RADemics

Fraud Detection in Insurance Claim Processing Using Predictive Analytics and AI Models

Neha Nivrutti Jamdar, A. Devi
Pune Institute of Computer Technology Pune,
Hyderabad Institute of Technology and Management

Fraud Detection in Insurance Claim Processing Using Predictive Analytics and AI Models

[1]Neha Nivrutti Jamdar, Pune Institute of Computer Technology Pune, Maharashtra, India. nehajamdar1@gmail.com

[2]A. Devi, Assistant Professor, Department of CSE, Hyderabad Institute of Technology and Management, Hyderabad, Telangana, India. devim.cse@hitam.org

Abstract

Insurance fraud has emerged as a critical operational and financial threat within modern insurance claim processing systems, driven by rapid digitization, increasing transaction volumes, and evolving cyber-enabled deception techniques. Traditional rule-based fraud detection mechanisms exhibit limited capability in identifying complex and adaptive fraudulent behaviors, necessitating the adoption of intelligent data-driven solutions. This book chapter explores the application of predictive analytics, machine learning, and deep learning models for robust fraud detection in insurance claim environments. Advanced analytical frameworks leveraging supervised learning, ensemble methods, neural networks, and multimodal data processing enhance the ability to detect hidden fraud patterns across structured, unstructured, and sequential insurance data. Emphasis remains on hybrid deep learning architectures, dimensionality reduction techniques, and feature engineering strategies that improve classification accuracy and computational efficiency in large-scale insurance datasets. The study further highlights real-time fraud detection systems, explainable artificial intelligence mechanisms, and privacy-preserving technologies such as federated learning and blockchain integration to address transparency, security, and regulatory compliance challenges. Comparative evaluation of predictive models demonstrates improved fraud identification performance through intelligent feature extraction and adaptive learning strategies, while also addressing limitations such as data imbalance, model interpretability, and evolving fraud tactics. The findings indicate that AI-driven predictive frameworks significantly strengthen insurance fraud prevention systems by enabling accurate, scalable, and automated decision-making in claim assessment processes.

Keywords: Insurance Fraud Detection, Predictive Analytics, Machine Learning, Deep Learning, Feature Engineering, Explainable AI

Introduction

The insurance sector operates as a fundamental pillar of global financial systems by providing risk coverage against uncertain events such as accidents, health emergencies, natural disasters, and property loss [1]. Rapid digital transformation within insurance operations has introduced automated claim processing systems, online policy management platforms, and data-driven customer engagement models [2]. Expansion of digital infrastructures has significantly improved operational efficiency, reduced processing time, and enhanced accessibility for policyholders. At the same time, increased dependency on digital ecosystems has expanded exposure toward

fraudulent claim activities [3]. Insurance fraud has evolved into a complex financial threat that impacts organizational profitability, increases operational burden, and reduces trust across stakeholders. Growing volume of insurance transactions, coupled with real-time claim submissions, demands advanced analytical mechanisms capable of handling large-scale heterogeneous datasets [4]. Conventional approaches struggle to manage this complexity, creating a strong need for intelligent fraud detection systems capable of adapting to dynamic insurance environments [5].

Insurance fraud encompasses intentional acts of deception designed to obtain unlawful financial benefits through manipulation of claim processes [6]. Fraudulent activities occur across multiple insurance domains including health, automobile, property, and life insurance [7]. Common fraudulent behaviors include staged accidents, inflated medical billing, fabricated documentation, and identity manipulation during claim submission [8]. Increasing digitization has enabled cyber-enabled fraud techniques involving synthetic identity creation, document forgery, and unauthorized access to insurance databases [9]. Organized fraud networks further intensify complexity by coordinating multiple actors such as policyholders, intermediaries, and service providers. Traditional rule-based systems exhibit limited adaptability in detecting such evolving fraud patterns due to static decision logic and dependency on predefined rules [10]. High variability in fraud strategies requires intelligent systems capable of learning from historical data and identifying hidden behavioral patterns across diverse claim attributes.