Applied Soft Computing 122 (2022) 108822

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Introducing macrophages to artificial immune systems for earthquake prediction

Wen Zhou [a,*], Yiwen Liang [b], Xinan Wang [c], Zhe Ming [b,a], Zhenhua Xiao [d], Xiying Fan [e]

[a] School of Computer Science, Hubei University of Technology, No. 28, Nanli Road, Hongshan District, Wuhan, China
[b] School of Computer Science, Wuhan University, No. 299 Bayi Road, Wuchang District, Wuhan, China
[c] School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, No. 2199, Nanshan District, Shenzhen, China
[d] School of Artificial Intelligence, Hubei Business College, No. 632, Xiongchu Street, Hongshan District, Wuhan, China
[e] School of Computer and Communication Engineering, University of Science and Technology Beijing, No. 30, Haidian District, Beijing, China

ARTICLE INFO

Article history:
Received 1 February 2020
Received in revised form 23 March 2022
Accepted 30 March 2022
Available online 12 April 2022

Keywords:
Artificial macrophage algorithm
Earthquake prediction
Artificial immune systems
Stochastic gradient descent

ABSTRACT

Earthquake prediction (EQP) is crucial for taking preemptive measures and accurately predicting damage. Several historical seismic-event-based EQP approaches have been proposed; however, these approaches only identify anomalies without distinguishing noise, reducing the prediction accuracy. Macrophages play an important role in the immune system by recognizing viruses, apoptotic cells, and normal cells, as well as performing immune responses and suppression to ensure homeostasis; that is, macrophages exhibit strong classification capabilities and self-adaptability. Therefore, in this study, a novel artificial macrophage algorithm (AMA) for EQP is proposed. More specifically, we first introduce the biological mechanism of macrophages to establish recognition and learning mechanisms to identify noise and anomalies. Second, we adopt a distance metric to denote the weights of the AMA, instead of using experience-based parameters. Finally, a stochastic gradient descent is introduced to ensure the adaptability of the AMA. The performance of the AMA was assessed through an analysis of historical seismic events in Sichuan and its surroundings. Our experimental results demonstrate that AMA outperforms state-of-the-art EQP algorithms. The parameters and statistical tests of AMA were further analyzed in this study.

1. Introduction

Sichuan has emerged as one of the world's most earthquake-prone regions, with more than 200 quakes of greater than 3 moment magnitude (MW) per year [1]. In 2008, an earthquake of 8.3 MW struck Sichuan, which was the largest earthquake in China in recent years, causing more than 60,000 deaths and 374,000 injuries [2]. On October 8, 2017, another large earthquake with magnitude 6.5 MW struck Sichuan, causing 29 deaths and 543 injuries [1]. Therefore, more accurate EQP approaches are urgently required.

EQP aims to predict the magnitude, occurrence time, location, and probability of future earthquakes [3]. The main challenge is determining the relationships between historical seismic data and earthquakes. To address this challenge, one approach is to treat EQP as a binary classification task (predicting the occurrence (1) or non-occurrence (0) of an earthquake with predetermined magnitude, time, and location), and then solve the EQP problem using existing binary classification methods [4–6]. However, owing to the insufficient data on medium-large seismic data, these methods perform poorly. For example, neural networks and deep learning (DL) methods require a large number of training sets; otherwise, their performance is unsatisfactory.

The artificial immune system (AIS) requires few training samples and has strong recognition ability; therefore, some researchers consider that it can be introduced into EQP [7]. Hence, it has been employed to solve EQP tasks [8,9][10]. In particular, in [11], a numerical differential-based dendritic cell algorithm (nd-DCA) was proposed to predict quakes with magnitudes greater than 4.5 over the next month. First, the Gutenberg–Richer inverse power-law (GR) and other seismic distributions were introduced for eight feature extractions. Second, first-order and second-order numerical differentials were adopted for signal acquisition. This approach yielded satisfactory results. However, it overlooked noise in seismic data, thereby reducing the model's prediction accuracy. We identified that this is because the numerical differential can only identify anomalies without distinguishing noise. In addition, it lacks a learning process, and its parameters are manually determined, which affects the performance of the EQP model.

* Corresponding author.
E-mail address: zw_mmwh@whu.edu.cn (W. Zhou).

https://doi.org/10.1016/j.asoc.2022.108822
1568-4946/

Accordingly, we were motivated to develop a more effective discrimination method that detects noise and automatically adjusts the model's parameters. Macrophages play an important role in the human immune system because they recognize viruses, apoptotic cells, and normal cells, as well as perform immune responses and suppression to ensure homeostasis [12]. This implies that macrophages allow us to effectively distinguish various data (including noise) and automatically obtain parameter regulation; subsequently, it may achieve better binary classification performance. Therefore, we attempted to adapt the properties of macrophages to the EQP scenario.

This paper systematically explores the simulation of biological macrophage recognition and regulation abilities, and presents a novel AIS called AMA. To avoid the effect of noise on EQP models, because macrophages can recognize various signals, we constructed an artificial macrophage recognition method, that combines seismological theory [3] and information gain [13] to replace the differentiation phase in [11], which can then distinguish noise from normal and abnormal data. Moreover, to ensure the adaptability of the proposed AMA, we applied the well-known stochastic gradient algorithm [14] to optimize the parameters of the AMA. Finally, the proposed approach is implemented for EQP of magnitudes greater than 4.5 MW within the next month in Sichuan and its surroundings.

The main contributions of this paper are as follows:

- We provide a new insight into AIS from a biological perspective and propose an artificial macrophage classification and optimization algorithm to adapt to the EQP scenario.
- We propose a novel EQP model, called AMA, which combines seismological theory and information gain to deal with noise to ensure the model's prediction accuracy.
- A distance-based dynamic signal fusion model is proposed, which not only avoids weights based on experience, but also enhances the processing ability of AMA.
- Inspired by stochastic gradient descent, we propose a two-level optimization strategy, including positive and negative learning, to ensure the adaptability of the AMA.

2. Related works

Artificial intelligence methods have been widely implemented for EQP [15–17]. For instance, Dehbozorgi and Farokhi [18] first investigated the neuro-fuzzy classifier for short-term EQP using historical seismic events. Various indicators, including magnitude, depth, time, location, statistics, and entropy parameters, were used to predict whether an earthquake would occur within the next five minutes. According to the spatial analysis of the magnitude distribution, an automatic clustering-based adaptive neuro-fuzzy inference system (ANFIS) was proposed for EQP [19]. However, these methods have difficulty summarizing effective rules for EQP tasks [20,21]. To address the EQP problem, many shallow machine learning studies, such as [22–24], have been proposed. Shi et al. [25] first introduced an artificial neural network (ANN) in EQP and established a relationship between earthquake magnitude and epicentral intensity. Then, a support vector regressive and hybrid neural network (SVR-HNN) was introduced to achieve EQP [26]; this work used maximum relevance and minimum redundancy criteria to extract the relevant indicators. Later, support vector machines and random forests were utilized to predict earthquakes in Cyprus [27]. In another study [6], a principal component analysis-based random forest (PCA-RF) was employed to reduce data dimension and generate new datasets for generalizing existing prediction models. The results show that the average accuracy of these generalized approaches was significantly improved compared to that of existing methods.

