



مقاله پژوهشی

ارائه الگوریتم جدید Fuzzy SARSA به منظور پیش بینی نوسانات سطح قند خون بیماران مبتلا به دیابت نوع یک

حسین فرضعلی وند^{۱*}

چکیده

مقدمه: یکی از عوارض خطرناک بیماری دیابت نوع یک افزایش و کاهش ناگهانی سطح غلظت قند خون می باشد که باعث بروز خطراتی چون اغماء و بیهوشی خواهد شد. بنابراین استفاده از روش هایی با درصد خطای پایین و هوشمند با اطلاعات موجود در جهت پیش بینی و در نهایت پیش گیری از این عوارض، گام مهمی در جهت کنترل بهینه بیماری محسوب می شود. به همین منظور در این مقاله ما از الگوریتم ترکیبی Fuzzy SARSA برای طراحی سیستم خبره در جهت پیش بینی نوسانات سطح قند خون بیماران مبتلا به دیابت نوع یک استفاده کرده ایم. هدف از این مقاله، ارائه و پیاده سازی یک الگوریتم ترکیبی جدید مبتنی بر یادگیری تقویتی و منطق فازی به نام نوع یک استفاده کرده ایم. هدف از این مقاله، ارائه و پیاده سازی یک الگوریتم ترکیبی جدید مبتنی بر یادگیری تقویتی و منطق فازی به نام FSA (Fuzzy SARSA Algorithm) برای طراحی سیستم خبره می باشد.

روش ها: داده های پزشکی مورد استفاده در این مقاله مربوط به ۳ بیمار زن ایرانی مبتلا به دیابت نوع یک است، که شامل الگوهای چون نوع و دز انسولین تزریقی، فاصله زمانی بین ساعات ثبت سطح غلظت قند خون و نیز سطح غلظت قند خون در ابتدای بازه زمانی می باشد. در این مقاله از الگوریتم معروف یادگیری تقویتی به نام SARSA استفاده شده و با منطق فازی پیاده سازی شده و الگوریتم Fuzzy SARSA را در سیستم خبره برای پیش بینی نوسانات سطح قند خون بیماران مبتلا به دیابت نوع یک به کار برده ایم.

یافته ها: نتایج به دست آمده برای سیستم خبره با الگوریتم پیشنهادی FSA برای عامل یادگیرنده با استفاده از پارامترهای، سطح غلظت قند خون در ابتدای بازه زمانی (mg/dlit)، دز انسولین کوتاه اثر تزریقی به بدن (unit)، دز انسولین طولانی اثر تزریقی به بدن (unit)، سطح استرس (unit)، سطح فعالیت بدنی (unit)، کربوهیدرات مصرفی (gr)، فاصله میان ابتدا و انتهای بازه زمانی (hour) نشان می دهد که عامل توانسته در ۸۵ درصد سطح قند خون را در بازه نرمال ۸۰ تا ۱۲۰ نگه دارد در ۱۵ درصد خطایی که عامل داشته مواقعی است که عامل اکتشاف درستی برای تزریق انسولین کوتاه اثر و طولانی اثر نداشته است.

نتیجه گیری: در پایان تحقیق با توجه به نتایج به دست آمده مشخص شد که الگوریتم پیشنهادی برای سیستم هوشمند با کاهش تقریباً ۱۵ درصدی خطای پیش بینی نسبت به دیگر روش های متداول و از جمله روش های مبتنی بر یادگیری تقویتی، می تواند به عنوان مدلی مناسب به منظور پیش بینی نوسانات سطح غلظت قند خون در بیماران مبتلا به دیابت نوع یک تلقی گردد، ضمن اینکه می توان از سایر روش های پیش بینی مبتنی بر هوش مصنوعی مانند شبکه های عصبی جهت پیش بینی دقیق استفاده کرد.

