Andrea Ciardiello

Machine Learning application to doctor-patient interaction

FiloBlu: Sentiment Analysis

Artificial Intelligence in Medicine- Roma 3.12.2018

Istituto Nazionale di Fisica Nucleare

Motivation: domiciliary care of oncologic patients

Domiciliary care for oncologic patients is preferred

- Higher standard of living
- Psychological wellbeing
- Cheaper than hospitalization

The patient is followed by a Caregiver, usually a non professional volunteer

- Spouse, parent (mother), son
- A professional nurse is also possible

To successfully follow therapy during the domiciliary care the patient/Caregiver is in constant contact with healthcare professionals

- Clinicians, psychologists , nurses have an active role
- Frequent monitoring on therapy quality and general health

Motivation: domiciliary care of oncologic patients

Patients are interested in actively collaborate to the management of their health, the use of ICT technologies.

The **FILO BLU** Project meets the citizens’ needs developing a tool to:

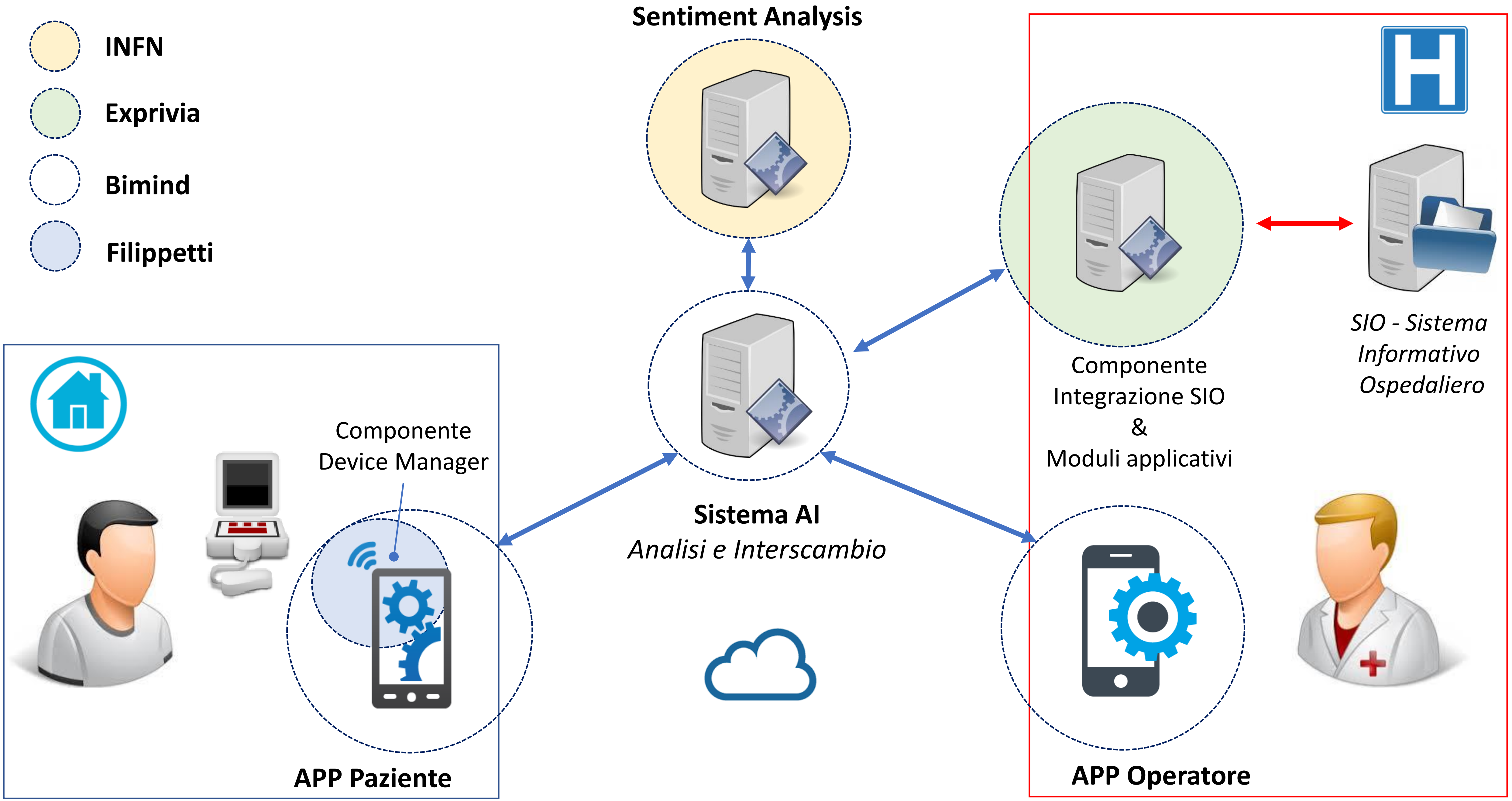
- optimize the efficiency and the effectiveness of care processes
- increase the involvement of the patient and the caregiver (empowerment)
- improve the healthcare provision and treatment of patients in cancer therapy

The **FILO BLU** system consists of:

- two APPs (Patient and Medical side) to support the doctor-patient communication
- a module for the interoperability with portable monitoring systems integrated to the APP
- A system for the analysis of medical-patient communications that aims to score the patient's status that analyzes the flow of communications in order to signal to doctors, through an "attention" score, potential critical situations (keeping into account both the written texts and any physiological values monitored).

