



Intelligent Diagnosis of Obstetric Diseases Using HGS-AOA Based Extreme Learning Machine



Ramesh Vatambeti^{1*}, Vijay Kumar Damera²

¹ School of Computer Science and Engineering, VIT-AP University, 522237 Vijayawada, India

² Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, 500090 Hyderabad, India

* Correspondence: Ramesh Vatambeti (ramesh.v@vitap.ac.in)

Received: 02-05-2023

Revised: 02-28-2023

Accepted: 03-10-2023

Citation: R. Vatambeti and V. K. Damera, "Intelligent diagnosis of obstetric diseases using HGS-AOA based extreme learning machine," *Acadlore Trans. Mach. Learn.*, vol. 2, no. 1, pp. 21-32, 2023. <https://doi.org/10.56578/ataiml020103>.



© 2023 by the authors. Licensee Acadlore Publishing Services Limited, Hong Kong. This article can be downloaded for free, and reused and quoted with a citation of the original published version, under the CC BY 4.0 license.

Abstract: This paper aimed to realize intelligent diagnosis of obstetric diseases using electronic medical records (EMRs). The Optimized Kernel Extreme Machine Learning (OKEML) technique was proposed to rebalance data. The hybrid approach of the Hunger Games Search (HGS) and the Arithmetic Optimization Algorithm (AOA) was adopted. This paper tested the effectiveness of the OKEML-HGS-AOA on Chinese Obstetric EMR (COEMR) datasets. Compared with other models, the proposed model outperformed the state-of-the-art experimental results on the COEMR, Arxiv Academic Paper Dataset (AAPD), and the Reuters Corpus Volume 1 (RCV1) datasets, with an accuracy of 88%, 90%, and 91%, respectively.

Keywords: Obstetric electronic medical records; Arithmetic optimization algorithm; Hunger games search; Chinese obstetric EMR; Kernel extreme machine learning

1. Introduction

AI-powered "intelligent diagnosis" aids medical professionals in making informed decisions in the clinic. Intelligent diagnosis is a powerful tool and has many real-world applications in the clinical setting [1]. It aids physicians in making a diagnosis of a patient's condition, which increases both speed and accuracy of the diagnosis process and serves as a valuable foundation for future diagnoses. As the sophistication of diagnosis and treatment tools increases, so does the complexity of medical data. Every day, doctors collect a mountain of clinical diagnostic data and use it to make informed decisions about their patients' condition [2]. In addition, doctors have a tough time when difficulties arise during pregnancy.

In recent years, EMRs have proliferated rapidly, allowing for a plethora of intelligent diagnosis techniques to be implemented. As a classification problem, early research on intelligent diagnosis largely used artificially constructed feature templates [3, 4] or single typical machine learning approach. EMRs are the most comprehensive and direct documentation of medical care provided to patients. Clinical diagnosis can be thought of as a physician's assessment of a patient's likelihood of having a particular disease, based on the affected organ's symptoms and examination results [5]. Affected roles may be identified with both "gestational diabetes mellitus" and "gestational hypertension" in an obstetric EMR, and these two diagnoses are strongly coupled to one another. If an EMR is considered as one sample, then there are several ways to classify individual samples. Multiple analytic outcomes in an EMR have distinct labels [6].

However, the distribution of EMR data is often imbalanced, with the number of rare diseases in the sample much smaller than that of prevalent ones. Uneven distribution of datasets led to poor performance of traditional classification techniques [7]. In classification, traditional algorithms frequently exclude some classes as noise or outliers [8]. If an EMR is imbalanced, the cost of a false negative exceeds that of a false positive by a significant margin. For instance, 99 out of 100 EMRs are deemed normal and 1 indicates the presence of malignancy. The EMR diagnostic results are predicted as normal if the standard organization technique is applied directly to this data. Even though the accuracy rate can reach 99 percent, the most important data concerning cancer is ignored [9].

Classifiers in neural networks are heavily impacted by characteristics of the input data. Rebalancing technique is effective because it allows classifier weights to be updated, which boosts the learning capacity of classifiers but hinders feature learning [10]. Furthermore, the multi-label rebalancing technique must consider EMRs. When an EMR with high-frequency analytic results is deleted, the EMR with low-frequency diagnostic results is also deleted. Conversely, when the EMR with low-frequency diagnostic results is cloned in order to add the high-frequency diagnostic penalties, the EMR with high-frequency analytic results increases.

To simulate intelligent diagnosis and associated activities, many researchers have paid attention to neural networks recently [11]. When the data is small in a conventional neural network, the value of incorporating it into external knowledge becomes more apparent. However, these approaches overlook the mutually beneficial interactions between neural networks and expert knowledge. Biomedical text resources have expanded rapidly in recent years because of computing and biology advances [12]. Information found in these fields can be put to good use to advance medical informatics. Because the accurate diagnosis ability of a physician relies on specific training and medical background knowledge, a solid foundation of medical knowledge is crucial throughout the diagnostic phase [13]. These initiatives have resulted in novel analysis approaches of medical data. However, the following issues continue to impede intelligent diagnosis using EMRs:

- Multiple diagnoses, such as "normal," "pathological," and "complications," are common in an EMR.
- External knowledge is poorly captured, and the methods require a huge number of calculations, which merely splice the knowledge with the model.
- It is important for doctors to know the final diagnosis. At the same time, they also need to know what specific medical expertise has been used in making that diagnosis.

Therefore, we created a whole new intelligent diagnosis model based on refined machine learning (ML). HGS-AOA determined the value used in the ELM kernel.

2. Literature Review

The unique information retrieval method (IKAR) of Zhao et al. [14] for extracting crucial information from reports, was used to automatically generate ultrasound diagnostic results. The reader had the ability to infer inexplicit information in the report and directly take information from it. IKAR had 90.23% F-score, 91.09% recall, and 89.38% accuracy using the dataset. In addition, the F-score was above 90% in half of 10 sections in the report. This study can serve as both a reference point for obstetrics and gynaecology information retrieval techniques and general resources in EMR field.