واژگان کلیدی: یادگیری تقویتی، الگوریتم Fuzzy SARSA، دیابت، پیش بینی نوسانات سطح غلظت قند خون

۱- باشگاه پژوهشگران جوان و نخبگان، دانشگاه آزاد اسلامی واحد ایذه، ایذه، ایران

***نشانی:** خوزستان، ایذه، خیابان زاگرس، میدان دانشجو، دانشگاه آزاد اسلامی واحد ایذه، باشگاه پژوهشگران و نخبگان تلفن:

۰۹۳۸۴۲۲۱۸۹۵، پست الکترونیک: farzalivand@gmail.com

تاریخ پذیرش: ۱۳۹۴/۰۲/۲۰

تاریخ درخواست اصلاح: ۱۳۹۴/۰۱/۲۴

تاریخ دریافت: ۱۳۹۳/۱۲/۱۸



مقدمه

بیماری دیابت

بیماری دیابت از جمله بیماری‌های شایع و بسیار خطرناک حال حاضر دنیا است که سالانه خسارات مادی و معنوی بسیار زیادی را متوجه جوامع و کشورها می‌سازد. در حال حاضر در ایران ۲ تا ۳ میلیون نفر به این بیماری مبتلا می‌باشند، که نیمی از این افراد حتی از ابتلاء خود به این بیماری نیز آگاهی ندارند. این بیماری را می‌توان به دو نوع تقسیم نمود، نوع یک یا نوع وابسته به انسولین که در این نوع لوزالمعده شخص مبتلا به دیابت قادر به ترشح انسولین نمی‌باشد و نوع دو یا نوع غیر وابسته به انسولین که در آن لوزالمعده شخص مبتلا به دیابت قادر به تولید و ترشح انسولین می‌باشد، اما میزان جذب آن در بافت‌های بدن بسیار پایین است. مشکل عمده‌ای که در حال حاضر در رابطه با این بیماری مخرب و خطرناک و به خصوص در نوع یک آن وجود دارد افزایش و یا کاهش بیش از حد و ناگهانی سطح قند خون است که از جمله عوارض بسیار خطرناک آن محسوب می‌گردد و می‌تواند حتی به بیهوشی و اغماء بیمار منتهی گردد. همچنین مشکل دیگری که در بیماران دیابتی وجود دارد تنظیم دز انسولین تزریقی به بدن بدون دانستن پاسخ سطح قند خون بدن به تزریق این مقدار انسولین می‌باشد که در بسیاری مواقع همین عدم اطلاع از میزان تزریق انسولین خود باعث به وجود آمدن مشکلاتی حاد و در برخی مواقع حتی غیر قابل جبران خواهد گردید.

با وجود این که دیابت از دیرباز برای بشر شناخته شده است اما دلایل کامل پدید آمدن آن همچنان ناشناخته باقی مانده است. افزایش شدید میزان ابتلا به دیابت، پیرو تغییر سبک زندگی و عوامل ناشناخته دیگر در سال‌های اخیر باعث شده سرمایه‌گذاری زیادی در جهت یافتن درمان‌های بهتر و حتی درمان قطعی برای دیابت انجام شود.

روش‌های درمانی برای دیابت می‌دانیم قند خون عادی، بین ۸۰ تا ۱۲۰ میلی‌گرم در دسی‌لیتر است. آنچه موجب عوارض بلندمدت بیماری دیابت است در حقیقت بالا بودن قند خون به مدت طولانی و همین‌طور نوسانات زیاد قندخون در بیمار است. پس هدف اصلی در هر روش درمانی برای دیابت، نگه داشتن قند خون در این محدوده است. شمار مبتلایان به دیابت روز به روز در حال افزایش است. این مسأله پژوهشگران را بر آن داشته

