International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

# Transforming CRM Warranty Management with Advanced Analytics and AI

**Savio Dmello**

Independent Researcher with Expertise in Business Intelligence
Email: *Saviodmo[at]gmail.com*

**Abstract:** ***Traditional Business Intelligence and Analytics in customer relationship management (CRM) and the service industry are essential for delivering exceptional customer service and ensuring prompt resolution of warranty claims. As customer service and warranty management processes evolve, maintaining the integrity of warranty claims becomes increasingly critical for preserving business trust and financial stability. This paper investigates the application of advance analytics and machine learning techniques to develop anomaly detection models that identify fraudulent patterns in warranty claims data. By analyzing customer service interactions and contract details within a CRM system, this study employs AI, data mining techniques, the Isolation Forest algorithm, and K-Means clustering to detect anomalies based on temporal and geographical patterns in claims. The methodology focuses on key performance indicators (KPIs) such as the timing of Annual Maintenance Contract (AMC) purchases and the clustering of claims by geographic locations associated with sales representatives and other business processes. The findings reveal significant correlations between fraudulent claims, submission timing, customer history, interactions and specific geospatial patterns, indicating potential collusion between sales representatives and customers. This paper contributes to the detection of malpractice in service-oriented industries and provides valuable insights for businesses seeking to enhance their warranty management processes through advanced data analytics.***

**Keywords:** Business Intelligence, Customer Relationship Management (CRM), Anomaly Detection, Artificial Intelligence (AI), Machine Learning, Fraud Detection, Warranty Claims, Geospatial Analysis

## 1. Introduction

In this paper on CRM analytics, the primary focus is to identify patterns associated with Annual Maintenance Contracts (AMC), particularly in relation to warranty start and end dates and customer location. This analysis uncovered instances of fraudulent claims and malpractice, often involving business partners colluding with customers or manipulating customer data related to household electronics items. However, many fraudulent patterns remain obscured within the vast volumes of customer data, making them difficult to detect using traditional methods. This raises a critical question: what additional strategies can be employed to identify fraudulent claims, malpractice, or anomalies within the warranty claim process? To address this challenge, it is first essential to understand the interlinked processes of CRM claims and warranty management, and how these systems collectively contribute to anomaly detection. Before leveraging AI capabilities, it is crucial to comprehend the structure of CRM data and the inherent limitations of CRM systems in detecting fraudulent activities.

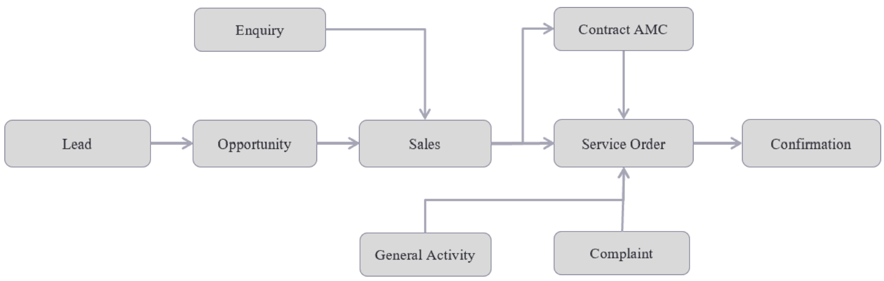



**Figure 1:** CRM Service and Warranty Process flow

Its important to understand the CRM service process.

a) Leads represent potential customers who have shown interest in the company’s offerings. The CRM system tracks and nurtures these leads to convert them into sales opportunities through targeted marketing efforts.
b) Opportunities arise when leads are qualified for potential sales. The CRM system enables tracking and managing these opportunities, helping sales teams prioritize efforts to convert prospects into customers.
c) The sales process encompasses all activities involved in converting opportunities into actual sales, including negotiation, contract signing, and payment processing, ensuring a smooth transition from interest to purchase.
d) Contracts, such as Annual Maintenance Contracts (AMC), are formalized agreements outlining the terms of service, pricing, and obligations between the company and the customer. This documentation helps ensure clarity and compliance.
e) A service order is created to manage customer claims related to services or products. This process includes details about the service requested, timelines, and the personnel responsible, facilitating efficient fulfillment of customer needs.
f) General activities encompass all customer interactions and engagements that are not tied to specific documents, including follow-ups and account management tasks,

providing a comprehensive view of customer relationship dynamics.