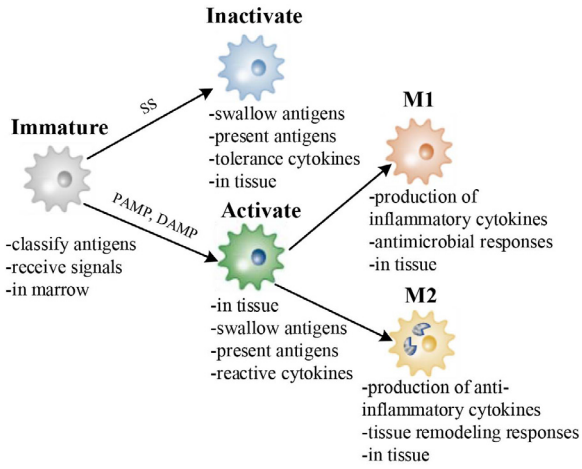


Fig. 1. Macrophage behaviors.

However, differences in geological structures hinder their universality. DL-based methods can autonomously calculate thousands of sophisticated indicators. Therefore, recurrent neural networks (RNNs) [28] and convolutional neural networks (CNNs) [29] have attracted the attention of several EQP researchers. For instance, DeVries et al. [30] introduced a static-stress-based criterion that could predict aftershock locations without presuming the fault orientation. In addition, it provided an improved aftershock location prediction method and identified the physical quantities that controlled the triggering of earthquakes when the earthquake cycle was active. Owing to the dynamic and unpredictable nature of earthquakes, Wang et al. [31] introduced long short-term memory (LSTM) to study the spatiotemporal relationship between earthquakes at different locations and further demonstrated the robustness and effectiveness of their method. However, because DL-based EQP approaches require a large amount of training data to ensure accuracy, it is difficult for them to make predictions based on historical data [16].

AIS has been widely utilized for privacy preservation [32], anomaly detection [33], and EQP [11]. In [8], because a negative selection algorithm (NSA) could identify self and non-self components, it was used to reduce the impact of inadequate earthquake data on training performance. Because DCA required few training samples, it was introduced for EQP [9]. When using a small sample set, the reliability of DCA is higher than that of the back propagation neural network (BPNN), indicating that it is more suitable for EQP with a small sample set. In [10], Haskell-based DCA was used for EQP (hDCA), which attenuated the imprecision of DCA and obtained satisfactory classification accuracy. However, these approaches only identify anomalies without distinguishing noise and are adaptively lacking, lowering prediction accuracy.

Moreover, owing to the strong identification ability of biological macrophages [12], it provides a novel binary classification method to process noise data using seismological theory and the information gain method. Because macrophages have the ability to adjust their state, a two-level optimization strategy can be inspired using stochastic gradient descent for the proposed AMA. Therefore, we can adapt the AMA to the EQP scenario.

3. Methodology

3.1. Overview of macrophages

Macrophages are key components in triggering humoral immune responses by engulfing antigens via various receptors,

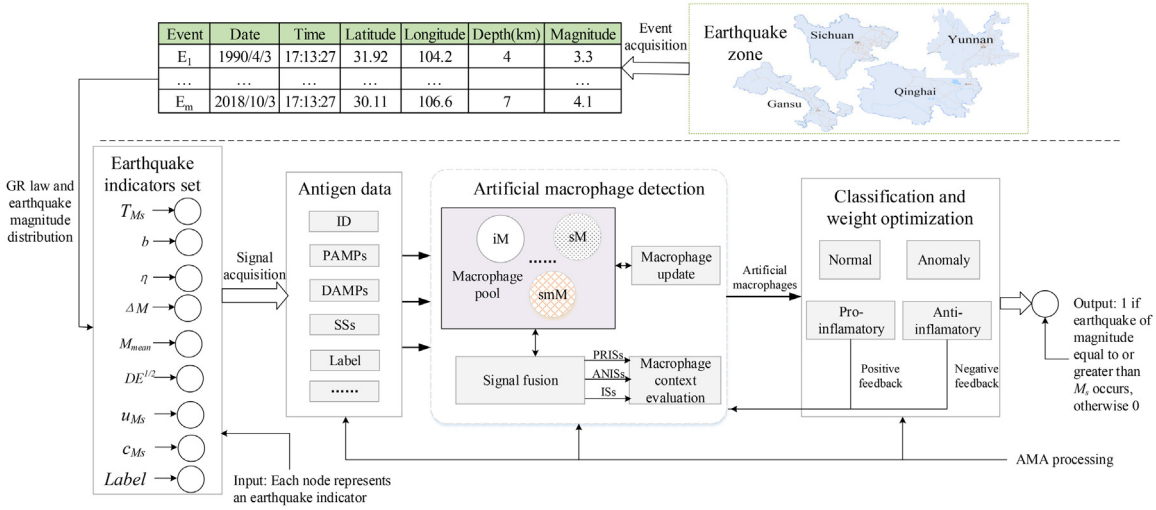


Fig. 2. Schematic of the proposed AMA model for EQP.

and they can interact with other immune cells to protect the immune system [12,34]. They have various functional properties that can be incorporated into an algorithm, as shown in Fig. 1, and include three steps: **(1) Initiation**: This describes the distribution of macrophage receptors resulting from the generation of macrophage inhibition receptors (SIRPα), activation receptors (toll-like receptor (TLR), and nucleotide-binding oligomerization domain-like receptors (NLR) through a stochastic procedure, which generates immature macrophages. **(2) Recognition phase**: This phase is designed to maximize the antigen recognition ability of immature macrophage receptors, including pathogen-associated molecular patterns (PAMPs), damage-associated molecular patterns (DAMPs), and safe signals (SSs). It generates semimature or mature macrophages. **(3) Regulation process**: Mature macrophages (activated) can interact with other cells and maintain the immune balance. They may be pro-inflammatory (M1) or anti-inflammatory (M2) [35]. If a macrophage is pro-inflammatory, it releases inflammatory cytokines to trigger a humoral immune response and regulates its concentration positively; otherwise, it releases anti-inflammatory cytokines and negatively regulates its concentration. In the context of engineering (i.e., in EQP), these properties are useful if it is required to detect an anomaly and analyze whether this is caused by noise or anomalous factors, and model adaptability can be achieved.

3.2. Proposed AMA model

The architecture of AMA for EQP, which is summarized in Fig. 2, comprises three parts: signal acquisition, artificial macrophage detection, and classification and weight optimization. The seismic event set $E = \{E_1, \ldots, E_i, \ldots, E_N\}$ is used to compute the feature matrix, which is defined as $F = \{F_1, \ldots, F_i, \ldots, F_N\}$, $0 < i < N$, where N is the number of seismic events, and $F_i = \{b_i, \eta_i, \Delta M_i, T_{Ms\,i}, \mu_{Ms\,i}, c_{Ms\,i}, dE_i^{1/2}, M_{meani}\}$, as explained in [11]. The output is the prediction result, which is one if an earthquake of magnitude equal to or greater than M_s (a predefined seismic magnitude threshold) occurs and zero otherwise [36].