است که به‌عنوان مشکلی فراگیر برای بهبود زندگی مبتلایان تلاش کنند. روش‌های درمانی دیابت مبتنی بر استفاده از تجهیزات پزشکی الکترونیکی، شامل چیپ‌های کنترلرلی، پمپ‌های تزریق کننده خودکار و حسگرهای پیوسته در زمان اندازه گیری قند خون است. مهندسان در تلاشند با ترکیب مجموعه‌ای از این اجزاء ساختاری تماماً خودکار را به‌وجود بیاورند که مثل لوزالمعده انسان فکر کرده و مقادیر مناسبی دارو را با کنشگرهای مکانیکی - الکتریکی به بیمار تزریق کند. در این روش برخلاف دو روش زیستی و ژن درمانی، انسولین تولید نمی‌شود بلکه باید به‌طور مرتب به مخزن مربوطه تحویل داده شود. هیچ سلول زیستی در این فرآیند وجود ندارد. این‌ها روش‌های درمان دیابت مبتنی بر تجهیزات پزشکی هستند. روش‌های زیستی در تلاشند سلول‌های ترشح کننده انسولین در لوزالمعده را که در بیماران دیابتی آسیب دیده یا به‌طور کامل از بین رفته‌اند، به بدن بیمار باز گردانند. این روش‌ها به‌صورت عمده شامل پیوند کامل لوزالمعده یا پیوند سلول‌های بتای لوزالمعده (سلول‌های ترشح کننده انسولین در لوزالمعده) به تنهایی می‌شوند.

به‌همین منظور استفاده از روش‌های هوشمند که بتوان از آن طریق ضمن کاهش تعداد اندازه‌گیری‌های قند خون در این بیماران و جلوگیری از اثرات متعاقب ناشی از این اندازه‌گیری‌ها به نوعی وقوع عوارض خطرناک احتمالی ذکر شده را پیش‌بینی نمایند می‌تواند مهمترین گام به‌منظور تشخیص، کنترل و حتی درمان مناسب این عوارض تلقی گردد.

تاکنون روش‌های متفاوتی برای رسیدن به این مهم مورد استفاده قرار گرفته‌اند که از آن جمله می‌توان به مدل‌های ریاضی و مینیمال [۱، ۲]. که عمدتاً به منظور تعیین حساسیت بدن بیمار مبتلا به دیابت در پاسخ به گلوکز تزریقی و نیز پیش‌بینی کوتاه مدت سطح قند و انسولین موجود در بدن مورد استفاده قرار گرفته‌اند، مدل‌های مبتنی بر رویکردهای احتمالاتی [۳]. که مشابه با مدل‌های ریاضی به‌منظور مدل‌سازی سیستم متابولیکی بدن مورد استفاده قرار گرفته‌اند، سیستم‌های خبره [۴، ۵]. اشاره نمود که به‌منظور پیش‌بینی وضعیت سطح قند خون بیمار مبتلا به دیابت در آینده استفاده شده‌اند اشاره نمود.

از جمله مشکلاتی که در مدل‌های ذکر شده به چشم می‌خورد



پرسپترون تک لایه و یادگیری تقویتی ارائه کردند، که توانستند با استفاده از کمترین اطلاعات برای پیش‌بینی حرکت شامل صرفاً قد، وزن و زوایای مفاصل سرعت یادگیری شبکه را افزایش دهند.

یکی از داستان‌های موفق در مورد کاربردهای یادگیری تقویتی، [Tesauro's TD-Gammon ۹] بود که برای بازی تخته نرد استفاده شد. TD-Gammon بر اساس TD(λ) توسعه داده شد و یک شبکه عصبی چند لایه‌ای برای تقریب تابع استفاده شد. آخرین نسخه TD-Gammon قادر به انجام بازی تخته نرد نزدیک به سطح بهترین بازیکن انسانی در دنیا بود.

Crites و [Barto ۱۰] از Q-Learning برای یک سیستم ۱۰ طبقه ۴ آسانسور استفاده کردند. هر آسانسور به‌طور مستقل با دیگر آسانسورها تصمیم می‌گرفت. تعدادی محدودیت نیز در مورد تصمیمات وجود داشت. این سیستم 10^{22} وضعیت داشت. شبیه به TD-Gammon، Crites و Barto نیز یک شبکه عصبی را برای نمایش تابع (وضعیت-عمل) به کار گرفتند.

