scientific reports

Check for updates

OPEN

Closed circuit artificial ıntelligence model named morgaf for childhood onset systemic lupus erythematosus diagnosis

Emil Aliyev[1,2]✉, Yagizhan Ugur[3], Veysel Cam[1], Yagmur Bayindir[1], Dilara Unal[1], Zeynep Balik[1], Seher Sener[1], Yelda Bilginer[1] & Seza Ozen[1]

Systemic Lupus Erythematosus (SLE) is a chronic, autoimmune disease characterized by multiple organ involvement and autoantibodies, and its diagnosis is not easy in clinical practice. Pediatric SLE (pSLE) is diagnosed using the SLICC 2012 criteria for adult SLE patients. Our study aims to develop a closed computer-based AI model to assist clinicians in diagnosis. Fifty pSLE patients followed up in Hacettepe University Pediatric Rheumatology Outpatient Clinics, and 50 healthy individuals similar to them in terms of age and gender were included in the study. Data sets, including clinical and laboratory findings of the individuals at the time of diagnosis, were given as input to the AI model. Python® software language and Tensorflow® AI library were preferred for model development. The concept of neural networks (NN) is increasingly common in AI studies. Morgaf (the name of the AI model) used the recurrent neural network (RNN) model, which remembers previous inputs during training, leading to lower error rates. Patient data was digitized and used to train the model, which consists of 1024 neurons. The model's error rate decreased from 0.5 to 0.028 during training, leading to successful predictions compared to expert interpretations. Thirty case data (data utterly foreign to the model) were given to Morgaf and simultaneously to expert pediatric rheumatologists, and their interpretations were compared. Prediction success was evaluated by performing regression analyses between both groups. Morgaf accepted 70% and above as a definitive diagnosis of pSLE, averaging 93% (78–98%). It defined 10% and below for a completely healthy case; the average was 1% (0–3%). To recognize diseases requiring follow-up, he set himself a range of 10–70% and estimated the mean to be 33% (15–47%). There was no difference between Morgaf's estimates and actual diagnoses ($p = 0.297$). Morgaf was 100% successful in recognizing lupus disease. This rate was the same as the diagnostic understanding of clinicians. Morgaf (93.3%) gave more precise recommendations for non-lupus cases than clinicians (70%) ($p = 0.034$). Regression analysis showed that Morgaf (y = 0.9264xi) was more successful in non-selective prediction than clinicians (y = 0.7322xi). Our study is the first in the literature to develop and test an AI model as a diagnostic tool for pSLE. In this cohort, AI model was at least as successful as pediatric rheumatologists in differentiating pSLE patients from healthy controls and non-pSLE patients. With this study, we have shown that Morgaf may help clinicians with diagnostic and differential diagnoses.

Keywords Pediatric systemic lupus erythematosus, Artificial ıntelligence, Diagnosis, Recurrent neural network

Systemic Lupus Erythematosus (SLE) is a systemic rheumatic disease with multiorgan involvement that takes a long time to diagnose and, therefore, causes delays in treatment[1,2]. The diagnostic process of pediatric SLE (pSLE) and all childhood rheumatologic diseases may take time due to the need for more objective findings and clinician dependence[1,3]. In recent years, Artificial Intelligence (AI)—based systems have been developed to support clinicians in many areas[4,5]. One of the most important reasons for this is the low number of pediatric

[1]Department of Pediatric Rheumatology, Ihsan Dogramaci Children's Hospital, Hacettepe University, Gevher Nesibe St., Hacettepe Dstr., Hacettepe Rg., 06230 Ankara, Turkey. [2]SEMBA Science, Education, Informatics Ltd. Co., Kaptanpasa St., Ap: 11, No: 6, F: 2, Buyukesat Dstr., Cankaya Rg., 06670 Ankara, Turkey. [3]Keytech Electronic and Software Ltd. Co., Ankara, Turkey. ✉email: sciencemil@gmail.com; sciedumedsemba@gmail.com

rheumatologists (PRs) compared to the population[6–8]. On the other hand, clinics that treat urinary tract infections and skin diseases, which are common in the pediatric age group, can often be confused with pSLE[6].

Another problem is that patients cannot be followed up in the triage system due to a lack of records. Even if patients reach PRs, the diagnostic process may take months because there are no specific diagnostic criteria for childhood SLE patients[9]. When all these problems are evaluated, the pSLE diagnosis process should be digitalized, and a model free of existing issues should be developed[4,10]. We aim to establish an AI-supported model that helps PRs and physicians in all fields diagnose pSLE.

