International Journal of Applied and Physical Sciences
volume 10 pp. 22-30 doi: https://dx.doi.org/10.20469/ijaps.10.50003

Accurate Uncertainty Dataset Classification Using Hybrid Deep Learning Models

Darbaz Marouf Hussein
IT Support and Maintenance Department
DelRaparin Technical and Vocational Institute
Ranya, Kurdistan Region, Iraq.

Kamal Al-Barznji*
Department of Computer Science
University of Raparin
Ranya, Kurdistan Region, Iraq.

Nergz Sattar Mohammed
IT Department
Rwandz Private Technical Institute
Rwandz, Kurdistan Region, Iraq.

Abstract: Data uncertainty can be produced by several variables, including measurement and sampling mistakes, sensor networks, environmental monitoring, and medical diagnostics. The goal of this study is to classify uncertain data. Classifying uncertain data is critical for maintaining data quality, improving decision-making, optimizing system efficiency, and increasing predictive accuracy. Addressing data uncertainty thoroughly ensures that systems and processes run smoothly and provide accurate, actionable insights. To discover uncertainty data, we proposed a hybrid model based on two well-known deep learning approaches (CNN and ANN). In this work, the classification of the Internet of Things (IoT) data has been done, especially healthcare data. According to the findings in this work, the outcome of the proposed hybrid model (CNN + ANN) has the best results and boasts the best success rate in comparison to the traditional machine learning-based methods in terms of performance. The results of the proposed hybrid model based on famous evolution metrics (Accuracy, Precision, Recall, and F-Score) are (97 %, 96 %, 94%, and 95 %) respectively. Machine Learning and Deep Learning, etc are one of the application of the study. Our work treats classification for uncertain data. For this purpose, a collection of people's Blood Glucose Levels (BGL) and numbers for some of their most noticeable body parts make up the dataset. Proposed a new hybrid model that combines CNN and ANN models. We have looked at measurements to see how the proposed method outperforms well-known machine learning algorithms. Finally, we have evaluated the proposed model, and based on conventional performance metrics, the tests also demonstrate that the suggested strategy finds ambiguous data more effectively than alternative approaches.

Keywords: *: Uncertain data, machine learning, convolutional neural network, artificial neural network.*

Received: 4 March 2024; ***Accepted:*** 2 April 2024; ***Published:*** 15 July 2024

I. INTRODUCTION

Information that deviates from the values that were intended for it or those that were initially stored due to the presence of noise is referred to as uncertain data. The degree of uncertainty or invalidity in data nowadays is one characteristic that distinguishes it. Data uncertainty can be caused by a wide range of reasons, including measurement error, sampling error, environmental monitoring, market research, sensor networks, and medical diagnostics. Additionally, many data mining programs may lose a significant amount of their fundamental performance and efficiency if data uncertainty is not well managed. Another challenge in many real-world applications is classifying and detecting uncertain stream data

*Correspondence concerning this article should be addressed to Kamal Al-Barznji, Department of Computer Science, University of Raparin, Ranya, Kurdistan Region, Iraq. E-mail: kamal.barznji@uor.edu.krd

[1].

Data classification of huge datasets is a critical issue in data mining. It is among the most important data mining techniques to use classifications. In data mining, researchers can figure out which examples of data belong to which groups or classes. Unpredictability of data is considered a major danger in Internet of Things scenarios [2]. Data uncertainty is caused by several variables, unpredictable environments, inaccurate sensors, including missing data, and transmission errors. Multiview learning is a fundamental technique in many Internet of Things (IoT) applications that combines several aspects to produce more thorough descriptions of data objects. The majority of earlier research on Multiview learning focused on improving prediction accuracy at the expense of decision reliability [3]. These uncertainties may have a significant impact on the accuracy and dependability of IoT applications, which could lead to less-than-ideal decision-making. Therefore, by enhancing the interpretability and dependability of IoT systems through the quantification of uncertainty, the framework enables better management of the unpredictable and variable data inputs that are fundamental to Internet of Things applications [4].

