



Health Analytics:

configurazione di un'architettura informativa per lo studio della correlazione tra esposizione ad agenti inquinanti e l'insorgenza di sintomi clinici.

Neural Networks 4 Archives:

sviluppo di un'applicazione basata sulle reti neurali per il riconoscimento facciale in ambito archivistico e iconografico.

Studente: **Marco Ruto**

Ai fini del conseguimento del Master in Analisi Dati per la Business Intelligence e Data Science, abbiamo svolto un project work presso il dipartimento di Economia e Statistica "Cognetti-De Martiis" dell'Università degli Studi di Torino, orientato su due progetti di ricerca in ambito accademico.

Il primo, *Health Analytics*, si è proposto di studiare la correlazione tra l'esposizione ad agenti inquinanti e l'insorgenza di sintomi clinici tramite l'utilizzo di modelli regressivi per serie temporali. Questo progetto è nato dalla disponibilità di due tipologie di dati: da una parte quelli di tipo sanitario relativamente alle analisi del sangue di un campione di individui della provincia di Torino, dall'altra quelli sull'inquinamento rilevati da sensori ARPA dislocati nella città. Da questa base di partenza sono stati individuati ulteriori dati di contesto a supporto dell'analisi, con lo scopo di evidenziare quanto assunto durante la fase di costruzione del business case di progetto, ovvero lo studio della correlazione tra l'esposizione ad agenti inquinanti e l'insorgenza di sintomi clinici.

Il secondo, *Neural Networks 4 Archives*, si è invece proposto di utilizzare modelli di Reti Neurali Convoluzionali per la classificazione di risorse archivistiche e iconografiche. A fronte della disponibilità di materiale fotografico attinto dall'archivio dell'Istituto di Studi Storici Gaetano Salvemini di Torino, e grazie ad un'attività di Web Scraping, finalizzata ad elevare la qualità del layer informativo, l'obiettivo si è tradotto nell'esigenza di implementare un motore di *back-end* che, previo l'utilizzo di Reti Neurali Artificiali, tecnica madre del Deep Learning nell'ambito della classificazione di immagini, ha consentito la costituzione di algoritmi specifici di riconoscimento facciale sviluppati in Python, nonché dall'azione fondamentale di tuning e training delle Reti Neurali alle quali afferiscono.

Il framework analitico utilizzato in entrambi i progetti è stato quello del *Data Science Life Cycle*, un ciclo di sette task che partendo dalla definizione dell'obiettivo di ricerca, permette di costruire un processo analitico basato su una serie di attività che porteranno ad avere un deployment finale dei risultati dell'analisi direttamente nei sistemi informatici che hanno prodotto gli output di progetto.