Liang et al. [15] suggested a disease prediction approach by combining several Chinese electronic health records (EHRs) encodings. In order to improve text representations, the model framework employed a multi-head self-attention instrument that considered both linguistic and numerical details. Entity extraction and embedding representations were accomplished with the help of the Bi-directional Long-Short Term Memory-Conditional Random Field (BiLSTM-CRF) and Text Convolutional Neural Networks (TextCNN) replicas. The entity and text representations in a text were combined and used to create an EHR representation. Experimental results using electronic devices surpassed prior baseline techniques with 91.92% F1 score.

Yang et al. [16] created and tested machine learning models to better understand and forecast the likelihood of type II diabetes in adults. After analyzing data in the medical records of all adults diagnosed with type II diabetes, eXtreme Gradient Boosting (XGBoost) and Natural Language Processing (NLP) were used to create the prediction model. Key metrics for gauging model efficacy included the F1 score, AUC, and DCA. In the sample of 29,843 people with type II diabetes, 2,804 (9.4%) had hypoglycemia. Generally speaking, the XGBoost embedded machine learning model achieved the best results, with 0.82 AUC and 0.93 precision. The XGBoost3 also had better performance than other competing models in DCA.

In order to separate representation learning from classifier learning, Zhang et al. [17] proposed an intelligent diagnosis model based on Double Decoupled Network (DDN), which was used to learn initial features of the data in representation learning stage. The study proposed to decouple the highly coupled diagnostic results and rebalance the datasets in classifier learning phase with the proposed Decoupled and Rebalancing Highly Imbalanced Labels (DRIL) procedure. The study tested the proposed DDN on the COEMR datasets and validated efficacy and generalizability of the model using these datasets and the Reuters Corpus Volume 1 (RCV1). Uneven obstetric EMRs proved that the proposed techniques worked. The DDN model outperformed state-of-the-art experimental results, with an accuracy of 84.17%, 86.35%, and 93.87% on the COEMR, AAPD, and RCV1 datasets, respectively.

The self-assessment app developed by Sridhar et al. [18] utilized machine learning to predict about 40 diseases based on users' reported symptoms. Keeping reasonably accurate prediction of the model was crucial to self-assessment. Accuracy decline is one of the problems in current machine learning methods, which was an important factor in such settings. Diseases were diagnosed using the proposed system. Data was first preprocessed, and then was run through a number of different machine learning classification models, including the K nearest neighbor. Vectors were split by the total number of models considered. Depending on the threshold values, the programme

displayed anywhere from one disease to all. The model returned its predicted disease(s) to the app, which displayed them to the user. The model was clearly the winner, compared with other models.

Pang et al. [19] created seven machine learning models to predict paediatric obesity for children aged 2 to 7 using EHR birth data. Researchers in Philadelphia accessed EHR data of 860,510 patients who had 11,194,579 healthcare encounters. After implementing severe quality control measures to filter out patients with implausible growth values and include only those having attended all recommended wellness visits before age 7, totally 27,203 participants (50.78% male) were left for model development. The prevalence of obesity, defined by the Prevention Act, was predicted using seven machine learning models. The study analyzed the differences between the models using Cochran's Q test and post hoc pairwise testing, and evaluated their performance using several standard classifier metrics.

Meng et al. [20] described a chronological deep learning model using a transformer architecture to predict future depression diagnoses based on EHR sequences. This algorithm forecasted chronic diseases at different time intervals by analyzing five different types of data from the EHR over time. The current pretraining and fine-tuning trend was applied to the EHR data for improvement. The model produced PRAUC values ranging from 0.70 to 0.76, which improved the best baseline model for depression prediction. To further enhance the model's interpretability, self-attention weights were included in each sequence to quantitatively display the inner relationship between different codes. These findings showed that the model made use of diverse EHR data to predict depression with high accuracy and interpretability, thus helping expand future clinical choice support systems for the screening and early detection of chronic diseases.

3. Proposed Model

This paper provided an intelligent diagnostic model based on optimal ML in order to lessen the high coupling of diagnosis outcomes and bolster the characteristics of input samples. In the presentation learning phase, the model with fixed parameters of presentation learning was used to learn the original properties $D=[d_1, d_2, \dots, d_n]$, where d_n is a sample identifier. An embedded vector sequence was produced from the provided text. Convolutional word was embedded in a vector sequence using a linear transformation function, which extracted indicative information from text. For each class of relevant data, this paper selected its maximum value from the feature mapping in the pooling layer. Finally, information was combined with disease diversity in fully connected convolutional or pooling layers. To decouple the high coupling diagnostic results, this paper proposed to devise an algorithm rebalance the datasets $D'=[d_1', d_2', \dots, d_n']$ from which the classifier would be learned. The classifier combined a full connection layer with a Softmax algorithm. Both phases made use of a similar network construction with shared weights (except the last full connection layer).

This section examined the effects of rebalancing methods on neural network training for both representation and classifier learning. Rebalancing strategies significantly enhanced classifier learning at the expenditure of the learnability of some features. Therefore, this paper proposed to decouple representation learning from classifier learning as a solution. Basic dataset features were uncovered in the representation learning stage. The rebalanced datasets were used for training in the classifier learning phase in order to achieve a better equilibrium between classifier learning, advance the generalisation capacity of low data, and boost the classification performance of imbalanced data.

3.1 Data Preprocessing

It was important to keep patient information private because EMRs used were from actual patients. It was expected that some noise was contained in them. Data must be deidentified and cleaned to begin processing the EMRs. In order to protect the privacy of those involved in analysis of the retrieved records, any references to specific patients, hospitals, doctors, patient identifiers, locations, or phone numbers had been scrubbed from the data. Then some procedures were performed on the EMR data as follows, such as data purification, data structuration, and word standardization.

3.1.1 Data cleaning

Due to flaws, the current hospital information system (HIS) has several problems, such as redundancy, missing information, and disorganisation. Automatic string matching was used to remove duplicates from the database. In particular, if two or more identical first-course records were found in a single EMR, the most reliable one was selected based on informational and temporal accuracy. For EMRs where the first course record was missing, they were removed from the database. The dataset was cleaned up by removing records with temporal faults. An algorithm was used to recognize records with temporal disorder according to the temporal logic of obstetric therapy. Finally, there were 11,303 first-semester records.