In addition, some heuristic and machine learning methods, such as Bayesian Belief Networks (BBNs), Fuzzy Sets [5], Naïve Bayes method [6]; Decision Trees [7], Rule-Based Classification Algorithms (RBCA) [8], Genetic models, Neural Networks (NNs) [9], and the Radial Basis Function (RBF) network with the particle-swarm optimization algorithm (PSO) [10], have been applied to classify data in the presence of uncertainty. Furthermore, the interests of academics in classifying ambiguous data have recently converged. This is limited by the necessity to apply clustering algorithms to particular types of data and the requirement for strategies that can handle ambiguous data.

The following are the main contributions (research highlights) we made to this study:

- Our work treats classification for uncertain data. For this purpose, a collection of people's Blood Glucose Levels (BGL) and numbers for some of their most noticeable body parts make up the dataset.
- Proposed a new hybrid model that combines CNN and ANN models.
- The uncertain data is classified using the Convolutional Neural Network (CNN). Based on the results of the tests, the CNN network works well and accurately to classify uncertain data.
- Using the Artificial Neural Network (ANN) method, we were able to expeditiously assemble the network, enhance its precision and accuracy, and better train the network.
- We have looked at measurements to see how the proposed method outperforms well-known machine learning algorithms.
- We have evaluated the proposed model, and based on conventional performance metrics, the tests also demonstrate that the suggested strategy finds ambiguous data more effectively than alternative approaches.

There are six sections to the work that are discussed in this study. The related work is in section two. The dataset utilized for experiments and the proposed methods of machine and deep learning with training models are described in sections three and four respectively. The results and discussion are found in section five. Section six includes a summary of the conclusion.

II. RELATED WORK

[11] Makes suggestions for semi-autonomous cars or recommendation systems. To empower a system to make wise choices when faced with the ambiguity of embedded AI/ML models and potential safety-related repercussions, as a result, this study offers a practical classification of the three main causes of uncertainty: scope compliance, data quality, and model fit. They notably highlight the importance of these classes in the development and testing of ML and AI models by creating ties to specific actions during development and testing as well as techniques for assessing and managing these distinct sources of uncertainty. This paper [1] proposes a unique calibration approach to enhance the lithofacies classification uncertainty analysis. Scaling is used on a pre-fit machine learning classifier based on validation data to modify the probability. Reliability diagrams show that, following Platt scaling, the predicted probability may more closely resemble the genuine probability found in the data. After the probability calibration, a lower value is produced for the Brier score, which is essentially a loss function that measures the multiclass accuracy in the probabilistic prediction.

To handle the uncertainty mitigation issue when humans are involved in a text categorization task, a DNN-based approach is put forth, in this paper [12] suggested using metric learning for the feature representation in conjunction with a dropout-entropy uncertainty measuring approach. Their suggested technique significantly outperforms competing methods, as demonstrated by extensive experiments on real-world datasets, which show a notable gain in accuracy when a very modest fraction of the uncertainty forecasts is delegated to domain experts. In this work [13] two well-known skin cancer imaging datasets were analyzed using a novel, straightforward, and incredibly effective uncertainty quantification methodology that they developed based on the Three-Way Decision (TWD) theory. In another work [14] the effects of classification uncertainty (using classification entropy or predictive variance from Bayesian approaches) and confidence calibration (utilizing post-process temperature scaling) on calibration errors and classification accuracy were investigated. The main conclusions were that: 1) confidence calibration reduces calibration errors, and 2) for Bayesian algorithms, removing test samples with high classification uncertainty and low classification confidence boosts average precision and classification accuracy. The findings imply that classification interpretation and accuracy are enhanced by calibration and uncertainty.

