

UNIVERSITÀ DEGLI STUDI DI MILANO selezione pubblica per n. 2 posto/i di Ricercatore a tempo determinato in tenure track (RTT) per il settore concorsuale 01/B1 - Informatica, settore scientifico-disciplinare INF/01 - Informatica presso il Dipartimento di INFORMATICA "GIOVANNI DEGLI ANTONI", (avviso bando pubblicato sulla G.U. n. 41 del 21/05/2024) - Codice concorso 5550

Jessica Gliozzo
CURRICULUM VITAE

INFORMAZIONI PERSONALI

Cognome: Gliozzo

Nome: Jessica

Data di nascita: 18/02/1991

ORCID ID: 0000-0001-7629-8112 <https://orcid.org/0000-0001-7629-8112>

SCOPUS ID: 57204172389 <https://www.scopus.com/authid/detail.uri?authorId=57204172389>

GOOGLE SCHOLAR: <https://scholar.google.com/citations?user=btVEevkAAAAJ&hl=it>

Indice

1	Titoli di studio e Posizioni lavorative	3
1.1	Titoli di studio	3
1.2	Posizioni lavorative	3
1.3	Ricercatore invitato	4
1.4	Altri corsi di approfondimento seguiti	4
2	Attività di ricerca	4
3	Indici di produttività scientifica	7
4	Partecipazione a gruppi di ricerca nazionali e internazionali	8
5	Progetti di ricerca	9
6	Attività didattica e seminari	10
6.1	Attività didattica	10
6.2	Seminari	10
6.3	Attività come correlatore di tesi	10
7	Presentatore a congressi e convegni	11
8	Attività di servizio per la comunità scientifica	11
8.1	Partecipazione all'organizzazione di conferenze e workshop	11
8.2	Attività di revisione per riviste internazionali	11
8.3	Partecipazione a Società ed Associazioni scientifiche	11
9	Borse di studio e software	11
9.1	Borse di studio	11
9.2	Progettazione, sviluppo e mantenimento di librerie software	12
10	Conoscenza delle lingue	12
11	Pubblicazioni	12
11.1	Pubblicazioni su Riviste Scientifiche	12
11.2	Proceedings di conferenze	14
11.3	Abstract a conferenza (Poster presentations)	15
11.4	Tesi di Dottorato	16

1 Titoli di studio e Posizioni lavorative

1.1 Titoli di studio

- **2020 - 2024:** Dottorato di ricerca in Informatica presso Università degli Studi di Milano, giudizio *cum laude* (dissertazione 17/04/2024). Dottorato svolto in collaborazione con il *Joint Research Centre JRC of the European Commission* attraverso il programma Collaborative Doctoral Partnership (CDP).
Titolo tesi: *Patient similarity networks-based methods for multimodal data integration and clinical outcome prediction* [T1], Relatore: Prof. Giorgio Valentini, Correlatore: Prof. Elena Casiraghi e Dr. Alex Patak.
- **2013 - 2016:** Laurea Magistrale in Biotecnologie Molecolari e Bioinformatica (Classe LM-8) presso Università degli Studi di Milano, voto di Laurea: 110/110 e lode (discussione 12/12/2016).
Titolo tesi: *Network-based methods for outcome prediction in the “sample space”*, Relatore: Prof. Giorgio Valentini, Correlatore: Prof. Alberto Paccanaro and Prof. Matteo Re.
- **2010 - 2013:** Laurea Triennale in Biotecnologie Mediche (Classe L-2) presso Università degli Studi di Milano, voto di Laurea: 104/110 (discussione 21/02/2014).
Titolo tesi: *Comparazione tra approccio biochimico e imaging con CCD camera per la valutazione dell’attività luciferasica in colture cellulari*, Relatore: Prof. Luisa Ottobrini.
- **2005 - 2010:** Diploma di scuola secondaria di secondo grado indirizzo Liceo Scientifico Tecnologico presso ITIS “Augusto Righi” (Corsico, MI), voto di maturità 100/100.

1.2 Posizioni lavorative

- **Febbraio 2024 - Presente:** Assegnista di ricerca tipo B presso Dipartimento di Informatica “Giovanni Degli Antoni”, Università degli Studi di Milano.
- **Novembre 2020 - Gennaio 2024:** Dottorando in Informatica presso Dipartimento di Informatica “Giovanni Degli Antoni”, Università degli Studi di Milano.
- **Febbraio 2020 - Ottobre 2020:** Ricercatore (Contratto Co.Co.Co.) presso Unità di Neuroradiologia, IRCCS Ospedale San Raffaele, Milano.
- **Maggio 2019 - Gennaio 2020:** Borsista di ricerca sul progetto “Studio genomico mediante Next Generation Sequencing (NGS) delle patologie proliferative emopoietiche primitive della cute”, presso U.O.C. Dermatologia, FONDAZIONE IRCCS CA’ GRANDA - Ospedale Maggiore Policlinico, Milano. (R.C. 2019 – 280/02 – borsa di studio sul Tema n. 27; R.C. 2020 – 280/02 – Borsa di studio sul Tema n. 51)
- **Aprile 2018 - Marzo 2019:** Borsista di ricerca sul progetto “Analisi genomica finalizzata alla definizione di una terapia molecolare personalizzata dei linfomi T aggressivi refrattari alla chemioterapia”, presso U.O.C. Dermatologia, FONDAZIONE IRCCS CA’ GRANDA - Ospedale Maggiore Policlinico, Milano. (Atti 1653/2017 All. FINANZIAMENTO PRIVATO / EB - Tema n. 1)
- **Marzo 2017 - Febbraio 2018:** Borsista di ricerca sul progetto “Analisi genomica alla definizione di una terapia molecolare personalizzata dei linfomi T aggressivi refrattari alla chemioterapia”, presso U.O.C. Dermatologia, FONDAZIONE IRCCS CA’ GRANDA - Ospedale Maggiore Policlinico, Milano. (1383/2014-All. CONVENZIONE LUMC / EB – Tema n. 1)
- **Marzo 2015 - Febbraio 2016:** Tirocinio laurea magistrale presso AnacletoLab, Computational Biology and Bioinformatics Lab, Dipartimento di Informatica “Giovanni Degli Antoni”, Università degli Studi di Milano. Sito web del laboratorio: <https://anacletolab.di.unimi.it/>.
- **Aprile 2013 - Settembre 2013:** Tirocinio laurea triennale presso Centro di Imaging Molecolare e Cellulare (IMAGO), L.I.T.A., Segrate (MI), Università degli Studi di Milano.

1.3 Ricercatore invitato

- **16-30 Ottobre 2018 / 15-30 Novembre 2019:** Ricercatore invitato presso KIRCHER LAB - Computational Genome Biology Group, Berlin Institute of Health, Charité, Berlin, Germany. Laboratory website: <https://kircherlab.github.io/>. Attività di ricerca finanziata dal “MIUR-DAAD Joint Mobility Program” per creare collaborazioni tra gruppi di ricerca Italiani e Tedeschi.
- **Marzo 2016 - Giugno 2016:** Ricercatore invitato presso PaccanaroLab, Centre for Systems Biology, Department of Computer Science – Royal Holloway, University of London, Egham, TW20 0EX UK. Collaborazione con PaccanaroLab (www.paccanarolab.org) per condurre parte della mia tesi magistrale all'estero, sovvenzionata dall'Università degli Studi di Milano.