Materials and methods

Study group and design

Fifty pSLE patients followed up in Hacettepe University Pediatric Rheumatology Outpatient Clinics, and 50 healthy individuals similar to them in terms of age, gender, and sociocultural characteristics were included in the study. Clinical data of the patients were obtained from the hospital automation system. The healthy control group was selected from the patients' peers (usually schoolmates or peers living in the same neighborhood). Since the clinical and laboratory data of lupus patients were evaluated at the time of diagnosis, blood samples were not taken from the patients and were not examined. The information of the individuals participating in the study was recorded in an offline computer-based data set. Access to the dataset was restricted except for those conducting the survey.

The study used fully anonymized pSLE patient data previously used by the Department of Pediatric Rheumatology of Hacettepe University Faculty of Medicine and had official approval from the relevant department. Written and verbal consent was obtained from all patients for these previously used data for different studies. For this purpose, patients were invited to Hacettepe University Faculty of Medicine, Department of Pediatric Rheumatology, for one-on-one information, and their signatures were obtained on the 'informed consent form' after verbal and written information was given. For participants who did not come to the department in person, the relevant consent files were sent to their e-mails after verbal information and approval by telephone, and their data were used if the signed files were scanned and delivered to us. All study methods were carried out following the relevant guidelines and regulations. In addition, an approval from the Ethics Committee of Hacettepe University Health Sciences Ethics Committee dated 09.24.2024 and numbered E-24742385-000-00003782004 was obtained.

The study design was planned to be double-blind. Thirty mixed-diagnosis cases (data entirely foreign for the model and the clinician) were submitted to the AI model, a concurrent expert pediatric rheumatologist (PR) and their interpretations were compared. The PR whose responses were compared with the AI model was utterly unfamiliar with the study and the data presented. Of these data, 15 were pSLE cases, 10 were healthy controls, and three were followed-up individuals with only anti-nuclear antibody (ANA) and anti-double stranded DNA (anti-ds-DNA) positivity and no other clinic. One was a newly diagnosed, untreated individual with acne rosacea who developed a skin infection due to secondary Staphylococcus epidermidis. Another patient had a urinary tract infection due to an adenovirus infection.

AI-based closed model

The reason for choosing the AI Closed Model is the law on 'Protection of Personal Data' in force in Turkey[11]. According to this law, data processing and evaluation cannot be shared with third parties, institutions, and organizations[12]. Online AI-based applications or correspondence sites cannot be used to process sensitive data such as patient information. Therefore, although the model created has autonomous features, it can never be processed in online systems. Python software and Tensorflow were preferred for model development due to their ease of use in medicine and healthcare today[13,14]. We also chose these resources for developing the model because they have been used in medicine and many other data-processing fields for years with almost negligible error rates. The model was named 'Morgaf' by Dr. Emil Aliyev, M.D., M.S. (hereafter, the AI model will be referred to as Morgaf). With this name, an application has been made to the Turkish Patent and Trademark Office, and the patenting process is ongoing[15].

The concept of a 'Neural Network' (NN) is becoming increasingly widespread in AI studies[14]. NNs consist of neurons and are structures that calculate and produce output with a predetermined mathematical function. In the development of Morgaf, the 'Recurrent Neural Network' (RNN) model was preferred (Fig. 1)[16].

The most important difference between the RNN and other networks is that while training the model and processing the inputs at Tn, it also remembers the inputs at Tn-1 (Fig. 2) (Tn is any time unit).

This feature makes the error rates on the network much lower than other models. The laboratory and clinical results of the patients were digitized as 1 (abnormal), 0 (normal range), and – 1 (abnormal). Data sets containing clinical and laboratory findings of individuals at the time of diagnosis were given as input to Morgaf. Morgaf is designed as an RNN model that constructs an NN with 1024 neurons (Fig. 1). The model needs many inputs to 'learn' the data. A total of 100 data sets is generally insufficient to build a closed-loop AI model.

For this reason, the 100 data sets were increased 15 times to 1500 with repetitions, 80% of which were reserved as training data and 20% as validation data. Training data is the data that Morgaf continuously 'repeats' during the learning process to 'train' itself. Validation data is the data that Morgaf checks itself after momentary training (Fig. 3).