FILO BLU has an interface similar to "WhatsApp“ and is equipped with features specifically designed for the healthcare applications :

- dispensation of precompiled questionnaires to the patient/caregiver;
- the transfer of the monitoring parameters of the medical devices eventually worn by the patient;
- the integration with electronic medical records



INFN
Exprivia
Bimind
Filippetti
Sentiment Analysis
Sistema AI
Analisi e Interscambio
Componente
Device Manager
APP Paziente
Componente
Integrazione SIO
&
Moduli applicativi
SIO - Sistema
Informativo
Ospedaliero
APP Operatore

Health Parameters Gathered by the App

- Pressione arteriosa : SISTOLICA (mmhg) DIASTOLICA(mmhg)
- Frequenza cardiaca (battiti per minuto, bpm)
- Atti respiratori al minuto
- Saturazione dell’Ossigeno
- Temperatura corporea
- Glicemia a digiuno (mg/dl)

Interactions now

- Patient/medic communication relies on IC technology
  - WhatsApp, Email, Social media

- Healthcare professional have many roles in the patient's life

Communications can have many purposes

- The medic as a confidant → Advices
- The medic as a friend → Greetings
- The medic as a professional → Therapy instructions, Health related questions

All of these are important but some message needs a different attention level

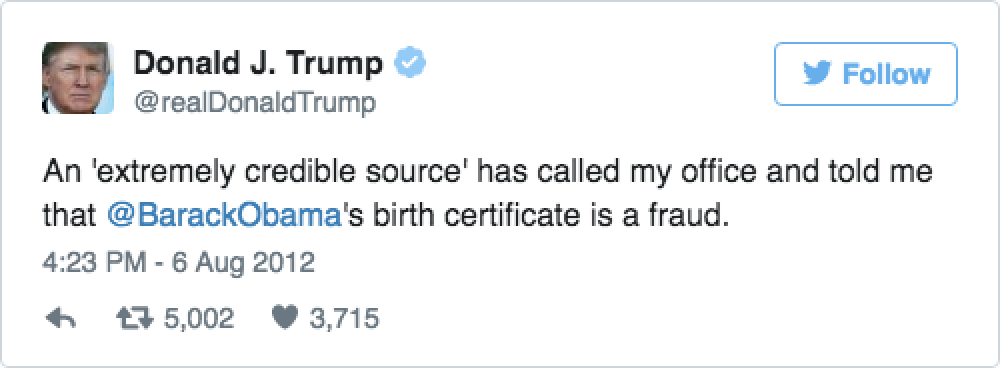
Report of a work in progress

Ongoing work

- App development
- Test in an hospital setting
- Integration with the hospital informatic center
- Development of Sentiment Analysis algorithm

and Machine Learning Application

Automated Text Analysis



Why Computational Text Analysis?

- There are costs to large-scale text analysis.
- Computers can lower these costs.

Most notable applications

- Newspapers, media attention and political events.
- Blogs and social media public opinion and communication.
- Brand awareness

Sentiment Analysis

Sentiment analysis (or opinion mining) involves classifying opinions in text into categories:

- "positive" , "negative" or "neutral“ is the most used classification.

There are many other approaches to sentiment. In particular many classification task are interesting

- Subjective (opinion) vs. Objective (fact) sentences
- Positive (favorable) vs. Negative (unfavorable) movie reviews

- Urgent vs. Not Urgent in a patient’s message?

Naïve Approach

First thing to try:
Dictionary methods

- Use lists of words to score documents, e.g. positive or negative.

Problem: How do generate dictionaries?

- Manually (a priori information on task)
- Crowd sourcing
- Statistical methods, discriminating words, Frequency analysis (from data)

