مجله دیابت و متابولیسم ایران. بهمن- اسفند ۱۳۹۴؛ دوره ۱۵ (شماره۳)۱۷۶-۱۷۲

مقاله پژوهشی

# بهبود تشخیص بیماری دیابت با استفاده از ترکیب سیستم استنتاج فازی سوگنو و الگوریتم کرم شب تاب

مهدی شیرالی[۱]، یعقوب مدملی[۲]، جمال روح افزا*[۲]، حمید کریمی[۳]، آرمان بابلی بهمئی[۲]، شریف ارتباطی[۲]

## چکیده

**مقدمه:** بیماری دیابت، بیماری مزمنی است که در نتیجه اختلال در تولید و عملکرد انسولین در بدن به وجود می آید.این بیماری یکی از خطرناکترین بیماریهای عصر حاضر است و تشخیص به موقع این بیماری نقش بهسزایی در درمان آن دارد. لذا این مطالعه باهدف بهبود تشخیص بیماری دیابت با استفاده از ترکیب سیستم استنتاج فازی سوگنو و الگوریتم کرم شب تاب انجام شد.

**روش‌ها:** در این مقاله با استفاده از سیستم استنتاج فازی سوگنو و الگوریتم هوشمند کرم شب تاب، روشی نوین برای تشخیص دیابت ارائه شده است. روش ارائه شده قادر است با استفاده از تعداد کمی قوانین ساده فازی با دقت مطلوبی بیماری دیابت را تشخیص دهد.

**یافته‌ها:** کارآیی ترکیب سیستم استنتاج فازی سوگنو و الگوریتم کرم شب تاب ۸۷/۲٤ درصد به‌دست آمد.

**نتیجه‌گیری:** نتایج تجربی نشان می‌دهند که این روش روی مجموعه داده استاندارد PID دقت بیشتری نسبتی به الگوریتم‌های موجود در این زمینه دارد.

**واژگان کلیدی**: تشخیص بیماری دیابت، سیستم استنتاج فازی، سوگنو، الگوریتم هوشمند، کرم شب تاب

Archive of SID

۱- گروه مهندسی برق، واحد بوشهر، دانشگاه آزاد اسلامی، بوشهر، ایران

۲- کمیته تحقیقات دانشجویی، دانشگاه علوم پزشکی دزفول، دزفول، ایران

۳- گروه غدد رشد و متابولیسم، دانشگاه علوم پزشکی دزفول، دزفول، ایران

* ***نشانی***: خوزستان، رامهرمز، کوی سلمان فارسی، کوچه شهید غلامرضا جلالی، کد پستی: ٦٣٨١٦٧٤٣١، تلفن: ۰۹۱٦۷۹۱۲۰٦۸، پست الکترونیک: roohafzaj@gmail.com

