

ANNOTAZIONE AUTOMATICA DELLE NOTE CLINICHE

uno strumento per migliorare il monitoraggio del percorso diagnostico e terapeutico nel cancro al seno: una prima esperienza

di E. Ervas¹, P. Berchiello², G. Rizzo³, J. Francois³, P. Giraud¹, G. Grilz⁴, R. Russo⁵,
G. Zanetta⁴, A. Ricotti⁵

RAZIONALE

La maggior parte delle Aziende Sanitarie si appoggia a sistemi di analisi che permettono di avere una visione dei dati dell'organizzazione per generare conoscenza utile nei processi decisionali, stimolare un rapido cambiamento ed eliminare le inefficienze¹.

Tradizionalmente per l'analisi dei percorsi diagnostico-terapeutici si utilizzano i flussi di rendicontazione amministrativi con la consapevolezza che presentano alcuni bias e lacune al loro interno, essendo stati creati per un altro scopo².

Molto più pertinente sarebbe l'**utilizzo dei dati clinici**, che spesso sono disponibili in formato non strutturato e quindi con evidenti limiti alla fruizione del dato stesso per la generazione di evidenze utili agli interlocutori clinici, scientifici e decisionali.

L'approccio tradizionale ha visto negli anni molti tentativi di strutturazione del dato clinico dando origine ad un elevato numero di cartelle cliniche informatiche eterogenee per la raccolta di dati clinici utili all'analisi di processi specifici, come ad esempio gli indicatori di percorso. La raccolta dei dati su più sistemi e la ricerca manuale di evidenze non strutturate nella documentazione del paziente, costituiscono uno dei maggiori distrattori del rapporto terapeutico³ con una conseguente duplicazione delle registrazioni ed un grande impiego di risorse, per via dell'approccio creato con valenza interna all'azienda e difficilmente ripetibile.

OBIETTIVO

Obiettivo del progetto è stato colmare le lacune informative relative all'analisi dei percorsi dei pazienti andando a integrare i dati provenienti da fonti strutturate e non strutturate e, in relazione a queste ultime, l'analisi di un corpus di referti clinici mediante strumenti di intelligenza artificiale per l'estrazione automatizzata di informazioni utili per la creazione di conoscenza, indispensabile per il governo dei processi.

METODI

Sono stati annotati, da parte di un clinico esperto di dominio (metodo supervisionato), circa 50 referti del Gruppo Interdisciplinare Cure (GIC) relativi alla patologia del Tumore della Mammella, identificando alcune **etichette** ritenute portatrici di informazioni utili per la conoscenza del percorso di cura del paziente. Le etichette sono elencate nella *Tabella 1* riportata in seguito.

Il referto del Gruppo Interdisciplinare di Cure delle Oncologie aderenti alla Rete Oncologica Piemonte e Valle d'Aosta riassume il percorso del paziente comprendendo l'iter diagnostico, l'intervento chirurgico e la terapia. Al fine di poter estrarre le etichette, è stato impostato un protocollo di annotazione con l'intento di definire le regole per l'annotazione manuale da parte dell'esperto di dominio, utilizzando un supporto informatico sviluppato ad hoc denominato cruscotto annotatore. Utilizzando tecniche di **intelligenza artificiale** è stato addestrato un algoritmo sul campione di 50 referti annotati manualmente.

¹Innovo s.n.c.; ²Università degli Studi di Torino - Dipartimento di Scienze Cliniche e Biologiche; ³LINKS Foundation; ⁴Piccola Casa della Divina Provvidenza - Ospedale Cottolengo; ⁵Università degli Studi di Torino - Dipartimento di Scienze della Sanità Pubblica e Pediatrica

La peculiarità dell'**algoritmo** risiede nella capacità di comprensione del referto e valutazione del peso di ciascuna parola per la determinazione dell'etichetta, che si contrappone a tecniche puramente deduttive basate sul concetto di "cerca una parola specifica nel testo". Successivamente, l'algoritmo così addestrato è stato utilizzato per analizzare 347 referti relativi all'anno 2018, popolando un database contenente il corredo di etichette ritrovate nel documento di ciascun paziente facente parte della coorte, effettuando un collegamento con gli eventi amministrativi estratti dai flussi, i quali sono stati precedentemente classificati in macroeventi per migliorare il processo di analisi.

