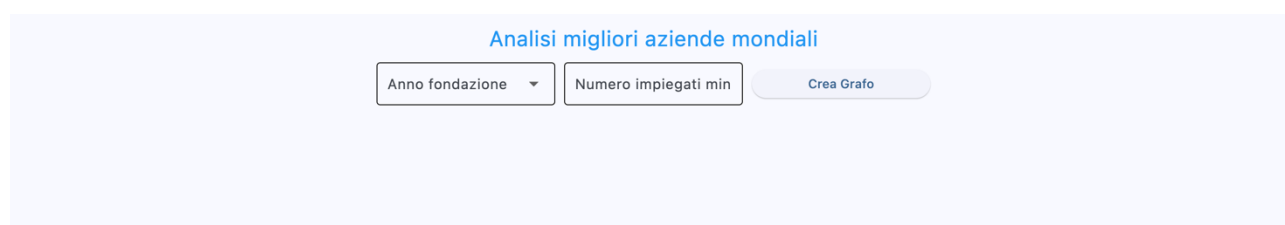
- Cartella *database*: contiene i file necessari per effettuare le interazioni con il dataset. In particolare, contiene il database in formato SQL (`companies_db.sql`), il file `DAO.py` contenente le query utilizzate per visualizzare i dati del database, il file `db_connect.py` per gestire la connessione al database e il file `connector.cnf` per configurare le credenziali d'accesso al database.
- Cartella *model*: contiene la logica del programma e gli algoritmi utilizzati per lo sviluppo. Al suo interno è collocato il file `model.py` per la scrittura di tutti gli algoritmi principali, il file `company.py` che definisce la classe `Company` e include i suoi attributi (`ID`, `OrganizationName`, `Industry`, `Country`, `YearFounded`, `CEO`, `Profits`, `Assets`, `Revenue`, `MarketValue`, `TotalEmployees`) e il file `connessa.py` con all'interno la classe `Connessa` formata dai 2 nodi e il peso dell'arco tra di essi.
- Cartella *UI*: si occupa dell'interfaccia utente e della gestione delle interazioni utilizzando Flet, un framework Python usato per lo sviluppo di interfacce grafiche. È formata dalla `view.py` responsabile della presentazione dei dati all'utente e dal file `controller.py` il quale si interpone tra il model e la view, gestendo gli eventi dell'interfaccia (come i click sui pulsanti) e invocando gli algoritmi definiti nel model.
- Cartella *grafici*: corrisponde ad una cartella d'appoggio. Al suo interno vengono collocati i grafici creati nel model in formato png, ma questi vengono distrutti in seguito all'utilizzo. La cartella rimane vuota e si evita di sovraccaricare la memoria.
- File *main.py*: rappresenta il punto di avvio dell'applicazione. Inizializza l'interfaccia Flet e crea le istanze di View, Controller e Model, collegandole tra loro.

6 – Videate dell'applicazione e link al video dimostrativo

6.1 Immagini con risultati

Dopo aver avviato il programma, viene visualizzata la schermata contenente unicamente la DropDown popolata da anni presenti nel database sotto la colonna YearFounded, una TextBox adibita all'inserimento di un numero di dipendenti minimo che l'azienda deve avere e il bottone per avviare la creazione del grafo (figura 6.1).

Figura 6.1: Schermata principale.



Analisi migliori aziende mondiali

Anno fondazione ▼ Numero impiegati min Crea Grafo

Nel caso in cui l'utente provi ad avviare la creazione di un grafo senza selezionare l'anno di fondazione o il numero minimo di dipendenti, l'applicazione darà errore e negherà l'avvio (figura 6.2). Questa azione si presenta per tutte gli altri metodi del programma nel caso in cui non siano completati tutti i campi.

Figura 6.2: Schermata errore.