In the early training phase of Morgaf's training, the error rate was approximately 0.5. In the later generations of the model, the error decreased to 0.028. This error rate was deemed sufficient, and Morgaf's training phase was terminated. Training continued autonomously in the latent phase. Autonomous training in RNN models such as Morgaf is essential for continuous self-evaluation and error outputs.

It is essential to interpret the outputs to understand the 'language' of Morgaf. According to the output Morgaf at the end of the training, a diagnosis was predicted to be between 0 and 100%. Accordingly, it was seen that

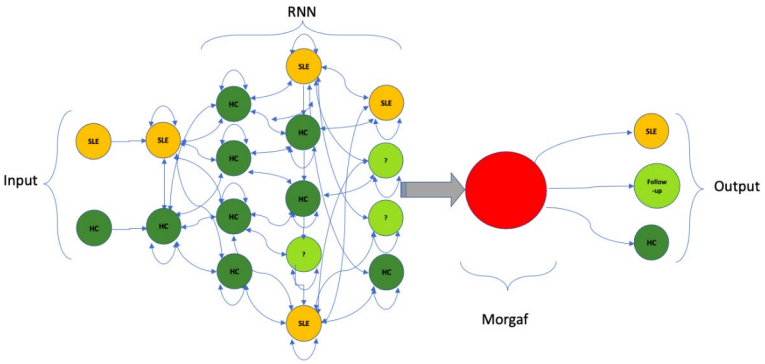

Fig. 1. Schematic description of Morgaf's RNN model. SLE: Systemic Lupus Erythematosus, HC: Healthy Control, RNN: Recurrent Neuron Network.

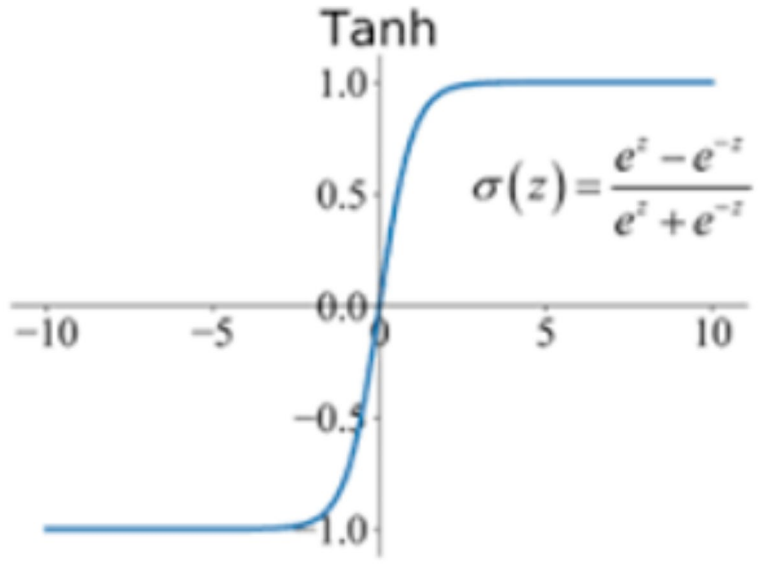

Fig. 2. Graphical description of the data processing interval of Morgaf's RNN in units of time. Tanh; hyperbolic tangent function of any unit of the time.

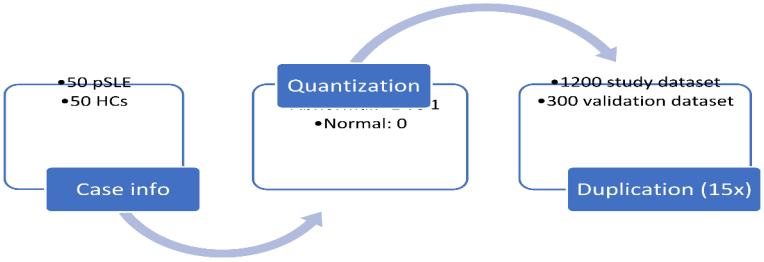

Fig. 3. Schematization of Morgaf's process of teaching itself from limited data. pSLE: Pediatric Systemic Lupus Erythematosus, HCs: Healthy Controls.

Morgaf determined the output as 70% and above to say definitive diagnosis pSLE, 10% and below to say healthy individual, and 10–70% in cases that require follow-up and in between (definitive pSLE and not healthy, but need to be followed up). Thus, the beta version of Morgaf was made ready.

