

# Automazione Diagnostica Guasti con LLM

March 27, 2025

## 1 Working flow

- Input testuale, Input diagnostica generico (è possibile che sia necessaria la conversione in csv)
- Analisi dell'input testuale con modelli di NLP (Natural Language Processing), per identificare parole chiave legate ai guasti : Esempio, il modello estrae "WC", "non funzionante", associato a TOI1 (impianto WC);
- Verifica nel file di diagnostica generica la presenza di eventuali guasti, con focus sulle parole chiave rilevate dal natural language processing: Esempio: Se viene trovato un guasto ("TOI1:WC Fuori servizio"=1), il sistema conferma il problema e identifica l'impianto coinvolto;
- Caricamento della diagnostica specifica dell'impianto coinvolto, così da analizzare i parametri specifici per trovare la causa esatta
- Output strutturato in linguaggio naturale basato sui dati rilevati: Esempio: Il WC non funziona perché il livello dell'acqua è inferiore al 10 percento, suggeriamo di controllare il riempimento del serbatoio TOI1 o verificare la pompa di alimentazione.

## 2 Strategia Tecnica

- Analisi NLP della Segnalazione
  - Estrarre le parole chiave dal plain text
  - Mappatura delle parole chiave alle colonne della diagnostica generale (identificare a seguito di un'analisi del dataset le colonne che rappresentano segnalazione di guasti)
  - Tecnologie : Regex/TF-IDF + KNN /BERT (se servono frasi più complesse) (Ricerca a proposito di ognuna di queste tecnologie, tempi e costi)
- Verifica del guasto nel report generale
  - Trova la colonna binaria corrispondente alla parola chiave individuata
  - Se il valore è 1, conferma il guasto e passa al file dettagliato dell'impianto (avrà dei nomi standard in modo che sia possibile prenderlo in input in automatico)
  - Tecnologie: Ricerca colonne + Random Forest

- Analisi del Report Dettagliato dell'impianto
  - Identifica la causa el guasto nei dati numerici associati all'impianto
  - Tecnologie : Decision Tree/Rule-based Analysis
- Generazione della spiegazione in linguaggio naturale
  - Produci un output testuale chiaro e leggibile per spiegare il problema
  - Tecnologie: Template-based NLP + GPT-4 per raffinamento

### 3 Step

- Analisi del dataset, identificando tutte le colonne binarie che possano identificare un guasto
- Ricerca a proposito delle tecnologie sopra menzionate: funzionamento, tempi, costi computazionali
- Scelta delle tecnologie da implementare
- Sviluppo