**تاریخ دریافت: ۱۳۹۴/۱۰/۴** **تاریخ درخواست اصلاح: ۱۳۹۴/۱۲/۴** **تاریخ پذیرش: ۱۳۹۴/۱۲/۱۷**

www.SID.ir

## مقدمه

در سراسر جهان، اکثر افرادی که در شهرها زندگی می‌کنند از فعالیت بدنی کمتری برخوردارند. جمعیت جهان رو به فزونی است و اکثر افراد نسبت به گذشته از طول عمر بیشتری برخوردارند. در عین حال مردم نسبت به گذشته از غذاهای ناسالم و پر حجم نیز استفاده می‌کنند. بنابراین ترکیب نامناسب (فعالیت بدنی کم و استفاده از غذاهای ناسالم) سبب افزایش غیرقابل کنترل شیوع دیابت در جهان شده است. بیماری دیابت، بیماری مزمنی است که در نتیجه اختلال در تولید و عملکرد انسولین در بدن به وجود می‌آید. انسولین هورمونی است که در لوزالمعده (پانکراس) تولید می‌شود و سلول‌ها را قادر می‌سازد که گلوکز را از خون گرفته و برای تولید انرژی استفاده کنند. لوزالمعده فرد مبتلا به دیابت، انسولین مورد نیاز بدن را تولید نمی‌کند (دیابت نوع یک) و یا انسولین کارایی لازم را در بدن این افراد ندارد (دیابت نوع دو). از این رو، افراد مبتلا به دیابت نسبت به افراد غیر دیابتی نمی‌توانند از گلوکز به نحو احسن در سوخت و ساز بدنشان استفاده کنند، در نتیجه قند خون افراد به مقدار قابل ملاحظه‌ای افزایش می‌یابد. در برآورد و تخمینی که در سال ۱۹۸۵ انجام شده است، مشاهده شده که ۳۰ میلیون نفر در سراسر جهان مبتلا به بیماری دیابت بودند. اگر جلوی پیشرفت این بیماری همه‌گیر گرفته نشود، تا سال ۲۰۲۵ تعداد مبتلایان به دیابت به ۳۳۳ میلیون نفر خواهد رسید. افزایش قند خون سبب ایجاد عوارض زودرس و دیررس دیابت در بدن می‌شود. اگر دیابت پیشگیری و درمان نشود، بسیاری از عوارض دیابت کشنده بوده (بیماری‌های قلبی، عفونت‌های مکرر، عارضه‌های چشمی و...) و یا حداقل سبب کاهش کیفیت زندگی فرد مبتلا به دیابت و خانواده وی می‌شود. بنابراین پیش‌بینی به موقع این بیماری نقش مهمی در جلوگیری از روی دادن عوارض آن دارد [۱]. مطالعات زیادی توسط پژوهشگران یادگیری ماشین روی داده‌های مربوط به دیابت و پیش‌بینی این بیماری انجام شده است [۴-۲]. همچنین مطالعاتی نیز در زمینه استفاده از الگوریتم‌های یادگیری ماشین برای مراقبت از بیماران دیابتی صورت گرفته است که از جمله این موارد می‌توان به پیش‌بینی مقدار گلوکز خون و تجویز مقدار مناسب انسولین برای آن [۵] و همچنین کشف پروتیین‌های تاثیرگذار در بیماری دیابت اشاره کرد [۶، ۷]. بیشتر کارهایی که در حوزه یادگیری ماشین در زمینه دیابت انجام شده است روی مجموعه داده PID[1] از مجموعه داده‌های UCI [۸] بوده است. در مطالعه تهامی و همکاران [۹] دو روش پیشنهاد شده است. در روش اول که از الگوریتم‌های تکاملی استفاده نموده به دقت ۷۶/۱۶ درصد رسیده است. روش دوم نیز از ANFIS[2] استفاده می‌کند که به دقت ۷۹/۶۹ درصد نائل شده است. Jiang و همکاران نیز [۱۰] از روش نزدیکترین همسایگی استفاده کرده‌اند و نتیجه آن دقت ۷۵/۵۵ درصد بوده است. Asma نیز [۱۱] از درخت تصمیم استفاده نموده و دقت ۷۸/۱۷۸ درصد گزارش کرده است. حدادنیا و همکاران [۱۲] نیز از ترکیب درخت تصمیم و الگوریتم خوشه‌بندی فازی استفاده کرده و دقت را تا ۸۴/۳۳ درصد افزایش داده‌اند. در تمام این روش‌ها سعی بر این بوده است که دقت تشخیص افزایش داده شود ولی با این حال توفیق در این زمینه چندان مناسب نبوده و دقت روش‌ها با توجه به اهمیت زیاد موضوع هنوز با حد ایده‌آل فاصله دارد. در این مقاله نیز روشی بر پایه ترکیب روش‌های فازی و الگوریتم کرم شب تاب ارائه شده است. روش ارائه شده سعی می‌کند تا با استفاده از الگوریتم کرم شب تاب پارامترهای یک سیستم فازی را تنظیم نماید.

## روش‌ها

### سیستم استنتاج فازی و الگوریتم کرم شب تاب

روش پیشنهادی در این مقاله از سیستم فازی سوگنو[3] [۱۳] برای کلاسه‌بندی کردن داده‌ها استفاده می‌کند. هر سیستم فازی از تعدادی قانون فازی IF-THEN تشکیل شده است. قوانین سیستم فازی سوگنو برخلاف قوانین سیستم ممدانی توابع عضویت نیستند، بلکه توابع تحلیلی و یا اعداد ثابت هستند.

[1] Pima Indian Diabetes
[2] Adaptive-Network-based Fuzzy Inference Systems
[3] Sugeno

Archive of SID
www.SID.ir

این ویژگی به سیستم اجازه می‌دهد تا دانش پیچیده را با تعداد قوانین کم بیان کند.