The authors [12] provide a new way of approaching AL. By merging the middle layer attributes of the basic learner with the direct features of the training data into a fully linked network, an uncertainty predictor is produced. The uncertainty learning model adopts a new loss function with rank learning. The candidate set is created through data sampling, and the set of chosen samples is expanded through data augmentation. The suggested approach does not employ handcrafted measures like entropy or marginal sampling; instead, end-to-end uncertainty learning is realized by the Uncertainty Learning Network. Three hyperspectral data sets are used to assess the suggested methodology. The experimental results validate that the suggested approach works well with deep models, like ResNet, serving as the basic learner.

Through automating the extraction of illness data from electronic text pathology reports from the US National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER) population-based cancer registries at the time of diagnosis and surgery, shown that how DNN-based classification tools can be used to improve cancer registries. Specifically, described several selective classification strategies that achieve a goal level of accuracy on several classification tasks while minimizing the rejection amount, or the number of electronic pathology reports for which the model's predictions are unreliable. They demonstrate that while both approaches are capable of identifying samples that require manual review and labeling by human annotators, The recently presented approaches outperform the in-house deep learning-based abstention classifier by maintaining a greater proportion without the need for retraining, resulting in a reduced computing cost [15].

The large number of parameters P makes it impossible to use the Delta methodology, a traditional method for evaluating epistemic uncertainty in statistical models, to deep neural networks in an easy way. Shown that when the least estimated eigenvalue of the Fisher information matrix is near the L2-regularization rate, the approximation error will be close to zero even when K P. Using the MNIST and CIFAR-10 datasets, illustrated the idea using a TensorFlow implementation and show how two LeNet and ResNet-based neural networks may generate meaningful image ranks based on prediction uncertainty. This implies that the categorization alone is not sufficient to capture additional information in the uncertainty estimate [16]; Skaug, & Brun, 2022).

III. MATERIALS AND METHODS

The goal of this study is to classify uncertain data. Figure 1. Shows the research's proposed architecture. We proposed the new hybrid model (CNN+ANN). First, start by taking an input sample dataset and cleaning up the dataset. Then, the data is put through a few processing steps. Finally, normalization has been done on the data to implement the proposed method. This has been done first by getting the results of different machine learning classifiers after running them on the data. Then, we performed feature selection through the CNN algorithm, and the features were merged and several new features were produced. Finally, these new features became inputs to the ANN algorithm, for which they were classified into certainty and uncertainty.

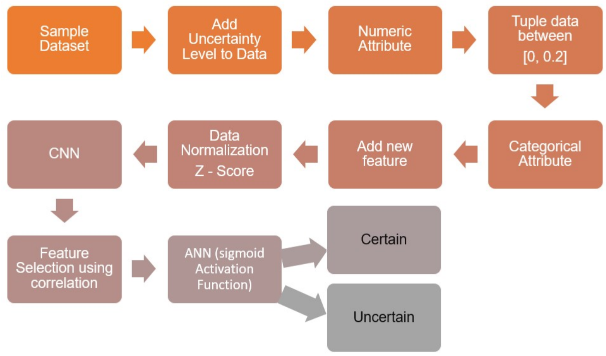

Fig. 1. The research's proposed architecture steps.

A. *Dataset*

In this work, the classification of the IoT data has been done, especially healthcare data. The information used includes the individuals' blood glucose levels (BGL) and numbers for some of their most visible body regions. The dataset offered includes a list of people of different ages who were either diagnosed with diabetes or not, together with their blood glucose readings and other basic health information like blood pressure, heart rate, body temperature, systolic blood pressure, and so on. Gaining insight into the effects a person's BGL has on their body is the main goal of the data collection process. There are 16971 records in the collection, with 10 characteristics. The dataset is available on the IEEE-Data Port website (Deepali & Sharmishta, 2022). As shown in Figure 2 below, the heat map shows the details of features, and the correlation between the features of the dataset is low, and all the features have the lowest correlation with each other.