1.4 Altri corsi di approfondimento seguiti

- **24-30 Luglio 2022:** Partecipato alla scuola estiva “Jacob T. Schwartz International School for Scientific Research - Lipari School on Computational Life Sciences” intitolata “Artificial Intelligence in Biomedicine” (24-30 Luglio, 2022), Isola di Lipari, Italia
- **16-19 Luglio 2018:** Seguito il corso “VIII Edition - NGS data analysis and applications in diagnostic (University of Pavia)”, 16-19 Luglio 2018

2 Attività di ricerca

Le attività di ricerca della candidata sono incentrate sull'analisi di dati biomolecolari e clinici in ambito medico. Nello specifico, la candidata si occupa dell'applicazione e sviluppo di metodi di apprendimento automatico (semi-)supervisionati e supervisionati a dati biologici eterogenei (per es. dati genomici, epigenomici, trascrittomici, proteomici, clinici, immagini) con l'obiettivo di fornire strumenti e metodologie a supporto della comunità scientifica per il progresso della medicina di precisione. Questo approccio innovativo alla medicina ha come scopo ultimo il miglioramento della prevenzione, diagnosi, prognosi e trattamento delle patologie tramite l'analisi di grandi moli di dati multi-modalità atte a caratterizzare dettagliatamente ogni individuo sulla base delle sue caratteristiche individuali. Sebbene i “Big Data” necessari all'avanzamento della medicina di precisione siano oggi disponibili grazie all'avanzamento delle cosiddette tecnologie high-throughput, la loro analisi rimane complicata e pone diverse sfide che la candidata ha avuto modo di investigare nelle diverse linee di ricerca da lei seguite:

1. Integrazione di dati multi-modalità tramite reti di similarità tra pazienti

La realizzazione degli obiettivi posti dalla medicina di precisione dipende fortemente dalla possibilità di analizzare in modo efficace dati medici multi-modalità, spesso ad elevata dimensionalità. Ne sono un esempio i dati multi-omici, ormai considerati il metodo standard per una esaustiva caratterizzazione molecolare dei pazienti, che prevedono la misurazione di tutte le biomolecole che descrivono uno specifico livello biologico. Nel panorama attuale, le Reti di Similarità tra Pazienti (RSP) sono considerate un potente strumento per rappresentare le relazioni tra i profili clinici e/o molecolari di insiemi di pazienti. Le RSP hanno ottenuto risultati allo stato dell'arte nell'ambito della medicina di precisione, come la sottotipizzazione dei pazienti e la previsione degli esiti clinici, attraverso tecniche di clustering e classificazione. La stessa candidata ha proposto *P-Net* [R10] un metodo unimodale basato su RSP per la predizione di fenotipi ed esiti clinici, pubblicato su Scientific Reports, Nature Publishing. *P-Net* è un algoritmo semi-supervisionato network-based basato su kernel per grafi, capace di sfruttare informazioni sulla topologia globale e locale della RSP per assegnare un punteggio ad ogni paziente relativo alla sua probabilità di mostrare l'esito/fenotipo di interesse. Inoltre, reti di similarità tra individui sono state sfruttate per interpretare le predizioni fornite da una squadra di “gemelli digitali” umani che simulano una squadra di atleti, tramite una serie di dati multi-modalità forniti da fitness tracker ed inseriti dall'utente tramite apps [R8]. Le reti di similarità sono state usate per il calcolo di spiegazioni controfattuali, con l'obiettivo di fornire suggerimenti per il miglioramento delle performance degli atleti predette dall'intelligenza artificiale dei gemelli digitali.

La crescente disponibilità di dati multi-omici, unita alle prestazioni dei modelli basati su RSP, sottolinea la necessità di sviluppare metodi sofisticati per integrare dati multimodalità per la costruzione di RSP. In questo contesto ho effettuato un'ampia revisione della letteratura [R5] che raccoglie ed analizza criticamente le misure di similarità, i metodi per la fusione di dati multimodalità per la costruzione di RSP e le numerose tassonomie proposte nel corso degli anni. Una nuova tassonomia degli approcci di integrazione

dei dati multimodali, specificatamente disegnata per l' integrazione, costruzione e analisi di RSP e stata inoltre proposta. In questo lavoro, sono stati considerati non solo metodi bioinformatici ma anche metodi proposti nella letteratura relativa all'apprendimento automatico, che potrebbero essere potenzialmente sfruttati per costruire RSP. Questo lavoro ha permesso di individuare diversi problemi aperti nell'ambito dell'integrazione di dati multi-omici.

In primo luogo, questi dataset sono caratterizzati da una elevata dimensionalità ed un limitato numero di campioni (problema altrimenti noto come "small-sample-size problem"). Questo porta a dati caratterizzati da una elevata sparsità che aumenta i costi computazionali e mina l'affidabilità delle analisi, sia supervisionate che semi-supervisionate. Questi problemi possono essere mitigati dall'applicazione di approcci non supervisionati di selezione ed estrazione delle caratteristiche per ridurre la dimensionalità delle singole viste, ma particolare attenzione va posta nella scelta dell'algoritmo di riduzione e della dimensione finale dello spazio di riduzione. Mentre la maggior parte dei lavori trascura l'importanza di questa fase di pre-elaborazione, la candidata ha proposto un metodo di riduzione della dimensionalità che fornisce una stima affidabile della dimensionalità intrinseca di ogni sorgente di dati tramite una analisi a blocchi e permette di valutare quanto un insieme di dati sia affetto dai problemi relativi alla small-sample-size [R2, P6]. Il metodo di riduzione della dimensionalità proposto sfrutta queste informazioni per adattarsi automaticamente alle caratteristiche dei dati in modo che le sorgenti di dati affette da una maggiore quantità di rumore e ridondanza siano sottoposte a due fasi consecutive di riduzione della dimensionalità (selezione delle caratteristiche e successiva estrazione delle caratteristiche), mentre modalità con una quantità trascurabile di rumore e ridondanza siano sottoposte ad un'unica fase di estrazione delle caratteristiche.

Un altro problema aperto, non considerato dalla maggior parte dei metodi di fusione di dati multi-omici, riguarda la frequente incompletezza delle fonti di dati -omiche, dove alcuni pazienti hanno una o più fonti di dati completamente mancanti. Questi metodi sono inoltre spesso progettati e/o valutati per affrontare problemi di apprendimento specifici, tra i quali il clustering non supervisionato è il più comune. La candidata ha proposto *miss-SNF* [P5], un nuovo approccio di fusione dei dati basato su RSP indipendente dalla natura del problema (ossia supervisionato/non supervisionato). *miss-SNF* sfrutta una strategia di apprendimento non lineare di message-passing per integrare insiemi di dati incompleti e ricostruisce parzialmente le similarità mancanti sfruttando i vicini dei pazienti. La RSP integrata calcolata può essere utilizzata per eseguire analisi non supervisionate e (semi-)supervisionate.

Parte dei lavori menzionati [R5, R2, P5] è stata sviluppata nell'ambito della tesi di dottorato della candidata [T1], che si è svolta tramite la collaborazione tra Università degli Studi di Milano e Joint Research Centre della Comunità Europea.

2. Metodi per l'integrazione, classificazione e visualizzazione di reti biomolecolari

La predizione automatica della funzione delle proteine a partire da reti di grandi dimensioni è di fondamentale importanza per elucidare nuove funzioni di queste fondamentali biomolecole. Una predizione efficiente di solito coinvolge l'integrazione di diverse reti di grandi dimensioni, dove reti diverse possono essere più o meno informative per la predizione della funzione di interesse. Questo richiede una integrazione delle reti che sia orientata alla funzione da predire e che tenga conto anche del grande sbilanciamento che esiste tra le poche proteine note per avere la funzione considerata (cioè gli esempi positivi) e le restanti proteine facenti parte della rete. Una volta ottenuta una rete integrata, un ulteriore problema riguarda l'esecuzione della predizione della funzione delle proteine che tenga conto del forte sbilanciamento delle classi. In questo contesto, ho contribuito allo sviluppo di ParCosNet [R17] una versione parallela dell'algoritmo CosNet che esegue la predizione della funzione delle proteine tramite reti neurali di Hopfield che tengono conto dello sbilanciamento tra le classi e di HEMDAG [R7] che comprende una famiglia di ensemble di classificatori gerarchici che permette di correggere le predizioni fornite da "classificatori flat" per ottenere delle predizioni che rispettino la True-Path-Rule che governa le ontologie basate su grafi aciclici diretti (come l'ontologia GO classicamente usata per la classificazione delle proteine).