Example dictionary
Hotel reviews:
Positive

- Tidy
- Calm
- Quite
- Clean
- Comfortable
- Swimming pool
- Spa
- …

Negative

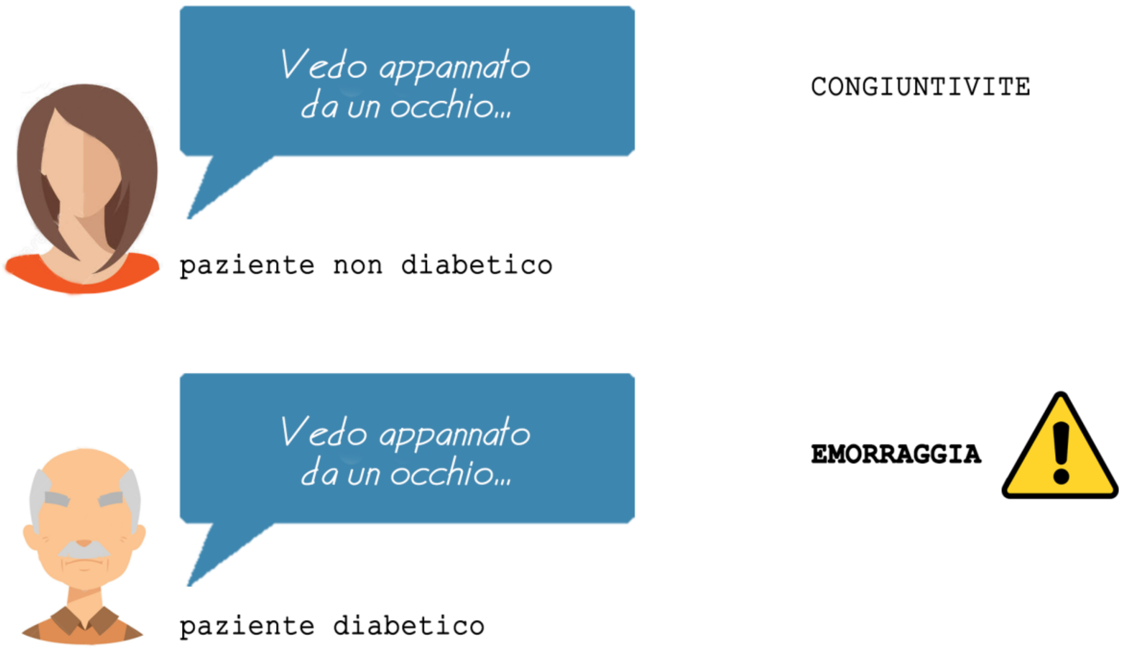
- Noisy
- Dirty
- Cockroaches
- Bedbugs
- Unfriendly
- …

Naïve Approach

The context is always relevant in text analysis.
But comunication in medical field may be completely different from patient to patient.

Finding correlation between the clinical state and the messages is essential

Dictionary based methods are not easy to implement



From text to Machine readable Data

Complexity Reduction: Making Assumptions

Remove capitalization, punctuation, numbers

Assumption: capitalization, punctuation does not provide useful information.

Stop Words: Common words that are not informative

- Remove sparse terms (rare words)
- Remove other terms (e.g. proper nouns).

Discard Word Order (Bag of Words), Tokenize

Assumption: Word Order Doesn't Matter.

•Bag of word models using character ngrams can be very efficient. **Do not underestimate them!**. They are relatively cheap to compute, and also easy to interpret.

Original message : Hi doc, my head hurt and I'am feeling very sick!

Stemming/lemming/SW : head, hurt, feel, very, sick,

Tokenize : 123, 5766, 2322, 4, 127

```
X =   [
        [1, 0, 0, 0, 0, 0, 0, ... ], #head
        [0, 1, 0, 0, 0, 0, 0, ... ], #hurt
        [0, 0, 1, 0, 0, 0, 0, ... ], #feel
        [0, 0, 0, 1, 0, 0, 0, ... ], #very
        [0, 0, 0, 0, 1, 0, 0, ... ], #sick
      ]
```