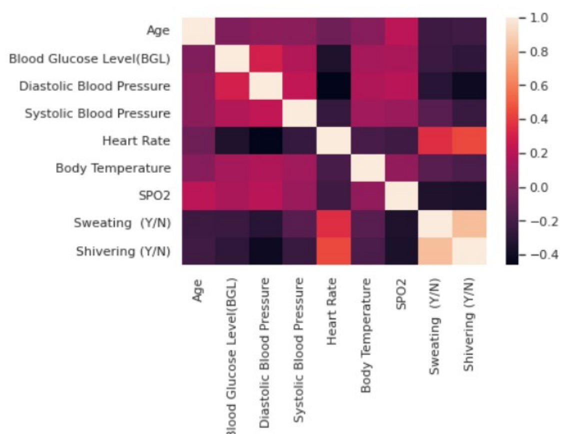

Fig. 2. Heat map and correlation between the features of the dataset.

1) Data preprocessing : In this section, the preprocessing of data is needed before it can be used for implementation. Text preprocessing refers to the cleaning and preparation of text data for use in a certain environment. Developers use it in almost every natural language processing (NLP) pipeline. In our study, we removed stopwords, sparse terms, special words, and accent marks before lemmatizing the text. We also remove accents, punctuation, and diacritics. The flowchart of the data preprocessing is shown in the figure above, step by step.

B. *Convolutional Neural Network (CNN)*

The next step in our suggested approach's growth is to combine and synthesize features to create unique and useful representations. Using a unique technique, the Convolutional Neural Network (CNN) is typically used for tasks such as picture classification. The network creates new feature maps by methodically scanning an image with different filters through complex procedures. The image's representation is enhanced by the several feature maps produced by this iterative procedure employing different filters. Three sets of numbers, ranging from 0 to 255, represent the green, blue, and red channels that make up an image. A grayscale picture, on the other hand, is made up of a single set of values that are all included in the same numerical range. The complex procedure is finished by arranging this numerical data into an image-like manner and feeding it into the CNN. The careful integration of nearby features results in the creation of new feature maps. This complex procedure captures the essence of CNN, a network that focuses on understanding complex patterns through iterative feature fusion and convolution. The goal is to extract many additional features that provide a detailed and comprehensive representation of the input data [17]. The architecture of the proposed CNN model is shown in Figure 3 below.

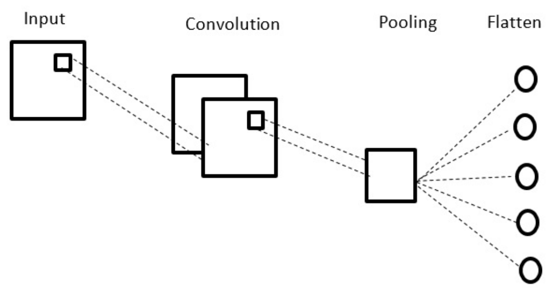

Fig. 3. The architecture of the proposed CNN model

Three layers make up the proposed CNN network: one pooling layer, two convolutional layers, with an activation function of "Relu," 12 filters of size (1, 3), and an input shape of (1, 9, 1) make up the first convolutional layer. The activation function "Relu" and 16 filters with sizes ranging from 1 to 3 are present in the second convolutional layer. A pooling layer with a size of (1, 2) makes up the third layer, which is where most of the parameters will be cut down. Flattening the data into a 1D array for the ANN network's input is the last stage in the CNN network's process.

1) Feature selection: By merging nine features, the CNN network creates nine additional features. Nine previous features were combined to create the new features. To train the classifier network, not every feature that is generated is helpful. Choosing the most crucial characteristics involves a useful process called feature selection. With each characteristic, we produce a correlation map. The characteristics that have little association are retained. One feature is maintained and the others are eliminated if there are strong relationships between them. Ultimately, nine characteristics are chosen at the end of this process and applied to the ANN classifier network.