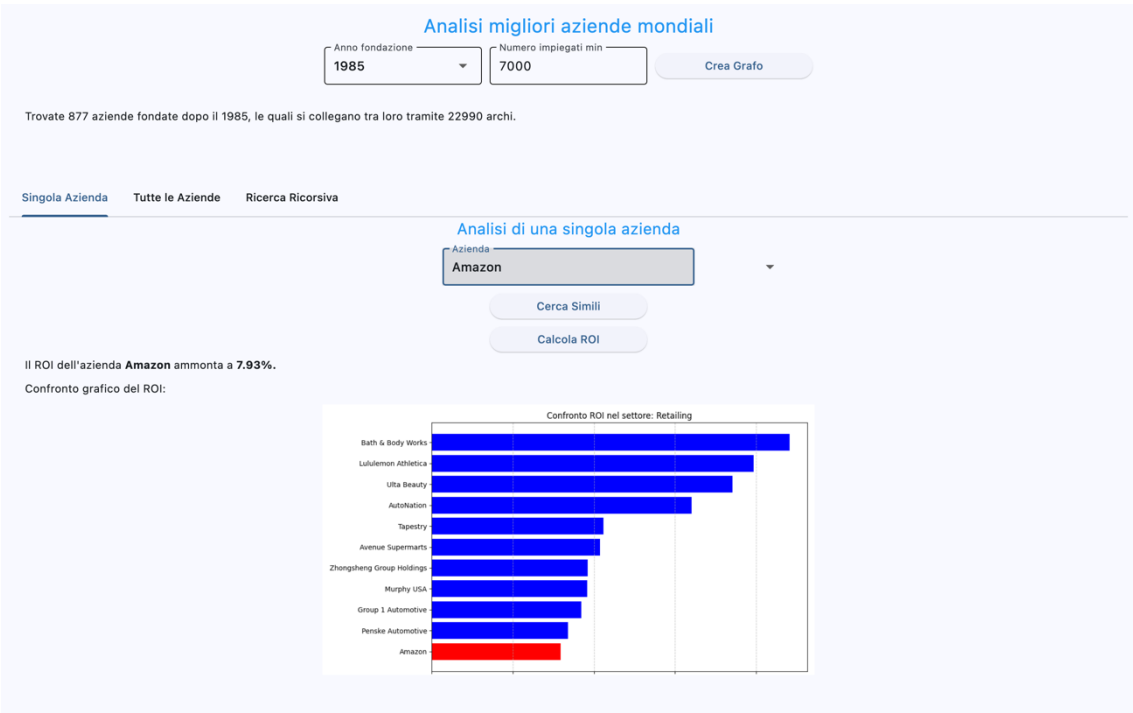
Se la creazione del grafo va a buon fine viene visualizzato la parte sottostante, contenente le tre schede per effettuare l'analisi.

Nella parte relativa alla singola azienda troviamo due output diversi. Il primo è relativo alla ricerca di aziende simili a quella selezionata nella DropDown (figura 6.3), il secondo è relativo al calcolo del ROI della compagnia, con un grafico annesso per rendere più facile la visualizzazione e permettere un confronto diretto (figura 6.4).

Figura 6.3: Aziende simili a quella selezionata (neighbors)



Figura 6.4: Calcolo del ROI dell'azienda selezionata



Il ROI delle altre aziende competenti nello stesso settore (Retailing):

- Azienda: **Bath & Body Works**, ROI totalizzato: **22.06%**.
- Azienda: **Lululemon Athletica**, ROI totalizzato: **19.84%**.
- Azienda: **Ulta Beauty**, ROI totalizzato: **18.54%**.
- Azienda: **AutoNation**, ROI totalizzato: **16.03%**.
- Azienda: **Tapestry**, ROI totalizzato: **10.59%**.
- Azienda: **Avenue Supermarts**, ROI totalizzato: **10.36%**.
- Azienda: **Zhongsheng Group Holdings**, ROI totalizzato: **9.62%**.
- Azienda: **Murphy USA**, ROI totalizzato: **9.59%**.
- Azienda: **Group 1 Automotive**, ROI totalizzato: **9.22%**.
- Azienda: **Penske Automotive**, ROI totalizzato: **8.40%**.
- Azienda: **Amazon**, ROI totalizzato: **7.93%**.
- Azienda: **Vipshop Holdings**, ROI totalizzato: **7.47%**.
- Azienda: **Diamond**, ROI totalizzato: **6.67%**.