3.2.1. Signal acquisition

During this phase, the matrix F is selected as the signal source of the AMA, and it is mapped to *PAMPs*, *DAMPs*, and *SSs* by information gain based on seismological theory. *PAMPs* indicate abnormal indicators, and that the earthquake activity is strong. *DAMPs* indicate a high possibility of noise data occurrence. *SSs* indicate that the probability of normal is relatively high, that is, earthquake activity is weak. Because information gain can characterize the amount of information carried by a feature, it can be introduced for signal mapping [13]. The information gain of attribute A relative to a collection of examples C, termed $Gain(C; A)$, is defined by Eq. (1), where $Values(A)$ represents the set of all possible values for attribute A, and C_v is the subset of C for which attribute A has a value v. The entropy of C relative to the binary classification, termed $Entropy(C)$, is defined by Eq. (2), where p_i is the proportion of C belonging to class i.

$$Gain(C; A) = Entropy(C) - \sum_{v \in Values(A)} (|C_v|/|C| Entropy(C_v)) \tag{1}$$

$$Entropy(C) = -\sum_{i=1}^{2} p_i \log_2 p_i \tag{2}$$

According to [3], μ_{Ms} and c_{Ms} describe the average interval between the occurrence of typical earthquake events and its coefficient of variation to determine whether an earthquake has occurred, respectively. Therefore, they serve as the source of the *SSs*. For their part, the indicators T, M_{mean}, and $dE^{1/2}$ describe phenomena, such as the rate of earthquake energy release, which is the internal factor of earthquake occurrence; therefore, they are adopted to describe *DAMPs*. Moreover, the indicators b, η, and ΔM describe the relationship between magnitude and frequency, which are objective descriptions of the occurrence of an earthquake; therefore, they are used to describe *PAMPs*. Therefore, for an instance F_i, *SS* value (SS_i) is calculated by the information gains and values of features $\mu_{Ms\,i}$ and $c_{Ms\,i}$, *DAMP* value ($DAMP_i$) is calculated by the information gains and values of features T_i, $M_{mean\,i}$ and $dE^{1/2}{}_i$, *PAMP* value ($PAMP_i$) is calculated by the information gains and values of features b_i, η_i, and ΔM_i. Inspired by biological immunity, macrophages can identify PAMP, DAMP and SS, where PAMP and SS are opposite signals. To make the AMA simple to understand, we assume that one of the signals in PAMP and SS should be zero. The values of PAMP and SS are determined by their information gains, described by $SS_G = Gain(C; \mu_{Ms_i}) + Gain(C; c_{Ms_i})$ and $PAMP_G = Gain(C; b_i) + Gain(C; \eta_i) + Gain(C; \Delta M_i)$, respectively. If SS_G is greater than $PAMP_G$, then $PAMP_G$ is equal to zero; otherwise, SS_G is equal to zero. The acquisition method for each signal of an instance is given in Eq. (3), and these signals are the input of the artificial

Table 1
The weighting calculation method for the signal processing function.

W	PRIS	ANIS	IS
PAMP	$dist(PAMP_i, mø_j.TLR)$	$dist(PAMP_i, mø_j.NLR)$	$dist(PAMP_i, mø_j.SIRP\alpha)$
DAMP	$dist(DAMP_i, mø_j.TLR)$	$dist(DAMP_i, mø_j.NLR)$	$dist(DAMP_i, mø_j.SIRP\alpha)$
SS	$dist(SS_i, mø_j.TLR)$	$dist(SS_i, mø_j.NLR)$	$dist(SS_i, mø_j.SIRP\alpha)$

macrophage detection phase.

$$\begin{cases} DAMP_i = Gain(C; T_i) * T_i + Gain(C; M_{mean\,i}) * M_{mean\,i} \\ \qquad + Gain(C; dE_i^{1/2}) * dE_i^{1/2}; \\ PAMP_i = 0, SS_i = Gain(C; \mu_{M_s\,i}) * \mu_{M_s\,i} + Gain(C; c_{M_s\,i}) * c_{M_s\,i}, \\ \qquad if \quad SS_G > PAMP_G, \\ SS_i = 0, PAMP_i = Gain(C; b_i) * b_i + Gain(C; \eta_i) * \eta_i \\ \qquad + Gain(C; \Delta M_i) * \Delta M_i, \quad if \quad SS_G \le PAMP_G. \end{cases} \tag{3}$$

3.2.2. Artificial macrophage detection

For each input data $ag = \langle ID, PAMP, DAMP, SS, Status, Label\rangle$, N_{subm} macrophages were selected from the M pool to sample it, where N_{subm} indicates the number of macrophages sampled for each ag in each iteration. *Label* and *Status* indicate the class and the prediction result of seismic events, respectively. An artificial macrophage is denoted as $mø = \langle ID, MR, PRIS, ANIS, Context, Fitness, IS\rangle$, where $MR = \langle type, weight, affinity\rangle$, $weight$ is the receptor value, and *affinity* denotes the affinity between the receptor and the ag. *PRIS*, *ANIS*, and *IS* are signs of earthquakes caused by seismic-related patterns, non-seismic-related patterns, and signs of an earthquake occurrence, respectively. *Context* is the current state of macrophages, whereas *Fitness* is the measurement of macrophages throughout the evolutionary process.

The artificial macrophage migration threshold, ρ, which is defined by Eq. (4), is adopted to determine whether a macrophage is in a semi-mature or mature state, where max_{PAMP}, max_{DAMP}, and max_{SS} represent the maximum values of *PAMP*, *DAMP*, and *SS*, respectively. For an ag_i, the fitness of a macrophage ($mø_j.Fitness$) is denoted by Eq. (5), where $dist(PAMP_i, mø_j.TLR) = \sqrt{PAMP_i^2 + mø_j.TLR.Weight^2}$, $dist(DAMP_i, mø_j.NLR) = \sqrt{DAMP_i^2 + mø_j.NLR.Weight^2}$, and $dist(SS_i, mø_j.SIRP\alpha) = \sqrt{SS_i^2 + mø_j.SIRP\alpha.Weight^2}$. If $mø_j.Fitness$ is greater than ρ, the state of the macrophage can be evaluated (mature, means normal or semi-mature, means abnormal). Once the state of a macrophage changes, the output signals $mø_j.PRIS$, $mø_j.ANIS$, and $mø_j.IS$ are calculated using the signal fusion method to perform anomaly detection, which are defined by Eq. (6), where the weights are listed in Table 1.

$$\rho = 0.5 * (max_{PAMP} * dist(PAMP, TLR) + max_{DAMP} * dist(DAMP, NLR) + max_{SS} * dist(SS, SIRP\alpha)) \tag{4}$$

$$mø_j.Fitness = dist(PAMP_i, mø_j.TLR) + dist(DAMP_i, mø_j.NLR) + dist(SS_i, mø_j.SIRP\alpha) \tag{5}$$

$$mø_j[PRIS, ANIS, IS] = \frac{W_{PAMP} * PAMP_i + W_{DAMP} * DAMP_i + W_{SS} * SS_i}{W_{PAMP} + W_{DAMP} + W_{SS}} \tag{6}$$

If the sum of the output $mø.PRIS$ and $mø.ANIS$ is greater than that of $mø.IS$, the macrophage is in an anomalous context (1); otherwise, it is in a normal context (0). Migrated artificial macrophages must be updated. To facilitate this operation, macrophages with the highest fitness value were selected from the original artificial macrophages to replace the migrated macrophages.