سیستم فازی پیشنهادی دارای m(تعداد ویژگی‌ها) ورودی و n قانون است. هر قانون $R_i$ به صورت زیر بیان می‌شود:

$$R_i: if(x_1 \ is \ f_{i1}) \ and \ (x_j \ is f_{ij}) \ and \ ....(x_m \ is \ f_{im}) \ then \ output = yi \quad (۱)$$

که در آن $x_j$ ، j امین ورودی و $f_{ij}$ تابع عضویت i امین قانون روی $x_j$ و $y_i$ خروجی قانون است. توابع عضویت $f_{ij}$ به صورت گوسی استفاده شده‌اند:

$$f_{ij}(x) = \exp\left(-\frac{1}{2}\left(\frac{x-c_{ij}}{\sigma_{ij}}\right)^2\right) \quad (۲)$$

از ضرب به‌عنوان عملگر "AND" استفاده شده است. در نتیجه خواهیم داشت:

$$\mu_i = \Pi_{j=1}^{m} f_{(ij)}(x_j) \quad (۳)$$

که در آن $\mu_i$ درجه فعال شده قانون را نشان می‌دهد.

خروجی سیستم نیز با رابطه مرکز ثقل به‌صورت زیر محاسبه می‌شود:

$$y(x) = \frac{\sum_{i=1}^{n} \mu_i y_i}{\sum_{i=1}^{n} \mu_i} \quad (٤)$$

پارامترهایی که در این سیستم باید تعیین شوند، مرکز و شعاع توابع گوسی و همچنین خروجی هر یک از قوانین است. برای تعیین این پارامترها از الگوریتم کرم شب تاب استفاده شده است. الگوریتم کرم شب‌تاب برای اولین بار توسط Yang [۱٤] معرفی و براساس سه قانون زیر تعریف گشت:

- ۱- کرم‌های شب‌تاب به دیگر کرم‌های شب‌تاب، صرف نظر از جنسیت آن‌ها، جذب می‌شود.
- ۲- جذابیت نسبی است، یعنی کرم‌های شب‌تاب با نور کمتر به سمت کرم‌های شب‌تاب با نور بیشتر جذب می‌شوند، در صورتی که هیچ کرم دیگری از یک کرم نورانی‌تر نباشد، این کرم به‌صورت تصادفی حرکت خواهد کرد.
- ۳- کرم‌های شب‌تاب با پخش کردن هرچه بیشتر نور طعمه را به سمت خود می‌کشند و در نهایت قربانی را با دیگر کرم‌ها به اشتراک می‌گذارند.

الگوریتم کرم شب‌تاب الگوریتمی تکاملی بر پایه جمعیت می‌باشد، که از رفتار کرم‌های شب‌تاب در جستجوی غذا و هوش جمعیتی آنها الهام گرفته شده است. در طبیعت کرم‌های شب‌تاب به‌صورت تصادفی حرکت می‌کنند و هرکدام که طعمه بهتری پیدا کند از خود نور بیشتری ساطع کرده، دیگران را به سمت خود جذب می‌کند. هر چه فاصله دو کرم از یکدیگر بیشتر باشد، درصد جذب آن‌ها به یکدیگر کمتر می‌شود. به‌عبارتی، فاصله با سرعت و میزان جذب نسبت عکس دارد. این الگوریتم از دو بخش اساسی تشکیل شده است:

الف) تغییرات شدت نور

میزان شدت نور به مقدار تابع هزینه بستگی دارد [۱۵]. لذا در مسائل کمینه‌سازی (بیشینه‌سازی)، کرم شب‌تاب با نور زیادتر (کمتر)، کرم‌های شب‌تاب با نور کم‌تر (زیادتر) را به خود جذب می‌کند. فرض کنید $n$ تعداد کرم‌ها، $x_i$ موقعیت ذره $i$ و $f(x_i)$ تابع هزینه باشد. لذا میزان درخشندگی هر کرم از برابر با مقدار تابع هزینه آن خواهد بود.

$$I_i = f(x_i), \quad 1 \le i \le n \quad (۵)$$

ب) حرکت به سمت کرم شب‌تاب با نور بیشتر

هر کرم دارای یک مشخصه جذب می‌باشد، که نشان می‌دهد کرم تا چه اندازه‌ای قوی است. این مشخصه مقداری نسبی بوده و با تغییر فاصله بین دو کرم $i$ و $j$ تغییر می‌کند. تابع جذب از رابطه زیر به‌دست می‌آید.

$$\beta(r) = \beta_0 e^{-\gamma r^2} \quad (٦)$$

که در آن $\beta_0$ میزان جذب به ازای $r = 0$ و $\gamma$ ضریب جذب نور می‌باشد. حرکت کرم $i$ با موقعیت $x_i$ به سمت کرم $j$ با موقعیت $x_j$ (با نور بیشتر) از رابطه زیر به‌دست می‌آید.

$$x_i(t+1) = x_i(t) + \beta_r(x_j - x_i) \quad (۷)$$

Archive of SID

www.SID.ir

شبه کد الگوریتم کرم شب‌تاب به‌صورت زیر می‌باشد:

۱- تعیین و مقداردهی اولیه کرم‌های (افراد) جمعیت
۲- به‌دست آوردن تابع هزینه برای هر فرد از جمعیت در موقعیت خود
۳- تا زمانی که شرایط توقف برقرار نشده است:
٤- تخصیص تصادفی شدت نور به هر فرد از جمعیت
۵- تعیین بهترین (پرنورترین) فرد در جمعیت به‌واسطه محاسبه توابع هزینه کل افراد جمعیت
٦- حرکت دیگر افراد جمعیت به سمت بهترین فرد و به روز رسانی شدت نور با توجه به آن
۷- پایان الگوریتم

عملکرد الگوریتم کرم شب‌تاب به‌طور کامل به تعداد جمعیت اولیه، تابع جذب و ضریب جذب وابسته است. هرچه ضریب جذب نور بیشتر باشد، سرعت جذب افراد جمعیت به سمت پرنورترین فرد بیشتر می‌شود.

همچنین MSE که به‌صورت زیر تعریف می‌شود به‌عنوان تابع برازش استفاده شد:

$$MSE = \sum_{i=1}^{k} (S_i - T_i)^2 \qquad (۸)$$

که در آن $k$ تعداد داده‌های آموزشی، $S_i$ خروجی دلخواه برای داده آموزشی i ام و $T_i$ خروجی سیستم فازی برای همان داده است.

## تشخیص بیماری دیابت با استفاده از سیستم استنتاج فازی و الگوریتم کرم شب تاب

روش ارائه شده با استفاده از نرم‌افزار Matlab پیاده‌سازی شد و الگوریتم‌های ارائه شده در سایر بررسی‌ها [۱۲-۹]، روی مجموعه داده PID مورد مقایسه قرار گرفت. این مجموعه داده در بردارنده ۸ ویژگی از ۷٦۸ زن است که حداقل ۲۱ سال سن داشته‌اند. از این میان ۵۰۰ نفر سالم و ۲٦۸ نفر مبتلا به دیابت هستند. ۸ ویژگی ثبت شده از افراد براساس تعریف سازمان جهانی بهداشت به این صورت است: الف: تعداد دفعات بارداری، ب: غلظت گلوکز پلاسمای خون در دو ساعت، ج: فشار خون، د: شاخص جرم بدن، ه: انسولین سرم دو ساعته و: ضخامت پوست ماهیچه سه سر بازویی، ز: سابقه بیماری دیابت، ک: سن.

نتایج به‌دست آمده از مقایسه روش‌ها در جدول ۱آمده است.

**جدول ۱- دقت روش های مختلف و روش ارائه شده**

| روش | دقت |
|---|---|
| ANFIS [۹] | ٪۷٦/٦۹ |
| الگوریتم تکاملی [۹] | ٪۷٦/۱۷ |
| K نزدیکترین همسایه [۱۰] | ٪۷۵/۵۵ |
| درخت تصمیم [۱۱] | ٪۷۸/۱۷۸ |
| فازی با خوشه‌بندی [۱۲] | ٪۸٤/۳۳ |
| روش پیشنهادی | ٪۸۷/۲٤ |

دقت‌های نشان داده شده در جدول ۱ نشان دهنده کارایی بالا و دقت بهتر روش پیشنهادی می‌باشد. با توجه به اینکه در روش پیشنهادی از سیستم استنتاج فازی و الگوریتم کرم شب تاب به‌طور هم‌زمان استفاده شده است، از هوشمندی بالایی برخوردار می‌باشد.

## نتیجه گیری

در این مقاله با استفاده از الگوریتم کرم شب تاب و سیستم استنتاج فازی روشی برای تشخیص بیماری دیابت ارائه شد. نتایج تجربی روی مجموعه داده PID نشان می‌دهند که روش ارائه شده دقت بیشتری نسبت به روش‌های موجود در این زمینه دارد. استفاده از سیستم های فازی با

Archive of SID

www.SID.ir

قدرت استنتاج قوی‌تر نظیر سیستم فازی ممدانی و سیستم فازی مرتبه دو نیز می‌تواند به دقت تشخیص به کمک روش‌های هوشمند بیافزاید. همچین الگوریتم‌های بهینه‌سازی دقیق‌تر و قدرتمندتری چون الگوریتم ممتیک و الگوریتم رقابت استعماری جهت استفاده در تشخیص بیماری دیابت پیشنهاد می‌گردد.

