



AI-Based Insurance Claim Processing and Verification System

Jayashri Kharat¹

Student, MBA Department
Patil College of Engineering, Pune

Prof. Kanifnath S Satav²

Project guide, MBA Department Dhole
Dhole Patil College of Engineering, Pune

Author Email: jkharat13897@gmail.com | ORCID: <https://orcid.org>

How to Cite this Article:

Kharat, J. (2026). AI-Based Insurance Claim Processing and Verification System. International Journal of Creative and Open Research in Engineering and Management, <i>02</i><i>(04)</i>. <https://doi.org/10.55041/ijcope.v2i4.927>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i4.927>

Abstract

Health insurance claim processing is one of the most important activities in the insurance industry, as it directly affects both insurance companies and policyholders. In the traditional claim processing system, claim verification, document checking, policy validation, and fraud investigation are mostly done manually. This process takes a lot of time, increases operational costs, and may lead to human errors. In addition, the increasing number of fraudulent claims has become a major challenge for insurance companies, resulting in financial losses and delays in genuine claim settlements. Therefore, there is a growing need for an intelligent and automated system that can improve the efficiency, accuracy, and security of health insurance claim processing.

This project presents an **AI-Based Health Insurance Claim Processing and Verification System** that uses Artificial Intelligence (AI) and Machine Learning (ML) techniques to automate and improve the claim management process. The main aim of the system is to verify claims, detect fraud, and predict claim approval or rejection automatically. The system uses health insurance claim data such as patient age, gender, disease type, treatment cost, policy limit, hospital type, and previous claim history for analysis and prediction.

The project follows a structured methodology that includes data collection, data cleaning, preprocessing, exploratory data analysis (EDA), model building, testing, and result evaluation. During preprocessing,

missing values are handled and categorical variables are converted into numerical form so that the machine learning model can process the data efficiently. For model implementation, the Random Forest algorithm is used because of its high accuracy, reliability, and ability to handle classification problems effectively.

The system is divided into three major modules: claim verification, fraud detection, and claim approval prediction. In the claim verification module, the treatment cost is checked against the policy limit to ensure claim eligibility. In the fraud detection module, the machine learning model identifies whether a claim is genuine or fraudulent by analysing patterns in the claim data. In the claim approval prediction module, the system predicts whether a claim should be approved or rejected based on the available claim information.

The results of the project show that the AI-based system performs effectively in improving the overall insurance claim process. The fraud detection model achieved good accuracy in identifying suspicious claims, while the claim approval model successfully predicted claim decisions with high reliability. The automated verification system also reduced the time required for manual checking and improved the speed of claim settlement.

Keywords: Artificial Intelligence, Machine Learning, Health Insurance, Fraud Detection, Claim Verification, Predictive Analytics



I. INTRODUCTION

The healthcare insurance industry handles millions of claims every year. Processing these claims manually creates several operational challenges, including document verification delays, claim fraud, and inconsistent decision-making.

Artificial Intelligence (AI) and Machine Learning (ML) provide opportunities to automate and optimize claim processing systems. AI can analyse claim patterns, validate policy conditions, and identify suspicious claims more effectively than manual systems.

This study proposes an AI-based system for health insurance claim processing and verification to improve efficiency, speed, and fraud detection.

Insurance claim processing is an essential function in the insurance sector. It includes receiving claims, verifying policy details, checking supporting documents, detecting fraud, and approving settlements.

Traditional claim processing faces several problems:

- Manual verification consumes time
- High operational cost
- Human errors in data entry
- Fraudulent claims
- Poor customer experience

AI technologies provide automation and intelligence in claim processing. Machine learning models can analyse claims data, identify fraud patterns, and verify policy compliance quickly.

The increasing digital transformation in insurance has made AI adoption essential for competitive advantage. AI-based claim systems can process claims faster and improve decision-making quality. Research highlights that AI systems use OCR, NLP, and predictive analytics to streamline workflows and reduce administrative burden.

II. LITERATURE REVIEW

Different researchers have explained that machine learning helps in analyzing large amounts of claim data and finding patterns that can identify fraudulent activities and predict claim outcomes. Tom Mitchell explained that machine learning systems improve automatically through experience and data analysis, which makes them useful for prediction tasks. Ian Goodfellow showed that deep learning methods are effective in identifying complex patterns in data, which can improve fraud detection systems. Research published by IEEE highlights that AI-based automation reduces processing time and improves claim verification accuracy. Studies from Springer Nature also show that predictive analytics helps in detecting suspicious claims and improving operational efficiency. Many organizations like IBM have developed AI-powered systems for automating document verification and policy validation. However, most existing studies focus mainly on fraud detection and do not provide a complete system that combines fraud detection, claim verification, and claim approval prediction together. This project focuses on filling that gap by developing an integrated AI-based Health Insurance Claim Processing and Verification System.

III. METHODOLOGY

The methodology follows these steps:

Step 1: Data Collection

Collect insurance claims dataset.

Step 2: Data Preprocessing

Clean missing values and normalize data.

Step 3: Feature Engineering

Extract important variables.