Nella seconda parte, una volta selezionato lo stato e il settore di appartenenza della compagnia, alla pressione del pulsante “Volume Vendite” vengono generati i due totali con le rispettive liste delle aziende che vi partecipano (figura 6.5). Cliccando sul bottone “Trova Cluster Aziende” si visualizzano le componenti connesse con relative caratteristiche e il numero totale delle componenti connesse trovate (figura 6.6).

Figura 6.5: Totale profitti generati dalle imprese nello stato e settore scelto.

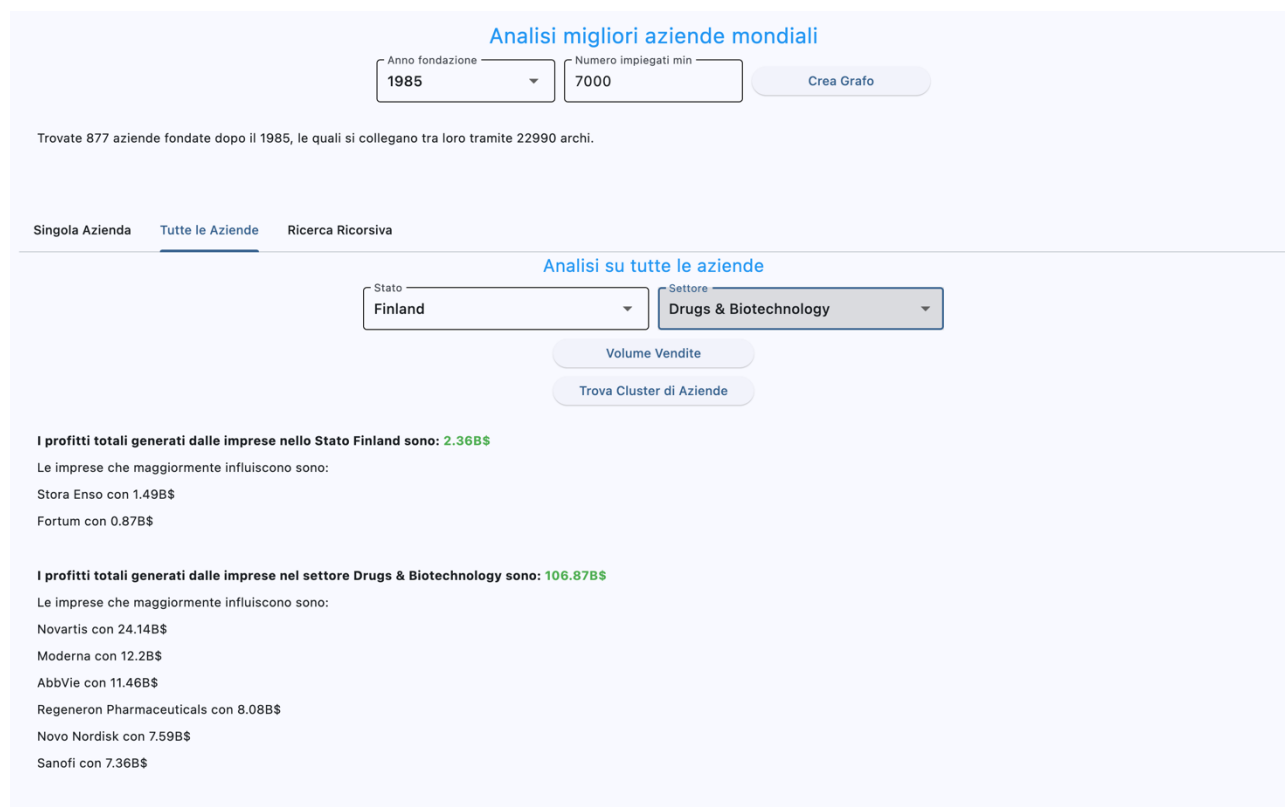
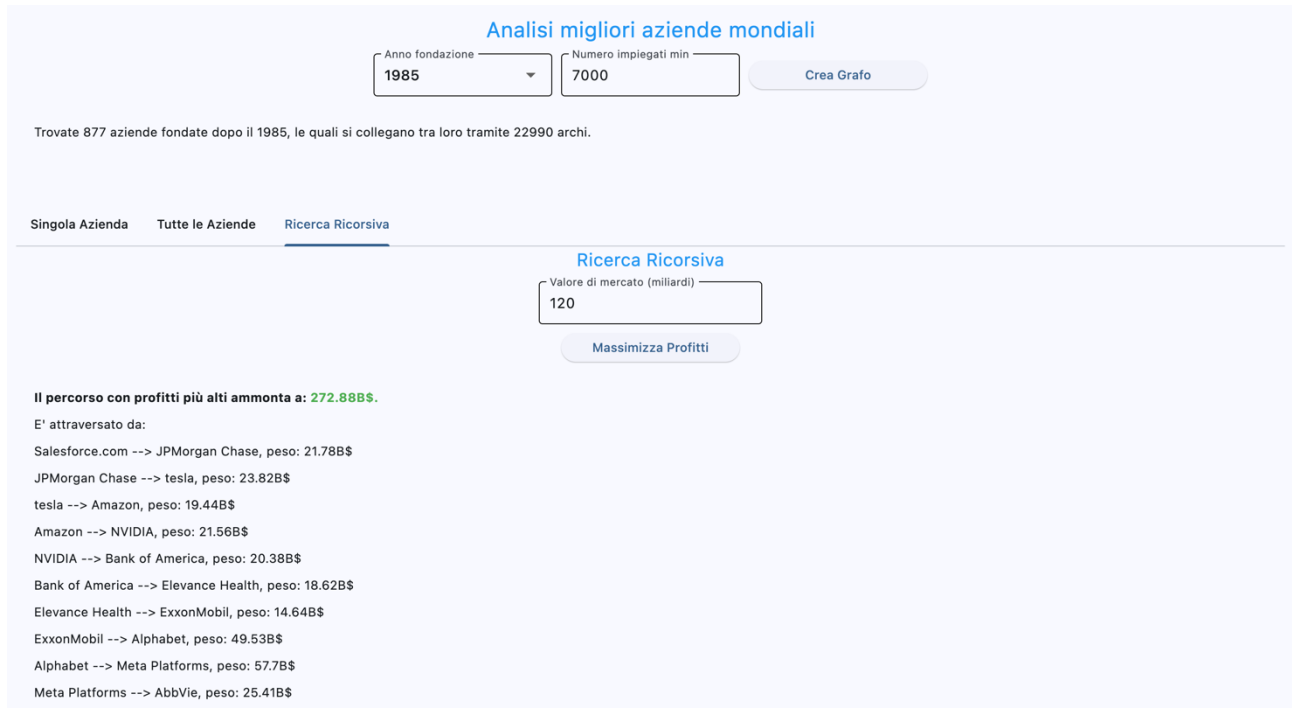