3.1.2 Data structuration

The original text of the EMRs was scrambled. The first-semester records were arranged in a way that made

analysis simple. The experimental dataset in this paper was based on several fields, such as primary complaints, admitting physical examinations, obstetric practise, auxiliary examinations, admitting diagnosis, diagnostic basis, differential diagnosis, and treatment plan.

3.2 Label Decoupling Module

Because the rebalancing technique did not rely on a specific classifier, it was used in more situations than the adaptive classifier. It was challenging to obtain reliable performance of multi-label data using conventional rebalancing techniques. The key problems included high degree coupling between samples and imbalance between labels in multi-label datasets.

According to the study of Khare and Kumari [21], it was possible to calculate both the imbalance ratio and the average imbalance ratio to Eq. (1). The imbalance ratio of a single label was first estimated using the Imbalance Ratio (IR) per label metric. Dataset D was a multi-label dataset if and only if it satisfied the subsequent conditions: $D = \{(X_i, Y_i) | 0 \leq i \leq n, Y_i \in L\}$, where X_i is the i -th sample in the dataset, Y_i is the label for that sample, and L is the label set for the datasets.

$$IR_{(l)} = \frac{\argmax_{l' \in L} (\sum_{i=1}^{|D|} h(l', Y_i))}{\sum_{i=1}^{|D|} h(l, Y_i)}, h(l, Y_i) = \begin{cases} 1, & l \in Y_i \\ 0, & l \notin Y_i \end{cases} \quad (1)$$

The Mean Imbalance average of the IR values across all labels in a multi-label dataset was:

$$MeanIR = \frac{1}{|L|} \sum_{l \in L} (IR(l)) \quad (2)$$

Frequencies were classified into high and low categories based on their MeanIR and IR value, respectively. It was label and then a high-frequency label when the IR value was less than the mean IR value. If the IR of label y was more than the mean IR , then it was part of the minor bag, and else it was part of the major bag.

3.3 Classification Using Optimized ML

3.3.1 Kernel Extreme Machine Learning (KEML)

The Extreme Machine Learning (EML) had three layers, namely, input, hidden and output layer, and considered a dataset of N training samples $(x_j, t_j) \in R_n \times R_m$. Let L be the number of implicit layer nodes of the EML and ϑ be the excitation function, then there was:

$$f(x_i) = \sum_{i=1}^L \beta_i \vartheta(\omega_i x_j + b_i) = t_j, j = 1, \dots, N \quad (3)$$

where, i is the weight vector in the i -th node in the output layer. To generate the output weights, EML used a least-squares approach to select the input weights and bias b of the understood layer nodes at random and then calculated the solution analytically. The computations aimed to minimize the number of mistakes during training and maximize generalisation performance.

According to EML theory, Eq. (3) was formulated in a condensed form below:

$$H\beta = T \quad (4)$$

In order to train the EML on training dataset, the following steps were taken once the number of implicit layers nodes were identified after the excitation function.

Step 1: Randomly yielded input weights ω_i and b_i , with $1 \leq i \leq N$;

Step 2: Computed the output matrix of layer H ;

Step 3: Computed the yield weight matrix $\beta = H^+ T$;

where, H^+ is the Moore- matrix H . When HHT was non singular, $H^+ = HT(HH^T)^{-1}$.

The least-squares solution output by the network was used in conjunction with a regularization factor to implement the ridge regression approach, thus removing mistakes from the "sick matrix."

$$\beta = H^T (HH^T + \eta I)^{-1} T \quad (5)$$

Therefore, the consistent EML output aimed at:

$$y(x) = h(x)\beta \quad (6)$$

A novel kernel-based EML (KEML) approach may be developed by including a kernel function into the EML when the feature was unknown. The kernel matrix QELM=HHT, which comprised the following components, was necessary for the KEML procedure:

$$Q_{ELM}(i, j) = h(x_i).h(x_j) = K(x_i, x_j) \quad (7)$$

Then, the network output may be written as follows:

$$y(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix} (\eta I + Q_{ELM})^{-1} T \quad (8)$$

The kernel function $K(x_i, x_j)$ in Eq. (8) is chosen to be the radial basis kernel function:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\gamma^2}\right) \quad (9)$$

where, the nucleus parameter of RBF kernel function is defined.

Many situations, including second major prediction, medical diagnosis, and financial stress prediction, revealed that the two essential factors had a significant influence on the performance of KEML.

3.3.2 Proposed HGS-AOA-KEML

In order to produce a better informed medical diagnosis, this paper used HGS-AOA-KEML to choose the best kernel characteristics of EML from the dataset.

Step 1: HGS-AOA should have a random starting population.

Step 2: The agent's binary value along each axis was used to depict its subset choice from the dataset (1 indicated that the subject was affected, and 0 indicated that the subject was normal).

Step 3: For each HGS-AOA, fitness of that particular collection of features was determined by using the formula below:

$$F(x) = a.error + \beta \cdot \frac{|R|}{|D|} \quad (10)$$

According to the proposed HGS-AOA, it was found that =0.97 and =0.03 were appropriate parameters for this investigation.

Step 4: Changed the agent population to reflect the HGS-AOA method.

Step 5: Took the person who has the lowest fitness score into consideration.

Step 6: Checked if the maximum number of iterations had been achieved, which was the termination condition. If so, moved on to Step 3; otherwise, returned to Step 2 and continued until the termination condition was reached.

Step 7: Gave back the best possible answer as the selected weight.

Step 8: The final classification result may be obtained by feeding the final weight value into KEML as an input parameter.

Step 9: Step 8's classification findings were used to determine the classification error accuracy, number of subsets used in the classification, sensitivity, specificity, and any other assessment criteria.