3.2.3. Classification and weight optimization

The antigen context value is used to determine the degree (K) of abnormality of an antigen by measuring the number of abnormal antigens, which is the ratio of mature antigens to total antigens. K is used for comparison with an abnormality threshold θ, which represents the ratio of abnormal seismic data instances to total instances in all samples. For an antigen, if K is greater than θ, it is abnormal; otherwise, it is normal. For a macrophage, if $mø.PRIS$ exceeds $mø.ANIS$, it is in a pro-inflammatory context, and the $mø.MR.Weight$ needs to be increased; in contrast, if it is in an anti-inflammatory context, the $mø.MR.Weight$ needs to be lowered. In this study, we adopt a stochastic gradient descent (defined in Eq.(7)) to implement weight optimization, where σ represents the learning rate, if *ANIS* exceeds *PRIS*, $-1 < \sigma < 0$, else $0 \le \sigma < 1$, and $F(ag, MR.Weight_i)$ is described in Eq. (8).

$$MR.Weight_{k+1} = MR.Weight_k + \sigma \bigtriangledown_{MR.Weight_k} F(ag, MR.Weight_k) \tag{7}$$

$$F(ag, MR.Weight_k) = \mathrm{Sgn}\left(\frac{\sum_{j=1}^{N_{subm}} \mathrm{Sgn}(mø_j.PRIS + mø_j.ANIS - mø_j.IS)}{N_{subm}} - \theta\right), \quad \text{if} \quad mø_j.Fitness > \rho \tag{8}$$

Algorithm 1 AMA

Input: S = a set of antigens, indicating data instances classified as safe or dangerous.
Output: G,L=set of antigens classified as normal/anomalous;W=weights.

```
Create M pool of N_m macrophage and macrophage initiation
while iterations ≤ ITE do
  for all antigens ag ∈ S do
    Randomly pick N_subm macrophages ∈ M
    Update PAMPs, DAMPs, and SSs according to Eq. (3)
    Calculate mø.Fitness, affinity, and ρ
    if mø.Finess> ρ then
      Compute mø.PRIS, mø.ANIS, mø.IS using Eq. (6)
    end if
  end for
  for all macrophages mø ∈ MatureM do
    if mø.PRIS + mø.ANIS > mø.IS then
      mø.context ← mature
      if mø.PRIS < mø.ANIS then
        W ← Weight optimization by Eq. (7) (−1 < σ < 0)
      else
        W ← Weight optimization by Eq. (7) (0 ≤ σ < 1)
      end if
    else
      mø.context ← semimature
    end if
  end for
  for all antigens ag ∈ S do
    Calculate K for all macrophages mø ∈ M that sampled ag
    Insert ag to L or G according to K and θ
  end for
end while
return G, L, W
```

In particular, the overall algorithm of AMA is outlined in Algorithm 1. The complexity of AMA is $O(N_m + ITE * (n + 2 * N_m + N_{subm} * n))$=$O(ITE * N_{subm} * n)$. The algorithm stops when the iteration number exceeds *ITE*, the time complexity of signal acquisition, artificial macrophage detection, and classification and weight optimization are $O(N_m)$, $O(N_m + n)$, and $O(N_{subm} * n)$, respectively.

Table 2
Comparison between AMA and other approaches when predicting DS1.

	PPV	NPV	Rn	S	FAR	MCC	AUC	Avg	R	ACC
AMA	0.41	0.83	0.71	0.35	**0.21**	0.10	0.56	0.58	0.49	0.60
DCA [9]	0.43	0.75	0.72	0.53	0.25	0.18	0.60	0.61	0.47	0.59
NSA [8]	0.21	0.52	0.21	0.51	0.38	0.02	0.51	0.36	−0.17	0.32
PCA-RF [6]	0.55	0.61	0.60	0.21	0.42	0.23	0.52	0.49	0.18	0.41
BPNN [22]	0.64	0.63	0.35	0.15	0.37	0.23	0.61	0.44	−0.02	0.45
RNN [28]	0.27	0.75	0.68	0.32	0.24	0.04	0.48	0.51	0.44	0.52
PNN [23]	0.32	0.64	0.13	**0.83**	0.33	−0.02	0.51	0.48	−0.20	0.39
hDCA [10]	0.32	0.74	0.70	0.37	0.26	0.06	0.53	0.53	0.44	0.53
LSTM [29]	0.42	0.81	0.69	0.31	0.23	0.12	0.54	0.56	0.46	0.56
SVR-HNN [26]	0.38	0.76	0.67	0.29	0.28	0.09	0.51	0.53	0.39	0.53
nd-DCA [11]	**0.73**	**0.88**	**0.90**	0.70	0.22	**0.31**	**0.73**	**0.80**	**0.78**	**0.71**

Table 3
Comparison between AMA and other approaches when predicting DS2.

	PPV	NPV	Rn	S	FAR	MCC	AUC	Avg	R	ACC
AMA	0.82	**1.00**	**1.00**	**0.95**	**0.00**	**0.90**	**0.98**	**0.94**	**1.00**	**0.96**
DCA [9]	0.17	0.82	0.50	0.52	0.18	−0.01	0.46	0.50	0.32	0.50
NSA [8]	0.00	0.51	0.49	0.00	0.42	0.00	0.45	0.25	−0.42	0.25
PCA-RF [6]	0.61	0.56	0.58	0.77	0.41	0.02	0.43	0.63	0.37	0.62
BPNN [22]	0.67	0.96	0.67	0.04	0.04	0.63	0.82	0.58	0.63	0.56
RNN [28]	**0.86**	0.98	0.90	0.91	0.01	0.87	0.90	0.91	0.89	0.87
PNN [23]	0.84	0.92	0.81	0.82	0.06	0.81	0.86	0.85	0.75	0.82
hDCA [10]	0.73	**1.00**	**1.00**	0.93	**0.00**	0.83	0.97	0.92	**1.00**	0.88
LSTM [29]	0.78	0.82	0.91	0.87	0.06	0.82	0.85	0.85	0.85	0.93
SVR-HNN [26]	0.69	0.80	0.86	0.79	0.12	0.68	0.66	0.79	0.74	0.75
nd-DCA [11]	0.79	0.92	0.92	0.81	0.02	0.01	0.90	0.86	0.93	0.89

Table 4
Comparison between AMA and other approaches when predicting DS3.