C. Artificial Neural Network (ANN)

As we begin the process of building an artificial neural network (ANN), we begin by importing the TensorFlow model. The Sequential class serves as our creative palette in the enchanted world of Keras, allowing us to create a harmonic model that is a linear stack of layers that are painstakingly placed in the order that they arrive. A symphony of rich layers, each one a purposefully built brain tapestry, unfolds on our canvas. Our design consists of three layers: two are decorated with the vivid colors of the 'Relu' activation function, while the third layer is dressed in the subdued grace of the sigmoid activation function. With 114 parameters, the first dense layer, including six units, establishes the framework. With 42 characteristics, the second layer tells its story in a unit count that is identical to the first. The third and last layer, a single unit covered in seven parameters, provides the grand finish. The generic Adam optimizer takes the conductor's baton and leads the learning rate through the complex score, setting off a moment of orchestration. The loss is like the beating heart of our creation: binary cross entropy, a pulse that drives our model to new heights of comprehension.

Our work is taking shape, and the big reveal is getting closer. The model is a work of art and science that has been painstakingly created and is ready to be fitted to the dataset. Under the vigilant supervision of a batch size of 32, the training set, a celestial input, converges to produce a magnificent work of 100 epochs. Our artificial neural network (ANN) is a remarkable example of the harmonious combination of careful planning and sophisticated computing. The architecture of the proposed ANN model steps is shown in Figure 4 below.

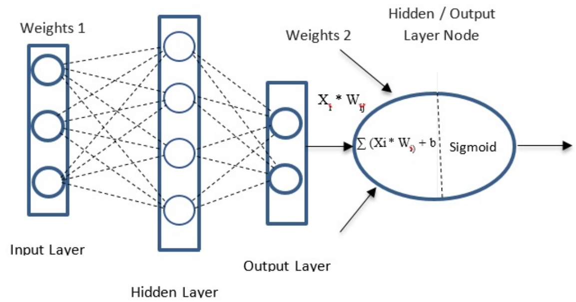

Fig. 4. The architecture of the proposed ANN model steps

IV. TRAINING AND EVALUATION MODELS

Supervised learning is now the most popular sub-branch in the field of machine learning. The investigation of supervised learning algorithms is frequently the first step in the process for aspiring machine learning enthusiasts. Deep learning, an intriguing area of machine learning that examines algorithms that are fashioned after the intricate structure and functions of the human brain, is built on artificial neural networks, or neural networks. This environment fosters the growth of both supervised machine learning and deep learning algorithms, which use labeled datasets to train and enhance algorithms that correctly classify data or predict outcomes. These datasets encompass a wide range of attributes, such as certainty and ambiguity, 30% of the dataset is used as a benchmark for algorithm testing, and the other 70% is distributed in a way that makes it easier for the model to be refined throughout training. In this section, we trained and evaluated a few supervised machine learning methods to compare the results of our proposed model with the dataset such as:

A. Support Vector Machine (SVM)

One of the most well-known supervised classes of learning algorithms is Support Vector Machines (SVM). SVM has seen significant expansion into a wide range of disciplines, along with numerous algorithmic and modeling variants. Simple support vector machine (SVM) models address scenarios in which the precise values of the data points are known. Its versatility is explained by its application in a variety of situations were knowing the degree of uncertainty surrounding forecasts is essential [18].

B. Logistic Regression (LR)

The well-known classification method known as logistic regression is frequently applied in the fields of statistics, machine learning, and data mining to learn binary responses. It works under the assumption that the data values are exactly predetermined, which isn't always the case. Many applications encounter uncertain data due to methods of data collection, such as repetitive measures, out-of-date sources, and imprecise measurements, such as in experimental physical settings [19].

C. Decision Tree (DT)

Decision trees are a popular method for classifying data. This research suggests a categorization technique for ambiguous data based on decision trees. Emerging applications include sensor networks, databases for moving objects, and biological and medical databases that frequently deal with data uncertainty. Many reasons, such as limited measurement precision, out-of-date sources, malfunctioning sensors, network latency, and transmission issues, can lead to uncertainty in data [20].

D. K-Nearest Neighbors (KNN)

In multi-source forest inventories, one of the primary techniques is the k-nearest neighbor estimation approach. It's a potent non-parametric technique with quite accurate and straightforward to compute estimates. This method's lack of an uncertainty estimate for anticipated values and areas of any size is one of its drawbacks [21].

