

# Abstract

Come punto di partenza la definizione di rischio è stata introdotta: il rischio è il verificarsi di un evento inaspettato o con esito negativo.

Dopo una breve introduzione della maggior parte delle categorie di rischio, come banche e autorità di vigilanza, la tesi si focalizza sui modelli di rischio di credito dove l'intero sistema finanziario sta investendo per evitare la prossima crisi finanziaria.

Tra le metriche relative al rischio di credito, i Risk Weighted Assets (RWAs) possono essere considerati una rilevante misura nel panorama del rischio di credito. Infatti, questi, rappresentano una misura aggregata di vari fattori di rischio che influenzano la valutazione dei prodotti finanziari.

L'accuratezza dei modelli di rischio di credito, come tutti i modelli, non dipende solo dalla parametrizzazione e la complessità di un modello, ma anche dai dati utilizzati come input. Ciò è spesso riassunto da "Garbage IN uguale a Garbage OUT". Per questo motivo l'attenzione è stata focalizzata sulle tecniche di data quality.

Nel secondo capitolo, alcune tecniche di machine learning per analizzare le anomalie nei dati sono state introdotte. Una particolare attenzione è stata riservata ai Local Outlier Factor (LOF) ed Isolation Forests (IF).

Nel terzo e quarto capitolo, questi algoritmi sono stati testati prima su di una serie di osservazioni generate artificialmente in modo da mostrarne le loro proprietà statistiche e poi su un dataset reale con informazioni relative al rischio di credito dove le anomalie relative agli RWAs sono state analizzate.