One-hot encoding

Head + very is not sick

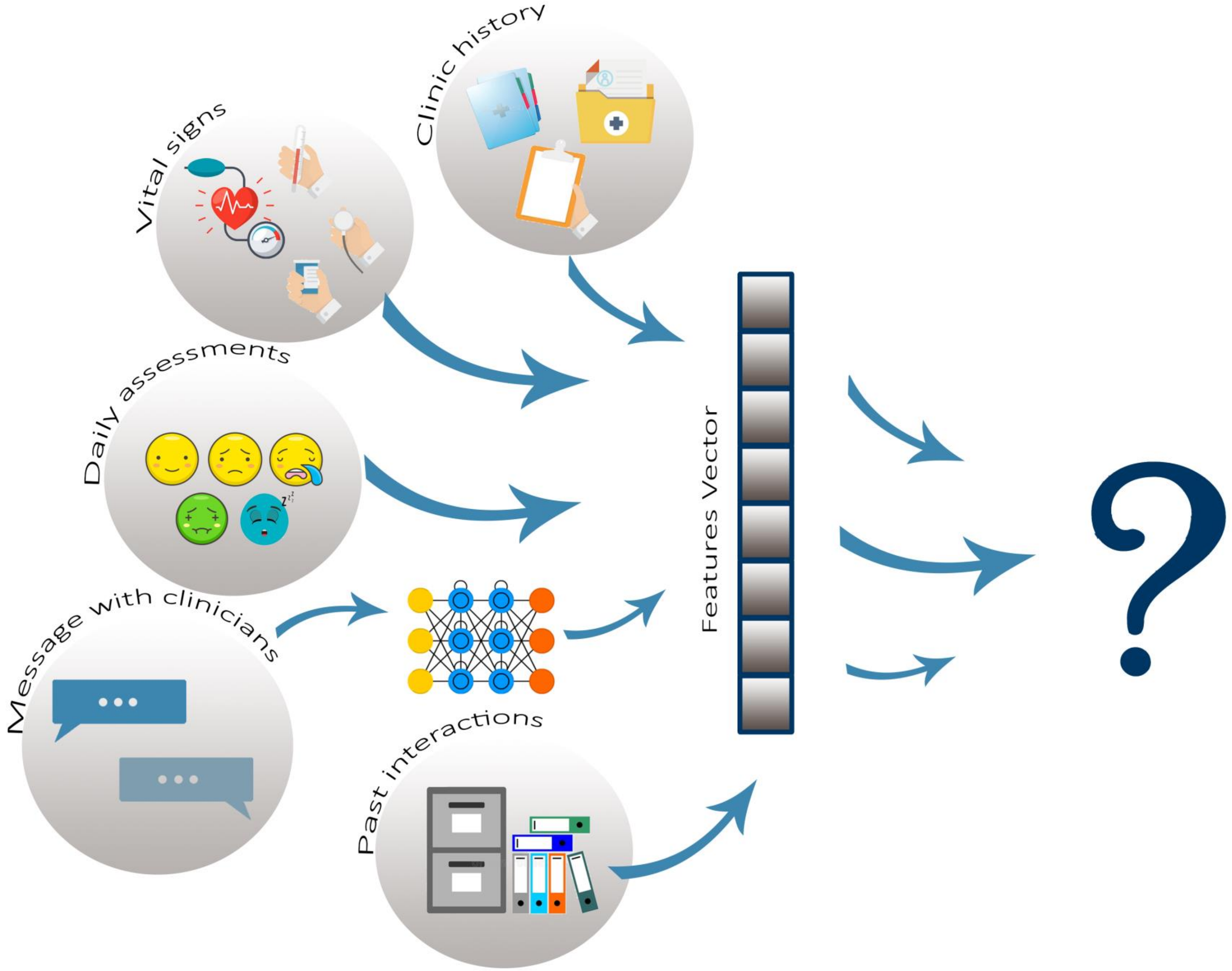
X = N x P matrix

- N = Number of elements in message
- P = Number of words in dictionary

X = main input for many computational text analysis applications.

Word embedding:
words from the vocabulary are mapped to vectors of real numbers.
one dimension per word → continuous vector space lower dimension
Word vectors: common context→ proximity in the space
Ex. Word2vec