E. Random Forest (RF)

Random Forest is a non-parametric ensemble approach where prediction error is not directly quantified. When employing model-based continuous maps as inputs to other modeling applications, like fire modeling, it's critical to comprehend forecast uncertainty [22].

F. XGBoost

A machine learning approach called XGBoost (extreme gradient boosting) is based on decision trees and sequential ensemble learning. Examine this brand-new enhancing feature. Particularly for big and complicated datasets, gradient boosting is a method that stands out for its speed and prediction accuracy. It is among the most effective machine learning methods for creating forecasting models [23].

V. EXPERIMENTAL AND DISCUSSION

The performance assessment of the suggested research approach is the main topic of this section. The BGL and values for some of the most obvious body portions of the subjects are included in the dataset. There are 16971 numbers in the dataset that have been labeled and organized. The model was built using the machine learning libraries for Python. The Python packages contain the libraries for Numpy, Pandas, and Sklearn. Utilizing a CNN, the uncertain data is categorized.

The ANN algorithm has been utilized to better train the CNN network, enhance its precision and accuracy, and expedite the network's integration. The CNN network is very accurate and efficient for uncertain data. Next, the outcomes of the suggested approach and those algorithms were compared using seven conventional machine learning models for classification, including KNN, SVM, RF, LR, DT, and XGBoost. Moreover, the trainset and test set are the two divisions of the dataset, 30% is in the test set and 70% is in the trainset. Using the ANN method, the CNN network classifies the ambiguous input. The model's performance has been assessed using the following metrics and measurements to determine how well the suggested approach performs:

Confusion Matrix: A table that shows the different kinds of accurate and inaccurate forecasts is called a confusion matrix [24]. The table of the confusion matrix is shown as follows.

TABLE 1
CONFUSION MATRIX

Predictions	Positive	Negative
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

Accuracy: Divide the total number of forecasts by the number of true positives and true negatives [25]; [24]. The equation accuracy metric is as follows:

Accuracy $= \frac{Tp+Tn}{Tp+Tn+Fp+Fn}$......(1)

Precision: The ratio of true positives to false positives and false forecasts is calculated. It is a metric used to assess a classifier's accuracy. Low accuracy is indicated by a high rate of false positives [25]; [24]. The precision metric is as follows:

Precision $= \frac{Tp}{Tp+Fp}$........(2)

Recall: Divide the total number of true positives by the total number of false negatives and positive values found in the test data. It goes by the name True Positive Rate (TPR) as well. It is a gauge of a classifier's comprehensiveness. A high proportion of false negatives indicates a low recall rate [25]; [24]. The equation for the recall metric is as follows:

Recall $= \frac{Tp}{Tp+Fn}$......(3)

F-score: the sum of the accuracy and recall weights of a measurement [25]; [24]. The equation for the f-measure metric is as follows:

$F -$ score $= 2x\frac{\text{Precision } x \text{ Recall}}{\text{Precision} + \text{Recall}}$.....(4)

A. *Results and Discussion*

According to the findings in this work as shown in Table 2, and Figure 5, the outcome of the proposed hybrid model (CNN + ANN) has the best results and boasts the best success rate in comparison to the traditional machine learning-based methods in terms of performance. The results of the proposed hybrid model based on famous evolution metrics (Accuracy, Precision, Recall, and F-Score) are (97 %, 96 %, 94 %, and 95 %) respectively. This shows that the proposed model better than traditional machine learning models at classifying uncertain data.