g) Confirmation documents are generated after services are delivered to acknowledge completion and verify customer satisfaction. This step is crucial for maintaining accurate service records and building customer trust.
h) Complaints are logged in the CRM system to capture customer dissatisfaction regarding products or services.

## 2. Evaluation & System integration

Since the traditional CRM systems are not capable of AI clustering, 1 year of sample data associated with each of the listed CRM process needs to be loaded to Google Cloud Vertex AI to leverage the AI capabilities.

1 year of sample data was split with 70% for training AI models, 15% for validation, and 15% for testing the inference.

To test the effectiveness of different AI model and algorithm, data was segregated based on the type of data. Mainly the time dependent data and location dependent data, Customer historical data, Sales person to Customer interactions etc.

Furthermore, Meaningful features from your dataset were created.

- Time Since AMC Purchase: Calculate the difference in days between the AMC purchase date and the claim date.
- Customer Complaint History: A count or binary indicator of prior complaints filed by the customer before the AMC purchase.
- Claim Frequency: Number of claims made by the customer in the last year.
- Geographical Information: Categorical features for the geographical location of the customer.
- Sales Representative Performance: Metrics indicating the sales representative's average claim approval rate.

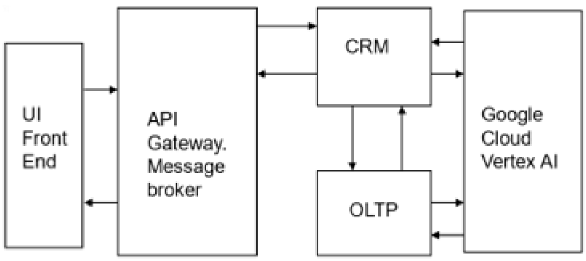



**Figure 2:** System Integration

## 3. Analysis

### 3.1 Timing Analysis

An important factor in identifying potential warranty fraud was timing. For example, if a customer purchased an AMC (Annual Maintenance Contract) and filed a warranty claim within 30 days, this was considered an indicator of potential fraud. A typical case involved a customer with a history of complaints suddenly purchasing a warranty and immediately filing a claim for a long-standing issue. This behavior raised red flags. To detect such patterns, an AI model was developed and trained to flag claims using time series analysis and anomaly detection algorithms. By analyzing historical claim data, customer enquires and interactions, customer credit risk, the model identified unusual patterns based on the time elapsed since the AMC purchase.

### 3.1.1 Model Selection and Training Approach

To address the timing analysis for detecting potential warranty fraud, a combination of machine learning techniques, particularly anomaly detection and classification models, was implemented. The following models were considered based on feedback from AI model repositories.

#### 3.1.1. Anomaly Detection Models

- Isolation Forest: This is effective for identifying outliers in the dataset. It works well when you have a large dataset with potentially many features.
- One-Class SVM: Suitable for cases where the majority of your data is normal, and need identification of anomalies.
- Local Outlier Factor (LOF): Useful for detecting anomalies based on the local density of data points.

#### 3.1.2. Classification Models

For predicting whether a claim was fraudulent based on historical data, the following models were considered:

- Logistic Regression: A good starting point for binary classification tasks.
- Random Forest Classifier: Effective in avoiding overfitting and handling feature importance.
- Gradient Boosting Machines (e.g., XGBoost, LightGBM): These models were ideal for structured data and capable of capturing complex patterns.