Data redundancy was avoided so that Morgaf could most accurately recognize pSLE. Based on the 'mental algorithms' of clinicians when diagnosing pSLE, Morgaf was taught the clinical and laboratory data that are key to the diagnosis (Table 1). For example, instead of the amount of protein in the urine, the clinician was taught

Parameters	Learning criterion	Learning inputs	Outputs	Presence in the patient
Sociodemographic data				
Age at diagnosis	Year; numeric value	Numeric	Numeric	Mandatory
Organ involvement at the time of diagnosis				
Skin	Yes/No	0 and 1	from 0 to 1	Not mandatory
Renal	Yes/No	0 and 1	from 0 to 1	Not mandatory
Bone Marrow	Yes/No	0 and 1	from 0 to 1	Not mandatory
Central Nervous System	Yes/No	0 and 1	from 0 to 1	Not mandatory
Joint(s)	Yes/No	0 and 1	from 0 to 1	Not mandatory
Liver	Yes/No	0 and 1	from 0 to 1	Not mandatory
Presence of painless aphthae	Yes/No	0 and 1	from 0 to 1	Not mandatory
Presence of serositis	Yes/No	0 and 1	from 0 to 1	Not mandatory
Eye	Yes/No	0 and 1	from 0 to 1	Not mandatory
Intestinal	Yes/No	0 and 1	from 0 to 1	Not mandatory
Systolic blood pressure (mm Hg)	80–110 is normal. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Not mandatory
Diastolic blood pressure (mm Hg)	60–80 is normal. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Not mandatory
Laboratory results at the time of diagnosis				
Hemoglobin (gr/dL)	Typical range: 13–17. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
Leukocyte (/mm^3)	Typical range: 5000–13,200. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
Lymphocyte value (/mm^3)	Typical range: 2000–6000. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
Neutrophil value (/mm^3)	Typical range: 2000–8300. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
Platelet value (/mm^3)	Typical range: 190,000–400,000. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
CRP value (mg/dL)	Typical range: 0–0,5. Other parameters are abnormal	0 and 1	from 0 to 1	Not mandatory
ESR value (mm/h)	Typical range: 0–20. Other parameters are abnormal	0 and 1	from 0 to 1	Not mandatory
Urine density	Typical range: 1003–1020. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Not mandatory
Presence of hematuria in urine	Yes/No	0 and 1	from 0 to 1	Not mandatory
Presence of ketones in urine	Yes/No	0 and 1	from 0 to 1	Not mandatory
Presence of pyuria in urine	Yes/No	0 and 1	from 0 to 1	Not mandatory
Presence of proteinuria in urine	Yes/No	0 and 1	from 0 to 1	Not mandatory
BUN (mg/dL)	Typical range: 5–18. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Not mandatory
Creatinine (mg/dL)	Typical range: 0,5–1,2. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Not mandatory
Complement C3 (mg/dL)	Typical range: 88–230. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
Complement C4 (mg/dL)	Typical range: 13–73. Other parameters are abnormal	−1, 0 and 1	from 0 to 1	Mandatory
ANA	If titer is < 1/40: negative, > 1/40 is positive	0 and 1	from 0 to 1	Mandatory
Anti-ds-DNA (IU/mL)	Typical range: 0–5. Other parameters are abnormal	0 and 1	from 0 to 1	Mandatory
B2GPA IgG (U/mL)	Typical expect result is zero. Other parameters are abnormal	0 and 1	from 0 to 1	Not mandatory
B2GPA IgM (U/mL)	Typical expect result is zero. Other parameters are abnormal	0 and 1	0 and 1	Not mandatory
ACLA IgG (U/mL)	Typical expect result is zero. Other parameters are abnormal	0 and 1	0 and 1	Not mandatory
ACLA IgM (U/mL)	Typical expect result is zero. Other parameters are abnormal	0 and 1	0 and 1	Not mandatory

Table 1. pSLE patient data used in the development and training of Morgaf and its digitalization. CRP, C-reactive Protein; ESR, Erythrocyte Sedimentation Rate; BUN, Blood Urine Nitrogen; ANA, Anti-nuclear antibody; Anti-ds-DNA, anti-double stranded DNA; B2GPA, Beta-2-glycoportein Antibody, ACLA, Anti-Cardiolipin Antibody.

whether it was present or absent, and instead of the titer of ANA. The clinician was taught whether it was positive or negative. As a matter of fact, in practice, clinicians continue their diagnostic algorithms by considering an ANA value above 1/40 as positive[17]. This is one of the most critical steps model builders take when developing an AI model. Because it is not the number of data given for a single patient, making it possible to reach the target with fewer data is the key to the model's success. In this context, the most important clinical and laboratory data of pSLE patients at diagnosis were used to create the model (Table 1).