TABLE 2
COMPARE MODELS PERFORMANCE FOR CLASSIFYING UNCERTAIN DATA ACCORDING TO METRICS

Models Metrics	Precision	Recall	F-Score	Accuracy
LR	95%	93%	94%	95%
SVM	93%	90%	92%	93%
DT	89%	88%	90%	91%
RF	90%	89%	91%	92%
KNN	95%	93%	93%	94%
Xgboost	89%	87%	88%	89%
CNN+ANN	96%	94%	95%	97%

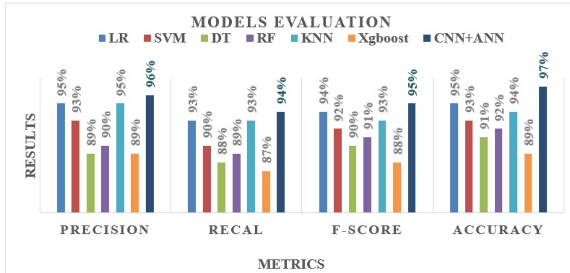

Fig. 5. Compare models evaluation according to metrics

VI. CONCLUSION

The degree of unpredictability or uncertainty in your data is referred to as data uncertainty, and it can affect the accuracy and dependability of your data analysis and decision-making. Measurement mistakes, missing numbers, outliers, ambiguity, inconsistency, and incompleteness are all potential sources of data uncertainty. Therefore, data uncertainty needs to be handled correctly so that different systems can work as well as possible. This work proposes a technique based on machine learning to categorize uncertain data. The proposed technique is a hybrid model of a CNN network and the ANN algorithm. Also, several tests were done to look at the proposed method and compare it to the K-NN classifier, SVM, DT, RF, LR, and XGBoost. In particular, according to all the popular measurement metrics as shown in Table 2, the results are a clear sign that the proposed model is better at classifying uncertain data than standard machine learning methods. In our future work, we will implement our proposed approach on more powerful big datasets to help make the best decisions in various areas.

REFERENCES

[1] R. Feng, “Improving uncertainty analysis in well log classification by machine learning with a scaling algorithm,” *Journal of Petroleum Science and Engineering*, vol. 196, p. 107995, 2021.

[2] C. Du, C. Du, and H. He, “Multimodal deep generative adversarial models for scalable doubly semi-supervised learning,” *Information Fusion*, vol. 68, pp. 118–130, 2021.

[3] W. Winiwarter and K. Rypdal, “Assessing the uncertainty associated with national greenhouse gas emission inventories:: a case study for austria,” *Atmospheric environment*, vol. 35, no. 32, pp. 5425–5440, 2001.

[4] C. Xu, W. Zhao, J. Zhao, Z. Guan, X. Song, and J. Li, “Uncertainty-aware multiview deep learning for internet of things applications,” *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1456–1466, 2022.

[5] G. Suresh, S. Shaik, O. Reddy, and B. Munibhadrayya, “Classification of uncertain data using fuzzy neural networks,” *World Comput. Sci. Inf. Technol. J.*, vol. 1, no. 4, pp. 124–131, 2011.

[6] J. Ren, S. Lee, X. Chen, B. Kao, R. Cheng, and D. Cheung, “Naive bayes classification of uncertain data,” 12 2009, pp. 944–949.

[7] S. Tsang, B. Kao, K. Y. Yip, W.-S. Ho, and S. D. Lee, “Decision trees for uncertain data,” *IEEE transactions on knowledge and data engineering*, vol. 23, no. 1, pp. 64–78, 2009.

[8] B. Qin, Y. Xia, S. Prabhakar, and Y. Tu, “A rule-based classification algorithm for uncertain data,” in *2009 IEEE 25th international conference on data engineering*. IEEE, 2009, pp. 1633–1640.

[9] J. Ge, Y. Xia, and C. Nadungodage, “Unn: A neural network for uncertain data classification,” in *Advances in Knowledge Discovery and Data Mining: 14th Pacific-Asia Conference, PAKDD 2010, Hyderabad, India, June 21-24, 2010. Proceedings. Part I 14*. Springer, 2010, pp. 449–460.

[10] N. S. Mohammed and H. Hakem Beitollahi, “Accurate classification in uncertainty dataset using particle swarm optimization-trained radial basis function,” *Available at SSRN 4203377*, 2022.