3.3.3 HGS optimization

Approach food

Following formulae meant to simulate the contraction mode and illustrate its approaching behaviour in mathematics:

$$\overrightarrow{X(t+1)} = \begin{cases} \overrightarrow{X(t)}.(1 + rand(1)), & r_1 < l \\ \overrightarrow{W_1} \cdot \overrightarrow{X_b} + \overrightarrow{R} \cdot \overrightarrow{W_2} \cdot |\overrightarrow{X_b} - \overrightarrow{X(t)}|, & r_1 > l, r_2 > E \\ \overrightarrow{W_1} \cdot \overrightarrow{X_b} - \overrightarrow{R} \cdot \overrightarrow{W_2} \cdot |\overrightarrow{X_b} - \overrightarrow{X(t)}|, & r_1 > l, r_2 < E \end{cases} \quad (11)$$

where, R is in the range $[-a, a]$, r^1 and r^2 are independent random variables likewise in the range $[0, 1]$, r^3 is also independent and random, t is the current iteration, W_1 and W_2 are hunger weights, X_b is the location information of a randomly selected individual from all the optimal individuals, $X(t)$ is the location information of each individual. The value of l was described in the parameter setting experiment.

The formula for E was:

$$E = \text{sech}(|F(i) - BF|) \quad (12)$$

where, $F(i)$ is the fitness value of an individual and $I = 1, 2, \dots, n$, and BF is the greatest fitness found so far in this iteration. To put it simply, Sech was a function $\left(\text{sech}(x) = \frac{2}{e^x + e^{-x}}\right)$.

The formula of \vec{R} was:

$$\vec{R} = 2 \times a \times \text{rand} - a \quad (13)$$

$$a = 2 \times \left(1 - \frac{t}{\text{Max_iter}}\right) \quad (14)$$

where, rand is a random sum in the range of $[0, 1]$, and Max_iter is the largest sum of repetitions.

Hunger role

Searching people's hunger experiences was modelled quantitatively. The expression for W_1 in Eq. (15) was:

$$\vec{W}_1(i) = \begin{cases} \text{hungry}(i) \frac{N}{SHun} \times r_4, & r_3 < l \\ 1, & r_3 > l \end{cases} \quad (15)$$

The formulation of $\vec{W}_2(i)$ in Eq. (16) was:

$$\vec{W}_2(i) = (1 - \exp(-|\text{hungry}(i) - SHungry|)) \times r_5 \times 2 \quad (16)$$

If each person's hunger was denoted by hungry, N is the total sum of people, and $SHungry$ is the aggregate hunger experienced by everyone, or the total number of people who were hungry. The values r_3 , r_4 , and r_5 are all completely arbitrary integers between zero and one.

There was $\text{hungry}(i)$:

$$\text{hungry}(i) = \begin{cases} 0, & \text{AllFitness}(i) = BF \\ \text{hungry}(i) + H, & \text{AllFitness}(i) < BF \end{cases} \quad (17)$$

where, $\text{AllFitness}(i)$ maintains the fitness of all individuals in the current iteration.

The expression for H may be written as:

$$TH = \frac{F(i) - BF}{WF - BF} \times r_6 \times 2 \times (UB - LB) \quad (18)$$

$$H = \begin{cases} LH \times (1 + r), & TH < LH \\ TH, & TH \geq LH \end{cases} \quad (19)$$

where i is a random integer between zero and one, $F(i)$ is the fitness of each person, and r_6 is a random number between zero and one. Top bound (UB) and lower bound (LB) reflected the upper and lower bounds of the search space, correspondingly, and BF was the greatest fitness degree attained so far in this iteration. H , the sensation of hunger, had a floor below which it could not fall.

3.3.4 AOA

By using elementary arithmetic operations for modelling, including division (D), addition (A), multiplication (M), and subtraction (S), AOA was a revolutionary meta-heuristic method to optimize a wide variety of search issues. Earlier substantial coverage was instances using search fields to avoid localised solutions. In conclusion, exploration yielded answers, which led to better performance.

Initial stage

Predefined subsets, denoted by A in Eq. (20), were used to kick off the optimization process. New random

optimum sets were produced and utilised at each cycle:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,j} & a_{1,1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,j} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & \cdots & \cdots & \cdots & \cdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ a_{N-1,1} & \cdots & \cdots & \cdots & a_{N-1,j} & \cdots & a_{N-1,n} \\ a_{N,1} & \cdots & \cdots & \cdots & a_{N,j} & a_{N,n-1} & a_{N,n} \end{bmatrix} \quad (20)$$

Exploration/Exploitation choices should be carefully considered before attempting AOA. The rate, at which the Mathematical Optimization Algorithm (MOA) operated, was determined using Eq. (21).

$$MOA(C_{iter}) = Min + C_{iter} \times \left(\frac{Max - Min}{M_{iter}} \right) \quad (21)$$

where, $MOA(C_{iter})$ is an estimate of the value of the repetition function, M_{iter} is the maximum iterations, Max & Min are the maximum and minimum accelerated values, C_{iter} is iteration number (within 1).

Exploration stage

The exploratory characteristics of AOA were discussed. According to the AOA, mathematical calculations paid to either operative contributed to an examination search strategy. Nonetheless, due to the widespread prevalence of S and A operators, these D and M workers may never expect to reach their goal with much ease. AOA exploration workers randomly used the search field over numerous places in quest of a superior alternative using two primary search strategies. Search strategies using M and D were illustrated by Eq. (22).

$$a_{i,j}(C_{iter} + 1) = \begin{cases} besta_j \div (MOP \div \varepsilon) \times ((UB_j - LB_j) \times \mu + LB_j), & r_2 < 0.5 \\ besta_j \times MOP \times ((UB_j - LB_j) \times \mu + LB_j), & otherwise \end{cases} \quad (22)$$

where $a_{i,j}(C_{iter}+1)$ is the j-th current position in next iteration, LB_j and UB_j are the lower and upper bound limits for j , which is the smallest integer number, $besta_j$ is the current j-th best option out of all possible ones.

