



Università degli Studi di Salerno

Laurea Triennale in Informatica
Corso di Fondamenti di Intelligenza Artificiale



FoodDiary-GA

<https://github.com/GuidaStefano/FoodDiary-GA>

Stefano Guida

matricola 0512115409

Anno Accademico 2023-2024

Sommario

1 - Introduzione	3
1.1 - Introduzione alla nutrizione	3
1.1.1 - Macronutrienti e calorie	3
1.1.2 - La dieta	3
1.1.4 - Obesità: un problema reale.....	4
1.2 – Scopo del progetto	4
1.3 – Specifica PEAS	4
1.4 - Caratteristiche dell’ambiente	5
2 - Analisi del Problema	6
2.1 - Algoritmi di Ricerca Locale	6
2.1.1 - Hill Climbing	6
2.1.2 - Simulated Annealing.....	6
2.2 - Algoritmi Genetici	6
2.3 - Considerazioni e Scelta dell'Algoritmo	7
3 - Struttura del progetto	7
3.1 - Gestione dei Dati	7
3.2 - Pulizia dei Dati	8
3.3 - Panoramica dei Packages.....	8
3.3.1 - Il package ‘ga’	8
3.3.2 – Il package obj.....	8
3.3.3 – Il package loader	8
3.3.4 – Il package util.....	8
4 – Algoritmo genetico	9
4.1 – Codifica dell’individuo	9
4.2 – Popolazione	9
4.3 – Funzione di Fitness	10
4.4 – Selezione	10
4.5 – Crossover.....	10
4.6 – Mutazione	11
4.7 – Stopping Condition.....	11
5 – Risultati	12
6 – Conclusione	12
6.1 - Prospettive Future.....	12

1 - Introduzione

1.1 - Introduzione alla nutrizione

La nutrizione è una disciplina scientifica che studia come gli alimenti influenzano la nostra salute e il funzionamento del corpo. In particolare, si concentra sull'analisi dei nutrienti e su come questi contribuiscano al nostro benessere complessivo. La base di ogni dieta equilibrata è l'apporto corretto di macronutrienti e micronutrienti, che sono fondamentali per fornire energia e sostenere le funzioni vitali.

1.1.1 - Macronutrienti e calorie

I macronutrienti principali sono i carboidrati, i grassi e le proteine:

- I carboidrati, che forniscono 4 calorie per grammo, sono la principale fonte di energia per il corpo e sono essenziali per mantenere l'attività fisica e mentale.
- I grassi, con 9 calorie per grammo, svolgono ruoli cruciali come l'assorbimento delle vitamine liposolubili e la produzione di ormoni.
- Le proteine, che forniscono anch'esse 4 calorie per grammo, sono fondamentali per la costruzione e la riparazione dei tessuti corporei.

1.1.2 - La dieta

Per garantire una salute ottimale, è importante che il nostro apporto calorico quotidiano sia ben bilanciato tra questi nutrienti. Le linee guida nutrizionali offrono indicazioni su come distribuire le calorie giornaliere tra i vari macronutrienti per sostenere un'alimentazione sana. Ad esempio, è consigliabile che circa il 60% delle calorie provenga dai carboidrati, il 25% dai grassi e il 15% dalle proteine. Questa ripartizione aiuta a garantire che il corpo riceva una quantità adeguata di energia e nutrienti essenziali.

Inoltre, la distribuzione delle calorie nei pasti della giornata è altrettanto importante. Una dieta ben progettata dovrebbe prevedere una suddivisione equilibrata delle calorie tra colazione, pranzo, cena e spuntini. Ad esempio, una parte significativa delle calorie dovrebbe

essere assunta durante il pranzo, mentre la colazione e la cena dovrebbero fornire quantità più moderate di energia. Questo approccio non solo aiuta a mantenere costanti i livelli di energia, ma supporta anche il metabolismo e la gestione del peso.

In sintesi, la scienza della nutrizione fornisce le basi per comprendere come gli alimenti e i nutrienti interagiscono con il nostro corpo. Una dieta equilibrata, che considera sia la quantità che la qualità dei nutrienti consumati, è fondamentale per mantenere una buona salute e prevenire malattie.