## مآخذ


۹. تهامی س ا، بامشکی س م، خلیل زاده م ع. *تشخیص بیماری دیابت نوع ۱ با استفاده از الگوریتم ANFIS , GA-NN*. اولین کنفرانس مشترک سیستم‌های هوشمند و سیستم‌های فازی، شهریور ۱۳۸٦، دانشگاه فردوسی مشهد.

10. Jiang Y. and Zhou Z. *Editing Training Data for kNN Classifiers with Neural Network Ensemble*, in Proc. ISNN (1), 2224, pp.396-361.
11. Asma A. *AlJarullah, King Saud University, Decision Tree Discovery for the Diagnosis of Type II Diabete*. intrnational Confrencee on Innovation in Information Technology, 2011.

۱۲. حداد نیا ج، وحیدی ج، قره‌خانی ا، فیوضی م. تشخیص فازی بیماری دیابت براساس قوانین و ویژگی‌های بهینه مبتنی بر ترکیب سیستم‌های داده‌کاوی و الگوریتم‌های هوشمند مصنوعی. کنفرانس بین المللی مدل سازی غیر خطی و بهینه سازی. آمل، دانشگاه شمال، ۲۰۱۲.

13. Haykin, Simon. *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.
14. Yang XS. *Nature-Inspired Metaheuristic Algorithms*, Luniver Press, 2008.

۱. وب‌سایت انجمن دیابت ایران، دیابت و عوارض آن: www.ir-diabetes-society.com

2. Shanker M. Using neural networks to predict the onset of diabetes mellitus. *J Chem Inform Computer Science* 1996; 36:35–41,.
3. Breiman L. Random forests. *Machine Learning* 2001; 45(1):5–32.
4. Dazzi D, Taddei F, Gavarini A, Uggeri E, Negro R and Pezzarossa A. The control of blood glucose in the critical diabetic patient: a neuro-fuzzy method. *Journal of Diabetes Complications* 2001; 15(2):80–87.
5. Patterson K and Sandham W. *Neural network and neuro-fuzzy systems for improving diabetes therapy*. In 20th Annual International Conference of the IEEE in Engineering in Medicine and Biology Society, Hong Kong, China, 1998.
6. Hsu W, Lee ML Liu B, and Ling TW. *Exploration mining in diabetic subjects databases: Findings and conclusion*. In 6th ACM SIGKDD international conference on Knowledge discovery and data mining, 2000.
7. Zorman M Masuda G Kokol P Yamamoto R and Stiglic B. *Mining diabetes database with decision trees and association rules*. In 15th IEEE Symposium on Computer-Based Medical Systems, 2002; pages 134–139.
8. Blake CL Merz CJ. UCI Repository of Machine Learning Databases, 1998; Retrieved from: www.ics.uci.edu/mlearn/MLRepository.html


Archive of SID

www.SID.ir

# IMPROVEMENT DIAGNOSIS OF DIABETES USING A COMBINATION OF SUGENO FUZZY INFERENCE SYSTEMS AND FIREFLY ALGORITHMS

Mehdi Shirali[1], Yaghoob Madmoli[2], Jamal Roohafza[2*], Hamid Karimi[3], Arman Baboli bahmaei[2], Sharif Ertebati[2]

1. *Student in Electrical Engineering, Bushehr Islamic Azad University, Bushehr, Iran*
2. *Student Research Committee, Dezful University of Medical Sciences, Dezful, Iran*
3. *Subspecialist of Endocrine, Growth and Metabolism, Faculty member of Dezful University of Medical Sciences, Dezful, Iran*

## ABSTRACT

**Background:** Diabetes is a chronic disease that comes into existence results in a disruption in the production and function of insulin in the body. Diabetes is a serious disease of modern societies and its on-time diagnosis have important role in the treatment of disease.So this study aimed to assess improvement diagnosis of diabetes using a combination of sugeno fuzzy inference systems and firefly algorithms.
**Methods:** In this paper, using Sugeno fuzzy inference systems and intelligent algorithms Firefly, a new method is presented for the detection of diabetes. The proposed method enables the use of a few simple fuzzy rules to detect diabetes with good accuracy.
**Results:** Sugeno fuzzy inference system and firefly algorithm combines efficiency was 87.24 percent.
**Conclusion:** Experimental results show that this method deals on the data set is more accurate than the standard PID algorithm in this field**.**

**Keywords:** Diagnosis of diabetes, Inference System, Sugeno, intelligent algorithms, Firefly

Archive of SID

[*] Khuzestan, Ramhormoz, Salman Farsi alley, Gholam Reza Jalali martyr alley, Tell: 09167912068, Postal Code: 638167431, Email: roohafzaj@gmail.com

www.SID.ir