Exploitation stage

We addressed the exploitative potential of AOA, because both addition (A) and subtraction (S) based AOA mathematical formulations yielded extremely dense outcomes. Operators using AOAs exploited the search field extensively throughout a wide variety of locations, in order to look for a better option using the two primary search strategies represented by Eq. (23), the A and S search strategies.

$$a_{i,j}(C_{iter} + 1) = \begin{cases} besta_j - (MOP) \times ((UB_j - LB_j) \times \mu + LB_j), & r_2 < 0.5 \\ besta_j + MOP \times ((UB_j - LB_j) \times \mu + LB_j), & otherwise \end{cases} \quad (23)$$

Several methods that take into account the population as a whole have been proposed lately. Although they have been widespread used in engineering, this paper investigated how to best apply them in practise. To speed up integration, strike a more consistent balance, and optimise for high quality, researchers must significantly adapt and improve their methodologies, based on fundamental evolutionary processes. Accordingly, this paper proposed a novel hybrid approach based on HGS-AOA.

4. Results and Discussion

4.1. Dataset Description

This paper used COEMR datasets to assess the proposed intelligent diagnosis model and validated its efficacy and generalizability on two benchmark datasets, AAPD and RCV1. Table 1 includes some basic statistics about the datasets that were used in the study. The data filter sum for each filter was 25 in the presentation learning stage. The 0.1 resampling rate was optimal because it applied to multi-label data. With 0.001 learning rate, 32-item batch size, and 0.3 dropout, Adam was used as the optimizer.

COEMR: 24,339 inpatient records were chosen from multiple hospitals for this dataset. Most of EMRs were structured and unstructured text data. Structured data included patient demographics and clinical data, including their age, race/ethnicity, and lab results. Unstructured data included patient statements, hospital records, results of

objective tests, and other similar information. Patients' names, ID numbers, and other identification details were deleted for privacy reasons. The full breakdown of obstetric COEMR is shown in Table 2.

Table 1. Statistical info of COEMR, AAPD and RCV1

Dataset	Total	Test	Label	Train	Scumble	MeanIR
COEMR	24,339	2434	73	21,905	0.3028	246.5693
AAPD	55,840	1000	54	54,840	0.1158	16.9971
RCV1	804,414	781,265	103	23,149	0.3497	279.6319

Table 2. Example of an obstetrical COEMR datasets

Title	Content
Sex	Female
Age	36 years old
Chief complaint	Taking "rest of June, vaginal bleeding for 4 hours" as the chief regular menstruation and self-measured HCG urine positive within 30 days after menorrhagia. The patient was found to have an ectopic pregnancy after a restful January. B-ultrasound test confirmed the diagnosis. And 40 days of menorrhagia revealed symptoms of early pregnancy, such as nausea, vomiting, and abdominal pain.
Admission	T: Blood pressure of 120 over 80
Physical examination	Medium nutrition, average growth, with a clear brain and an uninhibited mentality, you may enter the hospital and assume a physically independent stance.
Obstetric examination	Verify that your body is working together. The mucous membranes all over your body are immature and clear of yellow stains, rashes, and open wounds; you also shouldn't touch your swollen superficial lymph nodes
Auxiliary examination	The ec19.0cm to 9.0cm range is the extra pelvic measurement. Height of the uterus, in centimetres, is 29.0
	Fetal heart rate of 144 beats per minute, foetal estimated weight of 2600 grammes, and absence of contractions at 93 weeks.
	Color foetal sonography measurements; breech s/D 2.2 placenta, BPD 74.0 mm, FL 53.0 mm, AFI 165.0 mm
	Grade I
	Endangered preterm birth
	Placenta previa
Admission diagnosis	Intrauterine 28+2 weeks
	G3P1
	Breech presentation
	One week umbilicus
Diagnostic basis	Having a baby at a point in time between 28 and 37 weeks
	Distension of the endocervical OS and/or the presence of irregular or regular contractions reduced menstrual bleeding

Table 3. Distribution of COEMR

Label	Sum	Sum	Label	Sum	Label
Head position	18,139	1249	Induced labor	265	Fetal dysplasia
Threatened labor	6257	1112	RH negative blood	259	Threatened abortion
Pregnancy with uterine scar	5757	1033	Fetal distress	257	Placenta previa
Premature rupture of membranes	3239	1029	Pregnancy persuaded hypertension	251	Preeclampsia
Oligohydramnios	2897	819	Cervical insufficiency	217	
Gestational diabetes mellitus	2661	496	Pregnancy complicated with hysteromyoma	201	Precious child
Threatened preterm birth	2130	405	Diabetes complicated with pregnancy	189	Polyhydramnios
Umbilical cord around neck	2054	374	Pregnancy complicated with hyperthyroidism	182	
Breech	1806	335	Pregnancy complicated with anemia	178	Intrauterine fetal
Twin pregnancy	1329	287	Inevitable abortion	177	Growth restriction

Table 3 provides a visual representation of the prevalence of various diagnoses in the COEMR datasets. Among the diagnoses, "head posture" accounted for more than 90%, while "gestational hypertension" only 10%. All 24,339 samples were split into a training set (21,905) and a test set (2,343) using 91 rules based on the diagnostic outcomes of 73 diseases with different degrees.

Yang et al. [22] built the AAPD dataset, which was a dataset of scholarly articles from Arxiv and a sizable MLTC dataset, including 55,840 abstracts from the computer science section of Arxiv1.

Lewis et al. [23] supplied RCV1, an artificially tagged dataset of Reuters articles from 1996–1997, including a maximum of 103 possible categories of news articles.

4.2. Performance Metrics

This paper used four mutual rules based on confusion matrix to verify the classifier's accuracy. In order to keep the focus on the material covered by this paper, their formulations were not discussed.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (24)$$

$$Specificity = \frac{TN}{FP + TN} \quad (25)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (26)$$

where, TP is a positive result, TN a negative result, FP a false result, and FN a true result.

EHRs have already been used in numerous medical settings, but obstetric diseases are not the focus. Therefore, three datasets were used to evaluate the generic models, and the average results are presented in Tables 4-6.