1.1.4 - Obesità: un problema reale

Il sovrappeso e l'obesità rappresentano condizioni croniche che derivano da uno squilibrio energetico tra le kilocalorie consumate e quelle spese. Queste condizioni possono portare a una serie di gravi problemi di salute, come diabete, malattie respiratorie, malattie cardiovascolari, ipertensione, disordini della tiroide e disturbi psicologici. Questi problemi sono di crescente preoccupazione a livello globale e richiedono interventi efficaci e personalizzati.

In Italia, la prevalenza del sovrappeso e dell'obesità è altrettanto significativa. Secondo i dati più recenti, il 46% degli adulti italiani è affetto da sovrappeso e il 10% da obesità. Questi numeri si traducono in circa 14 milioni di adulti con sovrappeso e 3 milioni con obesità. Le complicazioni legate a queste condizioni includono una maggiore incidenza di malattie croniche e una riduzione della qualità della vita.

1.2 – Scopo del progetto

Il progetto **FoodDiary-GA** si propone di affrontare la complessità della pianificazione dietetica personalizzata attraverso l'uso di algoritmi genetici.

L'obiettivo principale è sviluppare uno strumento innovativo che aiuti a generare diete equilibrate e su misura per le esigenze individuali, tenendo conto dei fabbisogni nutrizionali specifici. Utilizzando algoritmi genetici, il progetto mira a esplorare un vasto spazio di possibili combinazioni alimentari per trovare soluzioni ottimali che rispettino le linee guida nutrizionali e le calorie giornaliere consigliate.

Questo approccio consente di creare piani alimentari personalizzati che non solo migliorano l'efficacia della dieta, ma offrono anche maggiore flessibilità e adattabilità rispetto ai metodi tradizionali di pianificazione dietetica.

1.3 – Specifica PEAS

Di seguito è riportata la specifica **PEAS**:

- **Performance:** Le performance dell'agente sono valutate in base alla qualità delle diete generate. Vengono considerate l'efficacia nel rispettare le linee guida nutrizionali e l'affinità del piano alimentare alle esigenze specifiche dell'utente.

- **Environment:** L'ambiente dell'agente è rappresentato dall'insieme di tutte le possibili combinazioni di alimenti disponibili nel database.
- **Actuators:** Gli attuatori dell'agente sono i moduli software che selezionano e combinano gli alimenti per formare piani dietetici. Questi moduli permettono di generare, modificare e ottimizzare le diete, rispondendo alle necessità e alle preferenze dell'utente.
- **Sensors:** I sensori acquisiscono i dati dall'ambiente necessari per la creazione delle diete. Questo include le informazioni nutrizionali sugli alimenti e le specifiche esigenze dietetiche dell'utente. I sensori raccolgono e interpretano questi dati per fornire piani alimentari adeguati e personalizzati.

1.4 - Caratteristiche dell'ambiente

Il sistema **FoodDiary-GA** opera in un ambiente caratterizzato da specifiche proprietà che ne definiscono il funzionamento e l'interazione dell'agente. Di seguito sono descritte le principali caratteristiche dell'ambiente:

- **Completamente Osservabile:** L'ambiente è completamente osservabile, poiché l'agente ha accesso a tutte le informazioni nutrizionali degli alimenti e alle specifiche esigenze dell'utente in ogni momento.
- **Singolo Agente:** Il sistema opera con un solo agente intelligente, il quale si occupa di selezionare e combinare gli alimenti per creare piani dietetici personalizzati. Non ci sono altri agenti che interagiscono direttamente con l'ambiente o influenzano le decisioni dell'agente.
- **Statico:** L'ambiente rimane statico mentre l'agente prende decisioni. Le informazioni sugli alimenti e le linee guida nutrizionali non cambiano durante il processo decisionale.
- **Stocastico:** C'è una componente randomica nell'ambiente che può influenzare lo stato successivo. Questa variabilità può derivare dalle preferenze alimentari degli utenti, dalle scelte casuali di alimenti nel database o da altre incertezze legate alla pianificazione dietetica.
- **Discreto:** L'ambiente è discreto in quanto le azioni e gli stati sono ben definiti e limitati. Ci sono un numero finito di alimenti disponibili e le possibili combinazioni di diete sono limitate.
- **Sequenziale:** Le decisioni dell'agente sono sequenziali e influenzano le decisioni future. Ogni scelta alimentare o modifica al piano dietetico può avere un impatto sulle successive combinazioni di alimenti e sull'efficacia complessiva della dieta.

