

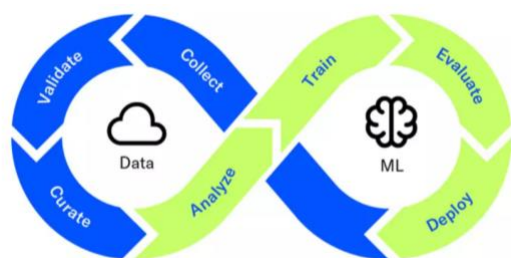
Progetto di Sistemi Distribuiti e Cloud Computing

Corso di Laurea Triennale in Ingegneria Informatica - a.a. 2024/25

Studente: _____Umberto Domenico Ciccia _____ Matricola: _____263844 _____

Data di assegnazione del progetto: _____09/05/2025 _____

Il progetto si propone di sviluppare una piattaforma cloud-based per la **previsione del prezzo di vendita di immobili** utilizzando un approccio MLOps. L'obiettivo è costruire un sistema automatizzato e scalabile che gestisca l'intero ciclo di vita di un modello di machine learning, dalla raccolta dei dati fino al deployment in produzione. Le previsioni si basano su caratteristiche come superficie, numero di stanze, localizzazione geografica e altre variabili, utilizzando dataset pubblici ampiamente riconosciuti come il [California Housing](#).



In accordo con lo schema illustrato, lo studente dovrà occuparsi della fase di acquisizione, validazione e preparazione dei dati. In questo contesto, dovranno essere impiegati strumenti per l'automazione delle pipeline di trasformazione, quali funzioni serverless o trigger basati su eventi. Il processo potrà essere implementato utilizzando servizi cloud per la creazione di pipeline ETL (es. AWS Glue, Azure Data Factory, Google Cloud Dataflow) oppure soluzioni open-source (es. Apache Airflow).

I dati post-processati saranno utilizzati per addestrare un modello di machine learning in grado di affrontare problemi di regressione, impiegando algoritmi come XGBoost, Random Forest o Regressione Lineare. Il training potrà avvenire tramite servizi gestiti che offrono funzionalità di tracciamento degli esperimenti (es. AWS SageMaker, Azure Machine Learning o Google Cloud Vertex AI), oppure mediante strumenti open-source scelti dallo studente (es. MLflow).

Le nuove versioni del modello dovranno essere messe automaticamente in produzione mediante l'adozione di una pipeline CI/CD (Continuous Integration e Continuous Deployment). Tale pipeline potrà essere implementata con strumenti nativi come AWS CodePipeline, Azure DevOps Pipelines o Google Cloud Build, oppure tramite soluzioni open-source a discrezione dello studente.

Infine, il modello addestrato dovrà essere esposto tramite un'API RESTful, consentendo l'esecuzione di predizioni in tempo reale.

Il sistema realizzato, tuttavia, dovrà utilizzare le soluzioni di calcolo, storage e virtualizzazione messe a disposizione da:

☐ Amazon AWS

Per il completamento del progetto, lo studente dovrà presentare una demo funzionante dell'applicazione sopra descritta.

Consegna e valutazione del progetto

La consegna del progetto consiste in:

- consegna mediante CD-ROM o link a spazio di Cloud storage contenente tutti i sorgenti necessari per il funzionamento del sistema e una relazione in formato pdf e/o dettagliata presentazione in PowerPoint;
- copia cartacea della relazione/lucidi da presentare il giorno della demo.

I principali criteri di valutazione del progetto saranno: 1) rispondenza ai requisiti; 2) originalità; 3) organizzazione del codice (es., leggibilità e modularità); 3) organizzazione e completezza della relazione.

Il progetto deve essere consegnato due giorni prima della data dell'esame.