ج- یادگیری تقویتی

یادگیری تقویتی یکی از حوزه‌های یادگیری ماشین است و اخیراً یکی از بیشترین موضوعات مطرح در هوش مصنوعی است. Arthur Samuel (۱۹۵۹) [۱۱] از جمله اولین کسانی بود که بر روی یادگیری ماشین کار می‌کرد، با برنامه چکر او از سیگنال‌های پاداش که یک مولفه کلیدی در یادگیری تقویتی پیشرفته هستند استفاده نمی‌کرد، اما [Sutton ۱۲] بر روی سیگنال‌های پاداش کار کرد و بعضی از تکنیک‌های Samuel را که شباهت زیادی به الگوریتم‌های تفاضل زمانی امروزی داشت را گسترش داد.

هدف اصلی از یادگیری، یافتن شیوه‌ای برای عملکرد در حالات مختلف است که این شیوه در مقایسه با سایرین، با در نظر گرفتن معیارهایی، بهتر است. هنگامی می‌توان گفت یادگیری اتفاق افتاده است که، عاملی بر اساس تجربیاتی که کسب می‌کند به نحوه دیگر، و به احتمال زیاد، بهتر عمل کند. در این صورت باید نحوه عملکرد عامل در اثر کسب اطلاعات جدید، متفاوت از نحوه عملکرد در زمان‌های قبل از کسب این اطلاعات و تجارب باشد. در دهه ۹۰ یادگیری تقویتی جذابیت زیادی را بین محققان کسب کرد. در مقابل رویکرد محبوب یادگیری با ناظر که عامل از طریق

عدم وجود اطلاعات کافی و کامل که بتوان از آن طریق به مدلی مشابه مدل واقعی دست یافت، زیرا مدل‌ها اگرچه ممکن است کاملاً صحیح عمل نمایند اما زمانی که با یکدیگر ترکیب می‌گردند خطا در خروجی مدل افزایش خواهد یافت چرا که تمام خواص تعاملی و برهمکنشی موجود میان ارگان‌های موجود در سیستم زنده بدن انسان در مدل لحاظ نشده‌اند. همچنین به این موارد می‌توان تعداد کم ویژگی‌ها و پارامترهای تاثیرگذار بر نوسانات سطح قند خون مورد استفاده در مدل‌های مبتنی بر هوش مصنوعی را نیز افزود که این خود باعث ایجاد خطاهایی در پیش‌بینی سطح قند خون بدن خواهد شد. در این مقاله، سعی کرده‌ایم با پیاده‌سازی الگوریتم معروف یادگیری تقویتی SARSA با منطق فازی برای سیستم خبره در فرآیند پیش‌بینی نوسانات سطح قند خون، بهترین رویکرد را نسبت به تحقیقات انجام شده قبلی انتخاب کنیم.

ب- کارهای مرتبط قبلی

قوچانی و تهامی [۶] مقایسه‌ای بین شبکه عصبی خود بازگشتی المن و شبکه عصبی پرسپترون سه لایه به‌منظور پیش‌بینی نوسانات سطح غلظت قند خون بیماران مبتلا به دیابت نوع یک انجام دادند که نتایج آن‌ها برای شبکه عصبی خود بازگشتی المن در کل فرآیند پیش‌بینی نوسانات سطح غلظت قند خون بسیار دقیقتر از شبکه‌های عصبی پرسپترون چند لایه عمل نمود.