Table 4 represents the experimental results of OKEML-HGS-AOA in COEMR datasets. Different models were used in comparison analysis, such as RF, NB, DT, EML, KEML with OKEML-HGS-AOA. RF reached 0.71 accuracy, 0.70 sensitivity and 0.73 specificity. NB reached 0.79 accuracy and 0.82 sensitivity. DT reached 0.82 accuracy, 0.86 sensitivity and 0.75 specificity. EML reached 0.85 accuracy, 0.81 sensitivity and 0.83 specificity. KEML reached 0.86 accuracy, 0.82 sensitivity and 0.84 specificity. OKEML-HGS-AOA reached 0.88 accuracy, 0.90 sensitivity and 0.94 specificity, respectively.

The experimental results of OKEML-HGS-AOA in AAPD datasets are represented in Table 5. In this comparison analysis, different models were used, such as RF, NB, DT, EML, KEML and OKEML-HGS-AOA. This paper first evaluated the RF and obtained an accuracy of 0.85, sensitivity of 0.81, and specificity of 0.73. Another method of NB reached the accuracy of 0.79 and also reached the sensitivity of 0.82 and specificity of 0.75; DT reached the accuracy of 0.87 and also reached the sensitivity of 0.84; and another method of EML reached the accuracy of 0.67 and also reached the sensitivity of 0.81 and specificity of 0.83. Another model of KEM reached an accuracy of 0.86, a sensitivity of 0.82, and a specificity of 0.84. Another model of OKEML-HGS-AOA reached an accuracy of 0.93, a sensitivity of 0.92, and a specificity of 0.96.

Table 4. Experimental consequences of OKEML-HGS-AOA in COEMR datasets

Models	Accuracy	Sensitivity	Specificity
RF	0.71	0.70	0.73
NB	0.79	0.82	0.75
DT	0.82	0.86	0.75
EML	0.85	0.81	0.83
KEML	0.86	0.82	0.84
OKEML-HGS-AOA	0.88	0.90	0.94

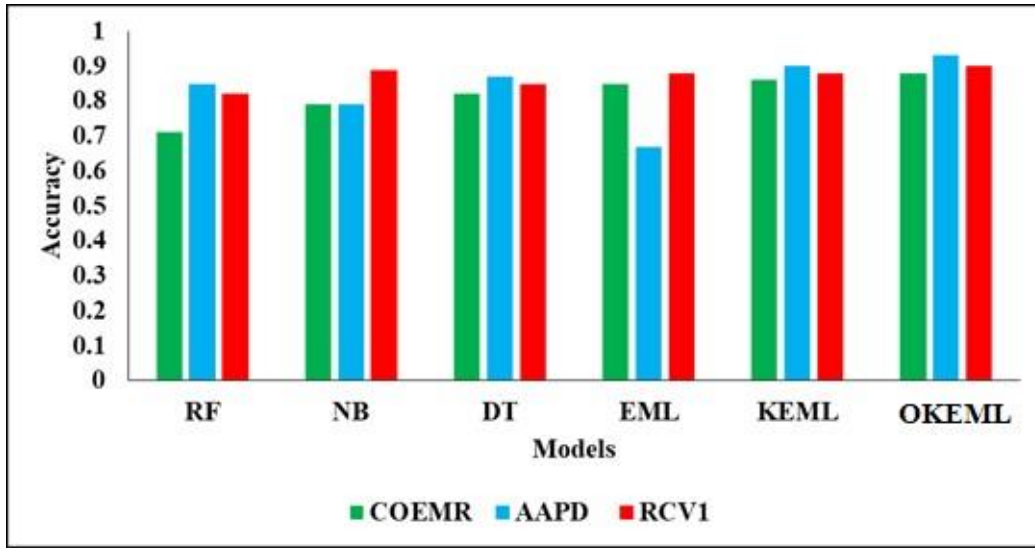
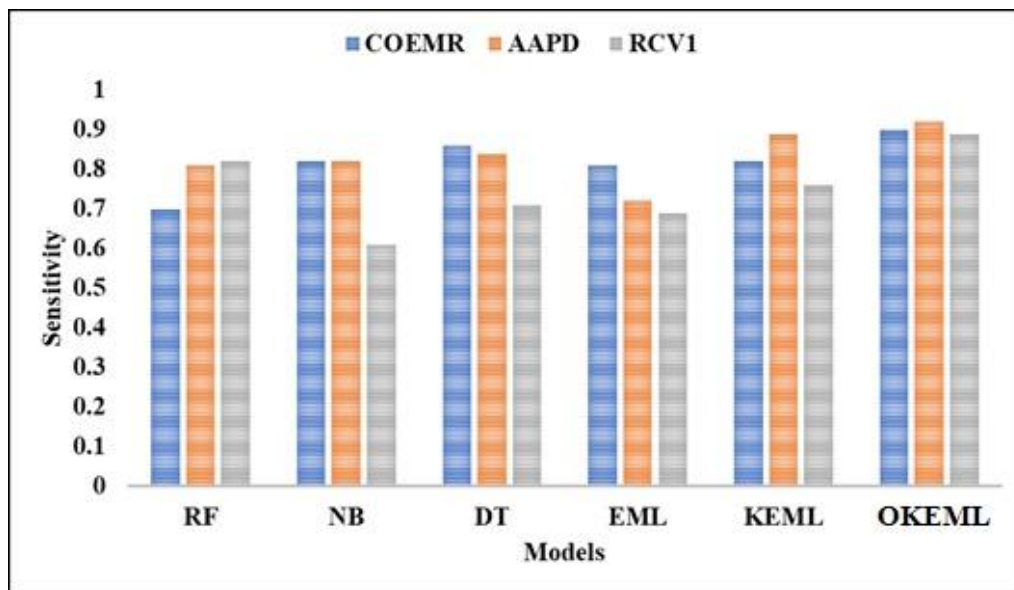
Table 5. Experimental results of OKEML-HGS-AOA in AAPD datasets