2 - Analisi del Problema

La seguente sezione esamina la natura del problema affrontato dal progetto **FoodDiary-GA** e valuta le strategie di risoluzione più adeguate.

Il problema centrale consiste nel generare piani dietetici ottimali utilizzando un ampio database di alimenti. L'obiettivo è selezionare e combinare gli alimenti in modo da soddisfare le esigenze nutrizionali specifiche degli utenti, rispettare le linee guida dietetiche e mantenere un equilibrio calorico e nutrizionale.

Questo rappresenta un classico problema di ottimizzazione, dove è necessario trovare la migliore combinazione di alimenti per massimizzare una o più funzioni obiettivo.

Tra gli algoritmi più adatti a risolvere questo tipo di problema ci sono gli Algoritmi di Ricerca Locale e gli Algoritmi Genetici.

2.1 - Algoritmi di Ricerca Locale

Gli algoritmi di ricerca locale sono tecniche di ottimizzazione combinatoria che operano modificando iterativamente una singola soluzione candidata per migliorare il risultato rispetto a una funzione obiettivo. Questi algoritmi esplorano soluzioni vicine a quella corrente nella speranza di trovare una soluzione migliore.

2.1.1 - Hill Climbing

L'algoritmo **Hill Climbing** monitora uno stato corrente e, ad ogni iterazione, passa allo stato vicino con il valore più elevato. Tuttavia, questo algoritmo può fermarsi in un massimo locale, dove non è possibile trovare una soluzione migliore anche se la soluzione complessiva non è ottimale.

2.1.2 - Simulated Annealing

Simulated Annealing prende ispirazione dal processo di indurimento dei metalli, che implica riscaldamento e raffreddamento graduale.

Questo algoritmo esplora inizialmente in modo più casuale e permette anche di accettare soluzioni peggiori temporaneamente per evitare di rimanere bloccato in un massimo locale. Con il tempo, la probabilità di accettare soluzioni peggiori diminuisce, concentrandosi su aree più promettenti.

2.2 - Algoritmi Genetici

Gli algoritmi genetici appartengono alla categoria degli algoritmi evolutivi e sono ispirati dai processi naturali di evoluzione.

Questi algoritmi utilizzano meccanismi di selezione, crossover e mutazione per evolvere una popolazione di soluzioni e migliorare progressivamente la qualità delle stesse rispetto a una funzione obiettivo.

La capacità di esplorare ampi spazi di ricerca e combinare soluzioni diverse rende gli algoritmi genetici particolarmente efficaci per problemi complessi di ottimizzazione.

2.3 - Considerazioni e Scelta dell'Algoritmo

Analizzando il problema di **FoodDiary-GA**, si osserva che l'ambiente presenta una vasta gamma di combinazioni possibili e potenziali massimi locali, che possono ostacolare la ricerca di una soluzione globale ottimale.

Sebbene gli algoritmi di ricerca locale, come Hill Climbing e Simulated Annealing, offrano buone capacità di esplorazione, essi possono rimanere intrappolati in massimi locali.

Per affrontare questa sfida, è stato scelto di utilizzare gli algoritmi genetici, grazie alla loro robustezza nell'esplorare ampi spazi di ricerca e alla loro capacità di combinare diverse soluzioni per generare piani dietetici ottimali. La scelta degli algoritmi genetici è stata inoltre influenzata dal desiderio di adottare un approccio progettuale innovativo e intuitivo, capace di adattarsi efficacemente alle complesse esigenze del progetto.

3 - Struttura del progetto

La seguente sezione offre una panoramica dettagliata dell'implementazione del progetto **FoodDiary-GA**, delineando i vari componenti software utilizzati, il ruolo e la funzionalità di ciascun package, nonché le tecnologie e gli strumenti impiegati durante lo sviluppo.

FoodDiary-GA è stato sviluppato utilizzando esclusivamente Java (JDK 17), che fornisce una base solida per l'implementazione dell'algoritmo genetico e la gestione delle componenti del sistema.

