International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0176
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1174 - 1179

A Study on Self-Uncertainty Detection in AI Systems Using Unsupervised Learning Techniques

Alisha Jabeen[1], Ansif Azeez[2], Abhinav S[3], Jenas Renjan Parakkal[4], Adhwaith K G[5], Anandha Krishnan[6]
[1]Assistant Professor, Department of Computer Science and Applications, Yenepoya Deemed to be University, Bangalore, Karnataka, India.
[2,3,4,5,6]BCA (AI, ML and Robotics), Yenepoya Deemed to be University, Bangalore, Karnataka, India.
Emails: *alishazabeen@gmail.com[1], ansifazeez1788@gmail.com[2], abhxn718@gmail.com[3], jenasrenjan2005@gmail.com[4], adhwaithkg1313@gmail.com[5], anandhakrishnant+work@outlook.com[6]*

Abstract
Artificial intelligence (AI) systems are being used more and more in practical applications like autonomous systems, healthcare diagnostics, and financial fraud detection. Many AI models are highly predictive, but they are unable to identify situations in which their forecasts could not be accurate. When the input data deviates from the distribution shown during training, this restriction may result in overconfident conclusions. This work looks at how we can detect uncertainty in data with unsupervised learning method. Here, Isolation Forest algorithm is used to identify unusual data points by providing the scores of anomalies. These scores basically tell how and much a transaction is different from normal behavior. For testing Credit Card Fraud Detection dataset was used, which has around 284,807 transactions with 30 features. Before using the model, the data was scaled to make all values are in a similar range. To see the output in a better way, PCA and the histogram plot were used. These helped in seeing how the scores of the anomalies are distributed and where the unusual points are located. From the results, it can be noticed that such methods are able to identify data points that don't follow the usual pattern. These points may represent cases where the predictions of the model are not very reliable. So, this approach can be useful to improve how much we can trust the AI systems, especially when dealing with real-world data.

Keywords: *Artificial Intelligence Reliability, Self-Uncertainty Detection, Unsupervised Learning, Isolation Forest, Anomaly Detection, and Credit Card Fraud Detection.*

1. Introduction
Artificial intelligence (AI) is currently being used. applied in practice in numerous fields like medical care, finance, and even cybersecurity. autonomous systems. In many of these cases, people use AI to make crucial decisions. decisions. Nevertheless, the majority of machine learning accuracy is the key reason why models are constructed, and they will not normally be seen when their. predictions might be wrong. This may turn out to be a problem when the input data is quite different to what the model possesses. seen during training. In such situations, the even when model gives a certain output we can be sure it is a good one. though it may not be correct. For example, in new types of transaction, financial systems. trends are emerging which were not included in the. training data. So, identifying such cases is vitally necessary to keep a faith in AI systems. A method of managing this is through data finding. Points that are not normally distributed. They are referred to as anomalies or outliers. The unsupervised learning techniques would come in handy in this situation since they are able to pick such patterns. and without labelled data. In this writing, we apply the Isolation Forest. abnormality detection algorithm. It gives an anomaly score per data point, a <|human|>where a is the anomaly score per data point, a. greater score implies data is even more unusual. The dataset used to detect the Credit Card Fraud was. used in testing and scaling of features was. applied before training. To understand the appear more successfully, PCA and histogram plots were being used. to picture the way the anomaly scores are. distributed. On the whole, this strategy assists in determining. grey cases and are capable of enhancing the. AI system reliability, particularly in real. world conditions [1].

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0176
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1174 - 1179

2. Related Work

Looking more specifically in the recent years, there was an increased interest in understanding uncertainty of AI systems, particularly on critical domains like healthcare, finance and cybersecurity. Though traditional machine learning models often improve accuracy, they do not always specify when a prediction may be unreliable. There are various suggestions on how to solve this problem. This was the objective to address, in an attempt with the Bayesians; by incorporating uncertainty through treating all model parameters as probabilities instead of fixed values. Other approaches such as Monte Carlo dropout and Bayesian neural networks are used, but they can be more complex to implement and involve additional computation. Ensemble methods take a different approach by combining multiple models. One possible sign of uncertainty is when model outputs vary for the same input. This improves performance but also adds complexity and increases costs. The unsupervised learning method is relatively easier than these. And rather than measuring uncertainty directly, it looks for unusual disparities in the data. While they don't provide specific values for uncertainty, they can be useful in pinpointing inputs on which the model might perform poorly [2].