La visualizzazione della rete integrata ottenuta, con le rispettive predizioni, porta frequentemente alla sovrapposizione dei nodi in una nuvola di punti di difficile interpretazione per via delle dimensioni della rete stessa. AnacletoLab ha precedentemente sviluppato degli algoritmi allo stato dell'arte che permettono l'integrazione delle reti (UNIPred) e la predizione della funzione delle proteine (CosNet), ma il loro uso richiede competenze informatiche non indifferenti per poter scaricare le reti ed usare gli algoritmi. Inoltre, la visualizzazione della rete integrata e delle predizioni non è implementata. Per rendere questi utili metodi accessibili ad una platea il più ampia possibile, la candidata è stata coinvolta nello sviluppo di

UNIPred-Web [P9, R16]. UNIPred-Web è una applicazione web che rende disponibili migliaia di reti di proteine provenienti da nove organismi diversi per eseguirne l'integrazione e la successiva predizione. Questo strumento supporta una visualizzazione "vertex centric" (centrata su un nodo) che permette di focalizzarsi su una sotto-rete che comprende una proteina di interesse, evitando l'effetto nuvola di punti, visualizzando contemporaneamente le predizioni dei nodi. La grande quantità di opzioni per modificare la visualizzazione della rete permette una analisi interattiva user-friendly. Successivamente è stata introdotta una nuova visualizzazione "community-based" che sfruttando un algoritmo per l'individuazione di comunità di nodi gerarchico offre una visualizzazione della rete a diverse risoluzioni permettendo contemporaneamente lo studio di gruppi di proteine altamente connessi che possono evidenziare pathway rilevanti [R11, P7].

Il problema dell'integrazione di reti biomolecolari non è limitato esclusivamente alle reti di proteine ma si estende più in generale a tutte le biomolecole presenti in un organismo, le cui interazioni funzionali possono essere rappresentate tramite grafi. Inoltre, reti di biomolecole possono essere integrate con ulteriori informazioni di interesse medico presenti in letteratura (per esempio l'associazione con malattie e fenotipi, i farmaci noti per interagire con le biomolecole). Tra le biomolecole, le molecole ad RNA sono di particolare interesse poiché si è dimostrato che RNA, sia codificanti che non codificanti, sono associati all'insorgenza di diverse patologie (per es. tumori, malattie genetiche, cardiache, neurologiche ed infettive) e rappresentano un target interessante per lo sviluppo di nuovi farmaci (come dimostrato dal successo dei vaccini ad mRNA per le infezioni da SARS-CoV-2). Tuttavia, le informazioni relative alle relazioni di molecole ad RNA con altri entità biologiche e informazioni mediche sono sparse in diversi database. Per offrire una visione complessiva della conoscenza disponibile riguardo alle molecole ad RNA e supportarne lo studio, la candidata è stata coinvolta nello sviluppo di RNA-KG [P3, R1, P4] che rappresenta il primo knowledge graph che integra in modo standardizzato più di 50 database pubblici relativi alle molecole a RNA sfruttando ontologie biomediche per definire le semantiche necessarie all'integrazione. L'articolo relativo ad RNA-KG è in fase di pubblicazione su Scientific Data, Nature Publishing. Questo studio sta proseguendo con l'applicazione di metodi di apprendimento automatico induttivi per la predizione di entità e relazioni tra le entità nel grafo, sfruttando metodi di Graph Representation Learning per ottenere delle rappresentazioni vettoriali che tengono conto della topologia della rete [P2].

3. Predizione di mutazioni ed attività di regioni non-codificanti nel genoma umano

L'importanza delle regioni regolatorie del genoma umano, come promotori ed enhancers, è ormai nota vista la loro attività nel controllo spazio-temporale della trascrizione genica che li porta ad essere coinvolti in processi essenziali come lo sviluppo delle diverse linee cellulari/tessuti e nel controllo dinamico di risposta agli stimoli fisiologici. Inoltre, l'occorrenza di mutazioni in queste regioni del genoma è fortemente associata ad effetti patogenici. È quindi fondamentale una predizione accurata e specifica per ogni tessuto/linea cellulare dell'attività di queste regioni regolatorie, necessaria per valutare l'impatto di varianti in queste regioni. In questo contesto, la candidata ha esplorato l'uso di reti neurali feed-forward addestrate su dati epigenomici e reti convoluzionali che sfruttano dati di sequenza per la predizione dello stato di attività delle regioni regolatorie in specifiche linee cellulari [R4, P10]. Lo studio si è concentrato su come il setup sperimentale influenza le predizioni, dimostrando che (I) l'uso della selezione automatica degli iperparametri dei modelli tramite ottimizzazione bayesiana è in grado di portare a un miglioramento delle performance, (II) il ri-bilanciamento dei dati ha un notevole effetto e dovrebbe essere evitato sul test set in quanto può portare a una valutazione dei risultati eccessivamente ottimista, (III) le buone performance delle reti convoluzionali su dati di sequenza suggeriscono che anche questi dati contengono informazioni sull'attività delle regioni regolatorie e che modelli multi-modal integranti entrambi i tipi di dati potrebbero portare a un miglioramento delle performance.

La candidata ha anche contribuito a lavori relativi alla predizione di mutazioni in regioni non-codificanti del genoma umano associate a rare malattie Mendeliane. Questo problema è notoriamente caratterizzato da un forte sbilanciamento tra le classi delle varianti neutrali (negativi) e le varianti associate a patologia (positivi), dove solo poche mutazioni in queste regioni sono associate in modo affidabile a malattie Mendeliane in letteratura. Per affrontare questo problema, sono state proposte tecniche per addestrare reti neurali che tengono conto di tale sbilanciamento sfruttando tecniche di ri-bilanciamento dei mini-batch e considerando mini-batch di dimensioni tali da assicurare con alta probabilità che almeno una variante patogena sia inclusa tra gli esempi del mini-batch [R15]. Inoltre, è stato proposto un algoritmo di apprendimento automatico ottimizzato per il calcolo su sistemi di High Performance Computing che sfrutta tecniche di ri-campionamento per gestire lo sbilanciamento delle classi e un ensemble di random forests per la classificazione [P8].

4. Metodi per l'analisi di immagini immunoistochimiche

La presenza della proteina ki67 è strettamente associata alla proliferazione cellulare ed è un indicatore di aggressività tumorale la cui espressione è utile nel valutare la prognosi di diversi tumori (per es. tumore al seno, polmone, prostata). La valutazione del numero di cellule che esprimono questa proteina viene solitamente eseguita in modo visivo dai medici anatomopatologi partendo da campioni di tessuto dove le cellule che presentano tale proteina vengono evidenziate tramite apposite colorazioni. Tuttavia, il conteggio di queste cellule da immagini ad alta risoluzione richiede molto tempo per cui viene eseguito solo su alcune parti delle immagini, è operatore dipendente ed è spesso carente di termini di riproducibilità ed accuratezza. Per ovviare a questi problemi ho proposto un nuovo algoritmo per il conteggio automatico dei nuclei che esprimono ki67 [R12], metodo che ha dimostrato risultati promettenti quando comparato con i conteggi eseguiti da clinici esperti. Inoltre, risultati allo stato dell'arte per quanto riguarda la conta di cellule da immagini istopatologiche sono stati ottenuti con metodi basati su reti neurali profonde. Tuttavia le risorse computazionali richieste per interrogare e talvolta anche immagazzinare tali modelli rappresentano un ostacolo alla loro applicazione in scenari caratterizzati da risorse limitate come smartphones, IoT-devices e dispositivi datati. Ho quindi proposto l'applicazione di approcci di compressione su reti neurali pre-addestrare per eseguire la conta di cellule (ki67 e linfociti infiltranti il tumore) in tessuti di tumore al seno al fine di ridurre il consumo di risorse ed energia di questi modelli [P1] mantenendo la loro accuratezza. Sebbene l'analisi di immagini provenienti da tumore al seno sia stata scelta per proporre un caso d'uso, questo approccio è più generale e abbiamo reso disponibile anche un framework per comprimere una qualsiasi rete neurale addestrata.

5. Analisi bioinformatiche di dati di sequenziamento e clinici

La candidata è stata coinvolta in progetti che hanno richiesto l'analisi di dati di sequenziamento di DNA, RNA e clinici a supporto di gruppi medici. Durante il suo primo incarico all'U.O.C Dermatologia del Policlinico di Milano Ospedale Maggiore si è occupata dell'analisi di dati di sequenziamento genomico e del RNA di dati da pazienti affetti da neoplasia a cellule dendritiche plasmacitoidi blastiche (BPDCN) e linfoma cutaneo primitivo a cellule T CD8+ epidermotropo aggressivo. Successivamente ha collaborato con il Dipartimento di Scienze della Salute dell'Università degli Studi di Milano all'analisi di dati di sequenziamento del RNA derivanti da Leucemia linfoblastica acuta T; lavoro che ha evidenziato nuovi target terapeutici per tale patologia tumorale [R6]. Questa collaborazione continua ad oggi con l'analisi non-supervisionata e supervisionata di dati clinici e mutazioni di pazienti affetti da Leucemia Mieloide Acuta core-binding factor. Attualmente collabora anche con il Dipartimento di Scienze Biomediche e Cliniche dell'Università degli Studi di Milano supportandoli nell'identificazione di una pipeline automatica per l'analisi di genomi virali di SARS-CoV-2 necessaria per monitoraggio delle varianti del virus e la tempestiva identificazione di quelle varianti di interesse per la Sanità Pubblica in quanto possono rappresentare un rischio epidemiologico per la popolazione.