Per la gestione dei dati relativi agli alimenti, fondamentali per il funzionamento del sistema, è stato utilizzato un file in formato .xlsx

3.1 - Gestione dei Dati

I dati sugli alimenti, comprendenti informazioni sui nutrienti e le calorie, sono stati organizzati in un file .xlsx. Per facilitare la lettura e l'interpretazione di questi file, è stata utilizzata la libreria **Apache POI**. Questa libreria è stata scelta per la sua efficacia nella gestione dei file Excel e per la sua capacità di manipolare e estrarre dati in modo efficiente.

3.2 - Pulizia dei Dati

Durante la fase di preparazione dei dati, è stata effettuata una pulizia mirata per semplificare il dataset. Sono state rimosse numerose colonne relative ai micronutrienti per ridurre la complessità e concentrare l'analisi sui principali parametri nutrizionali. Questa operazione ha permesso di mantenere il dataset più snello e gestibile.

3.3 - Panoramica dei Packages

3.3.1 - Il package 'ga'

Il package **ga** costituisce il cuore dell'implementazione dell'algoritmo genetico. All'interno di questo package, si trovano vari sub-package che contengono tutti gli operatori genetici e le classi necessarie per il funzionamento dell'algoritmo. Le classi astratte impiegate sono state adattate dagli esempi forniti sulla piattaforma e-learning del corso. Ecco un riepilogo dei sub-package:

- **Individuals:** contiene le classi per la rappresentazione e la codifica degli individui.
- **Population:** contiene le classi per la gestione della popolazione.
- **Initializer:** contiene le classi per l'inizializzazione della popolazione.
- **Fitness:** contiene le classi relative alla funzione fitness.
- **Operators:** contiene le classi che rappresentano gli operatori genetici di selezione, crossover e mutazione.
- **Metaheuristics:** contiene le classi che rappresentano le metaeuristiche utilizzate per guidare il processo di ricerca e ottimizzazione.
- **Results:** contiene la classe utilizzata per la gestione e l'analisi dei risultati.

3.3.2 – Il package obj

Il package **obj** contiene tutte le classi necessarie per la rappresentazione del piano alimentare. Include classi che modellano gli elementi chiave come gli alimenti, i pasti, i piani giornalieri e i piani settimanali.

3.3.3 – Il package loader

Il package **loader** contiene una singola classe, **FoodLoader**, che si occupa di gestire il file .xlsx. FoodLoader si occupa di leggere il file, interpretarlo utilizzando la libreria Apache POI e memorizzare tutti gli alimenti in memoria.

3.3.4 – Il package util

Il package **util** contiene alcune classi di utilità, utilizzate durante lo sviluppo del progetto.

4 – Algoritmo genetico

L'algoritmo genetico costituisce sicuramente il cuore di **FoodDiary-GA**, fornendo il meccanismo di ottimizzazione dei piani alimentari.

In questa sezione, verranno spiegate nel dettaglio le strategie adottate in termini di operatori genetici, come sono stati configurati i parametri ed altri aspetti rilevanti dell'algoritmo genetico.

4.1 – Codifica dell'individuo

Nel contesto di **FoodDiary-GA**, la codifica degli individui è un aspetto fondamentale dell'algoritmo genetico.

Un individuo è rappresentato come un piano alimentare settimanale, strutturato in una tabella in cui le colonne rappresentano i giorni della settimana e le righe i momenti della giornata (colazione, pranzo, spuntino, cena).

Ogni cella della tabella corrisponde a un pasto specifico e contiene una serie di coppie alimento-quantità (espressa in grammi).

Questo approccio, noto come "codifica basata su tabelle", sfrutta l'efficacia della rappresentazione tabellare per gestire e manipolare i dati alimentari. La struttura tabellare rende la rappresentazione degli individui chiara e intuitiva.

4.2 – Popolazione

La creazione e la gestione della popolazione iniziale sono aspetti fondamentali in un algoritmo genetico.

Per il progetto **FoodDiary-GA**, è stata adottata una popolazione di dimensione fissa, il che implica che il numero di individui rimane costante in ogni iterazione dell'algoritmo genetico. Sono state effettuate prove con diverse dimensioni della popolazione, inclusi 30, 60, 100 e 150 individui, e si è riscontrato che una popolazione iniziale di 100 individui offriva i migliori risultati.