Goal

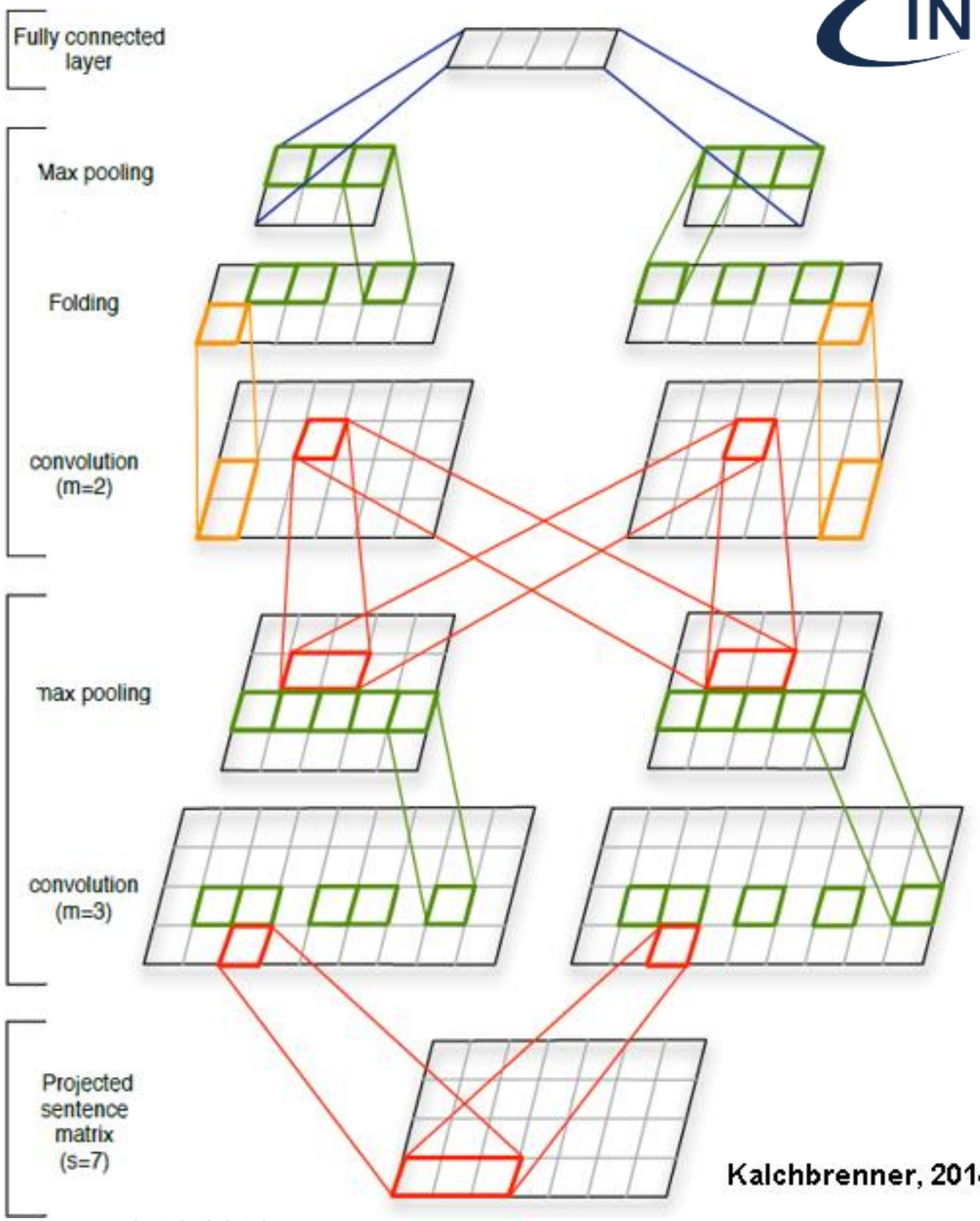


Clinic history
Vital signs
Daily assessments
Message with clinicians
Past interactions
Features Vector
?

CNNs

Convolutional feed-forward networks are specialized architectures that excel at extracting local patterns in the data:

- they are fed arbitrarily sized inputs (s)
- are capable of extracting local patterns sensitive to word order, regardless of where they appear in the input.
- these work very well for identifying indicative phrases or idioms of up to a fixed length (m).



Kalchbrenner, 2014

Recipe to Supervised Learning Method

Supervised methods: Hand coding is used to train statistical models to classify texts in pre-determined categories.

1) Set of known categories
    - Urgent, Not Urgent
    - Attention Level
2) Set of hand-labeled messages
    - Labeling done by clinicians
    - Training Set: messages we'll use to learn how to classify