3 Indici di produttività scientifica

Nella tabella 1 vengono riportati gli indici di produttività estratti il 12 Giugno 2024 dalla banca dati Scopus e da Google Scholar.

	h-index	i10-index	Numero di citazioni
Scopus	7	4	207
Google Scholar	7	5	321

Tabella 1: Indici di produttività della candidata.

Ho pubblicato 15 articoli su riviste scientifiche internazionali a revisione paritaria e 9 articoli in atti di conferenze. La tabella 2 riporta gli articoli pubblicati (o in revisione con minor revisions nel caso di [R1]) su riviste scientifiche indicizzate in Scopus, di classe Q1 e classificate nelle aree Computer Science o Multidisciplinary. I dati sono stati estratti il 12 Giugno 2024.

Paper	Journal	Numero Citazioni	IF journal	H-index journal	Quartile
[R1]	Scientific Data, Nature publishing	-	9.8	120	Q1 (Computer Science)
[R3]	Frontiers in Bioinformatics	0	new journal	new journal	new journal
[R4]	BMC Bioinformatics	2	3.0	242	Q1 (Computer Science)
[R5]	Briefings in Bioinformatics	15	9.5	142	Q1 (Computer Science)
[R7]	Bioinformatics	2	5.8	462	Q1 (Computer Science)
[R8]	IEEE Access	122	3.9	242	Q1 (Computer Science)
[R10]	Scientific Reports, Nature publishing	8	4.6	315	Q1 (Multidisciplinary)
[R11]	PLOS ONE	5	-	435	Q1 (Multidisciplinary)
[R12]	BMC Bioinformatics	12	3.0	242	Q1 (Computer Science)
[R16]	BMC Bioinformatics	8	3.0	242	Q1 (Computer Science)
[R17]	BMC Bioinformatics	2	3.0	242	Q1 (Computer Science)

Tabella 2: Indici relativi alle pubblicazioni nelle categorie Computer Science o Multidisciplinary. Il numero di citazioni dell'articolo è stato estratto da Scopus. La colonna "IF journal" si riferisce al 2-year impact factor della rivista considerata ricavato dal sito della rivista. Le colonne "H-index journal" e "Quartile" sono state estratte da Scimago.

4 Partecipazione a gruppi di ricerca nazionali e internazionali

- Membro di AnacletoLab, Laboratorio di Biologia Computazionale e Bioinformatica del Dipartimento di Informatica, Università degli Studi di Milano <http://anacletolab.di.unimi.it/>. Nell'ambito delle attività di ricerca di AnacletoLab collaboro con diversi gruppi di ricerca nazionali ed internazionali nelle aree della Biologia Computazionale, Biologia Molecolare e Medicina, tra cui:
 - Computational Biology Group della Charite - Universitaatmedizin della Humboldt Universitat di Berlino
 - il gruppo di Biologia Computazionale del Berlin Institute of Health
 - il Computer Science dept. della Royal Holloway, University of London
 - i Jackson Laboratory for Genomic Medicine, CT, USA,
 - la Division of Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
 - Il Dipartimento di Intelligenza Artificiale dell' Università di Granada
 - Altri gruppi di ricerca nazionali sia in ambito medico (Istituto Nazionale Tumori, Ospedale S.Raffaele di Milano).
- Collaborazione ad un progetto comune con il Berlin Institute of Health (BIH, che riunisce gruppi di ricerca della von Humboldt e della Freie Universität di Berlino e del Max Delbrück Center for Molecular Medicine) per lo studio sistematico delle variazioni genetiche nelle regione regolatorie del genoma umano e del loro impatto sulle patologie genetiche e tumorali. L'obiettivo è di decodificare il "codice genetico" alla base della regolazione genica nelle regioni non codificanti del genoma, e di studiarne le alterazioni per predire l'effetto patologico delle mutazioni genetiche alla base delle patologie tumorali e genetiche.
- Collaborazione con la Charite UniversitaatMedizine di Berlino, con i Jackson Lab for Genomic Medicine (USA), con il dipartimento di Intelligenza Artificiale, Università di Granada, con il dipartimento di Environmental Genomics and Systems Biology del Lawrence Berkeley National Laboratory (USA) per lo sviluppo di metodi di Graph Representation Learning per il supporto alla ricerca di nuovi farmaci basati su RNA nell' ambito del progetto PNRR finanziato da NextGenerationEU "National Center for Gene Therapy and Drugs based on RNA Technology".
- Collaborazione con School of Applied Mathematics (EMAp), Fundação Getúlio Vargas, Rio de Janeiro relativamente allo studio e sviluppo di metodi predittivi e di integrazione di dati multi-modalità basati su reti di similarità tra pazienti.

5 Progetti di ricerca

Partecipazione ai seguenti progetti di ricerca con eventuale indicazione di collaborazioni con gruppi nazionali ed internazionali:

- Partecipazione al progetto “Developing machine learning methods for the prioritization of regulatory variants in human diseases” in collaborazione con Università degli Studi di Milano e Berlin Institute of Health (BIH) per lo sviluppo di metodi di apprendimento automatico per analizzare sistematicamente l’impatto delle varianti genetiche che ricadono nelle regioni regolatorie del genoma umano, predicendo il loro effetto patologico. Il progetto è stato finanziato dal DAAD tedesco e dal MIUR tramite il “MIUR-DAAD Joint Mobility Program”.
(Periodo: 2018-2020)
- Lavoro presso l’U.O.C. Dermatologia, FONDAZIONE IRCCS CA’ GRANDA - Ospedale Maggiore Policlinico riguardante l’analisi bioinformatica di dati genomici e di espressione del RNA provenienti da Next Generation Sequencing (NGS) piattaforma Illumina, con lo scopo di individuare marcatori diagnostici/prognostici/terapeutici in pazienti con linfomi T aggressivi refrattari alla chemioterapia.
(Periodo: 2017-2020)
- Lavoro presso l’Unità di Neuroradiologia, IRCCS Ospedale San Raffaele, Milano al progetto “Biomarker Imaging and New Challenging Approaches to assess white matter disorders in developmental age”. Il progetto prevede l’applicazione di metodi statistici per decomporre il segnale multi-esponenziale T_2 proveniente da immagini di risonanza magnetica del cervello, allo scopo di ottenere delle mappe di frazione di acqua della mielina e studiare lo sviluppo della sostanza bianca in soggetti pediatrici sani ed affetti da leucodistrofia metacromatica.
(Periodo: Maggio 2019 - Gennaio 2020)
- Collaborazione al progetto “Multicriteria Data Structures and Algorithms: from compressed to learned indexes, and beyond” (PRIN no. 2017WR7SHH) con il quale è stato finanziato il mio primo anno di PhD. In questo contesto, ho sviluppato un caso di studio che mostra come l’applicazione di metodi di compressione su reti neurali pre-addestrate per eseguire la conta cellulare in patologia digitale sia un approccio efficace per ridurre il loro consumo di risorse computazionali (cioè spazio su disco e memoria) e di energia, preservandone o anche migliorandone l’accuratezza. I modelli compatti ottenuti possono essere usati su dispositivi aventi risorse limitate (per esempio smartphones, tablets, computer datati).
(Periodo: 2020 - 2021)
- Collaborazione al progetto “AI-driven data analysis and integration for bio-medical applications” finanziato dal Piano di Sostegno alla Ricerca (PSR2021, PSR2022) dove ho contribuito con lo sviluppo di (I) un metodo di integrazione di dati multimodali basato su reti di similarità tra pazienti per la fusione di diverse fonti di dati genomiche ad alta dimensionalità che presentano pazienti completamente mancanti, (II) un metodo multi-modale semi-supervisionato basato su reti di similarità tra pazienti che permette di integrare diverse sorgenti di dati (per esempio espressione genica, metilazione, dati clinici) per la predizione di esiti clinici.
(Periodo: 2021 - 2022)
- Collaborazione al progetto “MULTI-modal Data IntegratiON: dEvelopment and validation (MULTIONE)” finanziato tramite ore di calcolo sul HPC CINECA. Ho contribuito tramite lo studio di tecniche di riduzione della dimensionalità guidate dalla dimensionalità intrinseca per l’integrazione di dati multi-modalità.
(Periodo: 2023)
- Collaborazione al progetto “Multi-OneT: MULTI-modal Data IntegratiON for Explainable and Translational results” finanziato tramite ore di calcolo sul HPC INDACO dell’Università degli Studi di Milano. Ho contribuito a questo progetto tramite (I) lo studio di tecniche per l’integrazione di dati multi-modalità attraverso reti di similarità tra pazienti, (II) lo sviluppo di un metodo di fusione di dati multi-modale basato su reti di similarità tra pazienti capace di integrare datasets le cui viste presentano pazienti completamente mancanti.
(Periodo: 2023 - Presente)