	PPV	NPV	Rn	S	FAR	MCC	AUC	Avg	R	ACC
AMA	**0.87**	**1.00**	**1.00**	**0.95**	**0.00**	**0.93**	**0.97**	**0.96**	**1.00**	**0.95**
DCA [9]	0.33	0.78	0.24	0.15	0.22	0.10	0.62	0.37	0.02	0.37
NSA [8]	0.33	0.69	0.73	0.49	0.28	0.18	0.62	0.56	0.45	0.58
PCA-RF [6]	0.32	0.73	0.43	0.21	0.32	0.12	0.44	0.42	0.11	0.40
BPNN [22]	0.56	0.90	0.50	0.08	0.10	0.44	0.71	0.51	0.40	0.49
RNN [28]	0.45	0.66	0.66	0.55	0.33	0.02	0.52	0.58	0.14	0.59
PNN [23]	0.47	0.79	0.64	0.59	0.19	0.38	0.51	0.62	0.45	0.63
hDCA [10]	0.46	0.91	0.82	0.63	0.09	0.41	0.73	0.71	0.73	0.72
LSTM [29]	0.65	0.79	0.81	0.53	0.11	0.65	0.71	0.70	0.70	0.71
SVR-HNN [26]	0.59	0.62	0.73	0.49	0.19	0.59	0.63	0.61	0.54	0.65
nd-DCA [11]	0.73	0.87	0.75	0.69	0.16	0.59	0.69	0.76	0.85	0.84

Table 5
Comparison between AMA and other approaches when predicting DS4.

	PPV	NPV	Rn	S	FAR	MCC	AUC	Avg	R	ACC
AMA	0.73	0.96	**0.98**	**0.82**	**0.01**	**0.78**	**0.89**	**0.87**	**0.98**	**0.86**
DCA [9]	0.33	0.58	0.50	0.59	0.42	−0.09	0.43	0.50	0.08	0.50
NSA [8]	0.00	0.32	0.32	0.00	0.55	0.00	0.23	0.16	−0.55	0.20
PCA-RF [6]	0.31	0.59	0.53	0.23	0.43	0.02	0.50	0.41	0.10	0.40
BPNN [22]	**1.00**	0.10	0.15	0.00	0.90	0.12	0.59	0.31	−0.75	0.45
RNN [28]	0.63	0.67	0.59	0.60	0.28	0.55	0.57	0.62	0.31	0.60
PNN [23]	0.66	0.68	0.53	0.61	0.29	0.46	0.55	0.62	0.24	0.61
hDCA [10]	0.67	**0.98**	0.97	0.77	0.02	0.70	0.87	0.85	0.95	0.84
LSTM [29]	0.35	0.31	0.48	0.42	0.21	0.57	0.46	0.39	0.37	0.41
SVR-HNN [26]	0.37	0.56	0.61	0.59	0.16	0.65	0.56	0.53	0.45	0.55
nd-DCA [11]	0.91	0.89	0.88	0.75	**0.01**	0.76	0.87	0.86	0.92	0.81

4. Experiments

4.1. Datasets and baseline methods

We conducted experiments on four real-world seismic datasets from Yunnan (DS1), Gansu (DS2), Sichuan (DS3), and Qinghai (DS4), which were provided by the China Earthquake Networks Center [1]. The data preprocessing method was based on [11]. Moreover, the proposed AMA was validated and compared to various well-known EQP approaches: DCA [9], NSA [8], PCA-RF [6], BPNN [22], RNN [28], probabilistic neural network (PNN) [23], hDCA [10], LSTM [29], SVR-HNN [26], and nd-DCA [11].

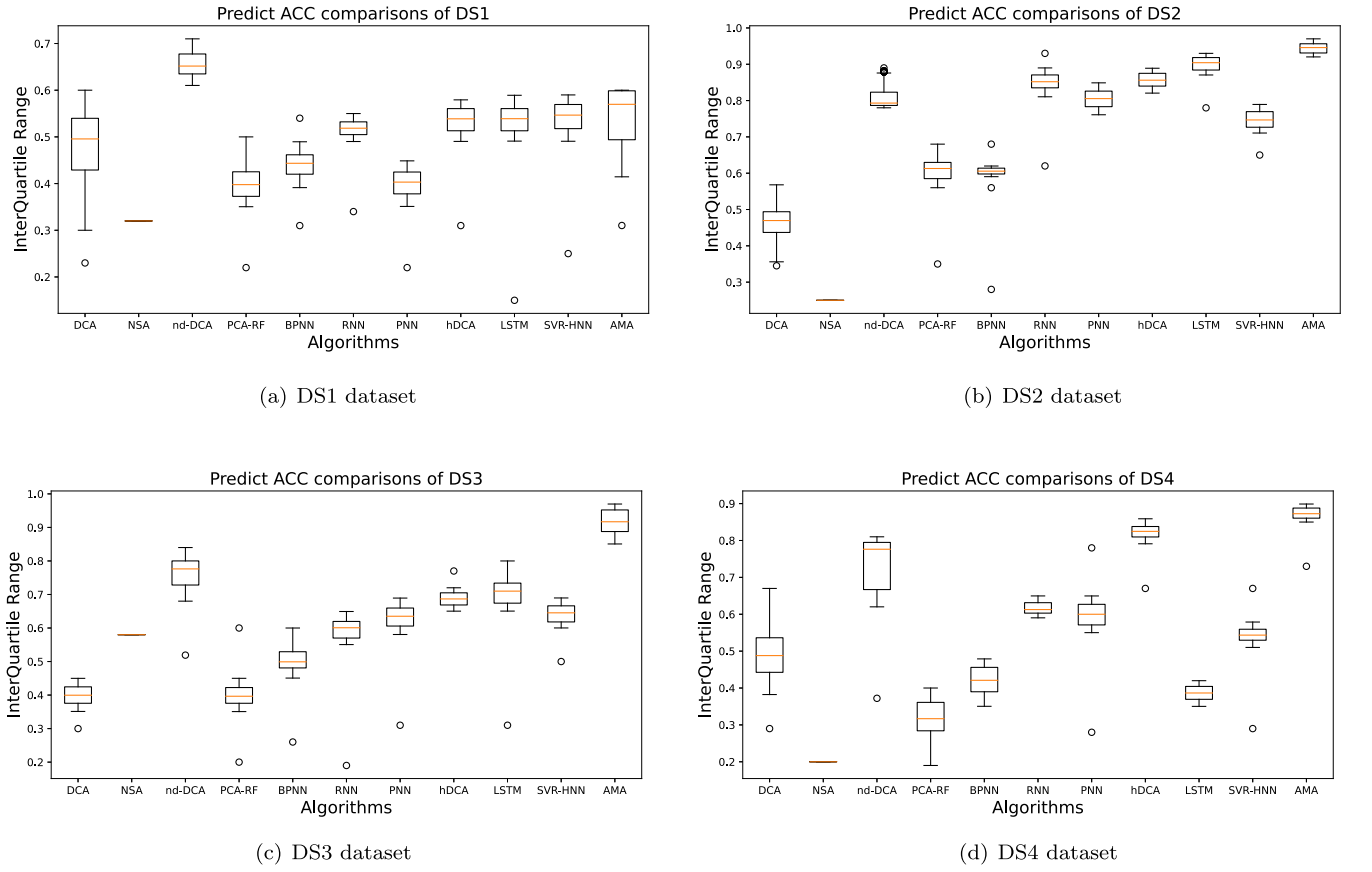
In the prototype implementation, the value of the input signals was normalized to be between 0 and 1; therefore, the weight of each macrophage receptor was also set to a random value between 0 and 1. The seismic magnitude threshold M_s was set to 4.5. Other parameters were set as follows: $N_m = 100$, N_{subm} =10, and $\sigma = \pm 0.03$ (+0.03 denotes positive regulation, and −0.03 denotes negative regulation). Moreover, the number of iterations, *ITE*, was set to 150. DCA and hDCA use PCA as the signal acquisition method. For DCA, hDCA and nd-DCA, 100 DCs were created for the pool, and 10 DCs were selected at random to sample each antigen. The NSA adopted the Euclidean distance to measure the similarity between antigens and antibodies, and the self-radius was set to 0.02. For the PCA-RF, the parameters were set as follows: maxDepth=0, numFeatures=0, and numTrees=10. The BPNN in our experiment had three layers: an input layer, a hidden layer, and an output layer. The learning rate was set to 0.0003, and the number of training epochs was set to 100. The weight in the multilayer perceptron was learned using the standard gradient descent. The RNN adopted a sigmoid as the activation function, and it had eight hidden layers. The batch size was set to 128, the learning rate was set to 0.001, and the number of training epochs was set to 100. The PNN in our experiment had an embedded layer, one product layer, and three hidden layers. In LSTM, the learning rate was set to 0.001; the momentum was used to accelerate model convergence, and was set to 0.9. We used a weight decay of 0.0005, and the dropout rate of the two dropout layers was 0.5. The settings of the other parameters of the compared approaches can be found in the related literature.