    - Validation Set: messages we'll use to learn how well we classify
3) Set of unlabeled messages that we want to classify
4) Method to extrapolate from hand coding to unlabeled messages
5) Validate by comparing predicted label to real (hand-coded) label.

How Could This Possibly Work?

- It might not: Validation is critical (task specific)
- Central Tendency in Text: Words often imply what a text is about
- Words create a theme for a document

Grimmer and Stewart (2013):

- There is no globally best method for automated text analysis.
- Validate, Validate, Validate.

We need Data, lots of Data

How to get data?
From App Usage

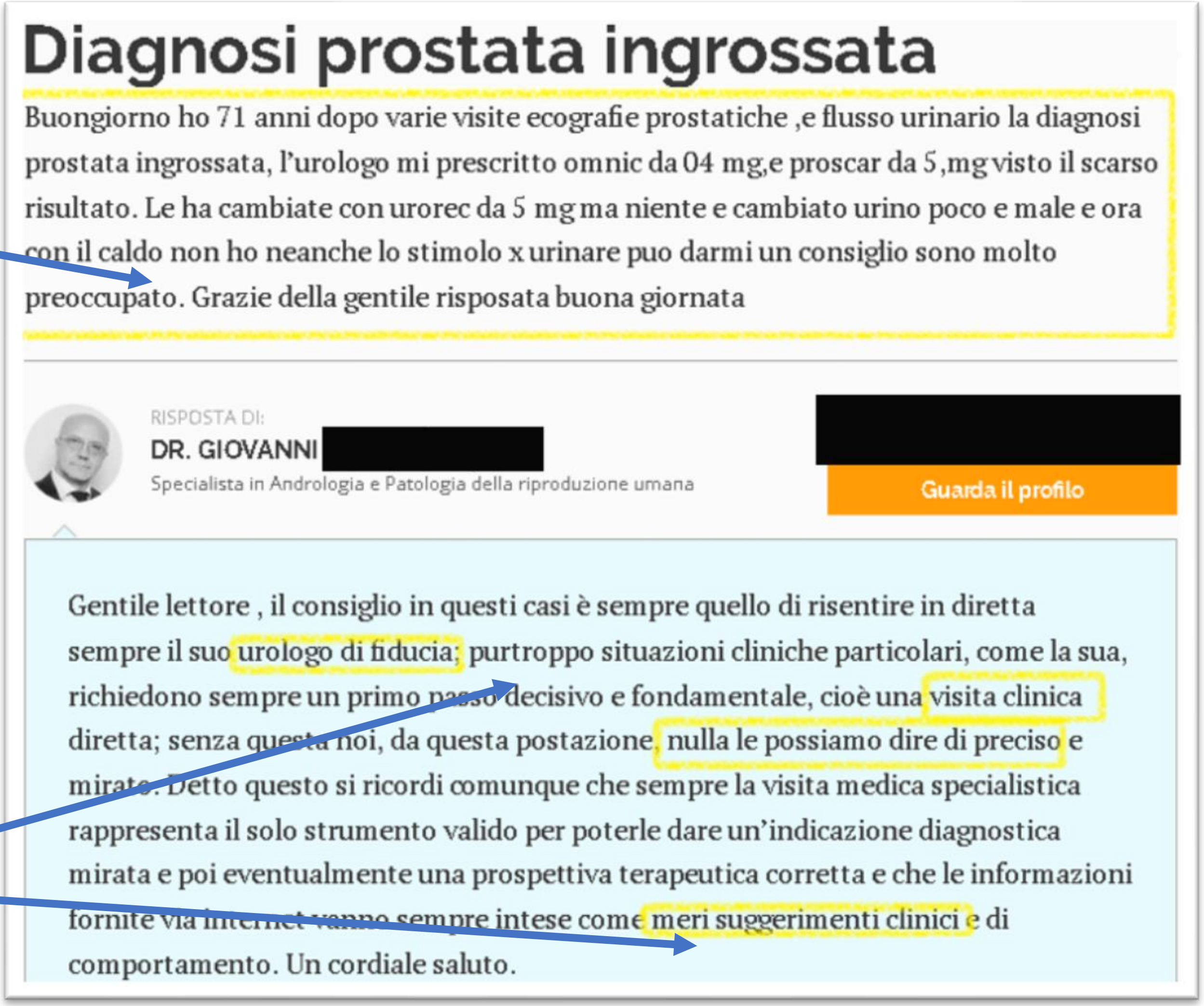
- Clinicians can select the attention level of an interaction directly from the APP
- The algorithm remember the choice and learn from the medic decision
- Similar cases (same context, same message) can be easily recognized