In addition to Morgaf and PR having a 100% case recognition rate, their median responses for each case ranged from 0 to 100%. For example, when predicting a patient with SLE, Morgaf said 83.21% for this patient. Since it was in the range of 70–100%, he correctly recognized the SLE patient. Since the clinician was answering whether the patient had SLE or not, we expected the clinician to express a percentage value for each case, so unlike Morgaf, his answers ranged from 0 to 100%. However, because it correctly predicted all SLE patients, the accuracy percentage was 100%. Another critical step is quantifying these data in a way Morgaf can learn. The inputs were then presented to Morgaf.

Data content	Developed limits for interpretation (%)	Outputs (%) (min–max (%))	Accurate prediction ability (%)
Definitive diagnosis pSLE	≥ 70	93 (78–98)	100
Healthy person	≤ 10	1 (0–3)	100
Disease requiring follow-up (non-pSLE)	10–70	33 (15–47)	100

Table 2. Case recognition boundaries and results of the Model's training. pSLE, Pediatric Systemic Lupus Erythematosus.

Case groups	Morgaf's response success (%)	Clinician response success (%)	*p* value
pSLE	100	100	0.297
non-pSLE	93	70	0.034
HCs	100	100	0.422

Table 3. Comparison of Morgaf and clinician responses. pSLE, Pediatric Systemic Lupus Erythematosus; HCs, Healthy Controls.

Statistical analysis

The study data were computerized and analyzed using the Statistical Package for Social Sciences for Windows version 22.0 (SPSS Inc, Chicago, IL)[18]. Visual (histograms and probability plots) and analytical methods (Kolmogorov–Smirnov/Shapiro–Wilk Test) showed that the variables measured in the study did not fit a normal distribution. Mann–Whitney U test was used to compare the variables between two independent groups. The statistical significance level was accepted as $p < 0.05$. Levene's correlation analysis tested for homogeneity of variances and non-normally distributed variables, and variables with a correlation coefficient < 0.1 were considered uncorrelated. Prediction success was evaluated by performing regression analyses between both groups.

Results

Morgaf accepted 70% and above as a definitive diagnosis of pSLE, and the mean percentage of diagnosis was 93% (78–98%). Morgaf received 10% and below to call it a completely healthy case; the average was 1% (0–3%). Morgaf assigned himself a range of 10–70% to recognize diseases requiring follow-up; the average was 33% (15–47%). Morgaf correctly recognized all cases of pSLE, healthy individuals, and definite non-pSLE cases (Table 2).

There was no difference between Morgaf's estimates and actual diagnoses ($p = 0.297$). It was 100% successful in recognizing lupus disease. This rate was the same as the diagnostic success rate of clinicians. Morgaf (93.3%) gave more precise recommendations for non-lupus cases than clinicians (70%) ($p = 0.034$) (Table 3).

Covariance analysis was performed to test the homogeneity of the variables ($p = 0.792$). Linear regression analysis was performed for both homogeneously distributed groups and tested with Levene ($p = 0.147$). Regression analysis showed that Morgaf ($y = 0.926x_1$) was more successful in predicting findings than clinicians ($y = 0.7322x_1$).

Discussion

AI-based models have been used in many areas of medicine[19]. AI models are increasingly invested in medicine as time goes by[20]. Although AI and deep learning models have been developed in medicine, unlike other fields, these models have not been implemented[4]. Especially in imaging, tools that evaluate radiological images through U-net software and help clinicians are familiar[21]. AI models that analyze and predict magnetic resonance imaging, direct radiography, and ultrasound images are standard[22–26]. However, all of these are applied in clinical practice in the clinics where the study was conducted, and they are not widely used[19,26]. The main obstacle to this is the legal barriers to sharing data with other clinics and that the data are generally online-based AI models[27]. As the name suggests, online AI models can use all online data[28,29]. Typically, applications that help diagnose symptom inputs are based on these models[29]. However, we know that patients and their complaints are individual. If the answers are not obtained from reliable sources or cannot be monitored, these applications may be biased and raise ethical issues such as violating personal data[30]. Therefore, our model is an offline model. This aligns with the data protection principle[31]. Apart from this, our model does not process data from online databases or online sources when making comments. The model can only work on the data we provide. That is, it can comment on the data of pSLE patients.