Tali dati sono stati inseriti all'interno di un cruscotto utile all'analisi dei risultati.

RISULTATI

L'analisi delle aree di testo annotate dall'esperto di dominio ha evidenziato una **elevata eterogeneità** nella modalità di stesura del referto, nonostante questi seguano un layout predefinito. Inoltre, essendo il referto GIC la collezione degli eventi occorsi al paziente durante il percorso di cura, risulta essere un documento molto completo, tuttavia ricco di ripetizioni che necessitano di essere disambiguate per verificare se sono tutte portatrici di nuove informazioni: la disambiguazione deve ad esempio distinguere un doppio esame nella stessa sede o esami su sedi diverse e deve distinguere le recidive.

Dall'analisi dei 50 referti supervisionati manualmente dall'esperto di dominio l'algoritmo ha restituito un **valore predittivo positivo** (VPP) media pesata per il numero delle occorrenze pari a **0,443** e una **sensibilità** (SE) media pesata di **0,767**; sebbene ci sia una notevole variabilità tra le diverse etichette.

Possiamo osservare che la sensibilità per tutte le etichette pesate per la loro frequenza è buona. Il valore predittivo positivo è da migliorare, nonostante comunque questa misura rappresenti un nuovo riferimento per la letteratura scientifica per l'estrazione di etichette come quelle utilizzate in questo studio. È altresì da notare come alcune etichette penalizzino severamente il risultato finale: ad esempio BRCA che occorre raramente nei referti e che l'AI non è in grado di riconoscere nei pochi casi presenti nel training set.

Dei 347 referti analizzati automaticamente, sono state estratte le etichette con i relativi valori, e tale risultato è stato incrociato con i dati provenienti dal

ETICHETTA	VPP	SE
Stadio di malattia	0,69	0,92
Visita oncogeriatrica	0,40	0,57
Terapia Neoadiuvante	0,14	0,71
Risonanza Magnetica Mammaria	0,47	0,77
Agobiopsia	0,20	0,41
Tumor (T)	0,44	0,77
Node (N)	0,55	0,81
Metastasi (M)	0,25	1,00
Micrometastasi	0,50	1,00
Macrometastasi	0,20	1,00
Recettori per gli estrogeni (ER)	0,69	0,82
Recettore per il progesterone (PGR)	0,70	0,79
Attività proliferativa (Ki67)	0,57	0,62
HER-2	0,39	0,88
Invasione vascolare	0,71	0,90
Margine	0,80	0,78
Grado Istologico	0,38	0,90
Breast Related Cancer Antigens (BRCA)	0,00	0,00
Media	0,44	0,77

Tabella 1. Valutazione delle performance dell'algoritmo: valore predittivo positivo (VPP) e sensibilità (SE) e media pesata per numero di occorrenze delle etichette.

flusso e rappresentato su un cruscotto di business intelligence. La **Figura 1**, seguente rappresenta la selezione di un singolo paziente (Paziente366) e l'associazione tra i dati rilevati dai flussi amministrativi e quelli rilevati dai referti.

CONCLUSIONI

Complessivamente i risultati di questo studio pilota si possono ritenere **soddisfacenti dal punto di vista tecnico** se confrontato con una letteratura scientifica in cui non vi è storico di estrazioni di rilevamento di tali etichette da referti clinici GIC. In una visione prospettica se i dati sono memorizzati in silos e in tassonomie non uniformi sono difficili

da trattare, il progetto ha tuttavia dimostrato che è possibile sfruttare le potenzialità dei dati non strutturati uniti con i dati amministrativi al fine di creare informazioni fruibili per il supporto decisionale.