[11] M. Kläs and A. M. Vollmer, “Uncertainty in machine learning applications: A practice-driven classification of uncertainty,” in *Computer Safety, Reliability, and Security: SAFECOMP 2018 Workshops, ASSURE, DECSoS, SASSUR, STRIVE, and WAISE, Västerås, Sweden, September 18, 2018, Proceedings 37*. Springer, 2018, pp. 431–438.

[12] X. Zhang, F. Chen, C.-T. Lu, and N. Ramakrishnan, “Mitigating uncertainty in document classification,” *arXiv preprint arXiv:1907.07590*, 2019.

[13] M. Abdar, M. Samami, S. D. Mahmoodabad, T. Doan, B. Mazoure, R. Hashemifesharaki, L. Liu, A. Khosravi, U. R. Acharya, V. Makarenkov *et al.*, "Uncertainty quantification in skin cancer classification using three-way decision-based bayesian deep learning," *Computers in biology and medicine*, vol. 135, p. 104418, 2021.

[14] G. Carneiro, L. Z. C. T. Pu, R. Singh, and A. Burt, "Deep learning uncertainty and confidence calibration for the five-class polyp classification from colonoscopy," *Medical image analysis*, vol. 62, p. 101653, 2020.

[15] A. Peluso, I. Danciu, H.-J. Yoon, J. M. Yusof, T. Bhattacharya, A. Spannaus, N. Schaefferkoetter, E. B. Durbin, X.-C. Wu, A. Stroup *et al.*, "Deep learning uncertainty quantification for clinical text classification," *Journal of Biomedical Informatics*, vol. 149, p. 104576, 2024.

[16] G. K. Nilsen, A. Z. Munthe-Kaas, H. J. Skaug, and M. Brun, "Epistemic uncertainty quantification in deep learning classification by the delta method," *Neural networks*, vol. 145, pp. 164–176, 2022.

[17] D. M. Hussein and H. Beitollahi, "A hybrid deep learning model to accurately detect anomalies in online social media," *Tikrit Journal of Pure Science*, vol. 27, no. 5, pp. 105–116, 2022.

[18] O. N. Manjrekar and M. P. Dudukovic, "Identification of flow regime in a bubble column reactor with a combination of optical probe data and machine learning technique," *Chemical Engineering Science: X*, vol. 2, p. 100023, 2019.

[19] A. Ghosh and P. Dey, "Flood severity assessment of the coastal tract situated between muriganga and saptamukhi estuaries of sundarban delta of india using frequency ratio (fr), fuzzy logic (fl), logistic regression (lr) and random forest (rf) models," *Regional Studies in Marine Science*, vol. 42, p. 101624, 2021.

[20] M. A. Hafeez, M. Rashid, H. Tariq, Z. U. Abideen, S. S. Alotaibi, and M. H. Sinky, "Performance improvement of decision tree: A robust classifier using tabu search algorithm," *Applied Sciences*, vol. 11, no. 15, p. 6728, 2021.

[21] Z. Lubis, P. Sihombing, and H. Mawengkang, "Optimization of k value at the k-nn algorithm in clustering using the expectation maximization algorithm," in *IOP Conference Series: Materials Science and Engineering*, vol. 725, no. 1. IOP Publishing, 2020, p. 012133.

[22] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6308–6325, 2020.

[23] S. Islam, A. Sholahuddin, and A. Abdullah, "Extreme gradient boosting (xgboost) method in making forecasting application and analysis of usd exchange rates against rupiah," in *Journal of Physics: Conference Series*, vol. 1722, no. 1. IOP Publishing, 2021, p. 012016.

[24] M. Heydarian, T. E. Doyle, and R. Samavi, "Mlcm: Multi-label confusion matrix," *IEEE Access*, vol. 10, pp. 19 083–19 095, 2022.

[25] K. Al-Barznji and A. Atanassov, "Collaborative filtering techniques for generating recommendations on big data," in *Proceedings of the International Conference Automatics and Informatics*, 2017, pp. 225–228.