Figura 6.6: Componenti connesse del grafo.



Nell'ultima parte, una volta selezionato la soglia minima di valore di mercato che deve essere rispettato, alla pressione del pulsante "Massimizza Profitto" viene generato un cammino minimo che massimizza il profitto delle imprese (calcolato tramite il peso delle imprese selezionate) (figura 6.7)

Figura 6.7: *Algoritmo ricorsivo*



Analisi migliori aziende mondiali

Anno fondazione: 1985 | Numero impiegati min: 7000 | Crea Grafo

Trovate 877 aziende fondate dopo il 1985, le quali si collegano tra loro tramite 22990 archi.

Singola Azienda | Tutte le Aziende | **Ricerca Ricorsiva**

Ricerca Ricorsiva

Valore di mercato (miliardi): 120 | Massimizza Profitti

Il percorso con profitti più alti ammonta a: 272.88B\$.

E' attraversato da:

- Salesforce.com --> JPMorgan Chase, peso: 21.78B\$
- JPMorgan Chase --> tesla, peso: 23.82B\$
- tesla --> Amazon, peso: 19.44B\$
- Amazon --> NVIDIA, peso: 21.56B\$
- NVIDIA --> Bank of America, peso: 20.38B\$
- Bank of America --> Elevance Health, peso: 18.62B\$
- Elevance Health --> ExxonMobil, peso: 14.64B\$
- ExxonMobil --> Alphabet, peso: 49.53B\$
- Alphabet --> Meta Platforms, peso: 57.7B\$
- Meta Platforms --> AbbVie, peso: 25.41B\$

6.2 Link video presentazione

Link video YouTube: <https://youtu.be/6ofkEqOvw5Q>

7 – Tabelle con risultati sperimentali

Al fine di effettuare l'analisi sono stati utilizzati i dati secondo il seguente criterio. Per ogni algoritmo svolto dall'applicazione sono inseriti in tabella i risultati ottenuti e il tempo impiegato per arrivare a tale risultato. Nella prima riga delle tabelle troviamo i dati maggiormente presenti (per quanto riguarda lo stato e il settore in cui sono stabilite le imprese) o i valori medi (per i valori numerici di fatturato, profitto, assets, valore di mercato e numero di dipendenti), mentre nelle righe seguenti vi sono risultati ottenuti con altri dati che si discostano dalla media.

7.1 Creazione del grafo

Per la creazione del grafo sono stati utilizzati i valori medi relativi all'anno di fondazione e al numero degli impiegati.

| Anno fondazione | Numero minimo dipendenti | Numero di nodi | Numero di archi | Tempo (secondi) |
|-----------------|--------------------------|----------------|-----------------|-----------------|
| 1959 | 48900 | 1250 | 2617 | 0.08161 |
| 1980 | 7500 | 979 | 24171 | 0.21641 |
| 2005 | 3000 | 202 | 2120 | 0.05008 |
| 1920 | 20000 | 1579 | 12344 | 0.14186 |

Da questo punto in poi i risultati ottenuti sono stati calcolati utilizzando il grafo relativo alla seconda riga della tabella, ossia grafo contenente i nodi le aziende fondate dopo il 1980, con numero di dipendenti minimo di 7500.

7.2 Analisi singola azienda

I risultati ottenuti dall'algoritmo utilizzato per il bottone *Cerca Simili* sono:

| Azienda | Numero simili | Tempo (secondi) |
|----------------------|---------------|-----------------|
| Amazon | 11 | 0.00023 |
| Sunac China Holdings | 28 | 0.00049 |
| Alibaba Group | 6 | 0.00027 |
| PayPal | 11 | 0.00019 |

I risultati ottenuti dall'algoritmo utilizzato per il bottone *Calcola ROI* sono:

| Azienda | ROI calcolato | Tempo (secondi) |
|----------------------|---------------|-----------------|
| Amazon | 7.93% | 0.10788 |
| Sunac China Holdings | 2.95% | 0.14721 |
| Alibaba Group | 3.68% | 0.10598 |
| PayPal | 5.50% | 0.11653 |

7.3 Analisi gruppo di aziende

Per gli algoritmi di analisi relativi ai gruppi di imprese, le prime due righe corrispondono ai campi più presenti di stato e settore.

Per il calcolo dei *Volumi di Vendita* sono stati conseguiti tali risultati:

| Stato | Settore | Totale stato (B\$) | Totale settore (B\$) | Tempo (secondi) |
|---------------|------------------------|--------------------|----------------------|-----------------|
| United States | Banking | 650.37 | 360.97 | 0.00039 |
| China | Diversified Financials | 530.48 | 189.69 | 0.00046 |
| Italy | Retailing | 1.15 | 73.24 | 0.00020 |
| Finland | IT Software & Services | 2.36 | 174.23 | 0.00031 |

Per l'algoritmo relativo ai *Cluster di Aziende* sono stati utilizzati più grafi differenti e sono conseguiti tali risultati:

| Anno fondazione | Numero minimo dipendenti | Componenti trovate | Tempo (secondi) |
|-----------------|--------------------------|--------------------|-----------------|
| 1959 | 48900 | 1007 | 0.01987 |
| 1980 | 7500 | 375 | 0.01689 |
| 2005 | 3000 | 71 | 0.00262 |
| 1920 | 20000 | 936 | 0.01601 |