Apart from AI models that analyze and interpret images, the second most frequently used area is the analysis of large data sets and models based on prediction development[24,26]. This is natural because we know that AI models require a lot of data in terms of data processing. However, this is a speed-limiting step in areas such as rheumatology, where the number of patients and data are scarce[32]. When we developed Morgaf, we tried to solve the problem of how to train a model from limited data. We replicated the data by fully randomized and subsequently validated by the clinician to verify that they were similar to accurate patient data. Of course, this can be considered a limitation of the model since we created synthetic data. Still, it should be remembered that data scarcity is a severe problem in most fields of medicine, and it is essential to prevent this. This interpretation

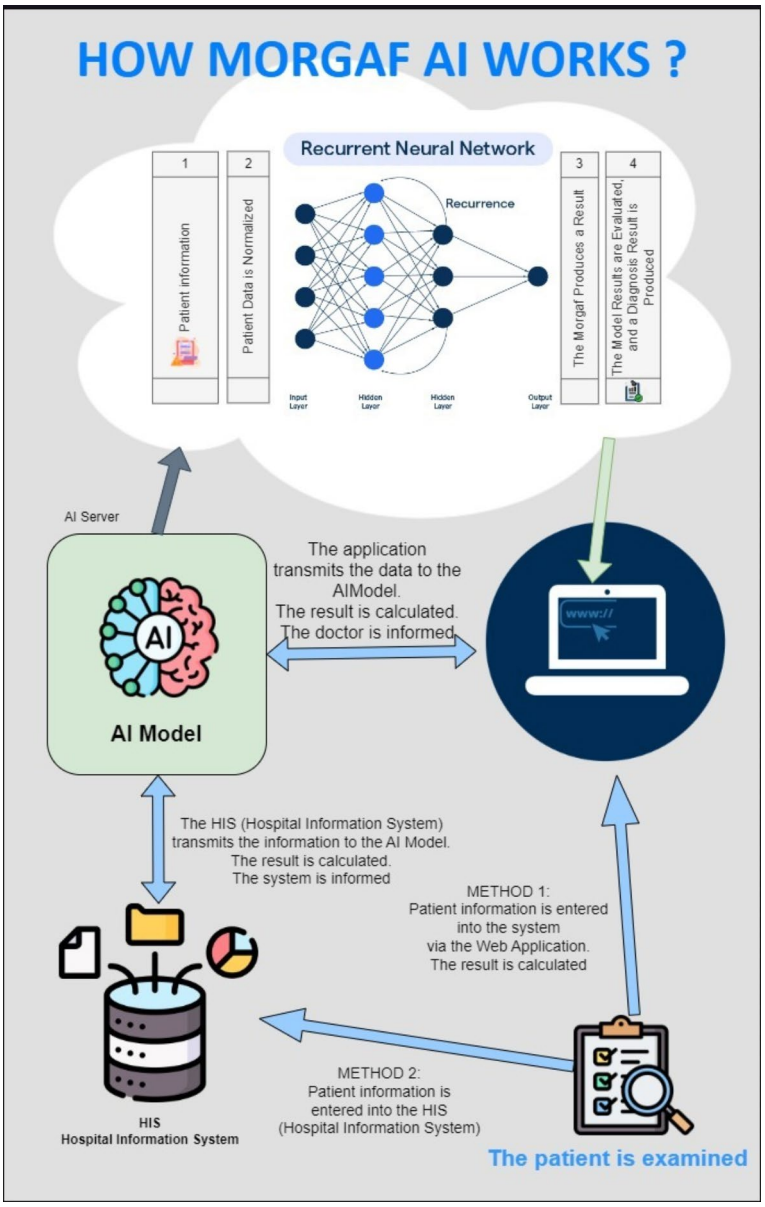

Fig. 4. Schematic description of Morgaf's functioning as an AI model.

of Morgaf makes it unique. Of course, we are aware that more real patient data would train the model better. Therefore, larger data sets and prospectively designed studies are needed.

Another research topic in rheumatology is applications that lead from symptom to diagnosis[29,33]. Since these models are also online, they cause severe ethical problems due to violating personal data and biased interpretation[30]. Morgaf, the model we developed, is an offline model.