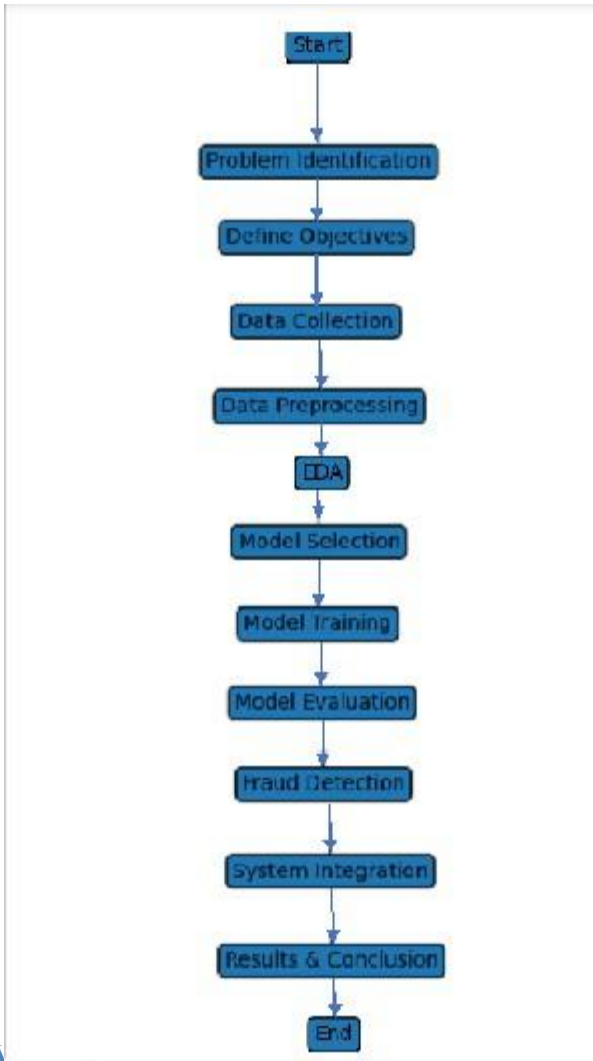
Step 4: Model Building

Train machine learning models. **Step 5: Fraud Detection** Identify suspicious claims.

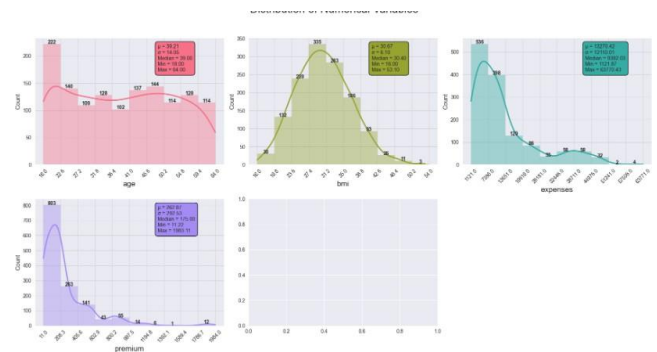
Step 6: Claim Verification

Validate documents and policy conditions. **Step 7: Claim Approval/Rejection** Generate decision.

Distribution of Numerical Variables



Variable	Observation
Age	Age values are fairly distributed across different groups, with slightly higher concentration in younger ages
BMI	Most BMI values are concentrated around the middle range (around 30)
Expenses	Expenses are highly concentrated at lower values, with fewer high-expense cases
Premium	Premium values are concentrated at lower levels, with very few high-premium cases



The dataset was collected and loaded into Python.

Categorical data (gender, disease, hospital type) were converted into numerical form.

The dataset was divided into training data (80%) and testing data (20%).

Random Forest was used to train the model. Two models were created:

- Fraud Detection Model
- Claim Approval Prediction Model

Final Results

- Faster claim processing
- Better fraud detection
- Improved accuracy
- Better customer satisfaction



V. CONCLUSION

The AI-Based Insurance Claim Processing and Verification System demonstrates how artificial intelligence can significantly transform traditional insurance operations. By automating claim analysis and verification, the system reduces processing time, minimizes human errors, and improves overall efficiency.

The implementation of Machine Learning techniques enabled accurate identification of fraudulent claims based on key factors such as claim amount, policy status, document submission, and claim history. The system not only enhances fraud detection but also supports faster and more reliable decision-making.

Although the model was developed using synthetic data, the results indicate strong potential for real-world application. With further improvements such as integration of real datasets, advanced algorithms, and continuous model training, the system can become even more robust and effective.

In conclusion, this project highlights the importance and future scope of AI in the insurance sector. It provides a scalable, efficient, and intelligent solution that can help insurance companies improve operational performance, reduce fraud, and deliver better customer service.

REFERENCES

- [1] AI-Powered Claims Processing Research Journal
- [2] IEEE Research Papers on Fraud Detection.
- [3] Tom Mitchell (1997). *Machine Learning*. McGraw-Hill Education.
- [4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. MIT Press.
- [5] IEEE. Research on AI-based Insurance Fraud Detection and Claim Automation.
- [6] Springer Nature. Studies on Predictive Analytics in Health Insurance.

- [7] IBM. AI-Based Claims Automation Solutions.
- [8] Fraud Detection in Healthcare Insurance Claims Using Machine Learning (2023), published in MDPI Risks Journal.
- [9] Fraud Detection in Healthcare Claims Using Machine Learning: A Systematic Review (2025), published in Artificial Intelligence in Medicine.