How to get data? Web Scraping

- Many web sites offer thousands of already written interaction
- Real word usage (spelling error occurrence, low level vocabulary)
- Week labelling (mainly for legal reason)

Message from patient

Label ?


Diagnosi prostata ingrossata

Buongiorno ho 71 anni dopo varie visite ecografie prostatiche ,e flusso urinario la diagnosi prostata ingrossata, l'urologo mi prescritto omnic da 04 mg,e proscar da 5,mg visto il scarso risultato. Le ha cambiate con urorec da 5 mg ma niente e cambiato urino poco e male e ora con il caldo non ho neanche lo stimolo x urinare puo darmi un consiglio sono molto preoccupato. Grazie della gentile risposata buona giornata

RISPOSTA DI:
DR. GIOVANNI [redacted]
Specialista in Andrologia e Patologia della riproduzione umana

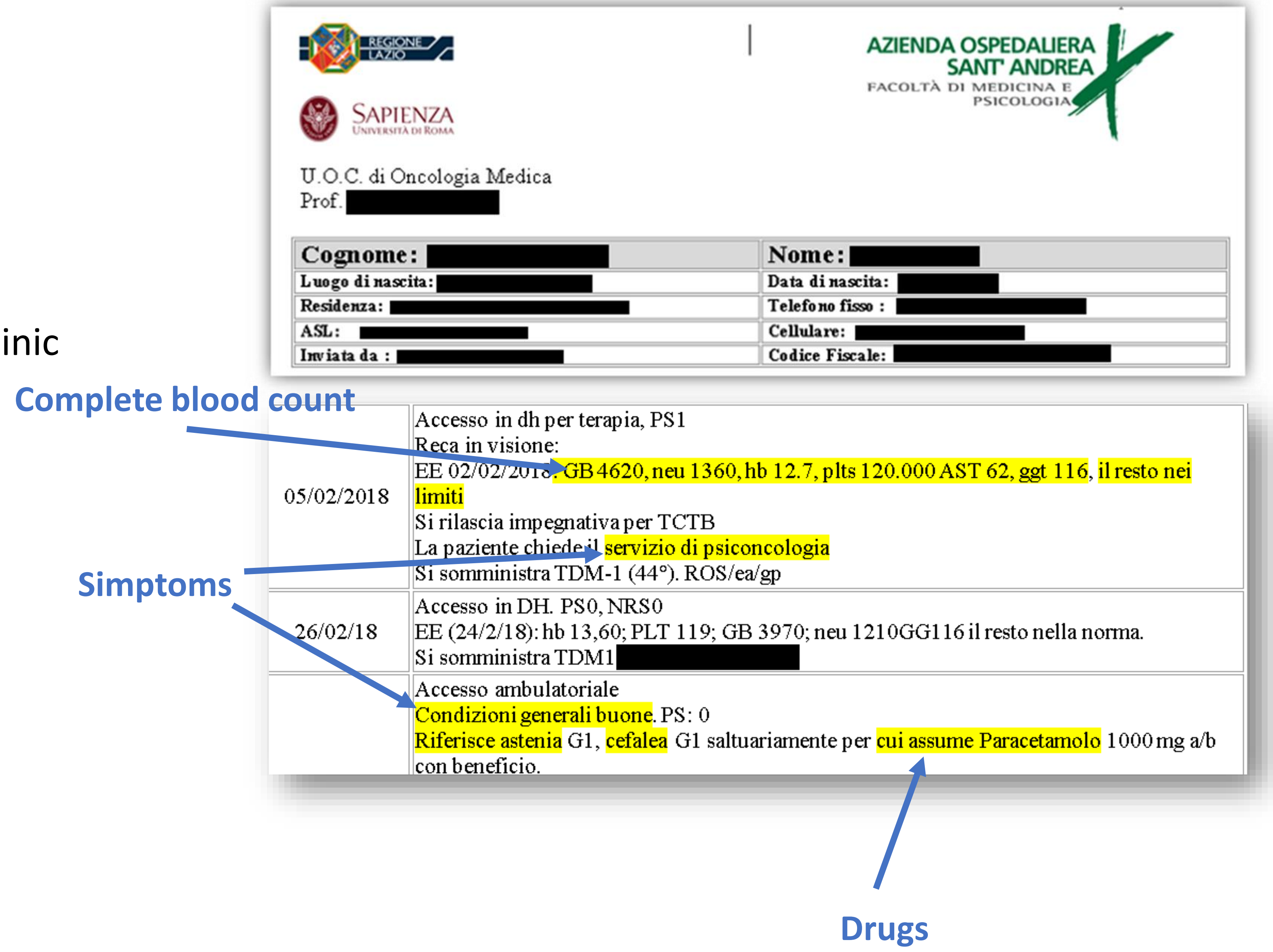
[redacted]
Guarda il profilo

Gentile lettore , il consiglio in questi casi è sempre quello di risentire in diretta sempre il suo urologo di fiducia; purtroppo situazioni cliniche particolari, come la sua, richiedono sempre un primo passo decisivo e fondamentale, cioè una visita clinica diretta; senza questa noi, da questa postazione, nulla le possiamo dire di preciso e mirato. Detto questo si ricordi comunque che sempre la visita medica specialistica rappresenta il solo strumento valido per poterle dare un'indicazione diagnostica mirata e poi eventualmente una prospettiva terapeutica corretta e che le informazioni fornite via internet vanno sempre intese come meri suggerimenti clinici e di comportamento. Un cordiale saluto.

How to get data?
Generating Fake Data

- Merging Questions with real clinic data
- Labelled by a clinician

- Low **throughput**, unless many undergrads are recruited



REGIONE LAZIO

SAPIENZA
UNIVERSITÀ DI ROMA

AZIENDA OSPEDALIERA SANT' ANDREA
FACOLTÀ DI MEDICINA E PSICOLOGIA

U.O.C. di Oncologia Medica
Prof. [redacted]

Cognome: [redacted]	**Nome:** [redacted]
Luogo di nascita: [redacted]	**Data di nascita:** [redacted]
Residenza: [redacted]	**Telefono fisso :** [redacted]
ASL: [redacted]	**Cellulare:** [redacted]
Inviata da : [redacted]	**Codice Fiscale:** [redacted]

05/02/2018	Accesso in dh per terapia, PS1 Reca in visione: EE 02/02/2018: GB 4620, neu 1360, hb 12.7, plts 120.000 AST 62, ggt 116, il resto nei limiti Si rilascia impegnativa per TCTB La paziente chiede il servizio di psiconcologia Si somministra TDM-1 (44°). ROS/ea/gp
26/02/18	Accesso in DH. PS0, NRS0 EE (24/2/18): hb 13,60; PLT 119; GB 3970; neu 1210GG116 il resto nella norma. Si somministra TDM1 [redacted]
	Accesso ambulatoriale Condizioni generali buone. PS: 0 Riferisce astenia G1, cefalea G1 saltuariamente per cui assume Paracetamolo 1000 mg a/b con beneficio.

Summary

- Project proposed and leaded by companies
- Our Task is Automated text Classification
- We are Applying Machine Learning methods

- Data gathering is on course

Extra:
Where we are so far:

- Tool e database for vectorialization text messages→ Done
- Neural Networks for classification:

  Phyton code (tensorflow) → Implementation word embedding layer
- Visualization e performace metrics → in progress
- Data Collection from APP → Less than we hope for
- Creation of a «toy sample» for development/test:
  - Web scraping tools for text mining→ in progress

Dati per il training

Queste reti imparano per esempi (Learning Supervised)

Per addesrare la rete a svolgere il suo compito servono

- Dati già classificati (Labeled)
- Esempi di testo + giudizio di urgenza del medico

Un algoritmo di classificazione testuale che raggiunga una error rate < 10% richiede

10 000 esempi per il training

Fake Data
Data Augmentation

Da utilizzare in attesa del campione reale di addestramento

- idea: utilizzo Q&A siti web che raccolgono domande di pazienti a medici/specialisti e le relative risposte che includono diverse specializzazione della medicina e patologie
    - Esempi **realistici**, con errori, typos, abbreviazioni
    - Facilmente reberibili
    - In grande quantità

Non facile estrarre una label dalla risposta del medico che per motivi legali non può esprimere un’opinione se non di attenersi alle istruzioni del medico curante, seguire la terapia o consultare uno specialista

Data Augmentation

Si può ridurre la quantità di esempi necessari.

Come aumentare artificialmente il campione:

- Estrazione di testo domande di pazienti + dizionario di termini tecnici, per esempio costruito dai consensi informati
- Generazione di testo sintetico con Reti Ricorrenti

Qualunque tecnica di Data Augmentation parte da un campione di Dati Reali !