- Collaborazione con il Joint Research Centre (JRC) della Commissione Europea per lo sviluppo della mia tesi di dottorato [T1].
(Periodo: 2020 - 2024)
- Collaborazione con il Dipartimento di Scienze Biomediche e Cliniche, Università degli Studi di Milano nell'ambito del progetto "Sistema integrato di Ateneo per lo studio, il monitoraggio e il controllo delle infezioni, delle emergenze epidemiche e della resistenza ai farmaci antimicrobici (IDEA)" finanziato attraverso le Grandi Sfide di Ateneo dell'Università degli Studi di Milano. Ho contribuito all'individuazione di pipeline automatiche per l'analisi di dati di sequenziamento di genomi virali SARS-CoV-2, allo scopo di monitorare e caratterizzare i lineages virali tramite l'identificazione di mutazioni di interesse biologico e per la salute pubblica.
(Periodo: 2022 - Presente)

Altre collaborazioni:

- Collaborazione con il Dipartimento di Scienze della Salute, Università degli Studi di Milano per l'analisi non-supervisionata e supervisionata di dati di pazienti affetti da Leucemia Mieloide Acuta core-binding factor.
(Periodo: 2024 - Presente)

6 Attività didattica e seminari

6.1 Attività didattica

- A.A. 2023-2024 Attività di tutoraggio per le lezioni pratiche del corso "Bioinformatics" (Laurea Magistrale in Informatica [LM-18], Università degli Studi di Milano)
- A.A. 2022-2023 Attività di tutoraggio per le lezioni pratiche del corso "Bioinformatics" (Laurea Magistrale in Informatica [LM-18], Università degli Studi di Milano)

6.2 Seminari

- Relatore invitato al Corso di perfezionamento in "FILOGENESI APPLICATA ALLA EPIDEMIOLOGIA E SORVEGLIANZA GENOMICA DEGLI AGENTI INFETTIVI" del Dipartimento di Scienze Biomediche e Cliniche, Università degli Studi di Milano. Presenta il suo lavoro intitolato "Pipelines automatiche per l'analisi di genomi virali SARS-CoV-2" (intervento programmato il 5 Luglio 2024).
- Relatore per il corso "Bioinformatics" (Laurea Magistrale in Informatica [LM-18], Università degli Studi di Milano) negli anni accademici 2020-2021 e 2021-2022. Presenta il suo lavoro intitolato "Kernels on graphs and the P-Net algorithm".
- Relatore per il corso "Principles and Models of Perception" (Laurea Triennale in Informatica per la comunicazione digitale [L-31], Università degli Studi di Milano) nell'anno accademico 2021-2022. Presenta il suo lavoro intitolato "Kernels on graphs and the P-Net algorithm".

6.3 Attività come correlatore di tesi

Correlatore delle seguenti tesi di Laurea Triennale:

- A.A. 2020/21 - Palladino Elisa, "Similarity measures for patient-networks based on clinical and genomic data" per il corso di laurea in Informatica per la Comunicazione Digitale [L-31]
- A.A. 2020/21 - Gnuva Mirco, "Integrazione di dati multimodali tramite fattorizzazione di matrici" per il corso di laurea in Informatica [L-31]
- A.A. 2019/20 - Raimondi Davide, "Classificazione del rischio di pazienti COVID-19 tramite uno score radiologico di esperti e uno score radiologico ottenuto tramite AI" per il corso di laurea in Informatica per la Comunicazione Digitale [L-31]

- A.A. 2019/20 - Minerva Gabriele Donato, “Imputazione di dati mancanti in ambito medico: uno studio su dati di pazienti COVID-19” per il corso di laurea in Informatica per la Comunicazione Digitale [L-31]

Correlatore delle seguenti tesi di Laurea Magistrale:

- A.A. 2022/23 - Mandredi Sara, “Sviluppo di una pipeline bioinformatica automatizzata per l’analisi di genomi virali SARS-CoV-2” per il corso di laurea in Informatica [LM-18]
- A.A. 2021/22 - Grimi Riccardo, “Deep Neural Network Multimodali per la predizione di regioni regolatorie nel genoma umano” per il corso di laurea in Informatica [LM-18]

7 Presentatore a congressi e convegni

- Presentatore alla conferenza “BIOINFORMATICS 2023 - 14th International Conference on Bioinformatics Models, Methods and Algorithms”. Ho presentato il mio lavoro “Patient Similarity Networks Integration for Partial Multimodal Datasets”.

8 Attività di servizio per la comunità scientifica

8.1 Partecipazione all’organizzazione di conferenze e workshop

Attualmente membro del Comitato Scientifico di Programma del workshop “SAI-Care 2024: Integrating Symbiotic AI in Biomedical Informatics, A Workshop on Advanced SAI-based applications in Healthcare”.

8.2 Attività di revisione per riviste internazionali

Ho svolto attività di revisione per le seguenti riviste internazionali:

- Revisore di un paper per il giornale Nature Scientific Reports (Sci. Rep.) a Novembre 2022.
- Revisore di un paper per il giornale Nature Communications (Nat. Commun.) ad Aprile 2024.

8.3 Partecipazione a Società ed Associazioni scientifiche

Membro di BITS - Società Italiana di Bioinformatica.

9 Borse di studio e software

9.1 Borse di studio

- **2015:** Vincitore della borsa di studio “BANDO PER BORSE DI STUDIO ALL’ESTERO AI FINI DELLA PREDISPOSIZIONE DELLA TESI DI LAUREA MAGISTRALE ANNO ACCADEMICO 2015/2016 I EDIZIONE”, sovvenzionato dall’ Università degli Studi di Milano.
- **2017:** Vincitore della borsa di studio di 1 anno sul progetto “Analisi genomica alla definizione di una terapia molecolare personalizzata dei linfomi T aggressivi refrattari alla chemioterapia” (1383/2014 – All.Convenzione LUMC/EB – Tema n.1) sovvenzionata da U.O.C. Dermatologia - Fondazione IRCCS Ca’ Granda – Ospedale Maggiore Policlinico (Milan, Italy).
- **2018:** Vincitore della borsa di studio di 1 anno sul progetto “Analisi genomica finalizzata alla definizione di una terapia molecolare personalizzata dei linfomi T aggressivi refrattari alla chemioterapia” (Atti 1653/2017 – All.Finanziamento privato/EB – Tema n.1) sovvenzionata da U.O.C. Dermatologia - Fondazione IRCCS Ca’ Granda – Ospedale Maggiore Policlinico (Milan, Italy).
- **2019:** Vincitore della borsa di studio di 1 anno sul progetto “Studio genomico mediante Next Generation Sequencing (NGS) delle patologie proliferative emopoietiche primitive della cute” (R.C.2019 – 280/02 – Borsa di Studio sul tema n. 27) sovvenzionata da U.O.C. Dermatologia - Fondazione IRCCS Ca’ Granda – Ospedale Maggiore Policlinico (Milan, Italy).

- **2020:** Vincitore della borsa di studio “Studio genomico mediante Next Generation Sequencing (NGS) delle patologie proliferative emopoietiche primitive della cute” (R.C. 2020 – 280/02 – Borsa di studio sul Tema n. 51) sovvenzionata da U.O.C. Dermatologia - Fondazione IRCCS Ca’ Granda – Ospedale Maggiore Policlinico (Milan, Italy).
- **2020:** Vincitore di una borsa di dottorato sovvenzionata da Università degli Studi di Milano e Joint Research Centre (JRC-Ispra) della Commissione Europea.