For anomaly detection, the model was trained on a "normal" subset of data (valid claims) to learn typical patterns and flag anomalies. In contrast, for classification tasks, historical claims were labeled as "fraudulent" or "non-fraudulent," allowing the models to learn from past cases.

#### 3.1.3. Isolation Forest for Fraud Detection

In the context of warranty fraud detection, where fraudulent claims were relatively rare compared to legitimate claims, the Isolation Forest model was particularly effective. It identified outlier claims based on historical patterns and behaviors, making it suitable for detecting anomalies in CRM analytics.

#### 3.1.4. Model Evaluation

After training, the models were evaluated using the following metrics:

- Confusion Matrix: This helped assess the model's performance in terms of true positives, false positives, true negatives, and false negatives.
- Precision and Recall: These metrics were crucial in fraud detection, where identifying fraudulent claims (precision) and ensuring that most fraudulent claims were caught (recall) were equally important.

#### 3.1.5. Data Representation and Visualization

The training data was organized and visualized using graphs and sample formulas to analyze warranty claims in relation to AMC purchases.

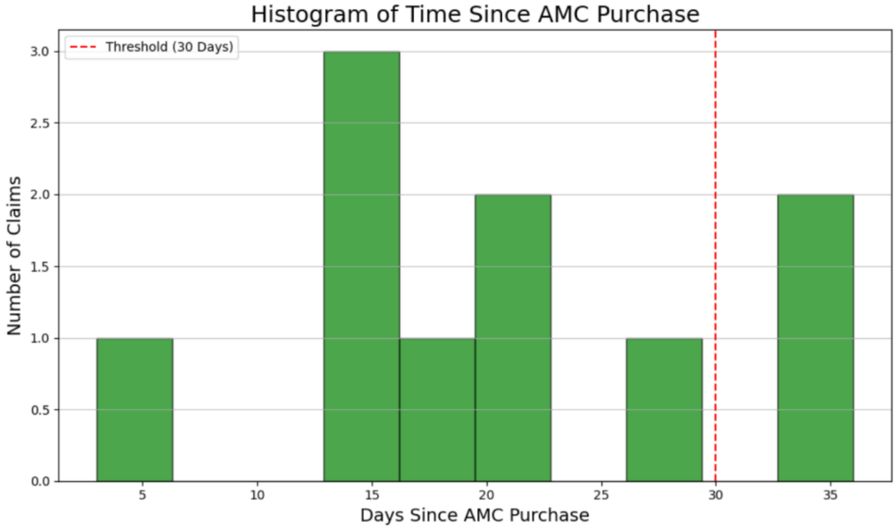
**Table 1:** Sample Claims data

| *Claim ID* | *AMC Purchase Date* | *Claim Date* | *Time Since AMC (Days)* | *Complaint History (Count)* | *Location* |
|---|---|---|---|---|---|
| 1 | 01/01/2023 | 01/15/2023 | 14 | 1 | Downtown Metro |
| 2 | 01/01/2023 | 02/05/2023 | 35 | 0 | North Suburban |
| 3 | 01/15/2023 | 01/30/2023 | 15 | 2 | South Urban |
| 4 | 02/01/2023 | 02/28/2023 | 27 | 3 | East Rural |
| 5 | 01/15/2023 | 01/18/2023 | 3 | 0 | West Small Town |
| 6 | 01/20/2023 | 02/10/2023 | 21 | 1 | Central City |
| 7 | 01/05/2023 | 01/25/2023 | 20 | 4 | Northeast Suburb |
| 8 | 02/10/2023 | 03/01/2023 | 19 | 2 | Southeast Urban |
| 9 | 01/10/2023 | 02/15/2023 | 36 | 1 | Southwest Rural |
| 10 | 02/01/2023 | 02/15/2023 | 14 | 0 | Northwest Small Town |