3. Methodology

This study proposes an unsupervised learning mechanism to examine self-uncertainty using the AI system. To detect data points that do not conform to the norm and therefore may suggest uncertain predictions. Executing the methodology is a few simple steps. Perhaps, the dataset is prepared and pre-processing has been done. Although its not part of the problem statement, it is done for better data fitting. Afterward, anomaly scores are produced for all the data points allowing us to recognize anomalies. The final step is to use visualization techniques (such as PCA) for a better understanding of the distribution of such anomalous points within the dataset. The first step is defining the problem and preparing the dataset. The next step is to use the Isolation Forest algorithm to detect anomalies in the data. The model generates anomaly scores, which indicate the rarity of the data point. Since the results can be more difficult to interpret, visualization methods like Principal Component Analysis (PCA) are applied. This aids in showcasing the distribution of anomalous data points compared to that of normal data. Overall, this method allows for uncertainty detection through emphasis on deviant behaviors.

3.1. Problem Definition

The main aim of this study is to understand how an AI system can identify situations where its predictions may not be reliable. In most cases, machine learning models are trained using past data and are expected to perform well on new inputs. However, real-world data often contains patterns that are different from what the model has seen before. In such situations, the model may still give confident predictions, even though they might not be correct. In this work, such cases are treated as uncertainty, especially when the data points significantly differ from the learned patterns. This study tackles the issue by finding anomalous observations in the dataset rather than directly computing probabilistic uncertainty. Predictions linked to a data point may also be less trustworthy if it seems out of the ordinary in relation to the bulk of the training data. As a result, identifying anomalies can offer helpful clues on possible ambiguity in AI systems [3].

3.2. Data Preprocessing

The dataset is preprocessed to make sure all input features are on a similar scale before the anomaly detection model is used. Numerical features that indicate various transaction aspects are included in the Credit Card Fraud Detection dataset. Feature scaling is used to normalize the data because various features could have varying value ranges. The dataset is transformed in this study using standard feature scaling so that each feature has a mean near zero and a standard deviation near one. By doing this step, you can stop some characteristics from taking over the model because of their higher numerical values. The dataset is more suited for training the anomaly detection algorithm after scaling.

3.3. Isolation Forest Model

The main unsupervised learning model is the Isolation Forest algorithm, which finds anomalous patterns in the dataset. Isolation Forest uses random partitioning to isolate data points in order to find anomalies. The algorithm creates several random

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0176
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1174 - 1179

decision trees, each of which divides the data according to split values and randomly chosen features. The algorithm's main tenet is that anomalous data points are simpler to identify than typical observations. They need fewer splits to be segregated in the tree structure because they are different from most of the data. In the Isolation Forest model, data points that are very different from normal patterns are separated quickly, so they get higher anomaly scores. On the other hand, normal data points require more steps to isolate. In this work, the model is trained using the scaled transaction dataset. It gives a data point an anomaly score, which tells us how different the data point is compared to the normal data. If the score is high, it implies that the data point is different from the normal data [4].

3.4. Uncertainty Estimation Using Anomaly Scores

This technique employs the anomaly score of the Isolation Forest to attain such insights. In case the data point is assigned a high rating, it implies that the point is significantly different from the expected patterns the model had encountered during the training phase. The forecast is not entirely correct in such cases. The advantage of this technique is that it doesn't demand labeled data. Instead, it attempts to attain insights based on the general structure of the dataset and identifies unusual patterns.

3.5. PCA Visualization

Principal Component Analysis (PCA) is used here as a visualization tool to better understand how anomalous data points are distributed in the dataset. It reduces the number of dimensions while keeping most of the important information (variation) in the data. Patterns and clusters are easier to see when the high-dimensional data is projected into a two-dimensional space. After calculating the anomaly scores, PCA is used in this investigation. Normal and abnormal observations can be visually examined thanks to the resulting two-dimensional representation. The Isolation Forest model frequently finds anomalous points in the plot's sparsely populated areas, suggesting that they deviate from the primary data clusters. This graphic aids in demonstrating how anomaly detection can draw attention to areas of possible data ambiguity. The overall workflow of the proposed self-uncertainty detection framework is illustrated in Figure 1.