9.2 Progettazione, sviluppo e mantenimento di librerie software

Le seguenti librerie software sono rese disponibili alla comunità scientifica tramite repository pubbliche.

- Sviluppatore della libreria software *miss-SNF*.
Libreria R disponibile su <https://github.com/GliozzoJ/missSNF>.
- Sviluppatore della libreria software *P-Net* (Patient-Net): Network-based ranking of patients with respect to a given phenotype/outcome.
Libreria R disponibile su <https://github.com/GliozzoJ/P-Net>.
- Sviluppatore della repository *pathonet_compression*: Reducing the Complexity of Deep Learning Models for Medical Applications in Resource-limited Contexts.
Repository Python disponibile su https://github.com/GliozzoJ/pathonet_compression.
- Collaboratore alla repository *PathoNet*: deep neural network backend for evaluation of Ki-67 and tumor-infiltrating lymphocytes in breast cancer.
Repository Python disponibile su <https://github.com/SHIDCenter/PathoNet>.
- Maintainer della libreria *NetInt*: Methods for Unweighted and Weighted Network Integration.
Libreria R disponibile su CRAN <https://cran.r-project.org/web/packages/NetInt/index.html>.
- Collaboratore alla libreria *mclustcomp*: Measures for Comparing Clusters.
Libreria R disponibile su <https://github.com/kisungyou/mclustcomp>.

10 Conoscenza delle lingue

Italiano: Madrelingua

Inglese: Fluente (Conoscenza professionale completa)

11 Pubblicazioni

11.1 Pubblicazioni su Riviste Scientifiche

- [R1] Emanuele Cavalleri, Alberto Cabri, Mauricio Soto-Gomez, Sara Bonfitto, Paolo Perlasca, Jessica Gliozzo, Tiffany J Callahan, Justin Reese, NR Peter, Elena Casiraghi, Giorgio Valentini e Marco Mesiti. «An ontology-based knowledge graph for representing interactions involving RNA molecules». In: *Scientific Data, Nature Publishing* (2024). [Accepted].
- [R2] Jessica Gliozzo, Valentina Guarino, Arturo Bonometti, Alberto Cabri, Emanuele Cavalleri, Mauricio Soto-Gomez, Justin Reese, Peter N Robinson, Marco Mesiti, Giorgio Valentini e Elena Casiraghi. «Intrinsic-Dimension analysis for guiding dimensionality reduction and data-fusion in multi-omics data processing». In: *Artificial Intelligence In Medicine* (2024). [Second step of revision]. DOI: 10.1101/2024.01.23.576822. URL: <https://www.biorxiv.org/content/early/2024/01/25/2024.01.23.576822>.
- [R3] Giorgio Valentini, Dario Malchiodi, Jessica Gliozzo, Marco Mesiti, Mauricio Soto-Gomez, Alberto Cabri, Justin Reese, Elena Casiraghi e Peter N. Robinson. «The promises of large language models for protein design and modeling». In: *Frontiers in Bioinformatics* 3 (23 nov. 2023). ISSN: 2673-7647. DOI: 10.3389/fbinf.2023.1304099. URL: <https://www.frontiersin.org/articles/10.3389/fbinf.2023.1304099>.

- [R4] Luca Cappelletti, Alessandro Petrini, Jessica Gliozzo, Elena Casiraghi, Max Schubach, Martin Kircher e Giorgio Valentini. «Boosting tissue-specific prediction of active cis-regulatory regions through deep learning and Bayesian optimization techniques». In: *BMC Bioinformatics* 23.2 (12 dic. 2022), p. 154. DOI: 10.1186/s12859-022-04582-5. URL: <https://doi.org/10.1186/s12859-022-04582-5>.
- [R5] Jessica Gliozzo, Marco Mesiti, Marco Notaro, Alessandro Petrini, Alex Patak, Antonio Puertas-Gallardo, Alberto Paccanaro, Giorgio Valentini e Elena Casiraghi. «Heterogeneous data integration methods for patient similarity networks». In: *Briefings in Bioinformatics* 23.4 (13 giu. 2022), bbac207. ISSN: 1477-4054. DOI: 10.1093/bib/bbac207. eprint: <https://academic.oup.com/bib/article-pdf/23/4/bbac207/45016588/bbac207.pdf>. URL: <https://doi.org/10.1093/bib/bbac207>.
- [R6] Adriana Cassaro, Giovanni Grillo, Marco Notaro, Jessica Gliozzo, Ilaria Esposito, Gianluigi Reda, Alessandra Trojani, Giorgio Valentini, Barbara Di Camillo, Roberto Cairoli e Alessandro Beghini. «FZD6 triggers Wnt–signalling driven by WNT10BIVS1 expression and highlights new targets in T-cell acute lymphoblastic leukemia». In: *Hematological Oncology* 39.3 (26 gen. 2021), pp. 364–379. DOI: <https://doi.org/10.1002/hon.2840>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/hon.2840>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hon.2840>.
- [R7] Marco Notaro, Marco Frasca, Alessandro Petrini, Jessica Gliozzo, Elena Casiraghi, Peter N Robinson e Giorgio Valentini. «HEMDAG: a family of modular and scalable hierarchical ensemble methods to improve Gene Ontology term prediction». In: *Bioinformatics* 37.23 (8 lug. 2021), pp. 4526–4533. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btab485. eprint: <https://academic.oup.com/bioinformatics/article-pdf/37/23/4526/50579686/btab485.pdf>. URL: <https://doi.org/10.1093/bioinformatics/btab485>.
- [R8] Barbara Rita Barricelli, Elena Casiraghi, Jessica Gliozzo, Alessandro Petrini e Stefano Valtolina. «Human Digital Twin for Fitness Management». In: *IEEE Access* 8 (4 feb. 2020), pp. 26637–26664. DOI: 10.1109/ACCESS.2020.2971576.
- [R9] Arturo Bonometti, Jessica Gliozzo, Chiara Moltrasio, Filippo Bagnoli e Emilio Berti. «Cutaneous-group histiocytoses associated with myeloid malignancies: A systematic review of 102 cases». In: *Australasian Journal of Dermatology* 62.2 (30 ott. 2020), e162–e169. DOI: <https://doi.org/10.1111/ajd.13491>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajd.13491>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/ajd.13491>.
- [R10] Jessica Gliozzo, Paolo Perlasca, Marco Mesiti, Elena Casiraghi, Viviana Vallacchi, Elisabetta Vergani, Marco Frasca, Giuliano Grossi, Alessandro Petrini, Matteo Re, Alberto Paccanaro e Giorgio Valentini. «Network modeling of patients’ biomolecular profiles for clinical phenotype/outcome prediction». In: *Scientific Reports, Nature Publishing* 10.1 (27 feb. 2020), p. 3612. ISSN: 2045-2322. DOI: 10.1038/s41598-020-60235-8. URL: <https://doi.org/10.1038/s41598-020-60235-8>.
- [R11] Paolo Perlasca, Marco Frasca, Cheick Tidiane Ba, Jessica Gliozzo, Marco Notaro, Mario Pennacchioni, Giorgio Valentini e Marco Mesiti. «Multi-resolution visualization and analysis of biomolecular networks through hierarchical community detection and web-based graphical tools». In: *PLOS ONE* 15.12 (22 dic. 2020), pp. 1–28. DOI: 10.1371/journal.pone.0244241. URL: <https://doi.org/10.1371/journal.pone.0244241>.
- [R12] Barbara Rita Barricelli, Elena Casiraghi, Jessica Gliozzo, Veronica Huber, Biagio Eugenio Leone, Alessandro Rizzi e Barbara Vergani. «ki67 nuclei detection and ki67-index estimation: a novel automatic approach based on human vision modeling». In: *BMC Bioinformatics* 20.1 (27 dic. 2019), p. 733. DOI: 10.1186/s12859-019-3285-4. URL: <https://doi.org/10.1186/s12859-019-3285-4>.
- [R13] Armando N. Bastidas Torres, Davy Cats, Hailiang Mei, Daniele Fanoni, Jessica Gliozzo, Laura Corti, Marco Paulli, Maarten H. Vermeer, Rein Willemze, Emilio Berti e Cornelis P. Tensen. «Whole-genome analysis uncovers recurrent IKZF1 inactivation and aberrant cell adhesion in blastic plasmacytoid dendritic cell neoplasm». In: *Genes, Chromosomes and Cancer* 59.5 (17 dic. 2019), pp. 295–308. DOI: <https://doi.org/10.1002/gcc.22831>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/gcc.22831>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/gcc.22831>.
- [R14] Arturo Bonometti, Jessica Gliozzo, Chiara Moltrasio, Filippo Bagnoli, Luigia Venegoni, Emanuela Passoni, Marco Paulli e Emilio Berti. «Disfiguring Nodular Cephalic Xanthoma Disseminatum: An Exceptional Variant of a Forgotten Entity». In: *Acta Dermato-Venereologica* 99.4 (4 feb. 2019), pp. 450–451. DOI: 10.2340/00015555-3111. URL: <https://medicaljournalssweden.se/actadv/article/view/3161>.