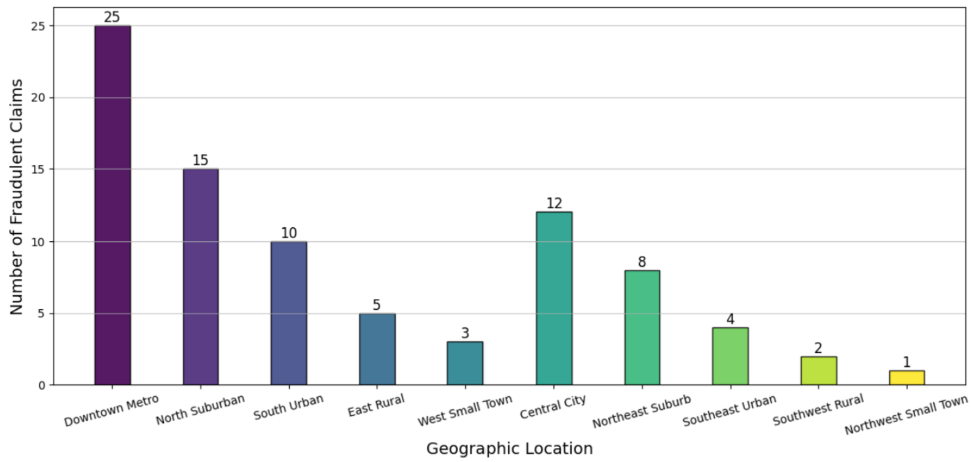
### 3.1.6. Graphs

- Histogram of Time Since AMC Purchase: This histogram showed the distribution of claims based on the time elapsed since the AMC purchase, allowing for visual identification of potential fraud patterns.
- Bar Chart of Fraudulent Claims by Geography: This bar chart visualized fraudulent claims by region, aiding in the identification of suspicious hotspots.

Sample data for histogram




Bar Chart of Fraudulent Claims by Geography
This bar chart visualizes the number of fraudulent claims by geographic region, helping to identify any suspicious hotspots.




### 3.1.7. Formulas

Some key formulas that were applied in the analysis:

- Time Since AMC Purchase:
  *Time Since AMC (Days)=Claim Date-AMC Purchase Date*
- Fraud Rate Calculation:
  *Fraud Rate=((Number of Fraudulent Claims)/Total Claims ) ×100*
- Sales Representative Performance:
  *Performance (%)=(Approved Claims/ Total Claims)×100*
- Anomaly Detection Score (using Isolation Forest): For each claim:
  *Anomaly Score=Isolation Forest(features)*

where features include Time Since AMC, Complaint History, etc.

The formulas provided the quantitative metrics necessary for deeper analysis, helping you to build a comprehensive AI model that flags suspicious warranty claims effectively.

### 3.2. Geographical Patterns

In the CRM process, multiple sales representatives sometimes overlap in their sales regions, resulting in several representatives being assigned to the same customer. A geographical analysis was conducted to reveal suspicious patterns. For example, when a region had a disproportionately high number of claims shortly after AMC purchases, this indicated potential collusion between sales representatives and customers. If a specific region or city has high number of customers claim rate identified from " Timing analysis". It pointed to possible malpractice by sales representative.

To analyze such patterns, clustering algorithms like K-means were applied to group claims based on geographical data and the timing of warranty claims. Anomaly detection models highlighted regions deviating from expected norms.

#### 3.2.1. K-means Clustering for Geographical Analysis

K-means clustering was used to group customer locations based on their geographical coordinates (latitude and longitude). The centroid of each cluster represented the likely location of the sales representative associated with these customers.