Interestingly, Morgaf, when comparing thirty completely unfamiliar data sets with the clinician's interpretations simultaneously, recognized non-pSLE cases (Fig. 1), unlike the clinician, even though we had yet to teach it to do so. This showed that Morgaf could easily differentiate between other clinics as it had 'learned' the data of pSLE and healthy individuals, so much so that it could make better comments than the clinician. It is possible to 'teach' Morgaf visual data, such as pathology preparations of pSLE patients, in the future so that it can be a more successful model. However, at this stage, we preferred to use a single programming language and library as it would reduce the model's chances of success when processing complex data. This can be considered a limitation of our model in fully recognizing pSLE disease. This model also needs to be tested in multiethnic cohorts and in other centers.

It is worth noting that Morgaf was well acquainted with childhood-onset SLE patients but not with other differential diagnoses. Furthermore, the model has not been tested in adult SLE patients. This was important to create the contrast between intact control and SLE that we needed when training the model. However, providing specific clinical and laboratory data on autoimmune, hematologic, infectious and many other diseases with

differential diagnosis of SLE during the training of the model could enable the model to become a robot that more sharply recognizes SLE in real time. This is an important limitation of our study.

Morgaf is the first AI model in the literature to be developed from limited data and successful in pediatric rheumatology and pSLE disease. The model is under development and is expected to be successful as a cost-effective model. If integrated into hospital automation systems (Fig. 4), it is predicted that the model may lead to earlier diagnosis of pSLE patients, including non-specialty clinicians, significantly reducing hospital costs.

Data availability

The data supporting the findings of this study were obtained from SEMBA Health Education Informatics Ltd. Co. (shortly SEMBA Ltd. Co.). However, restrictions apply to the availability of these data, which are used under license for this study and are not publicly available. However, the authors may contact Dr. Emil Aliyev, M.D., M.S., who is authorized to represent the company upon reasonable request and with the permission of SEMBA Ltd. Co. The relevant contact details are as follows: Address: Buyukesat Dsrt, Kaptanpasa St., Ap: 11, No: 6, F: 2, 06670, Cankaya Rg, Ankara, Turkey; Tel Number: + 905539032329; E-mail: sciedumedsemba@gmail.com.