4.2. Prediction verification

In this study, the following statistical indicators were used to evaluate 11 different models in terms of their EQP results: predictive positive value (PPV), negative predictive value (NPV), recall rate (Rn), specificity (S), false acceptance rate (FAR), Matthews correlation coefficient (MCC), R-score (R), accuracy (ACC), and area under the curve (AUC). In addition, the average (Avg) of PPV, NPV, Rn, and S was calculated to provide a measure of the overall prediction quality. The meanings and calculation methods of these indicators were outlined in a previous study [22] or [37].

4.3. Results and analysis

This section presents the experimental results of DS1, DS2, DS3 and DS4. The results for each area are discussed separately. Tables 2, 3, 4, and 5 present the results for DS1 to DS4, respectively. In all the tables, the number in bold represents the best result among the compared algorithms. Referring to DS1 (see Table 2), when using the performance evaluation indicators to evaluate the proposed AMA, multiple indicators were observed to be superior to NSA, PCA-RF, BPNN, and PNN. However, the results were not as good as those of nd-DCA in most indicators. Therefore, the results of the AMA were not competitive on this dataset. Table 3 summarizes the results for DS2. Notably, there was limited FAR in the predictions made by AMA, in contrast to the remaining methods, whereas R had a value of 1.00, which is a very desirable property in EQP approaches. The specific results for DS3 are presented in Table 4. The differences in Rn, MCC, and R between AMA and other algorithms were very clear, producing a difference exceeding 0.18 units, with the second-best classifier (hDCA) of Rn, a difference greater than 0.28 units with the second-best classifier (LSTM) of MCC, and a difference of more than 0.15 units compared to the second-best classifier (nd-DCA) of R. Finally, Table 5 presents the results for DS4. AMA provided better results



(a) DS1 dataset

(b) DS2 dataset

(c) DS3 dataset

(d) DS4 dataset

Fig. 3. Box plot of predicted ACC comparisons of each dataset.

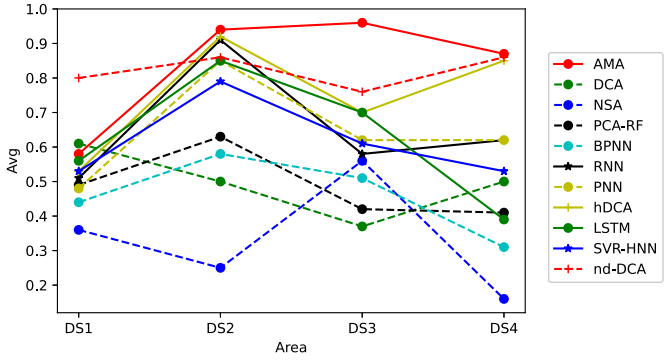


Fig. 4. Performance comparison between AMA and other algorithms.

for most evaluation indicators, with the smallest FAR, whereas R had a value of 0.98. When the other ten algorithms made predictions on DS4, the overall difference between AMA, nd-DCA, and hDCA was not particularly clear. BPNN and NSA were significantly weaker than the other algorithms in terms of the R indicator, which is a key indicator for an EQP algorithm. From a joint analysis of the four tables, because AMA achieved the best results on most indicators, such as R and Avg, we conclude that AMA is the most suitable classifier for these datasets.

The statistical characteristics of the ACC results of the 11 artificial intelligence approaches obtained from the repeated process with 150 runs are graphically shown in the box plot in Fig. 3. Notably in Fig. 3, the bottom and top of each box plot are the first and third quartiles of the sample, respectively, and the red band within the box denotes the median.

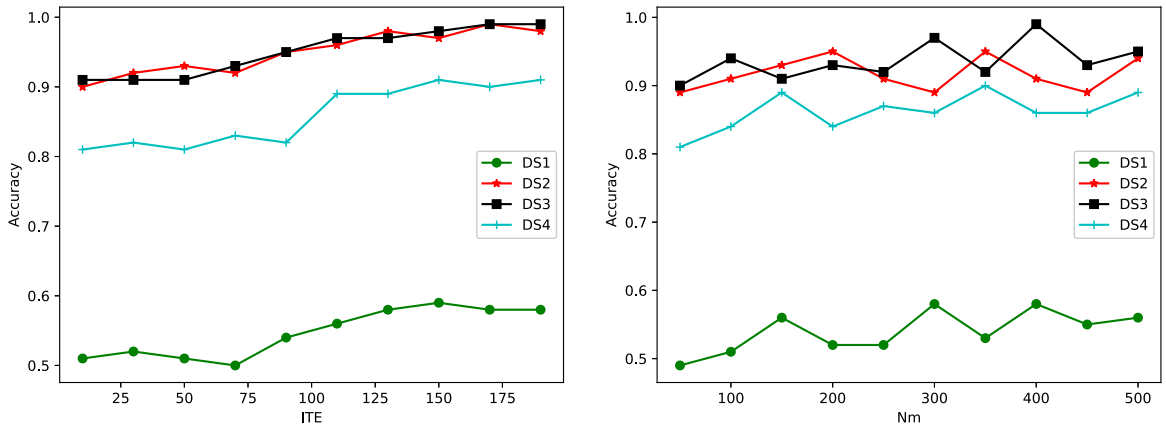
Fig. 4 presents the average values (Avg) of PPV, NPV, Rn, and S for the four datasets. It can be concluded that AMA achieved better results in terms of this parameter for the majority of datasets. The second-best algorithm was nd-DCA, which achieved better results on the four datasets compared to the proposed AMA; however, it had higher variability in its results. The DCA, PNN, RNN, LSTM, and SVR-HNN also achieved good results; these five algorithms could be considered in future studies. BPNN, PCA-RF, and DCA achieved values of approximately 50%, indicating that these methods are to some extent less stable than desired and thus, may not be selected for investigation in future studies. Finally, the method that achieved the worst result, with an average Avg lower than 50%, was the NSA.

