

Relazione – Settimana 1

Gruppo di Lavoro

26 marzo 2025

Introduzione e Obiettivi

Durante la prima settimana di progetto, l'obiettivo principale è stato quello di organizzare in modo sistematico le informazioni necessarie, definire il piano di lavoro e iniziare l'approfondimento teorico. Abbiamo perciò istituito un ambiente di lavoro condiviso su *Notion*, strutturato in diverse sezioni:

- **Link e documentazione:** per raccogliere e categorizzare articoli, paper e video utili alla comprensione dei concetti chiave;
- **Domande e idee:** uno spazio dedicato al confronto su eventuali dubbi, spunti e proposte di miglioramento;
- **Pianificazione e scadenze:** una *schedule* chiara, con date di consegna e priorità per ogni singola task, in modo da monitorare facilmente l'avanzamento del lavoro.

Ricerca e Studio Preliminare

Per comprendere le basi teoriche del nostro progetto, abbiamo dedicato parte della settimana allo studio di:

1. **Machine Learning e reti neurali:** analizzando il funzionamento di reti *feed-forward*, reti ricorrenti e, in particolare, architetture di tipo *Transformer*.
2. **Word Embedding:** abbiamo revisionato i concetti di *Word2Vec* e altri modelli associati, prendendo atto dei vantaggi e dei limiti di questi metodi rispetto ad approcci più recenti.

3. **Sentiment Analysis:** approfondendo la letteratura e i casi di studio che utilizzano modelli moderni, tra cui *BERT* di Google, come base per l'analisi del testo e la classificazione in positivo/negativo.

Questa fase di ricerca è stata essenziale per capire i punti di forza di BERT rispetto a Word2Vec: in particolare, la capacità di cogliere i contesti linguistici in modo più ricco e la possibilità di *fine-tuning* su dataset specifici.

Scelta di BERT e Struttura del Valutatore

Una volta compresi i limiti dei metodi più classici, abbiamo optato per *BERT* come modello principale per la *Sentiment Analysis* e per la generazione di uno *score* di affidabilità. A questo scopo, abbiamo realizzato lo schema di flusso del nostro *valutatore*, illustrato in Figura 1.

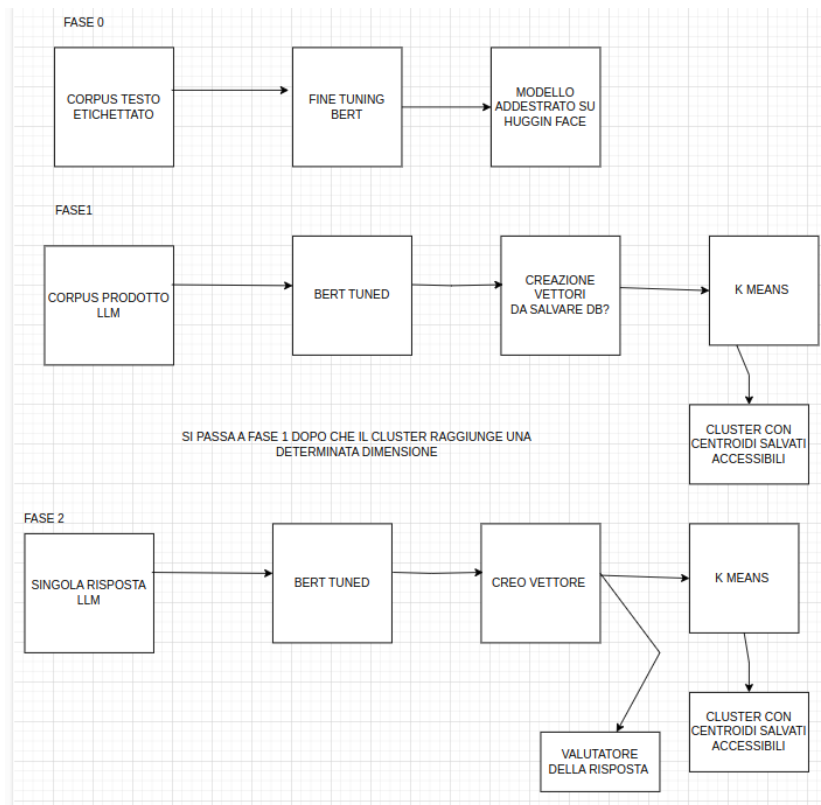


Figura 1: Schema del valutatore basato su BERT e clustering *k-means*.

L'idea di fondo è la seguente:

1. **Input dal LLM:** il testo prodotto da un *Language Model* (LLM) viene passato a BERT, il quale fornisce in uscita:
 - Un’etichetta di *Sentiment* (positivo/negativo), che codificheremo come +1 o -1;
 - Uno *score* compreso tra 0 e 1, rappresentativo del livello di affidabilità della classificazione.
2. **Clustering con *k-means*:** raccolte numerose risposte, utilizziamo l’algoritmo *k-means* per creare due cluster principali (positivo e negativo), definiti dai loro centroidi. Ogni nuova risposta, una volta classificata e vettorizzata, sarà assegnata a uno dei due cluster in base alla distanza dal rispettivo centroide.

Fine-Tuning e Dataset

Abbiamo compreso che, per ottenere risultati efficaci, *BERT* andrà *raffinato* (*fine-tuning*) su un dataset *ad hoc*. Poiché vogliamo applicare il modello a un caso d’uso specifico (ovvero la valutazione di risposte generate da LLM), sarà necessario:

- Creare un dataset di frasi o risposte etichettate manualmente (positivo/negativo) e arricchite di eventuali metadati utili.
- Addestrare BERT su questi dati, in modo che apprenda le peculiarità del linguaggio e dei contesti specifici del nostro dominio.

La costruzione del dataset avverrà nelle prossime settimane, durante le quali definiremo le linee guida di etichettatura per garantire coerenza e uniformità.

Prossimi Passi

Al termine di questa prima settimana, abbiamo discusso l’idea con il professor Ferretti, che ha approvato la direzione presa. Ci ha inoltre suggerito di effettuare i primi test su *Google Colab* utilizzando la versione base di BERT, per poi passare ai *cluster* del dipartimento, dotati di maggiori risorse computazionali.

Resta ancora aperta la questione relativa alla **dinamicità dei cluster**: non è chiaro se, in fase di utilizzo a regime, ogni nuova risposta debba aggiornare il modello di clustering oppure se il clustering debba essere “statico”

dopo una prima fase di addestramento. Questo aspetto sarà oggetto di approfondimento nelle prossime iterazioni, quando potremo misurare l'impatto di un aggiornamento dinamico sulla precisione e sulla stabilità del sistema.

Conclusioni

In sintesi, la prima settimana di lavoro ci ha permesso di:

- Organizzare la collaborazione e la gestione dei materiali tramite *Notion*;
- Studiare i concetti chiave di *Machine Learning*, *reti neurali* e *Transformer*;
- Identificare *BERT* come modello di riferimento per la *Sentiment Analysis* e per l'estrazione di uno *score* di affidabilità;
- Definire la struttura complessiva del valutatore, basato sul *fine-tuning* di BERT e sul clustering *k-means*.

Nei prossimi giorni, procederemo alla creazione del dataset di addestramento e ai primi test su *Google Colab*, in modo da validare le nostre ipotesi progettuali e valutare l'efficacia di *BERT* nel contesto specifico del nostro sistema di valutazione.