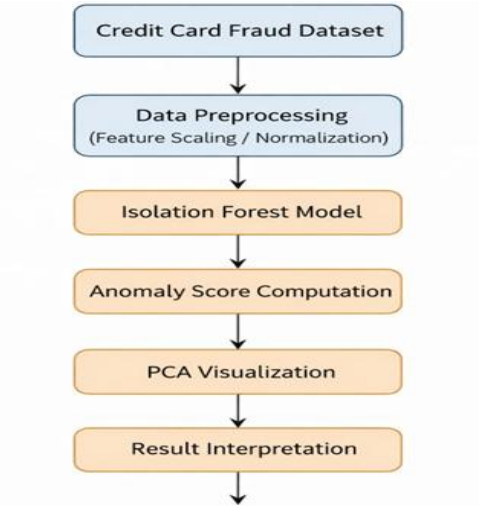

Figure 1 Architecture of the Proposed Self-Uncertainty Detection Framework

4. Dataset and Experimental Setup

The Credit Card Fraud Detection dataset, which is accessible to the public, was used for the experimental evaluation. Each of the 284,807 financial transactions in the dataset is represented by 30 numerical attributes that were obtained from actual transaction data. The dataset is appropriate for anomaly detection tasks because the features are mainly composed of anonymized variables that represent various aspects of the transactions. The dataset was preprocessed to guarantee consistency across all input features before the anomaly detection model was used. To normalize the data and lessen the influence of differences in feature magnitudes, feature scaling was carried out. During model training, this phase makes sure that each feature contributes more fairly. The main unsupervised learning model for identifying abnormal observations was the Isolation Forest method. The model was trained on the scaled dataset, and it generated an anomaly score for each transaction. These scores show how much a particular data point differs from the overall data pattern. All experiments were implemented using Python with common libraries such as Scikit-learn, NumPy, and Pandas. After calculating the anomaly scores, the results were analyzed using simple statistical plots and

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0176
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1174 - 1179

dimensionality reduction techniques to better understand how unusual data points are distributed in the dataset. To find unusual observations in the dataset, the Isolation Forest model was built up with 100 estimators and automatic contamination detection. The preprocessed dataset was used to train the model, and anomaly scores were calculated for every transaction. These scores were then utilized to examine the distribution of typical and unusual occurrences and to employ dimensionality to depict the data's structure [5 -6].

5. Results and Observations

Anomaly scores were produced for every transaction after the Isolation Forest model was trained on the preprocessed dataset. The degree to which a specific observation deviates from the distribution of the data as a whole is indicated by these scores. Anomaly scores are often lower for transactions that closely match typical data patterns and higher for unique findings. Figure 2 shows the distribution of anomaly scores generated by the Isolation Forest model. The majority of transactions fall into the lower range of anomaly scores, suggesting that they follow the dataset's predominant trends. On the other hand, anomaly scores are significantly greater with fewer observations. These points can be regarded as possible anomalies since they show transactions that deviate greatly from the bulk of the data. From the standpoint of uncertainty detection, these situations might relate to inputs where more examination of the model's predictions is necessary.

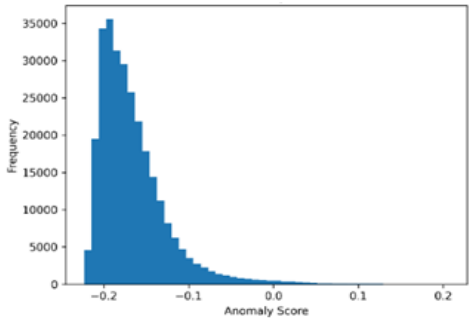

Figure 2 Isolation Forest Anomaly Score Distribution

Principal Component Analysis (PCA) was used to visualize the dataset in order to further examine how these aberrant observations appear inside the data space. This makes it easier to visualize how the data points are distributed relative to one another by reducing the complexity of the data while retaining significant variance patterns. Figure 3 shows the final visualization. In the two-dimensional visualization of the data points, the majority of the data points are grouped into clusters that represent normal patterns of transactions. On the contrary, abnormal data points are represented by areas that are not close to the clusters and are also sparsely populated. Such a distribution of data points illustrates how anomaly detection can identify abnormal patterns within the data set. These findings offer credence to the notion that anomaly scores from unsupervised learning can be effective indicators of unclear or unknown data patterns in AI systems [7].