عمادی و همکاران [۷]، یک الگوریتم جدید مبتنی بر خواص غیرخطی درون‌یاب- برون‌یاب شبکه عصبی ارائه کردند. هدف آن‌ها پیش‌بینی الگوی حرکت بر خاستن از صندلی بر مبنای کمترین داده‌ها اندازه‌گیری شده به‌عنوان ورودی شبکه برای هر شخص (به‌طور جداگانه) بوده است. آن‌ها سه شبکه عصبی را مورد ارزیابی قرار داده‌اند. در این روش در مقایسه با روش‌های قبلی با درصد خطای کمتری نسبت به داده‌های اندازه‌گیری شده حرکت پیش‌بینی گردیده است. شبکه عصبی به‌کار رفته یک شبکه پرسپترون دولایه می‌باشد. روش یادگیری شبکه عصبی به‌کار رفته پس انتشار خطا بود که یک روش با سرپرستی است و پارامترهای شبکه بر اساس مشاهده اطلاعات ثبت شده از فرد تعلیم می‌یابند.

عمادی و همکاران [۸]، یک الگوریتم جدید برای پیش‌بینی حرکت برخاستن از صندلی با استفاده از یک شبکه عصبی



مثال‌هایی که توسط ناظر خارجی [۱۳] آماده شده است آموزش می‌بیند. یادگیری تقویتی نیاز دارد که عامل به‌طور مستقیم با تعامل با سیستم آموزش ببیند و به پاداش‌ها یا جریمه‌های مبتنی بر اثر هر عمل در سیستم پاسخ دهد.

در یادگیری تقویتی، هدف اصلی از یادگیری، انجام دادن کاری و یا رسیدن به هدفی است، بدون آنکه عامل یادگیرنده، با اطلاعات مستقیم بیرونی تغذیه شود. در این روش، تنها مسیر اطلاع رسانی به عامل، از طریق یک سیگنال پاداش یا جریمه می‌باشد. تنها چیزی که از طریق سیگنال پاداش به عامل فهمانده می‌شود، این است که آیا تصمیم مناسبی گرفته است یا نه؟

اگرچه کاربردهای یادگیری تقویتی بسیاری وجود دارد که سودمندی آن را توصیف می‌کنند اما کاربرد آن در سیستم‌های مختلف به‌طور کامل کشف نشده است [۱۴، ۱۵]. موضوع این مطالعه تحقیق در مورد طرز کار الگوریتم ارائه شده Fuzzy SARSA بر روی مسئله پیش‌بینی نوسانات سطح قند خون در بیماران مبتلا به دیابت نوع یک است. در این مقاله به‌طور خاص از الگوریتم معروف یادگیری تقویتی به نام SARSA که با منطق فازی پیاده سازی شده است، برای پیش‌بینی نوسانات سطح قند خون و انتخاب بهترین قانون برای تزریق به بیمار استفاده شده تا تعیین کند که این الگوریتم با روش فازی چه عملکرد بهینه‌ای را در هنگام پیش بینی نوسانات در زمان‌های مختلف دارد و انتخاب سیاست بهینه در کل فرآیند پیش‌بینی و سرعت آموزش عامل در این الگوریتم نسبت به دیگر الگوریتم‌های یادگیری تقویتی به‌کار رفته در این قبیل مسائل ایجاد می‌کند.

به علاوه تجربیات به‌دست آمده از این تحقیق می‌تواند پایه‌ای برای تحقیقات بیشتر در به‌کارگیری تکنیک‌های یادگیری تقویتی برای مسائل پیش بینی نوسانات سطح قند خون باشند.

این مقاله به‌صورت زیر سازماندهی شده است: بخش بعدی معرفی یادگیری تقویتی، بخش ۳ داده و روش، بخش ۴ انتخاب قانون اعزام در سیستم خبره، بخش ۵ نتایج شبیه‌سازی و بخش آخر نتیجه گیری.