Punti Cruciali

- Attenti! Servono 10^4 esempi:

  Si possono limitare con tecniche avanzate ma si parte da un dataset labeled

- Profondità del dizionario Italiano disponibile:

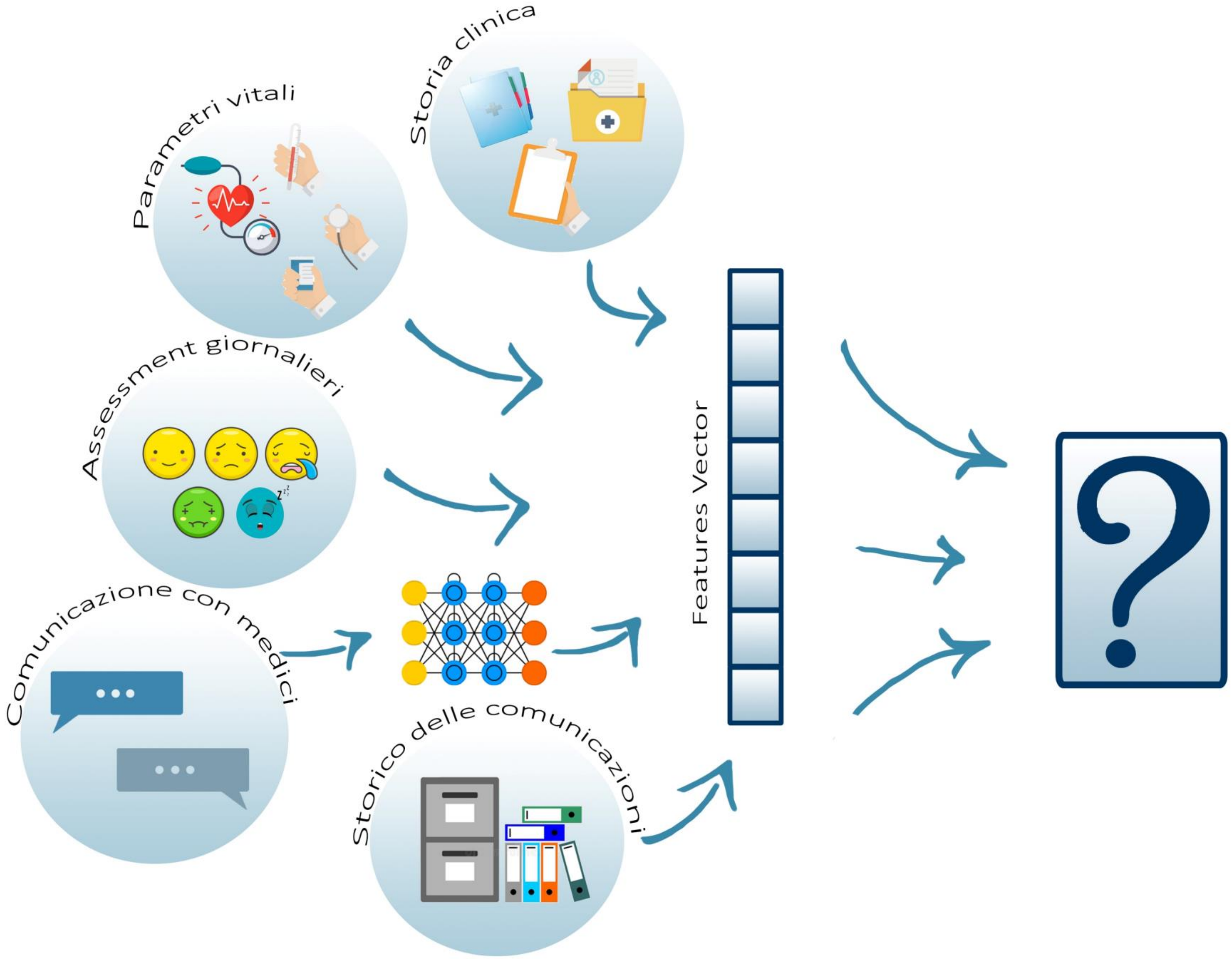
  Al momento ~900k-parole, che includono termini medico/scientifici. Un test su 5 Q&A testuali dal db www.paginemediche.it ha permesso di riconoscere i 100% delle parole utilizzate.

  - può essere ampliato introducendo i termini specialistici ricorrenti nel campione di training iniziale ottenuto dalla app Filo Blu, e successivamente aggiornato mano a mano che la statistica a disposizione aumenta

- Errori testuali nei messaggi:

  - Ci aspettiamo una rate di errore nelle parole contenute nei messaggi scritti sull'app, che potrebbe essere significativa dovuta a typo, abbreviazioni, T9, scrittura frettolosa del messaggio
  - Utilizzo di tecniche di adversarial attack per irrobustire la rete rispetto a questo tipo di errori: introduzione di un "rumore" nel testo durante il training indicando al contempo alla rete il risultato giusto, in modo che la rete apprenda il comportamento del rumore

Goal



Parametri vitali
Storia clinica
Assessment giornalieri
Comunicazione con medici
Storico delle comunicazioni
Features Vector
?

Goal