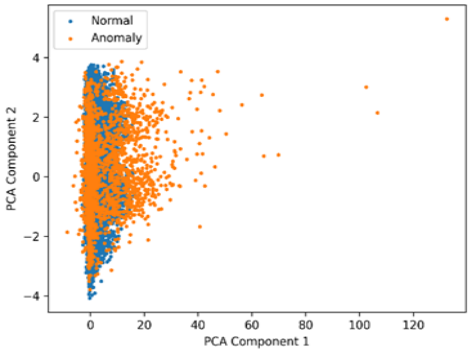

Figure 3 PCA Visualization of Detected Anomalies

From the PCA plot above, it is clear that the model is able to differentiate the normal data points from the abnormal ones to a certain level. Some of the data points are considered abnormal when they are away from the main cluster. Therefore, in summary, it shows that the Isolation Forest algorithm is able to help in the identification of the data points that are different from the normal pattern. The model is able to help in pointing out the situations when the predictions are not clear by identifying the abnormal data points. This shows that, especially in real situations, anomaly detection is a helpful concept in understanding uncertainty in machine learning models.

6. Discussion

The results of this study show that anomaly detection methods can be used to gain valuable information

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0176
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1174 - 1179

regarding the accuracy of predictions made by AI models The Data The points that are significantly different from the normal patterns can be identified based on the anomaly scores calculated by the Isolation Forest model These scenarios may represent situations in which the model is responding to unknown values and the predictions need to be confirmed. An important finding was that the majority of the transactions have low anomaly scores, which indicates that the transactions are following the overall trends identified in the instruction. This indicates that the model can. The model does a good job of capturing the dataset's general pattern. However, only a small percentage of transactions exhibit higher anomaly scores; these are known as outliers. These points are crucial because they can point to potential areas of uncertainty. This concept is further supported by the PCA visualisation. These anomalous points can be seen far from the main cluster when the data is reduced to fewer dimensions. This demonstrates unequivocally that their behaviour differs from to normal transactions. In real-life situations, These situations might indicate unusual occurrences that require additional attention from the system or a human specialist. The fact that this approach does not rely on labelled data is another benefit. This type of unsupervised method becomes more useful since there are many datasets that do not contain a large number of labelled data points. It simply looks for data points that are not following a normal pattern. But there are some limitations to this method. It can detect unusual patterns but cannot always explain them. It might also depend on the type of dataset and how it is configured. In the future, it might be possible to get better results by combining this method with other types of uncertainty estimation. This research has proved that anomaly detection is a useful tool for understanding AI uncertainty. In the future, many machine learning models can be made more reliable by considering data points that are not following normal patterns [8].

Conclusion and Future Works

In the paper, an attempt has been made to comprehend the concept of unsupervised learning and its application in identifying the uncertainties present in the AI systems. Instead of focusing on the measurement of the uncertainties, more importance should be given to the identification of abnormal patterns in the data set. Isolation Forest algorithm has been applied to the Credit Card Fraud Detection data set to generate the scores for the identification of abnormal patterns in the data set. From the results obtained, it is clear that most of the transactions are within the normal range, and a few can be identified with high levels of anomaly score. These are considered the outlier points. The PCA visualization also made it more clear how these points were located away from the main cluster of points. This is also indicating how anomaly detection could be helpful in identifying the data for which the AI model might not perform well. This method could be helpful in enhancing the reliability of AI models, especially in real-life scenarios where the data is constantly changing. In terms of future work, it could be enhanced by trying different anomaly detection techniques or even combining it with different uncertainty estimation techniques.

References

[1]. Y. Gal and Z. Ghahramani, “Dropout as a Bayesian approximation: Representing model uncertainty in deep learning,” Proceedings of the 33rd International Conference on Machine Learning (ICML), pp. 1050– 1059, 2016.
[2]. A. Kendall and Y. Gal, “What uncertainties do we need in Bayesian deep learning for computer vision?” Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 5574–5584, 2017.
[3]. F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation forest,” Proceedings of the IEEE International Conference on Data Mining, pp. 413–422, 2008.
[4]. J. MacQueen, “Some methods for classification and analysis of multivariate observations,” Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297, 1967.
[5]. M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density- based algorithm for discovering clusters in large spatial databases with noise,” Proceedings of the Second International Conference on Knowledge Discovery and

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0176
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1174 - 1179

Data Mining (KDD), pp. 226–231, 1996
[6]. G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, no. 5786, pp. 504–507, 2006
[7]. C. M. Bishop, Pattern Recognition and Machine Learning Springer, 2006.
[8]. T. Dietterich, "Ensemble methods in machine learning," International Workshopon Multiple Classifier Systems, pp. 1–15, 2000.