4.4. Results of Kruskal–Wallis H test

Because some of the algorithms used were not deterministic, a hypothesis testing was adopted in our study. We should compare the performance of multiple algorithms on multiple datasets, and the error values on different datasets do not obey the normal distribution around the average accuracy. Therefore, Kruskal–Wallis H test was adopted to determine whether a statistically significant difference existed between the accuracy of AMA and the other algorithms on datasets DS1, DS2, DS3, and DS4 [38]. The significance level α was set to 0.05. Formally, letting the distribution function of ACC be of the form $F_i(ACC) = F(ACC - \mu_i)$ for group i, where μ_i indicates the ACC of an algorithm in our experiments. The hypothesis listed as follow:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_{11} \tag{9}$$

$$H_1 : not \quad all \quad \mu_i \quad are \quad equal. \tag{10}$$



(a) ITE-related accuracy of AMA. (b) N_m-related accuracy of AMA.

Fig. 5. The parameter-related accuracy of AMA.

Table 6
The classification accuracy of Mann–Whitney test between AMA and the other algorithms.

Datasets	Statistic	p-values
AMA - DCA [9]	16.0	**0.029**
AMA - NSA [8]	16.0	**0.030**
AMA - PCA-RF [6]	15.0	0.059
AMA - BPNN [22]	16.0	**0.029**
AMA - RNN [28]	13.5	0.146
AMA - PNN [23]	13.0	0.193
AMA - hDCA [10]	12.0	0.312
AMA - LSTM [29]	13.0	0.193
AMA - SVR-HNN [26]	14.0	0.112
AMA - nd-DCA [11]	11.0	0.470

The p-value acquired from the experimental results was 0.006, it was below 0.05, indicating that the H_0 was not true (that is, the proposed AMA was different from the other algorithms). In addition, Mann–Whitney test was adopted for pairwise comparisons to obtain the H statistics and p-values between each pair of algorithms. As observed in Table 6, in terms of classification accuracy, AMA exhibited significant differences on DCA, NSA, and BPNN. However, there was no significant difference between AMA and the other seven algorithms.

4.5. Parameter analysis

The AMA parameters included the iterations ITE and macrophage population size N_m; therefore, we conducted the following experiments to analyze their effect on algorithm performance. To assess the AMA performance in terms of ITE, ITE was set to 10, 30, 50, …,200, while other parameters were set to default values. N_m takes the values of 50, 100, …, 500, to evaluate the AMA performance in relation to N_m. The DS1, DS2, DS3, and DS4 datasets were used to implement these experiments. Fig. 5(a) illustrates the average accuracy of AMA with ITE, and it can thus be concluded that ITE controls the accuracy stability of AMA. When the number of iterations reached a certain value, the accuracy of the AMA on each dataset tended to be stable. For example, in DS1, AMA achieved lower accuracy when ITE varied from 10 to 130; however, accuracy was better when ITE was greater than 130. In DS4, AMA achieved lower accuracy when ITE varied from 10 to 110; however, accuracy improved when ITE was greater than 110. Fig. 5(b) shows the average algorithm accuracy related to N_m. This shows that the prediction accuracy of AMA fluctuated slightly as N_m changed, although the overall trend was relatively stable. Clearly, the parameter N_m does not strongly affect the prediction accuracy.

5. Conclusions and future works

In this study, a classification method utilizing artificial macrophages was proposed and implemented in the EQP scenario. This method was derived from the function of macrophages in the immune system; in contrast to traditional EQP methods, which only identified anomalies, the proposed approach identified anomalies caused by noise, updated the population of artificial macrophages, implementing positive and negative adjustments based on the recognition results. In the AMA, the GR law and other seismic magnitude distributions were used to obtain seismic indicators. Then, a seismological theory and information gain method was implemented in the AMA to process the noise data. Moreover, artificial macrophages were used to perform signal fusion to determine the normality and abnormality of the input data. Gradient descent and iteration conditions were adopted to analyze the parameters and optimize the weights of the artificial macrophages, thereby facilitating the adaptability of the AMA. To verify the performance of the AMA, we conducted experiments on four seismic events, and the results demonstrated that AMA obtained better results on the evaluation indicators of most regions than the comparison EQP approaches. Meanwhile, the parameters ITE and N_m in AMA were analyzed, and the results showed that the number of iterations ITE had a greater impact on algorithm performance, whereas the number of macrophage populations N_m had no significant impact on the performance of the algorithm. However, from the analysis of multiple sets of experimental results, it was observed that the method proposed in this study lacked robustness, and the adaptive adjustment ability must be strengthened for changes in application scenarios.

Future study will focus on enriching the AMA and applying it to other earthquake-prone areas. We will further test the AMA to improve the AIS detection performance through subsequent experiments. Moreover, we will adopt the methods in Refs. [39–41] to address disturbances and uncertainties in practical applications in future studies.

CRediT authorship contribution statement

Wen Zhou: Conceptualization, Methodology, Software, Validation, Formal analysis, Resources, Data curation, Writing – original draft, Visualization, Investigation. **Yiwen Liang:** Supervision, Writing – reviewing and editing, Funding acquisition. **Xinan Wang:** Supervision, Reviewing and editing, Funding acquisition. **Zhe Ming:** Reviewing and editing. **Zhenhua Xiao:** Reviewing and editing. **Xiying Fan:** Reviewing and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank NSFC, China- http://www.nsfc.gov.cn/ for their support through Grant Number 61877045, Fundamental Research Project of Shenzhen Science and Technology Program, China for their support through Grant Number JCYJ20160428153956266, Scientific Research Project of Hubei Provincial Department of Education, China for their support through Grant Number D20191406, and Hubei Provincial Natural Science Foundation of China for their support through Grant Number 2015CFB507.

