



UNIVERSITÀ DEGLI STUDI
DI SALERNO

CORSO DI LAUREA IN INFORMATICA

Unisa Mortgage-AI

Autore:

Docenti:

Antonio Clavino

Polese Giuseppe

Caruccio Loredana

Prof.

Prof.ssa

Contents

1	Introduzione al problema	3
1.1	Il Machine Learning e l'accesso al credito	3
1.2	Il Contesto Italiano dei Mutui	3
1.3	Il problema degli approcci IA in ambito finanziario	3
1.4	Il problema affrontato	4
1.4.1	Il problema dell'explainability nei modelli di Machine Learning	4
1.4.2	In ambito creditizio	4
1.4.3	Il problema dei dati mancanti	4
2	Data Understanding and Data Preparation	5

1 Introduzione al problema

1.1 Il Machine Learning e l'accesso al credito

Il tema della questione finanziaria legata a mutui con l'ausilio dell'intelligenza artificiale è un tema che negli ultimi e nei prossimi anni potrà rivestire un ruolo sempre più importante. Lo scopo principale di molti studi è l'applicazione di metodi di Machine Learning per definire un modello computazionale di classificazione che selezioni una richiesta di mutuo accettabile o non accettabile basandosi su dati storici.

1.2 Il Contesto Italiano dei Mutui

Nel contesto italiano, la concessione di un mutuo rappresenta un processo articolato che tiene conto del reddito del richiedente, della stabilità lavorativa, degli impegni finanziari in corso, dell'importo richiesto e dello storico creditizio. L'obiettivo è garantire la sostenibilità del finanziamento nel tempo, nel rispetto delle normative vigenti.

Nel sistema bancario italiano, la valutazione di una richiesta di mutuo si basa su diversi fattori, tra cui il reddito annuo del richiedente, la stabilità e la tipologia del contratto lavorativo, la presenza di impegni finanziari pregressi, l'importo del mutuo richiesto, la durata del finanziamento e lo storico creditizio.

In questo scenario, la possibilità di disporre di strumenti intelligenti in grado di fornire una valutazione preliminare e personalizzata della richiesta di mutuo risulta particolarmente rilevante, sia per gli istituti bancari sia per i potenziali clienti.

1.3 Il problema degli approcci IA in ambito finanziario

Come qualsiasi approccio di Machine Learning e IA, il problema principale individuato è la reperibilità dei dati. La discrepanza tra i dati rende molti dei modelli inutilizzabili nell'effettivo, soprattutto perché c'è la necessità di poter utilizzare dati reali e non solo dati di sintesi. Inoltre, l'uso di modelli di IA poco interpretabili può ridurre la fiducia degli utenti e rendere difficile la giustificazione delle decisioni.

Inoltre, l'utilizzo di dati storici può introdurre bias sistematici, legati a pratiche decisionali passate o a squilibri presenti nel dataset. Tali bias possono tradursi in decisioni discriminatorie, con conseguenze rilevanti sia dal punto di vista etico sia normativo.

Infine, la presenza di dati incompleti o mancanti rappresenta un ulteriore problema, in quanto una gestione non adeguata di tali informazioni può compromettere l'affidabilità delle predizioni.

1.4 Il problema affrontato

Partendo da una curiosità personale, l'intenzione di questo progetto è quella di definire un modello che analizzi i dati per classificare correttamente se un mutuo può essere approvato da una banca oppure no, cercando di porre maggiore attenzione sui dati presenti nel dataset, prestando maggiore attenzione sull'explainability, cercando di trovare un approccio migliore per affrontare il problema utilizzando approcci studiati.

1.4.1 Il problema dell'explainability nei modelli di Machine Learning

Uno dei problemi principali del Machine Learning, è l'explainability del modello, ossia quanto il processo decisionale su cui si basa il modello sia comprensibile a chi lo utilizza. Oltre a migliorare la fiducia verso un approccio di IA, migliorare l'explainability di un modello porta benefici in ambito legale e di miglioramento del modello stesso. La capacità di spiegare le decisioni automatiche è fondamentale non solo per garantire la fiducia degli utenti finali, ma anche per rispettare i requisiti normativi in materia di trasparenza e accountability, migliorando l'explainability del modello si può capire il processo decisionale che sta dietro una decisione, rendendo il processo stesso trasparente e affidabile, permettendo di modificare decisioni che portano in errore il modello stesso.

1.4.2 In ambito creditizio

In ambito creditizio, la mancanza di spiegazioni chiare può generare diffidenza nei confronti dei sistemi automatici e limitare la loro adozione. Ad esempio se 2 clienti chiedono un mutuo e ad uno con le stesse caratteristiche dell'altro viene accettato il mutuo, l'altro vorrà capire le motivazioni del declino, quindi il sistema di valutazione del mutuo deve essere in grado di fornire indicazioni comprensibili sui fattori che hanno influenzato l'esito della richiesta.

1.4.3 Il problema dei dati mancanti

I dataset finanziari sono spesso caratterizzati dalla presenza di dati mancanti, dovuti a informazioni non dichiarate o non disponibili. La gestione di tali valori rappresenta una fase critica del processo di preprocessing, in quanto scelte errate possono introdurre bias o ridurre le prestazioni del modello.

2 Data Understanding and Data Preparation