7.4 Analisi ricorsiva

Per effettuare questo punto non è stato utilizzato il valore medio nella prima riga (ossia 38), ma un valore più alto, in quanto avviare la ricorsione usando un valore soglia così basso causa enormi rallentamenti su grafi di tale grandezza. I risultati ottenuti sono:

| Valore di mercato | Totale fatturato (B\$) | Nodi attraversati | Tempo (secondi) |
|-------------------|------------------------|-------------------|-----------------|
| 100 | 349.83 | 18 | 1921.27277 |
| 120 | 126.14 | 16 | 121.39302 |
| 180 | 308.47 | 13 | 9.67025 |
| 200 | 302.02 | 12 | 5.11434 |

Avendo scelto un grafo molto grande (979 nodi e 24171 archi), i tempi per eseguire la ricorsione sono molto elevati, dovuti dal tipo di lavoro che deve essere svolto.

8 – Valutazione risultati ottenuti e conclusioni

8.1 Valutazione risultati ottenuti

L'applicazione sviluppata è un efficace strumento di analisi per studiare le connessioni tra le principali aziende nel mondo e per mettere in relazione tra loro le proprie caratteristiche e performance.

I punti di forza del programma sono la flessibilità dell'analisi, in quanto è possibile modellare e personalizzare i vincoli da imporre sulle aziende, la facilità con cui i dati possono essere visti grazie all'impaginazione chiara ed intuitiva e alla suddivisione in schede. Un altro aspetto importante è che gli algoritmi sono efficienti e permettono di analizzare una grossa mole di dati, come il database utilizzato per le ricerche.

I punti di debolezza dell'applicazione riguardano principalmente la velocità d'esecuzione della ricerca ricorsiva. Siccome si utilizza un data-set di grosse dimensioni, questo fa sì che si crei un grafo contenente migliaia di archi ed un numero elevato di nodi, pertanto, se non si applicano abbastanza vincoli durante la ricorsione, questa potrà dare problemi di inefficienza e rallentamento nell'esecuzione del programma. Come abbiamo visto nel capitolo 7.4, per un grafo molto ampio anche se l'algoritmo viene ottimizzato, l'elaborazione di tutte le possibili combinazioni resta molto intensa.

8.2 Conclusioni

Nel complesso l'applicazione fornisce un grande supporto alle decisioni finanziarie contribuendo in diversi ambiti:

- **Analisi della concorrenza:** le aziende possono confrontare le proprie performance con quelle dei competitors, analizzando i collegamenti nel grafo per individuare potenziali opportunità di collaborazione o strategie di mercato.
- **Supporto per gli investitori:** grazie all'indice di ROI calcolato e alle classifiche in cui vengono visualizzate le imprese, un potenziale investitore riesce a comprendere meglio quali sono le aziende sulle quali è meno rischioso e più remunerativo investire il proprio capitale.

- Grazie l'algoritmo ricorsivo è possibile fare un'ottimizzazione dei profitti attraversando il minor numero di archi possibili; quindi, questo stesso algoritmo può essere riutilizzato, anche inserendo dati differenti, per migliorare le strategie aziendali.

I risultati sperimentali confermano la correttezza e l'efficienza degli algoritmi, con tempi di esecuzione ottimizzati nella maggior parte dei casi. Tuttavia, alcuni limiti computazionali emergono nelle analisi su larga scala, soprattutto nella ricerca ricorsiva. Per migliorare ulteriormente il sistema, si potrebbero introdurre tecniche che vadano a ridurre il numero di combinazioni da valutare.

Oltre all'applicazione specifica sviluppata, il codice implementato è altamente modulare e riutilizzabile per altre applicazioni con esigenze simili. Gli algoritmi di creazione e analisi del grafo, il calcolo del ROI e la ricerca ricorsiva di percorsi ottimali possono essere adattati a diversi contesti, come il calcolo di altri indici, l'analisi per investimenti o la ricerca di un ottimo di fatturato con utilizzo di diversi prodotti e l'inserimento di costi variabili e fissi.

In definitiva, l'applicazione sviluppata rappresenta un solido strumento di analisi che unisce metodologie di data analysis e ottimizzazione, dimostrando un forte valore pratico sia per le aziende che per gli investitori.