



UNIVERSITÀ DI PISA

Corso di Laurea in Informatica classe L-31

TESI DI LAUREA TRIENNALE

**AlphaFold e le prospettive della
bioinformatica**

Relatore

Paolo?? Milazzo

Candidato

Ludovico Venturi

Correlatore

Mario Pirchio??

ANNO ACCADEMICO 2020/2021

Riassunto

Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe.

Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe. Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe.

Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe. Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe.

Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe.

Va posto al centro della seconda pagina e non dovrebbe superare le 20 righe.

Indice

Riassunto	1
Introduzione	4
1 Bioinformatica	5
1.1 Di cosa si occupa	5
1.2 Background filosofico	5
2 Predizione della struttura delle proteine	6
2.1 Background biologico	6
2.2 CASP	6
2.3 Prima di AlphaFold	6
3 AlphaFold	7
3.1 AlphaFold 1	7
3.2 AlphaFold 2	7
3.3 Uso	7
3.4 Visualizzazione 3D del risultato	7
4 Sperimentazione di AlphaFold	8
4.1 proteina BFG-54g????	8
4.1.1 Confronto con altri metodi	8
Conclusioni	9

Trasformare l'esperienza dell'università in qualcosa di positivo, di progressivo, che può alimentare il fuoco delle mie passioni. Fai qualcosa di specifico, renditi esperto.

Guida il lettore da 0 ad AlphaFold facendolo meravigliare davanti alla bellezza della bioinformatica, e della vita.

Medita e poi scrivi qui: non passare da fonti terze. Non perdere il flusso. Tu stai scrivendo qualcosa per te, non per il mondo. Scrivi, poi confrontati. Se ti confronti è normale che ti vedi inferiore. Come puoi invece essere inferiore a te stesso?

Ciò che conta è fare, fare, fare, mettere in pratica.

Hai scelto tu di uscire dall'informatica. Hai paura di risultare ignorante in biologia? Hai paura di esserti immischiato in un campo a te esterno e di sembrare "capiscione"?

1. Non ne sa quasi nulla nessuno dei prof 2. Non interessa loro 3. ho Mario Pirchio a cui chiedere aiuto 4. voglio uscire dall'informatica pura. Non mi fido. Non mi interessa. Qui mi interessa 5. affronta la responsabilità. Ho la responsabilità di creare la mia strada e crederci, di laurearmi per mio padre e la mia famiglia.

Mentre disegnavo ho notato che ciò che mi spingeva a migliorare il disegno era riuscire a intravedere il risultato finale in quello che stavo facendo. Non stavo tracciando una linea su un foglio. Stavo facendo piccoli passi per mettere su carta ciò che vedevo dentro di me (non nella mente, ma nel cuore).

Realizzavo una piccola parte di me al di fuori di me. E vedere che ciò che stavo creando si stava avvicinando a ciò che avevo in mente mi dava una soddisfazione immensa. E questa felicità mi spingeva tantissimo a continuare e a migliorarmi.

Voglio scrivere questo documento per realizzare una piccola parte di me all'esterno di me. L'obiettivo del disegno era realizzare un ritratto di Thich Nhat Hanh, per esprimere la mia gratitudine nei suoi confronti.

Obiettivo finale: realizzare un documento riguardante il background della bioinformatica e lo studio di AlphaFold per esprimere la mia speranza che l'informatica possa essere usata per il bene della Vita, che ci possa avvicinare ad una comprensione maggiore di essa e di quanto ogni fenomeno sia interrelato.

La tesi serve a dimostrare una ipotesi che avete elaborato dall'inizio, non a mostrare che voi sapete tutto

Introduzione

L'informatica: un potente strumento

Illustrare il mio obiettivo e la suddivisione del lavoro, dopo aver esposto la mia posizione sui rischi e le prospettive positive aperte dall'informatica.

Capitolo 1

Bioinformatica

1.1 Di cosa si occupa

Una parte importante della bioinformatica si occupa dell'utilizzo di strumenti informatici finalizzati a manipolare, archiviare e confrontare stringhe e sequenze di caratteri.

La bioinformatica tuttavia non si ferma all'analisi delle sequenze. Tra le più interessanti applicazioni bioinformatiche odierne vi sono quelle incentrate sull'analisi strutturale.

Difatti la bioinformatica pone le sue fondamenta nel campo della *structural bioinformatics*: per portare un esempio il database PDB (*Protein Data Bank*) nasce nel 1977 per archiviare coordinate atomiche e legami derivati dagli studi cristallografici sulle proteine (Bernstein et al. [1977](#)).

1.2 Background filosofico

Buttaci un po' di filosofia della scienza e di quali cambiamenti potrebbe apportare alla struttura delle rivoluzioni scientifiche. Cita Fleck in qualche modo! Trova casi di cambi di paradigma e a "riscoperte" tornate alla ribalta grazie all'informatica. Magari l'informatica è un modo, analizzando tanti dati, di contrastare i bias nella scienza?

Capitolo 2

Predizione della struttura delle proteine

L'analisi della struttura delle proteine è intrinsecamente complessa: "nessun'altra classe di molecole (piccole o grandi) esibisce una varietà di forme, dimensioni, struttura e movimenti comparabili a quelli delle proteine" (Baxevanis, Bader e Wishart [2020](#)).

2.1 Background biologico

2.2 CASP

2.3 Prima di AlphaFold

Capitolo 3

AlphaFold

3.1 AlphaFold 1

3.2 AlphaFold 2

3.3 Uso

3.4 Visualizzazione 3D del risultato

Capitolo 4

Sperimentazione di AlphaFold

4.1 proteina BFG-54g????

4.1.1 Confronto con altri metodi

Conclusioni

considerazioni sulle porte aperte dalla bioinformatica Soddisfazione
Il problema del protein folding è risolto? (No).

Bibliografia

- Baxevanis, Andreas D, Gary D Bader e David S Wishart (2020). *Bioinformatics*. John Wiley & Sons.
- Bernstein, Frances C. et al. (1977). “The protein data bank: A computer-based archival file for macromolecular structures”. In: *Journal of Molecular Biology* 112.3, pp. 535–542.