Ogni individuo viene generato in modo pseudo-casuale: per ogni giorno della settimana, le calorie totali giornaliere vengono suddivise tra i quattro pasti principali (colazione, pranzo, spuntino, cena) seguendo le linee guida nutrizionali. Successivamente, ogni pasto viene composto cercando di avvicinarsi il più possibile alla quantità di calorie assegnata.

In questa fase iniziale, l'ottimizzazione si concentra solo sulle calorie, mentre i macronutrienti non vengono considerati, permettendo di stabilire una base di partenza efficace per ulteriori ottimizzazioni.

4.3 – Funzione di Fitness

La funzione di fitness è il criterio secondo cui viene valutata la bontà dei diversi individui. In **FoodDiary-GA** si utilizza una singola funzione di fitness che è effettivamente combinazione lineare di quattro sotto-funzioni di fitness:

$$f = f_1 + f_2 + f_3 + f_4$$

In particolare, questa funzione di fitness viene calcolata per ogni giornata del piano alimentare. Il valore di fitness del piano alimentare settimanale è dato dalla somma dei valori di fitness di tutte le giornate.

Scomponendo la formula, abbiamo:

- Funzione di fitness relative alle **calorie giornaliere** (f_1)
- Funzione di fitness relativa ai **carboidrati giornalieri** (f_2)
- Funzione di fitness relativa alle **proteine giornaliere** (f_3)
- Funzione di fitness relativa ai **grassi giornalieri** (f_4)

Anche se si tratta di quattro funzioni differenti, l'idea alla base è unica: per ogni parametro (calorie, carboidrati, proteine, grassi) viene assegnato un bouns (+ 2 punti), un punteggio neutro (0 punti) o un malus (- 2) a seconda della distanza da un valore target, stabilito in fase di configurazione.

4.4 – Selezione

L'operatore genetico di Selezione rappresenta il meccanismo secondo il quale gli individui più promettenti della popolazione vengono selezionati per formare il mating pool.

In **FoodDiary-GA**, si è scelto di utilizzare la strategia **Roulette Wheel**, una strategia in cui ogni individuo può essere selezionato con una probabilità proporzionale al suo punteggio di fitness relativo a l'intera popolazione.

4.5 – Crossover

L'operatore genetico di Crossover rappresenta un componente cruciale dell'algoritmo genetico, mirato a generare nuove soluzioni combinando le caratteristiche di due individui esistenti.

In **FoodDiary-GA**, si è scelto di adottare la tecnica del **Single Point Crossover**. Questo approccio prevede che venga scelto un singolo punto di taglio all'interno delle rappresentazioni degli individui, consentendo di scambiare parte del loro patrimonio genetico per creare nuovi individui.

In termini pratici, il Crossover viene eseguito selezionando, in modo casuale, un giorno della settimana come punto di taglio. A partire da quel giorno, le informazioni alimentari dei due individui vengono scambiate.

Ad esempio, se due piani alimentari rappresentano i pasti per tutta la settimana, il Crossover comporta lo scambio dei pasti da un giorno in poi fino alla fine della settimana.

Questo metodo consente di preservare la variabilità genetica tra le soluzioni e favorisce l'esplorazione di nuove combinazioni alimentari che potrebbero portare a risultati migliori in termini di ottimizzazione delle calorie e dei nutrienti.

4.6 – Mutazione

Nel contesto di **FoodDiary-GA**, la Mutazione è un elemento chiave dell'algoritmo genetico che introduce variazioni nei piani alimentari per garantire la diversità genetica e prevenire il fenomeno della convergenza prematura.

Inizialmente, è stata adottata una strategia di **Random Resetting**, in cui un giorno del piano alimentare veniva selezionato casualmente e generato nuovamente. Tuttavia, questa strategia si è rivelata inefficace: i risultati ottenuti non mostravano un miglioramento significativo nella funzione di fitness e non contribuivano a ottimizzare i piani alimentari in modo soddisfacente.

Di conseguenza, è stata effettuata una revisione della strategia di Mutazione. Invece di scegliere un giorno in modo casuale, è stato adottato un approccio mirato che prevede la selezione del giorno con il peggior valore di fitness della settimana.

Questa scelta consente di concentrarsi su aree specifiche del piano alimentare che necessitano di miglioramenti, apportando modifiche mirate per ottimizzare le calorie e i nutrienti in modo più efficace. Il cambiamento nella strategia di Mutazione ha portato a prestazioni decisamente migliorate.