To calculate the centroid (midpoint) of a set of geographic points, following formula was used.

$$\textit{Given a set of points } (x1, y1), (x2, y2), \dots, (xn, yn)$$
$$\textit{the centroid } (C_x, C_y) \textit{ is calculated as}:$$

$$C_x = \left(\frac{1}{n}\right) * \sum_{i=1}^{n} x_i$$

$$C_y = \left(\frac{1}{n}\right) * \sum_{i=1}^{n} y_i$$

*Where:*

$$C_x = x - \textit{coordinate of the centroid (latitude)}$$
$$C_y = y - \textit{coordinate of the centroid (longitude)}$$
$$n = \textit{number of customer locations}$$

**Table 2:** Coordinates for Sample customer addresses were identified

| *Customer Location* | *Latitude* | *Longitude* |
|---|---|---|
| A | 34.0522 | -118.2437 |
| B | 34.0525 | -118.244 |
| C | 34.052 | -118.243 |
| D | 34.053 | -118.245 |
| E | 34.054 | -118.246 |

Calculate Centroid:

$$Cx = (34.0522 + 34.0525 + 34.0520 + 34.0530 + 34.0540)/5 = 34.05254$$
$$Cy = (-118.2437 + -118.2440 + -118.2430 + -118.2450 + -118.2460)/5 = -118.24384$$

The centroid would be approximately at Latitude: 34.05254 Longitude: -118.24384

This point can be considered the likely location of the sales representative related to the identified fraudulent patterns.

## 4. Learning

By combining machine learning techniques like Isolation Forest for anomaly detection and K-means clustering for geographic analysis, a robust system for detecting fraudulent warranty claims was developed. These methods provided an effective means of identifying suspicious patterns and behaviors, aiding in the prevention of fraud within CRM analytics.

## 5. Other AI solutions

Other AI Solutions that are already proved by other researcher in other fields that can be further evaluated and applied to improve accuracy in CRM warranty process.

### 5.1 Customer History Analysis:

Example: Customers with frequent complaints or repair history that subsequently purchase an AMC and file warranty claims should be scrutinized. For instance, if a customer had ongoing issues with their equipment and then bought a warranty, followed by a claim, it indicates a pattern.

AI Approach: Use classification models (like decision trees or random forests) to identify customers likely to commit fraud based on their historical service records. Features could include previous complaint frequencies, types of repairs, and the timing of AMC purchases.

### 5.2 Duplicate Claims:

Example: A customer or business partner might file multiple claims for the same issue. For instance, if a customer files a warranty claim for a faulty machine, and the same issue is reported again shortly after, it could indicate manipulation.

AI Approach: Implement natural language processing (NLP) techniques to analyze text data from claims and match descriptions. Machine learning models can help identify duplicate or similar claims based on descriptions, timestamps, and involved parties.

### 5.3 Unusual Sales Representative Activity:

Example: If a sales rep has a higher-than-average claim approval rate, it may suggest unethical practices. For instance, if one sales representative has an unusually high percentage of their customer claims getting approved shortly after AMC purchase, it may warrant further investigation.

AI Approach: Use regression analysis to evaluate the relationship between individual sales representative performance and the claims filed by their customers. Identify outliers who have significantly higher approval rates than their peers.

**5.4 Non-compliance with Policy:**

Example: Claims made on products that clearly fall outside the warranty scope or after the warranty period. For instance, if a customer claims a warranty on equipment that was not covered under the AMC terms, it needs to be flagged.

AI Approach: Develop rule-based systems that check compliance against the predefined warranty terms and conditions. Machine learning can also be applied to predict the likelihood of claims being valid based on historical data.

**5.5 Analysis of Customer Interactions:**

Example: Review interactions between customers and sales representatives to identify suspicious communications. For example, if a sales rep discusses warranty terms in a manner suggesting manipulation, it could be a sign of collusion.

AI Approach: Implement sentiment analysis and text mining on interaction logs (emails, calls) to uncover potential collusion or manipulation language. Anomaly detection on communication patterns can further highlight unusual behaviors.

## 6. Continuous Monitoring and Improvement

Once deployed, the model's performance was continuously monitor and updated with new data. A feedback loop was setup where the flagged cases can be manually reviewed and used to retrain the model, improving its accuracy over time.
By following this structured approach, an effective AI model was developed to identify suspicious warranty claims based on timing analysis and potentially flag fraudulent activities efficiently.

