



UNIVERSITÀ DEGLI STUDI DI PADOVA
CORSO DI LAUREA IN INFORMATICA (L-31)

Corso di Ingegneria del Software
Anno Accademico 2025/2026

Verbale di Riunione

Verbale Interno del 13 Gennaio 2026

Gruppo: NightPRO

swe.nightpro@gmail.com

Data: 2026-01-13

Indice

1	Informazioni Generali	3
1.1	Componenti del Gruppo	3
1.2	Dettagli Riunione	3
2	Ordine del Giorno (Agenda)	4
3	Diario della Riunione	5
3.1	Stato Tecnico e Architettura del Sistema	5
3.2	Scelta Architetturale: RAG vs Fine-tuning	5
3.3	Infrastruttura e Hosting: Gestione Railway e Repository	6
3.4	Analisi dei Requisiti e Design dell'Interfaccia Utente	6
3.5	Preparazione della Riunione Esterna con Ergon	6
4	Decisioni Prese	8
5	Attività da Svolgere	8

1 Informazioni Generali

1.1 Componenti del Gruppo

Elenco dei membri del gruppo di lavoro NightPRO.

Cognome	Nome	Matricola
Biasuzzi	Davide	2111000
Bilato	Leonardo	2071084
Zanella	Francesco	2116442
Romascu	Mihaela-Mariana	2079726
Ogniben	Michele	2042325
Perozzo	Samuele	2110989
Ponso	Giovanni	2000558

Tabella 1: Componenti del Gruppo NightPRO.

1.2 Dettagli Riunione

- **Data:** 2026-01-13
- **Ora:** 18:30 - 21:00
- **Luogo:** Online Meet
- **Partecipanti:** Francesco Zanella, Samuele Perozzo, Michele Ogniben, Giovanni Ponso, Davide Biasuzzi, Leonardo Bilato, Mihaela-Mariana Romascu.
- **Assenti:** Nessuno
- **Redatto da:** Davide Biasuzzi
- **Verificato da:** Giovanni Ponso
- **Versione:** 1.0

2 Ordine del Giorno (Agenda)

1. Stato di avanzamento collegamento backend e frontend.
2. Gestione infrastruttura e chiavi API OpenAI.
3. Scelta tecnica tra Fine-tuning e RAG (Retrieval-Augmented Generation).
4. Analisi dei requisiti e design dell'interfaccia (Analisi finale).
5. Preparazione domande per l'incontro con Ergon/Carlesso.

3 Diario della Riunione

3.1 Stato Tecnico e Architettura del Sistema

Il gruppo ha verificato lo stato di avanzamento del collegamento tra backend (FastAPI) e frontend (React). Michele Ogniben si è occupato della parte iniziale di backend con integrazione LLM, mentre Samuele Perozzo e Francesco Zanella si sono occupati della connessione backend-frontend-database.

In particolare, è stata discussa l'implementazione della logica di recupero dati fondamentale per il sistema RAG (Retrieval-Augmented Generation). Durante questa fase è emersa la necessità di implementare un meccanismo di ricerca "fuzzy" per gestire eventuali errori di battitura da parte degli utenti. Il gruppo ha approfondito l'utilizzo dell'algoritmo **Soundex**, che permette di identificare parole foneticamente simili, migliorando così l'esperienza utente e la robustezza del sistema.

L'infrastruttura di backend è stata realizzata utilizzando **Docker** e **FastAPI**, garantendo modularità e facilità di deployment. Tuttavia, durante la fase di test è emersa una criticità relativa alla chiave API OpenAI fornita dall'azienda proponente: il credito associato alla chiave risulta esaurito, impedendo il completamento dei test. Temporaneamente, un membro del gruppo ha messo a disposizione la propria chiave personale con credito residuo per consentire lo sviluppo e i test necessari. Questa problematica sarà discussa nella prossima riunione esterna con l'azienda per richiedere una nuova chiave funzionante.

3.2 Scelta Architetturale: RAG vs Fine-tuning

Una parte significativa della riunione è stata dedicata alla discussione della scelta tecnica tra **Fine-tuning** e **RAG** (Retrieval-Augmented Generation) per l'implementazione del sistema di intelligenza artificiale.

Il gruppo ha analizzato le specifiche del capitolato e ha rilevato che non vi è alcuna menzione esplicita riguardo all'obbligo di utilizzare il fine-tuning del modello linguistico. Dopo un'attenta valutazione, il team ha concluso che il fine-tuning non sia la soluzione più adatta al caso d'uso specifico del progetto, per le seguenti motivazioni:

- **Dati insufficienti:** il fine-tuning richiede un dataset molto ampio e ben strutturato, di cui il gruppo non dispone attualmente;
- **Costi elevati:** il processo di fine-tuning comporta costi computazionali e temporali significativi, non giustificati dalle necessità del progetto;
- **Inadeguatezza al dominio:** per cataloghi aziendali specifici e in continua evoluzione, il RAG risulta più flessibile e manutenibile, permettendo di aggiornare le informazioni senza dover riaddestrare il modello.

È stata quindi presa la **decisione unanime** di utilizzare esclusivamente l'approccio RAG, che permette al sistema di recuperare informazioni contestuali dal database aziendale e fornirle al modello linguistico per generare risposte accurate e aggiornate.

Durante la discussione sono stati approfonditi anche i concetti di **token** ed **embedding vettoriale**, fondamentali per ottimizzare le performance di sistema.