Received: 7 October 2024; Accepted: 4 March 2025
Published online: 01 July 2025

References

1. Kısaarslan, A. P. et al. Neuropsychiatric involvement in juvenile-onset systemic lupus erythematosus: A multicenter study. *Joint Bone Spine* **90**(4), 105559 (2023).
2. Aliyev, E. et al. Assessing cognitive functions in non-neuropsychiatric childhood systemic lupus erythematosus: Cross-sectional study. *J. Psychosom. Res.* **189**, 112027 (2025).
3. Demirkaya, E. et al. Neuropsychiatric involvement in juvenile systemic lupus erythematosus. *Turk. J. Pediatr.* **50**(2), 126 (2008).
4. McMaster, C. et al. Artificial intelligence and deep learning for rheumatologists. *Arthritis Rheumatol.* **74**(12), 1893–1905 (2022).
5. Widmer, G., Horn, W. & Nagele, B. Automatic knowledge base refinement: learning from examples and deep knowledge in rheumatology. *Artif. Intell. Med.* **5**(3), 225–243 (1993).
6. Lythgoe, H. et al. Evaluation of the ACR and SLICC classification criteria in juvenile-onset systemic lupus erythematosus: A longitudinal analysis. *Lupus* **26**(12), 1285–1290 (2017).
7. Brunner, H. I. et al. Validation of the pediatric automated neuropsychological assessment metrics in childhood-onset systemic lupus erythematosus. *Arthritis Care Res.* **65**(3), 372–381 (2013).
8. Koker, O. et al. The emerging paradigm in pediatric rheumatology: harnessing the power of artificial intelligence. *Rheumatol. Int.* **44**, 2315–2325 (2024).
9. DiFrancesco, M. W. et al. Functional neuronal network activity differs with cognitive dysfunction in childhood-onset systemic lupus erythematosus. *Arthritis Res. Therapy* **15**(2), 1–12 (2013).
10. Shi, Y. et al. Advancing precision rheumatology: Applications of machine learning for rheumatoid arthritis management. *Front. Immunol.* **15**, 1409555 (2024).
11. Küzeci, E., *Kişisel verilerin korunması.* (2010).
12. Kurumu, T.P.V.M. & Başkanliği, P.D *Robotik teknolojileri ve otomasyon sektöründeki patent uygulamalari, dünya'daki ve türkiye'deki genel durum.*
13. Harrison, C. J. & Sidey-Gibbons, C. J. Machine learning in medicine: A practical introduction to natural language processing. *BMC Med. Res. Methodol.* **21**(1), 158 (2021).
14 Pires, J. G. Machine learning in medicine using JavaScript: Building web apps using TensorFlow. js for interpreting biomedical datasets. *medRxiv* https://doi.org/10.1101/2023.06.21.23291717 (2023).
15. Ozsunay, E. The new turkish trademark law. *Fordham Int'l LJ* **19**, 1542 (1995).
16. Yu, Y. et al. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **31**(7), 1235–1270 (2019).
17. Tan, E. M. et al. Range of antinuclear antibodies in "healthy" individuals. *Arthritis Rheum.* **40**(9), 1601–1611 (1997).
18 Cronk, B. C. *How to use IBM SPSS statistics: A step-by-step guide to analysis and interpretation* (Routledge, 2016).
19. Stoel, B. C. et al. Deep learning in rheumatological image interpretation. *Nat. Rev. Rheumatol.* **20**(3), 182–195 (2024).
20 Khanna, N. N. et al. Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. *Healthcare (Basel)* **10**(12), 2493 (2022).
21 Binvignat, M. et al. Use of machine learning in osteoarthritis research: a systematic literature review. *RMD Open* **8**(1), e001998 (2022).
22 Hoffmann, T. et al. Artificial intelligence-based quantification of pulmonary HRCT (AIqpHRCT) for the evaluation of interstitial lung disease in patients with inflammatory rheumatic diseases. *Rheumatol. Int.* **44**, 24983–32496 (2024).
23. Nie, D. et al. Artificial intelligence differentiates abdominal Henoch-Schönlein purpura from acute appendicitis in children. *Int. J. Rheum. Dis.* **26**(12), 2534–2542 (2023).
24. Bhat, C. S. et al. Artificial intelligence for interpretation of segments of whole body MRI in CNO: Pilot study comparing radiologists versus machine learning algorithm. *Pediatr. Rheumatol. Online J.* **18**(1), 47 (2020).
25. Erjiang, E. et al. Modelling future bone mineral density: Simplicity or complexity?. *Bone* **187**, 117178 (2024).
26. Adams, L. C. et al. Artificial intelligence to analyze magnetic resonance imaging in rheumatology. *Joint Bone Spine* **91**(3), 105651 (2024).
27. Kedra, J. et al. Machine learning approaches to improve disease management of patients with rheumatoid arthritis: Review and future directions. *Expert. Rev. Clin. Immunol.* **17**(12), 1311–1321 (2021).
28. Kassani, P. H. et al. Artificial intelligence for nailfold capillaroscopy analyses—A proof of concept application in juvenile dermatomyositis. *Pediatr. Res.* **95**(4), 981–987 (2024).
29. Goodman, R. S. et al. Accuracy and reliability of chatbot responses to physician questions. *JAMA Netw. Open* **6**(10), e2336483 (2023).
30. Karimian, G., Petelos, E. & Evers, S. M. The ethical issues of the application of artificial intelligence in healthcare: A systematic scoping review. *AI Ethics* **2**(4), 539–551 (2022).
31. Kidambi, R. et al. Morel: Model-based offline reinforcement learning. *Adv. Neural Inform. Process. Syst.* **33**, 21810–21823 (2020).
32. Norgeot, B. et al. Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis. *JAMA Netw. Open* **2**(3), e190606 (2019).
33. Segú-Vergés, C. et al. Application of systems biology-based in silico tools to optimize treatment strategy identification in Still's disease. *Arthritis Res. Ther.* **23**(1), 126 (2021).

Author contributions

Dr. Emil Aliyev was responsible for the hypothesis and methodology of the study, designing the study, interviewing the participants and obtaining written and verbal consent, writing the statistics (formal analysis) and results of the study, and literature review. Engineer Yagizhan Ugur created the AI model, provided inputs, ran the model, and provided the outputs to the corresponding author, Dr. Emil Aliyev. Both authors have full access to the study data. Dr. Veysel Cam, Dr. Zeynep Balik, and Dr. Seher Sener compiled the sample data to be submitted to the model and compared the model with the responses of the pediatric rheumatologist. Dr. Yagmur Bayindir, Dr. Dilara Unal, and Dr. Emil Aliyev were writing the formal analysis. Dr. Emil Aliyev and Dr. Seza Özen finalized the manuscript. Dr. Yelda Bilginer supported making the data compatible with the AI model.

Funding

No funding was used for the study.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to E.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.