- [R15] Luca Cappelletti, Jessica Gliozzo, Alessandro Petrini e Giorgio Valentini. «Training Neural Networks with Balanced Mini-batch to Improve the Prediction of Pathogenic Genomic Variants in Mendelian Diseases». In: *Sensors & Transducers* 234.6 (30 giu. 2019), pp. 16–21. ISSN: 2306-8515. URL: https://www.sensorsportal.com/HTML/DIGEST/P_3087.htm.
- [R16] Paolo Perlasca, Marco Frasca, Cheick Tidiane Ba, Marco Notaro, Alessandro Petrini, Elena Casiraghi, Giuliano Grossi, Jessica Gliozzo, Giorgio Valentini e Marco Mesiti. «UNIPred-Web: a web tool for the integration and visualization of biomolecular networks for protein function prediction». In: *BMC Bioinformatics* 20.1 (14 ago. 2019), p. 422. ISSN: 1471-2105. DOI: 10.1186/s12859-019-2959-2. URL: <https://doi.org/10.1186/s12859-019-2959-2>.
- [R17] Marco Frasca, Giuliano Grossi, Jessica Gliozzo, Marco Mesiti, Marco Notaro, Paolo Perlasca, Alessandro Petrini e Giorgio Valentini. «A GPU-based algorithm for fast node label learning in large and unbalanced biomolecular networks». In: *BMC Bioinformatics* 19.10 (15 ott. 2018), p. 353. ISSN: 1471-2105. DOI: 10.1186/s12859-018-2301-4. URL: <https://doi.org/10.1186/s12859-018-2301-4>.

11.2 Proceedings di conferenze

- [P1] Jessica Gliozzo, Giosuè Marinò, Arturo Bonometti, Marco Frasca e Dario Malchiodi. «Resource-Limited Automated Ki67 Index Estimation in Breast Cancer». In: *Proceedings of the 2023 10th International Conference on Bioinformatics Research and Applications, Barcelona, Spain, September 22-24, 2023*. IC-BRA '23. New York, NY, USA: Association for Computing Machinery, 27 feb. 2024, pp. 165–172. DOI: 10.1145/3632047.3632072. URL: <https://doi.org/10.1145/3632047.3632072>.
- [P2] Francesco Torgano, Emanuele Cavalleri, Jessica Gliozzo, Emanuele Saitto, Marco Mesiti, Elena Casiraghi e Giorgio Valentini. «RNA Knowledge Graph analysis via embedding methods». In: *BIOMEDICINE 2024 - 4th International Conference on Biology and Biomedicine, Heraklion, Crete Island, Greece. July 19-22, 2024*. [Accepted]. Lug. 2024.
- [P3] Emanuele Cavalleri, Sara Bonfitto, Alberto Cabri, Jessica Gliozzo, Paolo Perlasca, Mauricio Soto-Gomez, Gabriella Trucco, Elena Casiraghi, Giorgio Valentini e Marco Mesiti. «A Meta-Graph for the Construction of an RNA-Centered Knowledge Graph». In: *Bioinformatics and Biomedical Engineering, 10th International Work-Conference, IWBBIO 2023, Meloneras, Gran Canaria, Spain, July 12–14, 2023*. A cura di Ignacio Rojas, Olga Valenzuela, Fernando Rojas Ruiz, Luis Javier Herrera e Francisco Ortuño. Cham: Springer Nature Switzerland, 29 giu. 2023, pp. 165–180. ISBN: 978-3-031-34953-9. DOI: 10.1007/978-3-031-34953-9_13. URL: https://doi.org/10.1007/978-3-031-34953-9_13.
- [P4] Emanuele Cavalleri, Sara Bonfitto, Alberto Cabri, Jessica Gliozzo, Paolo Perlasca, Mauricio Soto-Gomez, Gabriella Trucco, Elena Casiraghi, Giorgio Valentini, Marco Mesiti et al. «Towards the Construction of an RNA-based Knowledge Graph». In: *CEUR WORKSHOP PROCEEDINGS, SEBD 2023 : 31st Symposium of Advanced Database Systems, 2 July 2023 - 5 July 2023*. Vol. 3478. CEUR-Ws. Galzignano Terme, Italy, lug. 2023, pp. 173–180. URL: <https://ceur-ws.org/Vol-3478/paper30.pdf>.
- [P5] Jessica Gliozzo., Alex Patak., Antonio Puertas-Gallardo., Elena Casiraghi. e Giorgio Valentini. «Patient Similarity Networks Integration for Partial Multimodal Datasets». In: *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023), Lisbona, Portugal, February 16–18, 2023 - BIOINFORMATICS*. INSTICC. SciTePress, 2023, pp. 228–234. ISBN: 978-989-758-631-6. DOI: 10.5220/0011725500003414.
- [P6] Valentina Guarino., Jessica Gliozzo., Ferdinando Clarelli., Béatrice Pignolet., Kaalindi Misra., Elisabetta Mascia., Giordano Antonino., Silvia Santoro., Laura Ferré., Miryam Cannizzaro., Melissa Sorosina., Roland Liblau., Massimo Filippi., Ettore Mosca., Federica Esposito., Giorgio Valentini. e Elena Casiraghi. «Intrinsic-Dimension Analysis for Guiding Dimensionality Reduction in Multi-Omics Data». In: *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023), Lisbona, Portugal, February 16–18, 2023 - BIOINFORMATICS*. INSTICC. SciTePress, 2023, pp. 243–251. ISBN: 978-989-758-631-6. DOI: 10.5220/0011775200003414.

- [P7] Paolo Perlasca, Marco Frasca, Cheick Tidiane Ba, Jessica Gliozzo, Marco Notaro, Mario Pennacchioni, Giorgio Valentini e Marco Mesiti. «Integration and Visual Analysis of Biomolecular Networks Through UNIPred-Web». In: *Current Trends in Web Engineering ICWE 2022. Communications in Computer and Information Science, International Workshops, BECS, SWEET and WALS, Bari, Italy, July 5–8, 2022*. A cura di Giuseppe Agapito, Anna Bernasconi, Cinzia Cappiello, Hasan Ali Khattak, InYoung Ko, Giuseppe Loseto, Michael Mrissa, Luca Nanni, Pietro Pinoli, Azzurra Ragone, Michele Ruta, Floriano Scioscia e Abhishek Srivastava. Vol. 1668. Cham: Springer Nature Switzerland, 2 feb. 2023, pp. 192–197. ISBN: 978-3-031-25380-5. DOI: https://doi.org/10.1007/978-3-031-25380-5_15.
- [P8] Alessandro Petrini, Marco Notaro, Jessica Gliozzo, Tiziana Castrignanò, Peter N. Robinson, Elena Casiraghi e Giorgio Valentini. «ParSMURF-NG: A Machine Learning High Performance Computing System for the Analysis of Imbalanced Big Omics Data». In: *Artificial Intelligence Applications and Innovations. AIAI 2022 IFIP WG 12.5 International Workshops. AIAI 2022. IFIP Advances in Information and Communication Technology. Hersonissos, Crete, Greece, June 17–20, 2022*. A cura di Ilias Maglogiannis, Lazaros Iliadis, John Macintyre e Paulo Cortez. Vol. 652. Cham: Springer International Publishing, 10 giu. 2022, pp. 424–435. ISBN: 978-3-031-08341-9. DOI: https://doi.org/10.1007/978-3-031-08341-9_34.
- [P9] Cheick Tidiane Ba, Elena Casiraghi, Marco Frasca, Jessica Gliozzo, Giuliano Grossi, Marco Mesiti, Marco Notaro, Paolo Perlasca, Alessandro Petrini, Matteo Re e Giorgio Valentini. «A Graphical Tool for the Exploration and Visual Analysis of Biomolecular Networks». In: *Computational Intelligence Methods for Bioinformatics and Biostatistics. CIBB 2018. Lecture Notes in Computer Science. 15th International Meeting, CIBB 2018, Caparica, Portugal, September 6–8, 2018*. A cura di Maria Raposo, Paulo Ribeiro, Susana Sérgio, Antonino Staiano e Angelo Ciaranella. Vol. 11925. Cham: Springer International Publishing, 23 gen. 2020, pp. 88–98. ISBN: 978-3-030-34585-3. DOI: 10.1007/978-3-030-34585-3_8.
- [P10] Luca Cappelletti, Alessandro Petrini, Jessica Gliozzo, Elena Casiraghi, Max Schubach, Martin Kircher e Giorgio Valentini. «Bayesian Optimization Improves Tissue-Specific Prediction of Active Regulatory Regions with Deep Neural Networks». In: *Bioinformatics and Biomedical Engineering, IWBBIO 2020. Lecture Notes in Computer Science, 8th International Work-Conference, IWBBIO 2020, Granada, Spain, May 6–8, 2020*. A cura di Ignacio Rojas, Olga Valenzuela, Fernando Rojas, Luis Javier Herrera e Francisco Ortuño. Vol. 12108. Cham: Springer International Publishing, 30 apr. 2020, pp. 600–612. ISBN: 978-3-030-45385-5. DOI: 10.1007/978-3-030-45385-5_54.