Models	Accuracy	Sensitivity	Specificity
RF	0.85	0.81	0.73
NB	0.79	0.82	0.75
DT	0.87	0.84	0.83
EML	0.67	0.72	0.72
KEML	0.90	0.89	0.92
OKEML-HGS-AOA	0.93	0.92	0.96

Table 6. Experimental results of OKEML-HGS-AOA in RCV1 datasets

Models	Accuracy	Sensitivity	Specificity
RF	0.82	0.82	0.83
NB	0.89	0.61	0.84
DT	0.85	0.71	0.83
EML	0.88	0.69	0.91
KEML	0.88	0.76	0.92
OKEML-HGS-AOA	0.90	0.89	0.95

The above Table 6 represents the experimental results of OKEML-HGS-AOA in RCV1 datasets. This comparison analysis used different models, such as RF, NB, DT, EML, KEML and OKEML-HGS-AOA. Initially this paper evaluated the RF, which reached the accuracy of 0.82 and the sensitivity of 0.82 and specificity of 0.83. Another method of NB reached the accuracy of 0.89 and the sensitivity of 0.61. DT reached the accuracy of 0.85 and sensitivity of 0.71. Another method of EML reached the accuracy of 0.88 and sensitivity of 0.69 and specificity of 0.91. Another model of KEML reached the accuracy of 0.88 and sensitivity of 0.76 and specificity of 0.92. OKEML-HGS-AOA reached the accuracy of 0.90 and sensitivity of 0.89 and specificity of 0.95 respectively. These results are represented in Figure 1, Figure 2 and Figure 3.

**Figure 1.** Analysis of proposed models for three datasets in terms of accuracy**Figure 2.** Analysis of proposed models for three datasets in terms of sensitivity

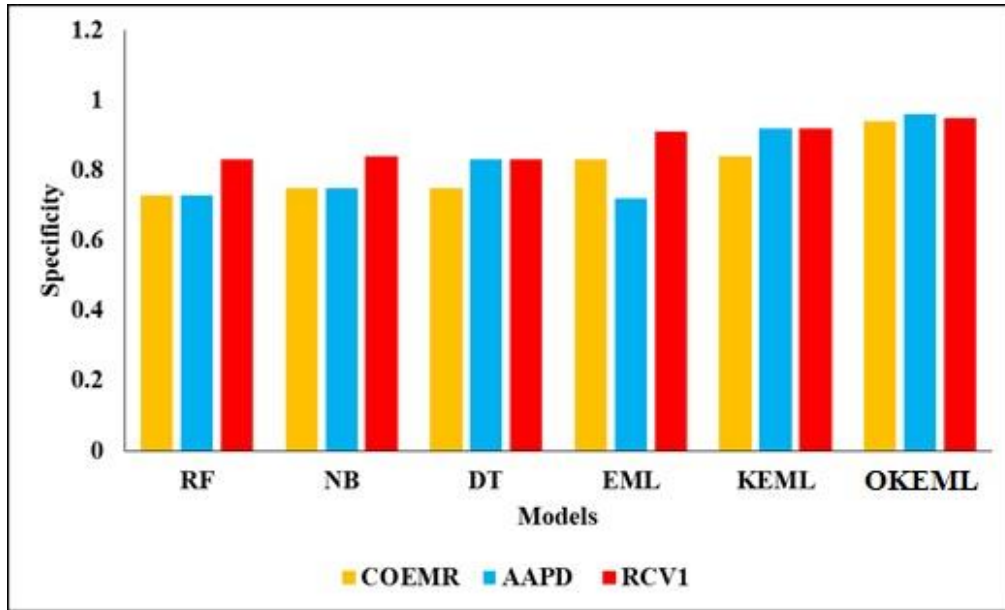


Figure 3. Analysis of proposed models for three datasets in terms of specificity

5. Conclusions

This paper proposed the OKEML paradigm to facilitate intelligent diagnosis from imbalanced EMRs. A two-stage training approach was proposed to separate the two. In the representation learning stage, the diagnostic outcomes of EMRs were taken into consideration to balance the data distribution. This paper used KEML for classification and the HGS-AOA hybrid model for selecting the most appropriate kernels. Experiments were conducted on COEMR datasets, which validated OKEML significantly enhanced the accuracy of intelligent diagnosis based on imbalanced EMRs, particularly for diagnosing diseases that occurred less frequently. As discussed in this study, the experimental outcomes were affected to varying degrees by the characteristics and classification algorithms used. Our next efforts will focus on incorporating clinicians' feedback with the extracted indicators, thus further enhancing the performance of the model. We will continue our theoretical investigation of multi-label classification performance gaps and provide new approaches for improving results. It is hoped that the diagnostic assistant will provide clinicians with a useful tool. The application of deep learning intelligent diagnosis will target at more complicated diseases like diabetes.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] W. C. Lin, J. S. Chen, M. F. Chiang, and M. R. Hribar, "Applications of artificial intelligence to electronic health record data in ophthalmology," *Translational Vision Science & Technology*, vol. 9, no. 2, pp. 13-13, 2020. <http://doi.org/10.1167/tvst.9.2.13>.
- [2] M. Gollapalli, S. A. Kudos, M. A. Alhamad, A. A. Alshehri, H. S. Alyemni, M. O. Alali, R. M. Mohammad, M. A. A. Khan, M. M. Abdulqader, and K. M. Aloup, "Machine learning models towards prediction of COVID and non-COVID 19 patients in the hospital's intensive care units (ICU)," *Math. Model. Eng. Probl.*, vol. 9, no. 6, pp. 1471-1480, 2022. <https://doi.org/10.18280/mmep.090605>.
- [3] S. M. Lauritsen, M. E. Kalør, E. L. Kongsgaard, K. M. Lauritsen, M. J. Jørgensen, J. Lange, and B. Thiesson, "Early detection of sepsis utilizing deep learning on electronic health record event sequences," *Artif. Intell. Med.*, vol. 104, pp. 101820-101820, 2020.
- [4] G. M. Mangipudi, S. Eswaran, and P. B. Honnavalli, "Quantum cryptography and quantum key distribution protocols: A survey on the Concepts, Protocols, Current Trends and Open Challenges," *Procedia Computer*

