

# Machine Learning Canvas

Product:

Generative AI for Medical Imaging

Authors:













Andrea Basile

Date:

26/06/2024

Version:

1.0

<div>Background</div> <div>Describe the customer’s goals and pains. Aumentare le dimensioni del dataset di partenza attraverso tecniche di intelligenza artificiale generativa per la sintesi di scansioni MRI pesate in T1 relative alla malattia della 'bvFTD' al fine di migliorarne la generalizzabilità.</div>	<div>Solution</div> <div>Define the solution, including features, integration, constraints and what's out-of-scope  La possibile soluzione è l'utilizzo di Generative Adversarial Networks, Variational Autoencoders e Diffusion Model pre-esistenti. Il modello dovrà essere addestrato sulle risonanze magnetiche sulle quali verranno considerate le features come: valori dei voxel, texture, colore, orientamento, pattern strutturali, etc. al fine di capirne la distribuzione di probabilità dei voxel stessi e poter riprodurre nuove risonanze in maniera sintetica. Il modello potrà essere integrato e mantenuto nel sistema che ospita già il classificatore attraverso pratiche e strumenti di MLOps. Il principale limite al corretto funzionamento del modello generativo è il dataset ridotto. la soluzione permetterà di generare nuove immagini a partire dal dataset fornito in input pertanto, altri utilizzi differenti sono da considerarsi out-of-scope.</div>	<div>Data</div> <div>Identify the training and production data sources, as well as the labeling process and decisions.  I dati necessari per addestrare il modello dovranno essere MRI scan. La fonte potenziale è il PACS (Picture archiving and communication system) dell'attività di riferimento.</div>	<div>Modeling</div> <div>List the iterative approach to model our task.  Il processo di modeling sarà costituito da: 1) utilizzo di librerie python per definire l'architettura del/dei modelli 2) addestramento dei vari modelli sul dataset in input 3) registrazione degli esperimenti e versionamento del modello 4) valutazione delle performances 5) rilascio del modello più prestante</div>	<div>Feedback</div> <div>Outline sources of feedback from our system to use for iteration. -</div>
<div>Value proposition</div> <div>Propose the product with the value it creates and the pains it alleviates.  Il modello generativo permetterebbe di ottenere nuovi dati rispetto ad una malattia rara senza coinvolgere pazienti, apparecchiature costose, tempi lunghi e esperti del dominio che effettuino il labeling</div>	<div>Feasibility</div> <div>Discuss the feasibility of the solution and if we have the required resources.  La fattibilità del progetto è fortemente dipendente da: 1) risorse computazionali a disposizione 2) tecnologie esistenti adatte al contesto  Entrambi questi fattori potrebbero essere superati tramite l'affitto di GPU in cloud e utilizzando modelli che riescono a generalizzare meglio su pochi dati (VAE)</div>	<div>Metrics</div> <div>Prioritize key metrics that reflect the objectives.  Inception Score, Frechet Inception Distance e Kernel Inception Distance</div>	<div>Inference</div> <div>Decide whether we want to do batch (offline) or real-time (online) inference. L'inferenza sarà eseguita in modalità offline.</div>	<div>Project</div> <div>Define the required team members, deliverables and projected timelines.  Il team richiesto al fine di portare avanti tale progetto deve essere costituito da: - data scientist - bioinformatici - ingegneri informatici</div>
<div>Objectives</div> <div>Breakdown the product into key objectives that need to be delivered. - recupero del dataset - creazione del modello - tracciamento degli esperimenti - sperimentazione con varianti e combinazioni - selezione del modello migliore - deploy del modello</div>	<div>Evaluation</div> <div>Design offline and online evaluation criteria.  La valutazione offline verrà fatta usando le metriche apposite e MLFlow per capirne le performances e selezionare il modello migliore.</div>			