È stato deciso di inserire questa scelta tecnica tra le domande da porre all'azienda proponente Ergon durante la prossima riunione esterna, per ottenere conferma e validazione dell'approccio scelto.

3.3 Infrastruttura e Hosting: Gestione Railway e Repository

Il gruppo ha discusso sul confermare **Railway** come piattaforma di hosting della webapp, e che avremmo chiesto conferma all'azienda proponente durante la prossima riunione esterna.

Un'altra criticità emersa riguarda la pulizia del database: sono stati identificati ordini e articoli con chiavi esterne mancanti o inconsistenti, che devono essere rimossi per garantire l'integrità referenziale e il corretto funzionamento del sistema. Giovanni Ponso e Samuele Perozzo sono stati incaricati di procedere con la pulizia del database e la configurazione corretta di Railway.

Inoltre, è stata sottolineata l'importanza di implementare un sistema di **logging delle chiamate API** per facilitare il debugging, specialmente per gestire eventuali errori 429 (rate limit) che potrebbero verificarsi durante l'utilizzo intensivo delle API di OpenAI.

3.4 Analisi dei Requisiti e Design dell'Interfaccia Utente

Mihaela-Mariana Romascu e Leonardo Bilato hanno presentato i mockup dell'interfaccia utente realizzati con **Penpot**, illustrando sia l'interfaccia destinata al cliente finale sia quella per l'operatore aziendale.

Durante la presentazione sono emerse diverse proposte di miglioramento e funzionalità da integrare:

Gestione degli ordini ambigui È stato discusso come gestire l'ambiguità negli ordini ricevuti dal sistema. Il gruppo ha deciso di implementare una **percentuale di confidenza** (affidabilità) per ogni ordine, visualizzata nell'interfaccia dell'operatore. Questo permetterà di identificare rapidamente gli ordini che necessitano di revisione manuale.

Interfaccia operatore Per migliorare l'usabilità dell'interfaccia dell'operatore, sono state decise le seguenti modifiche:

- **Colonna ID univoca** per ogni ordine nella tabella principale, facilitando l'identificazione e la tracciabilità degli ordini;
- **Sistema di filtri per data** con ordinamento crescente/decrescente, implementato tramite icone dedicate invece di una barra di ricerca complessa;
- **Barra di ricerca dinamica** per cercare ordini specifici per cliente o prodotto;
- **Pulsante di richiesta assistenza umana** per permettere al cliente di segnalare esplicitamente situazioni in cui il sistema automatico non è sufficiente.

Mihaela-Mariana Romascu è stata incaricata di aggiornare i mockup includendo le modifiche discusse, in particolare l'aggiunta della colonna ID e dei filtri per data.

3.5 Preparazione della Riunione Esterna con Ergon

Il gruppo ha raccolto e discusso i dubbi e le domande da porre durante la riunione esterna prevista con i rappresentanti dell'azienda Ergon e il professor Carlesso.

Le principali questioni da chiarire riguardano:

- **Dati mancanti nel dataset:** verificare se sia possibile ottenere informazioni aggiuntive quali prezzi, quantità disponibili e disponibilità dei prodotti, attualmente non presenti nel dataset fornito;
- **Validazione dell'approccio RAG:** confermare la scelta tecnica di utilizzare RAG invece del fine-tuning, spiegando le motivazioni tecniche e ottenendo l'approvazione dell'azienda;

- **Processo di assistenza dell'operatore:** discutere del processo di assistenza dell'operatore all'utente;
- **Infrastruttura di deploy:** richiedere informazioni sui server aziendali per pianificare il deployment finale del sistema in produzione.

Per facilitare la raccolta e la sincronizzazione delle domande tra tutti i membri del team, è stato deciso di creare un **file Google Doc condiviso** dove ciascun membro potrà inserire le proprie domande prima della riunione. Questo permetterà di avere una visione completa e organizzata di tutti i punti da discutere.

4 Decisioni Prese

1. Adozione dell'approccio **RAG** (Retrieval-Augmented Generation) come soluzione tecnica principale, escludendo il fine-tuning.
2. Pulizia del database da ordini e articoli con chiavi esterne inconsistenti.
3. Futura implementazione di un sistema di logging per le chiamate API.
4. Aggiunta nell'interfaccia operatore di:
 - Percentuale di confidenza per identificare ordini ambigui;
 - Filtri per data con ordinamento cronologico tramite icone dedicate;
 - Barra di ricerca dinamica per cliente/prodotto;
 - Pulsante per richiesta assistenza umana.
5. Creazione di un file Google Doc condiviso per sincronizzare le domande prima della riunione esterna.
6. Discussione con Ergon nella prossima riunione esterna riguardo a:
 - Chiave API OpenAI non funzionante;
 - Validazione dell'approccio RAG;
 - Dati mancanti nel dataset;
 - Conferma piattaforma Railway come hosting della webapp;
 - Informazioni sui server aziendali per il deploy.

5 Attività da Svolgere

Le attività discusse si intendono implicitamente assegnate ai membri che già se ne occupavano, mantenendo la continuità con il lavoro corrente. La presente riunione ha avuto funzione di disambiguazione e allineamento interno in vista dell'incontro con l'azienda previsto per **il 14 Gennaio 2026**. La pianificazione formale delle attività per il prossimo sprint avverrà durante la riunione di fine sprint del **16 Gennaio 2026**.