4.7 – Stopping Condition

Nell'ambito degli algoritmi genetici, la determinazione di condizioni di arresto adeguate è fondamentale per garantire che l'algoritmo concluda l'ottimizzazione in modo efficiente e senza sprechi di risorse.

Dopo un'attenta valutazione, si è deciso di basare le condizioni di arresto principalmente sul numero di iterazioni eseguite. Sebbene il limite possa sembrare elevato, si è stabilito un massimo di **2000 iterazioni** per garantire che l'algoritmo abbia sufficienti opportunità di esplorare e ottimizzare il piano alimentare.

In aggiunta, è stata introdotta una condizione di arresto basata sui miglioramenti consecutivi: l'algoritmo si fermerà se non si riscontrano miglioramenti per **400 iterazioni consecutive**. Questa scelta è stata motivata dal desiderio di evitare che l'algoritmo continui a funzionare senza apportare progressi significativi.

Nonostante il limite di iterazioni abbastanza elevato, il processo di ottimizzazione risulta essere comunque rapido.

5 – Risultati

Di seguito un esempio di risultato riscontrato dall'algoritmo, con il seguente valore target (1900 calorie, 165g carboidrati, 160g proteine, 60g grassi):

```
Gen 1080) 42.0 vs 46.0 (NewAvg vs BestAvg)
Gen 1081) 42.0 vs 46.0 (NewAvg vs BestAvg) ==> Early Stop
Ricerca terminata in 1081/2000 iterazioni.
SATURDAY : 1957 calorie (C: 174,29 - P: 167,65 - F: 67,34)
TUESDAY : 1767 calorie (C: 178,55 - P: 156,03 - F: 48,36)
THURSDAY : 1972 calorie (C: 161,77 - P: 168,08 - F: 73,40)
FRIDAY : 1972 calorie (C: 178,88 - P: 171,78 - F: 71,04)
WEDNESDAY : 1831 calorie (C: 158,32 - P: 150,50 - F: 72,55)
SUNDAY : 1904 calorie (C: 165,43 - P: 173,98 - F: 59,11)
MONDAY : 1911 calorie (C: 130,36 - P: 178,97 - F: 75,49)
```

L'ottimizzazione di è fermata alla generazione 1081, dopo 400 iterazioni senza miglioramenti, riuscendo ad ottenere degli ottimi risultati se rapportati al valore target.

6 – Conclusione

Il progetto **FoodDiary-GA** è stato sviluppato per scopi didattici e rappresenta una versione semplificata di un generatore di piani alimentari. Sebbene la soluzione attuale offra una base funzionante per la generazione di diete bilanciate a livello nutrizionale, essa non tiene conto di tutte le variabili che potrebbero influenzare la personalizzazione e l'efficacia dei piani alimentari.

6.1 - Prospettive Future

L'idea futura è quella di estendere e migliorare il progetto per renderlo più completo e adattabile alle esigenze individuali. I principali miglioramenti previsti includono:

- **Controllo sulla scelta degli alimenti:** Implementare un sistema che permetta di selezionare e variare gli alimenti in base a preferenze specifiche, fornendo così maggiore flessibilità e personalizzazione nella pianificazione dei pasti.
- **Integrazione con i gusti dell'utente:** Aggiungere la capacità di integrare i gusti e le preferenze alimentari dell'utente per garantire che i piani alimentari non solo soddisfino i requisiti nutrizionali, ma siano anche gradevoli e accettabili per chi li segue.
- **Supporto per intolleranze e allergie:** Implementare funzionalità per gestire intolleranze e allergie alimentari, assicurando che i piani alimentari siano sicuri e adatti a utenti con esigenze dietetiche particolari.

- **Ottimizzazione avanzata:** Utilizzare tecniche di ottimizzazione più sofisticate e personalizzate, per migliorare ulteriormente la qualità delle soluzioni generate e adattarle meglio alle specifiche necessità dell'utente.
- **Interfaccia utente:** Sviluppare un'interfaccia utente intuitiva e interattiva che permetta agli utenti di inserire facilmente le loro preferenze e requisiti, e di visualizzare e modificare i piani alimentari in modo semplice.