References

[1] China Earthquake Networks Center, National earthquake data center, 2020, http://data.earthquake.cn.
[2] SichuanNews, "5.12" Wenchuan Earthquake, more than 5,300 students died in sichuan, 2009, http://scnews.newssc.org/system/2009/05/07/011881366.shtml.
[3] C.R. Allen, Responsibilities in earthquake prediction, Bull. Seismol. Soc. Am. 66 (6) (1976) 2069–2074.
[4] M.J. Fernández-Gómez, G. Asencio-Cortés, A. Troncoso, F. Martínez-Álvarez, Large earthquake magnitude prediction in Chile with imbalanced classifiers and ensemble learning, Appl. Sci. 7 (6) (2017) 625–632.
[5] Y.Y. Kagan, D.D. Jackson, Y. Rong, A testable five-year forecast of moderate and large earthquakes in southern California based on smoothed seismicity, Seismol. Res. Lett. 78 (1) (2007) 94–98.
[6] G. Asencio-Cortés, F. Martínez-Álvarez, A. Morales-Esteban, J. Reyes, A. Troncoso, Improving earthquake prediction with principal component analysis: application to Chile, in: International Conference on Hybrid Artificial Intelligence Systems, HAIS 2015, Springer, 2015, pp. 393–404.
[7] Z. Chelly, Z. Elouedi, A survey of the dendritic cell algorithm, Knowl. Inf. Syst. 48 (3) (2016) 505–535.
[8] J. Wu, Y. Liang, C. Tan, W. Zhou, Method of earthquake prediction based on negative selection, Appl. Res. Comput. 36 (4) (2019) 1097–1100.
[9] Y. Gan, Y. Liang, C. Tan, W. Zhou, J. Wu, Earthquake prediction method based on danger theory, Comput. Eng. 45 (1) (2019) 278–283.
[10] W. Zhou, H. Dong, Y. Liang, The deterministic dendritic cell algorithm with Haskell in earthquake magnitude prediction, Earth Sci. Inform. 13 (2) (2020) 447–457.
[11] W. Zhou, Y. Liang, Z. Ming, H. Dong, Earthquake prediction model based on danger theory in artificial immunity, Neural Netw. World 30 (4) (2020) 231–247.
[12] Y. Lv, B. Zhen, L. Shi, S. Niu, Loss of cell surface CD47 clustering formation and binding avidity to SIRPα facilitate apoptotic cell clearance by macrophages, J. Immunol. 195 (2) (2015) 661–671.
[13] F. Gu, J. Greensmith, U. Aickelin, Further exploration of the dendritic cell algorithm: Antigen multiplier and time windows, 5132, ISBN: 978-3-540-85071-7, 2010, pp. 1–12, http://dx.doi.org/10.1007/978-3-540-85072-4_13.
[14] G. Evgin, Diagnosis of Alzheimer's disease with Sobolev gradient-based optimization and 3D convolutional neural network, Int. J. Numer. Methods Biomed. Eng. 35 (7) (2019) 1–16.
[15] F. Martínez-Álvarez, J. Reyes, A. Morales-Esteban, C. Rubio-Escudero, Determining the best set of seismicity indicators to predict earthquakes. Two case studies: Chile and the Iberian Peninsula, Knowl.-Based Syst. 50 (2013) 198–210.
[16] M.H.A. Banna, K.A. Taher, M.S. Kaiser, M. Mahmud, M.S. Rahman, A.S. Hosen, G.H. Cho, Application of artificial intelligence in predicting earthquakes: State-of-the-art and future challenges, IEEE Access 8 (2020) 192880–192923.
[17] A. Panakkat, H. Adeli, Recent efforts in earthquake prediction (1990–2007), Nat. Hazards Rev. 9 (2) (2008) 70–80.
[18] L. Dehbozorgi, F. Farokhi, Effective feature selection for short-term earthquake prediction using neuro-fuzzy classifier, in: Second IITA International Conference on Geoscience & Remote Sensing, no. 43, IITA-GRS 2010, in: 2, 2010, pp. 205–230.
[19] A. Pandit, K.C. Biswal, Prediction of earthquake magnitude using adaptive neuro fuzzy inference system, Earth Sci. Inform. 12 (4) (2019) 513–524.
[20] R. Tehseen, M.S. Farooq, A. Abid, Earthquake prediction using expert systems: a systematic mapping study, Sustainability 12 (6) (2020) 1–32.
[21] R. Kamath, R. Kamat, Earthquake magnitude prediction for Andaman-nicobar islands: Adaptive neuro fuzzy modeling with fuzzy subtractive clustering approach, J. Chem. Pharm. Sci. 10 (3) (2017) 1228–1233.
[22] A. Panakkat, H. Adeli, Neural network models for earthquake magnitude prediction using multiple seismicity indicators, Int. J. Neural Syst. 17 (01) (2007) 13–33.
[23] H. Adeli, A. Panakkat, A probabilistic neural network for earthquake magnitude prediction, Neural Netw. 22 (7) (2009) 1018–1024.
[24] S. Luis, Predict the magnitudes of seismic events using Bayesian methods, Soil Dyn. Earthq. Eng. 129 (2020) 1–10.
[25] C.G. Shi, X.L. Liu, Application of neural network to earthquake engineering, Earthq. Eng. Eng. Vib. 11 (2) (1991) 39–46.
[26] K.M. Asim, A. Idris, T. Iqbal, F. Martínez-Álvarez, Earthquake prediction model using support vector regressor and hybrid neural networks, PLoS One 13 (7) (2018) 1–22.
[27] K.M. Asim, S.S. Moustafa, I.A. Niaz, E.A. Elawadi, T. Iqbal, F. Martínez-Álvarez, Seismicity analysis and machine learning models for short-term low magnitude seismic activity predictions in Cyprus, Soil Dyn. Earthq. Eng. 130 (2020) 1–11.
[28] A. Panakkat, H. Adeli, Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators, Comput.-Aided Civ. Infrastruct. Eng. 24 (4) (2009) 280–292.
[29] J.P. Huang, X.A. Wang, Y. Zhao, C. Xin, H. Xiang, Large earthquake magnitude prediction in Taiwan based on deep learning neural network, Neural Netw. World 28 (2) (2018) 149–160.
[30] P.M. DeVries, F. Viégas, M. Wattenberg, B.J. Meade, Deep learning of aftershock patterns following large earthquakes, Nature 560 (7720) (2018) 632–634.
[31] Q. Wang, Y. Guo, L. Yu, P. Li, Earthquake prediction based on spatio-temporal data mining: An LSTM network approach, IEEE Trans. Emerg. Top. Comput. 8 (1) (2020) 148–158.
[32] H. Jiang, W. Luo, Z. Zhang, A privacy-preserving aggregation scheme based on immunological negative surveys for smart meters, Appl. Soft Comput. 85 (2019) 1–15.
[33] W. Luo, M. Pavone, C.A.C. Coello, L. Jiao, R. Mehr, Foreword: Some advances in immune computation and applications, Swarm Evol. Comput. 50 (2019) 100596.
[34] P.J. Murray, Macrophage polarization, Annu. Rev. Physiol. 79 (1) (2017) 541–566.
[35] D.G. Russell, L. Huang, B.C. Vanderven, Immunometabolism at the interface between macrophages and pathogens, Nature Rev. Immunol. 19 (5) (2019) 291–304.
[36] G. Asencio-Cortés, F. Martínez-Álvarez, A. Morales-Esteban, A. Troncoso, Medium-large earthquake magnitude prediction in Tokyo with artificial neural networks, Neural Comput. Appl. 28 (5) (2017) 1043–1055.
[37] T.R.C. Read, N.A.C. Cressie, Goodness-of-Fit Statistics for Discrete Multivariate Data, Springer Science & Business Media, 1991.
[38] W.H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis, J. Amer. Statist. Assoc. 47 (260) (1952) 583–621.
[39] H. Tao, J. Li, Y. Chen, V. Stojanovic, H. Yang, Robust point-to-point iterative learning control with trial-varying initial conditions, IET Control Theory Appl. 14 (19) (2020) 3344–3350.
[40] S.D. Bolboacă, L. Jäntschi, Sensitivity, specificity, and accuracy of predictive models on phenols toxicity, J. Comput. Sci. 5 (3) (2014) 345–350.
[41] V. Stojanovic, S. He, B. Zhang, State and parameter joint estimation of linear stochastic systems in presence of faults and non-Gaussian noises, Internat. J. Robust Nonlinear Control 30 (16) (2020) 6683–6700.