11.3 Abstract a conferenza (Poster presentations)

- [A1] Mirco Gnuva, Jessica Gliozzo, Alberto Paccanaro, Giorgio Valentini e Elena Casiraghi. «Comparison of early integration approaches for cancer survival prediction». In: *BITS 2022, 18th Annual Meeting of the Bioinformatics Italian Society, Verona, Italy* (giu. 2022).
- [A2] Silvia Alberti-Violetti, Daniele Fanoni, Chiara Moltrasio, Jessica Gliozzo, Luigia Venegoni, Valentina Merlo, Giorgia Saporiti, Francesco Onida e Emilio Berti. «Blastic Plasmacytoid Dendritic Cell Neoplasm: clinic-pathologic and molecular data from a single medical center». In: *24th World Congress of Dermatology (WCD), 10-15 June 2019, Milan, Italy* (giu. 2019). URL: <https://www.wcd2019milan-dl.org/abstract-book/documents/abstracts/39-skin-cancer/blastoc-plasmocytoid-dendritic-cells-neoplasm-4804.pdf>.
- [A3] Arturo Bonometti, Chiara Moltrasio, Jessica Gliozzo, Filippo Bagnoli, Gianluca Nazzaro, Silvia Alberti-Violetti, Emanuela De Juli, Emanuela Passoni, Marco Paulli e Emilio Berti. «Myeloid Leukemia and Cutaneous Histiocytosis: fortuitous encounter?» In: *24th World Congress of Dermatology (WCD), 10-15 June 2019, Milan, Italy* (giu. 2019). URL: <https://www.wcd2019milan-dl.org/abstract-book/documents/abstracts/10-dermatopathology/myeloid-leukemia-and-cutaneous-histiocytosis-3840.pdf>.
- [A4] Arturo Bonometti, Chiara Moltrasio, Jessica Gliozzo, Filippo Bagnoli, Emanuela Passoni, Gianluca Nazzaro, Silvia Alberti-Violetti, Emanuela De Juli, Marco Paulli e Emilio Berti. «Uncommon variants of Non-Langerhans Cell Histiocytosis». In: *24th World Congress of Dermatology (WCD), 10-15 June 2019, Milan, Italy* (giu. 2019). URL: <https://www.wcd2019milan-dl.org/abstract-book/documents/abstracts/10-dermatopathology/uncommon-variants-of-non-langerhans-3772.pdf>.

- [A5] Sebastiano Recalcati, Chiara Moltrasio, Gianluca Nazzaro, Emanuela Passoni, Jessica Gliozzo, Emilio Berti e Simona Muratori. «Effects of Polydeoxyribonucleotide in the treatment of Scleroderma». In: *24th World Congress of Dermatology (WCD), 10-15 June 2019, Milan, Italy* (giu. 2019). URL: <https://www.wcd2019milan-dl.org/abstract-book/documents/abstracts/07-autoimmune-connective-tissue-diseases/effects-of-polydeoxyribonucleotide-in-the-2428.pdf>.
- [A6] Silvia Alberti-Violetti, Mirco Virgilio Pozzi, Daniele Fanoni, Chiara Moltrasio, Jessica Gliozzo, Luigia Venegoni, Valentina Merlo e Emilio Berti. «Indolent cytotoxic cutaneous lymphomas with clinical and histological features of atypical lymphoid proliferation not otherwise specified.» In: *European Journal of Cancer, EORTC Cutaneous Lymphoma Task Force Meeting 2018 abstract book* 101 (set. 2018), S22–S23. DOI: <https://doi.org/10.1016/j.ejca.2018.07.265>.
- [A7] Arturo Bonometti, Jessica Gliozzo, Chiara Moltrasio, Bagnoli Filippo, Emanuela DeJuli, Emanuela Passoni e Emilio Berti. «RETICULOHISTIOCYTOSIS, GENERALIZED ERUPTIVE HISTIOCYTOSIS AND MYELOID NEOPLASM: A SYSTEMATIC REVIEW». In: *Abstracts from the 34th Annual Meeting of the Histiocyte Society Lisbon, Portugal, October 22–23, 2018, Pediatric Blood & Cancer* 66.S1 (13 nov. 2018), e27548. DOI: <https://doi.org/10.1002/pbc.27548>.
- [A8] Jessica Gliozzo, Paolo Perlasca, Marco Mesiti, Juan Caceres Silva, Alessandro Petrini, Elena Casiraghi, Marco Frasca, Giuliano Grossi, Matteo Re, Alberto Paccanaro e Giorgio Valentini. «Patients’ networks for clinical phenotype/outcome prediction». In: *Grand BIMS Opening Symposium - 11th Berlin (Late) Summer Meeting, 25-27 October 2018, Berlin, Germany*. Ott. 2018.
- [A9] Jessica Gliozzo, Marco Notaro, Alessandro Petrini, Paolo Perlasca, Marco Mesiti, Elena Casiraghi, Marco Frasca, Giuliano Grossi, Matteo Re, Alberto Paccanaro e Giorgio Valentini. «Modeling biomolecular profiles in a graph-structured sample space for clinical outcome prediction with melanoma and ovarian cancer patients». In: *BITS 2017, Bioinformatics Italian Society Meeting, Cagliari, Italy. 5-7 July 2017* (5 lug. 2017).
- [A10] Paolo Perlasca, Marco Mesiti, Marco Notaro, Alessandro Petrini, Jessica Gliozzo, Giorgio Valentini e Marco Frasca. «A Web Graphical Tool for the Integration of Unbalanced Biomolecular Networks». In: *BITS 2017, Bioinformatics Italian Society Meeting, Cagliari, Italy. 5-7 July 2017* (lug. 2017).
- [A11] Alessandro Petrini, Marco Notaro, Jessica Gliozzo, Giorgio Valentini, Giuliano Grossi e Marco Frasca. «Speeding up node label learning in unbalanced biomolecular networks through a parallel and sparse GPU – based Hopfield model». In: *BITS 2017, Bioinformatics Italian Society Meeting, Cagliari, Italy. 5-7 July 2017* (5 lug. 2017).

11.4 Tesi di Dottorato

- [T1] Jessica Gliozzo. «PATIENT SIMILARITY NETWORKS-BASED METHODS FOR MULTIMODAL DATA INTEGRATION AND CLINICAL OUTCOME PREDICTION». In: *Dipartimento di Informatica Giovanni Degli Antoni, Università degli Studi di Milano* (17 apr. 2024). A cura di A. Patak ; coordinator: R. Sassi. supervisor: G. Valentini ; co-supervisor: E. Casiraghi. PhD Thesis. 36. ciclo, Academic Year 2022/2023. URL: <https://hdl.handle.net/2434/1040030>.

Data: 19 giugno 2024

Luogo: Cesano Boscone (MI)