## 7. Conclusion

This paper demonstrated the effective application of AI and machine learning techniques, specifically anomaly detection models such as Isolation Forest and K-means clustering, in enhancing warranty management within CRM systems. By analyzing temporal patterns like the timing of Annual Maintenance Contract (AMC) purchases and geographical clustering of claims, the study successfully identified fraudulent behaviors and potential collusion between sales representatives and customers. The results indicated a strong correlation between fraudulent claims, early submission of warranty claims post-AMC, and suspicious regional patterns.

The findings also emphasized the importance of integrating AI models into traditional analytics systems for better insights. The models provided valuable insights, highlighting hidden patterns that traditional methods struggled to uncover. Additionally, the geographical clustering of claims helped identify regions and sales representatives involved in questionable activities, offering actionable insights for management and process improvement.

This paper not only reinforced the potential of AI in detecting fraudulent warranty claims but also paved the way for future enhancements in warranty management. Businesses can use these insights to develop more transparent, efficient, and secure CRM processes, reducing financial losses and maintaining trust with customers.

## References


[1] Santos, C., & Isaias, P. (2016). After-sales and CRM: Their role as a differentiation strategy for clients. In *2016 Multi Conference on Computer Science and Information Systems, International Association for Development of the Information Society* (pp. 66-72).

[2] Hasan, I., Rizvi, S. (2022). AI-Driven Fraud Detection and Mitigation in e-Commerce Transactions. In: Gupta, D., Polkowski, Z., Khanna, A., Bhattacharyya, S., Castillo, O. (eds) Proceedings of Data Analytics and Management . Lecture Notes on Data Engineering and Communications Technologies, vol 90. Springer, Singapore. https://doi.org/10.1007/978-981-16-6289-8_34

[3] M. Alkathiri, J. Abdul and M. B. Potdar, "Kluster: Application of k-means clustering to multidimensional GEO-spatial data," *2017 International Conference on Information, Communication, Instrumentation and Control (ICICIC)*, Indore, India, 2017, pp. 1-7, doi: 10.1109/ICOMICON.2017.8279080.

[4] AI and Machine Learning Techniques Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation forest. *IEEE International Conference on Data Mining*, 413-422. https://doi.org/10.1109/ICDM.2008.17

[5] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of statistics*, 1189-1232. https://doi.org/10.1214/aos/1013203451

[6] Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing*, 69(4), 167-176. https://doi.org/10.1509/jmkg.2005.69.4.167

[7] Wang, C. H., & Chang, C. H. (2007). A case study of integrating data mining techniques and CRM system for customer relationship management in a retail store. *Expert Systems with Applications*, 33(1), 2-3. https://doi.org/10.1016/j.eswa.2006.04.007

[8] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, 41(3), 1-58. https://doi.org/10.1145/1541880.1541882

[9] Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235-255. https://doi.org/10.1214/ss/1042727940

[10] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, 31(8), 651-666. https://doi.org/10.1016/j.patrec.2009.09.011

[11] Murthy, D. N. P., & Blischke, W. R. (2006). Warranty management and product manufacture. *Springer Science & Business Media*. https://doi.org/10.1007/978-1-84628-289-5

[12] SAP SE. (2023). *SAP Customer Relationship Management (Version 7.0) [Software]*. SAP SE. https://www.sap.com

[13] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). *Isolation forest*. In 2008 Eighth IEEE International Conference on Data Mining (pp. 413-422). IEEE. https://doi.org/10.1109/ICDM.2008.17

[14] Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). *Estimating the support*

*of a high-dimensional distribution*. Neural Computation, 13(7), 1443-1471. https://doi.org/10.1162/089976601750264965

[15] Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). *LOF: Identifying density-based local outliers*. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (pp. 93-104). ACM. https://doi.org/10.1145/342009.335388

[16] Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). ACM. https://doi.org/10.1145/2939672.2939785