La metodologia proposta presenta i seguenti punti di miglioramento utili per trasformare le informazioni in conoscenza:

la variabilità nella modalità di stesura dei referti richiede di aumentare la dimensione del corpus su cui addestrare l'algoritmo;

la presenza nei referti GIC di ripetizioni di informazioni necessita una disambiguazione anche attraverso un'analisi della dimensione temporale della struttura del testo;

la necessità di miglioramento delle performance dell'algoritmo, in particolare per quanto riguarda la misura di precisione e quindi la riduzione dei falsi positivi.

Il prossimo obiettivo è quello di migliorare il valore predittivo in modo da poter utilizzare i risultati per effettuare una analisi dei pazienti stratificandoli e impostare quindi una conseguente valutazione differenziata dei percorsi e degli esiti (es. stadio di malattia, ki67, ...)

Nello stesso modo, il progetto ha messo in luce come sia possibile andare ad attaccare alcune aree che necessitano di maggiori dati al fine di colmare le lacune informative presenti sui flussi amministrativi per le aziende ospedaliere; in questo modo è possibile andare ad identificare snodi fondamentali nei percorsi non reperibili dai flussi (es. biopsia).

Nei progetti di intelligenza artificiale in ambito sanitario, dove è possibile che i risultati degli algoritmi supportino il clinico nel processo decisionale, i canoni di valutazione delle performance sono più stringenti rispetto ad altri domini; risulta quindi rilevante l'introduzione di processi supervisionati che richiedono la collaborazione tra la parte clinica che ricopre il ruolo di esperto di dominio e la parte tecnica, sia per la fase più tecniche di definizione dei criteri che per la fase di validazione dei risultati.

La complessità di progetti come questo richiede inoltre la **collaborazione tra gruppi di lavoro multidisciplinari** affinché la condivisione delle corrette risorse e competenze permetta di favorire l'**aumento di conoscenza**; il progetto è un esempio di partnership tra pubblico privato e ha visto la collaborazione tra i seguenti partner:

- Università degli Studi di Torino per la parte di Epidemiologia statistica e Machine Learning;
- LINKS Foundation per il Processamento del Linguaggio Naturale e Intelligenza Artificiale;
- Innovo s.n.c. per l'analisi dei dati clinici;
- Piccola Casa della Divina Provvidenza – Ospedale Cottolengo per il supporto sull'area clinica;
- Roche per l'esperienza in ambito Real World Evidence e di Networking.

BIBLIOGRAFIA

- ¹ DIKW pyramid: https://en.wikipedia.org/wiki/DIKW_pyramid
- ² Osservatorio Innovazione digitale in Sanità - I dati in Sanità: L'utilizzo attuale nelle aziende sanitarie – Prof. Paolo Locatelli
- ³ Cancer Real World from needs to challenge: BIG DATA IN SANITA', falsi miti e opportunità dalla trusted artificial intelligence - Dott. Daniela Scaramuccia, Dott. Robert Alexander

Dati Processo da Flussi				
Paziente_ID	Episodio	Griglia_Data	Griglia_Evento...	Griglia_Evento
Paziente366	Episodio 1	12/10/2017	01 - Biopsia	85.11.1 - Biopsia Eco-Guidata Della Mam...
		27/10/2017	10 - CAS	89.05 - Visita Presso Il Centro Di Accoglie...
		17/01/2018	20 - GIC Pre	89.07 - Consulto Definito Complesso
		06/03/2018	40 - Intervento	257 [C] - Mastectomia totale per neoplas...
		28/03/2018	60 - GIC Post	89.07 - Consulto Definito Complesso

Dati da Referti			
Paziente	Etichetta	Valore	re...
Paziente366	Biopsia	Biopsia	1
	Stadio	IA	1
	N	N0	1
	Neoadiuvante	Neoadiuvante	1
	Ki67	2	1
	ER	99	1
	Visita oncoger...	Visita oncogeriatrica	1
	Margine	indenni	1
	Invasione vas...	Invasione vascolare	1
	T	pT1b	1
	Grado	2	1
	HER 2	SCORE 3+	1
	PGR	0	1
	Macrometastasi	Macrometastasi	1

Figura 1. Selezione della Paziente366 dal cruscotto di Business Intelligence