- Science, vol. 2019, 2019.
- [5] S. C. Huang, A. Pareek, R. Zamanian, I. Banerjee, and M. P. Lungren, "Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: A case-study in pulmonary embolism detection," *Sci. Rep.*, vol. 10, no. 1, pp. 1-9, 2020.
 - [6] S. B. Satukumati, S. Satla, and R. Kogila, "Feature extraction techniques for chronic kidney disease identification," *Ingenierie des Systemes d'Information*, vol. 24, no. 1, pp. 95-99, 2019. <https://doi.org/10.18280/isi.240114>.
 - [7] R. Vatambeti and V. K. Damera, "Gait based person identification using deep learning model of Generative Adversarial Network (GAN)," *Acadlore Trans. Mach. Learn.*, vol. 1, no. 2, pp. 90-100, 2022. <https://doi.org/10.56578/ataiml010203>.
 - [8] B. Amutha, K. Ramana, D. Gaurav, A. Gokul, K. Sandeep, Y. Kusum, and N. P. J. Maruthi, "A personalized eccentric cyber-physical system architecture for smart healthcare," *Secur. Commun. Networks*, vol. 2021, Article ID: 1747077, 2021. <https://doi.org/10.1155/2021/1747077>.
 - [9] J. R. A. Solares, F. E. D. Raimondi, Y. Zhu, F. Rahimian, D. Canoy, J. Tran, A. C. P. Gomes, A. H. Payberah, M. Zottoli, M. Nazarzadeh, N. Conrad, K. Rahimi, and G. Salimi-Khorshidi, "Deep learning for electronic health records: A comparative review of multiple deep neural architectures," *J. Biomed. Inform.*, vol. 101, pp. 103337-103337, 2020. <http://doi.org/10.1016/j.jbi.2019.103337>.
 - [10] R. Vatambeti, S. V. Mantena, K. V. D. Kiran, M. Manohar, and C. Manjunath, "Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique," *Cluster Comput.*, vol. 2023, 2023. <https://doi.org/10.1007/s10586-023-03970-7>.
 - [11] J. S. Obeid, J. Dahne, S. Christensen, S. Howard, T. Crawford, L. J. Frey, T. Stecker, and B. E. Bunnell, "Identifying and predicting intentional self-harm in electronic health record clinical notes: deep learning approach," *JMIR Med. Inform.*, vol. 8, no. 7, Article ID: e17784, 2020. <https://doi.org/10.2196/17784>.
 - [12] V. G. Biradar, H. Nagaraj, S. Mohan, and P. K. Pareek, "Industrial Fluids Components Health Management Using Deep Learning," In *Artificial Neural Networks - Recent Advances, New Perspectives and Applications*, Croatia, Rijeka: IntechOpen, Article ID: 107929, 2022.
 - [13] H. Goodrum, K. Roberts, and E. V. Bernstam, "Automatic classification of scanned electronic health record documents," *Int. J. Med. Inform.*, vol. 144, Article ID: 104302, 2020. <https://doi.org/10.1016/j.ijmedinf.2020.104302>.
 - [14] Y. M. Zhao, L. Hu, and L. Chi, "IKAR: An interdisciplinary knowledge-based automatic retrieval method from Chinese electronic medical record," *Information*, vol. 14, no. 1, pp. 49-49, 2023. <https://doi.org/10.3390/info14010049>.
 - [15] Z. Y. Liang, Z. C. Zhang, H. Y. Chen, and Z. Q. Zhang, "Disease prediction based on multi-type data fusion from Chinese electronic health record," *Math. Biosci. Eng.*, vol. 19, no. 12, pp. 13732-13746, 2022. <https://doi.org/10.3934/mbe.2022640>.
 - [16] H. Yang, J. X. Li, S. R. Liu, X. L. Yang, and J. L. Liu, "Predicting risk of hypoglycemia in patients with Type 2 Diabetes by electronic health record-based machine learning: Development and validation," *JMIR Medical Informatics*, vol. 10, no. 6, Article ID: e36958, 2022. <https://doi.org/10.2196/36958>.
 - [17] K. L. Zhang, S. Zhang, Y. Song, L. K. Cai, and B. Hu, "Double decoupled network for imbalanced obstetric intelligent diagnosis," *Math. Biosci. Eng.*, vol. 19, pp. 10006-10021, 2022. <https://doi.org/10.3934/mbe.2022467>.
 - [18] A. Sridhar, A. Mawia, and A. L. Amutha, "Mobile application development for disease diagnosis based on symptoms using machine learning techniques," *Procedia Comput. Sci.*, vol. 218, pp. 2594-2603, 2023.
 - [19] X. Q. Pang, C. B. Forrest, F. Lê-Scherban, and A. J. Masino, "Prediction of early childhood obesity with machine learning and electronic health record data," *Int. J. Med. Inform.*, vol. 150, Article ID: 104454, 2021.
 - [20] Y. W. Meng, W. Speier, M. K. Ong, and C. W. Arnold, "Bidirectional representation learning from transformers using multimodal electronic health record data to predict depression," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 8, pp. 3121-3129, 2021. <https://doi.org/10.1109/JBHI.2021.3063721>.
 - [21] V. Khare and S. Kumari, "Performance comparison of three classifiers for fetal health classification based on cardiotocographic data," *Acadlore Trans. Mach. Learn.*, vol. 1, no. 1, pp. 52-60, 2022. <https://doi.org/10.56578/ataiml010107>.
 - [22] P. C. Yang, X. Sun, W. Li, S. M. Ma, W. Wu, and H. F. Wang, "SGM: Sequence generation model for multilabel classification," pp. 3915-3926, vol. 2018, 2018. <https://doi.org/10.48550/arXiv.1806.04822>.
 - [23] D. D. Lewis, Y. M. Yang, T. Russell-Rose, and F. Li, "Rev1: A new benchmark collection for text categorization research," *Mach. Learn. Res.*, vol. 2004, pp. 361-397, 2004.