FONDAZIONE BRUNO KESSLER
MAGAZINE

L’intelligenza artificiale per la medicina d’emergenza: la ricerca di Pietro Ferrazzi

6 Agosto 2026

Trasformare il linguaggio naturale della documentazione clinica in dati strutturati per far progredire la ricerca

Trasformare milioni di documenti clinici in informazioni strutturate, utili per la ricerca e per migliorare il lavoro dei medici nei pronto soccorso. È questo l’obiettivo del lavoro di **Pietro Ferrazzi**, dottorando presso l’unità Natural Language Processing – NLP del centro Augmented Intelligence di FBK e iscritto al corso di dottorato Brain, Mind and Computer Science dell’Università di Padova. Al centro della sua ricerca c’è una delle sfide ancora aperte del Natural Language Processing in medicina: sviluppare modelli capaci di funzionare sulla documentazione clinica reale, e non solo sui dataset derivati dalla letteratura scientifica.

Arrivato in Trentino proprio per il dottorato, Ferrazzi ha costruito un percorso interdisciplinare che unisce statistica, big data e intelligenza artificiale. Dopo la laurea in Statistica e la magistrale in Big Data all’Università di Padova – arricchita da due esperienze all’Université Libre de Bruxelles e l’Universitat Politècnica de Catalunya (Barcellona) – ha scelto di proseguire il proprio percorso in Fondazione Bruno Kessler, dove si occupa di elaborazione del linguaggio naturale applicata alla medicina.

“La mia ricerca riguarda metodi di post-training per modelli di Natural Language Processing in ambito medico. Questi modelli vengono generalmente sviluppati e valutati su dataset derivati dalla letteratura scientifica. Il mio obiettivo è capire come migliorarne le prestazioni quando vengono applicati alla documentazione clinica del pronto soccorso, che presenta caratteristiche molto diverse.”

Uno dei casi di studio attraverso cui Ferrazzi mette alla prova queste metodologie è il Case Report Form Filling, sviluppato nell’ambito del progetto europeo eCREAM, che punta a ottimizzare la ricerca nella **medicina d’emergenza** valorizzando le informazioni contenute nella **documentazione clinica**.

Ogni giorno i professionisti clinici raccolgono una grande quantità di informazioni sui pazienti in forma testuale. Per renderle utilizzabili nella **ricerca**, devono però trasferirle manualmente in

forma strutturata. Per farlo, uno strumento utile sono i Case Report Form, le schede utilizzate per raccogliere in modo standardizzato i dati clinici. Un'attività essenziale, ma lunga e ripetitiva.

"L'obiettivo è automatizzare questo processo. Partendo dalla documentazione clinica, vogliamo identificare automaticamente informazioni come patologie pregresse, terapie in corso o parametri clinici e trasformarle in dati strutturati", spiega Pietro Ferrazzi.

Per sviluppare e valutare questi sistemi, il gruppo di ricerca ha utilizzato un patrimonio di dati particolarmente prezioso: **circa due milioni di documenti clinici**, raccolti nell'arco di tre anni presso l'Ospedale San Giovanni Bosco di Torino e annotati manualmente dai clinici.

"Poter contare su annotazioni realizzate dai medici è già insolito. Ancora più raro è disporre e poter pubblicare dati provenienti direttamente da un ospedale , che richiedono un enorme investimento di tempo e competenze."

A partire da questo patrimonio di dati, Ferrazzi e i ricercatori dell'unità Natural Language Processing di FBK hanno organizzato il **Case Report Form Filling Shared Task**, ospitato nell'ambito del workshop CL4Health della conferenza internazionale LREC 2026, con l'obiettivo di colmare il divario tra le esigenze della medicina d'emergenza e la comunità del Natural Language Processing.

La prima edizione della competizione ha coinvolto dodici gruppi di ricerca internazionali, che hanno sviluppato e confrontato sistemi di intelligenza artificiale utilizzando il dataset messo a disposizione da FBK. Oltre alla presentazione dei risultati durante la conferenza, l'iniziativa ha favorito la nascita di nuove collaborazioni scientifiche. Per la prossima edizione il gruppo sta lavorando per coinvolgere nuovi partner europei, con l'obiettivo di ampliare ulteriormente la comunità di ricerca attorno a questa sfida.

Il Case Report Form Filling rappresenta anche un tassello importante del percorso di dottorato di Ferrazzi. La sua **tesi** si concentra infatti su **metodologie di post-training** per il Natural Language Processing applicato alla medicina e questo caso di studio gli permette di verificare come tali metodologie si comportino nel passaggio dai dataset derivati dalla letteratura scientifica alla documentazione clinica reale.

"È un banco di prova importante perché i dati clinici presentano caratteristiche molto diverse rispetto ai dataset tradizionali. Inoltre, per l'italiano esistono ancora pochissime risorse di questo tipo, mentre per l'inglese la disponibilità è sensibilmente maggiore", spiega Ferrazzi.

L'obiettivo finale è sviluppare strumenti capaci di alleggerire il lavoro dei professionisti sanitari e, allo stesso tempo, valorizzare l'enorme patrimonio informativo custodito negli archivi ospedalieri. Rendere accessibili queste informazioni significa non solo semplificare attività oggi svolte manualmente, ma anche mettere a disposizione della ricerca dati strutturati che potranno contribuire allo sviluppo di una medicina d'emergenza sempre più supportata dall'intelligenza artificiale.

LINK

https://magazine.fbk.eu/it/news/lintelligenza-artificiale-per-la-medicina-demergenza-la-ricerca-di-pietro-ferrazzi/

TAG

- #big data
- #dataset
- #dati clinici
- #FBKPhDProgram
- #intelligenzaartificiale
- #intelligenzaaumentata
- #linguaggio
- #medicina
- #modelli
- #NLP
- #statistica

AUTORI

- Michela Antino