معرفی یادگیری تقویتی

در چارچوب یادگیری تقویتی یک عامل یادگیرنده باید اطلاعات را از محیطش درک کند. اطلاعات درک شده برای تشخیص وضعیت فعلی محیط استفاده می‌شود. سپس عامل یک عمل را

بر اساس وضعیت درک شده انتخاب می‌کند تا اجرا کند. عمل انتخاب شده ممکن است بر اثر یک تغییر در وضعیت محیط حاصل شده باشد. براساس وضعیت جدید یک تقویت سریع وجود دارد که برای پاداش دادن یا جریمه کردن عمل انتخاب شده، در نظر گرفته شده است. این فعل و انفعالات بین عامل و محیطش تا زمانی که عامل یک استراتژی تصمیم‌گیری را یاد می‌گیرد تا مجموع پاداش‌ها را حداکثر کند ادامه می‌یابد. Sutton و Barto [۱۴] چهار عنصر کلیدی را برای برخورد با مسائل یادگیری تقویتی عنوان می‌کنند: یک سیاست، یک تابع پاداش، یک تابع مقدار و یک مدل برای محیط. یک سیاست، رفتار عامل را در یک وضعیت داده شده تعریف می‌کند. یک تابع پاداش هدف کلی عامل را توصیف می‌کند که عامل را برای به دست آوردن هدف، آموزش می‌دهد. یک تابع مقدار، مقدار یک وضعیت را تعیین می‌کند یا یک جفت (وضعیت-عمل) که مشخص می‌کند عامل چقدر خوب است. یک مدل محیط، وضعیت بعدی را با توجه به وضعیت فعلی داده شده و عمل پیشنهاد شده پیش‌بینی می‌کند. علاوه بر چهار عنصر بالا یک فرض کلیدی در چارچوب یادگیری تقویتی وجود دارد. اینکه هر تصمیم که الگوریتم یادگیری تقویتی می‌گیرد براساس وضعیت فعلی است که تمام موارد مهم را درباره دنباله کاملی از وضعیت‌های قبلی که منجر به آن می‌شوند خلاصه می‌کند. تعدادی از اطلاعات درباره دنباله کامل ممکن است گم شوند ولی همه آن‌هایی که برای آینده اهمیت دارند، شامل سیگنال وضعیت جاری هستند. این خاصیت مارکوف نامیده می‌شود. بنابراین اگر یک محیط دارای خاصیت مارکوف باشد وضعیت بعدی آن می‌تواند وضعیت فعلی و عمل فعلی داده شده را پیشگویی کند. این فرض اصلی وضعیت فعلی را قادر می‌سازد که پایه خوبی برای پیش بینی وضعیت بعدی باشد. تحت این فرض فعل و انفعالات یک عامل و محیطش می‌تواند یک فرآیند تصمیم مارکوف نامیده شود. برای یک مسئله یادگیری تقویتی کوچک، تخمین‌های توابع مقدار می‌تواند به‌عنوان یک جدول با یک ورودی برای هر وضعیت یا برای هر جفت (وضعیت-عمل) باشد. بنابراین برای یک مسئله بزرگ با تعداد زیادی از وضعیت‌ها و عمل‌ها به روز کردن اطلاعات به‌طور دقیق در چنین جدول بزرگی ممکن است یک مشکل باشد. تقریب تابع معمولاً یک روش مرسوم برای حل این مسئله است. تقریب تابع یک رویکردی است که تجربیات را از یک زیر

- سطح فعالیت بدنی (unit)
- کربوهیدرات مصرفی (gr)
- فاصله میان ابتدا و انتهای بازه زمانی (hour)

طول مدت زمان ثبت که در این تحقیق در برخی بیماران ۱۳۵ و در برخی دیگر ۷۵ روز می باشد. از این تعداد ۷۵٪ تعداد داده‌های ثبت شده برای هر نفر به منظور آموزش عامل یادگیرنده و ۲۵٪ به منظور اکتشاف و استخراج عامل مورد استفاده قرار می گیرند.

لازم به ذکر است که میزان فعالیت بدنی و نیز میزان سطح استرس با انتخاب یکی از اعداد ۱ تا ۴ که به ترتیب بیان کننده وضعیت معمولی (۱)، کم (۲)، متوسط (۳) و زیاد (۴) می باشد، در زیر جدول ثبت مشخص می گردد. نمونه‌ای از داده‌های ثبت شده در جدول ۱ قابل مشاهده می باشد.

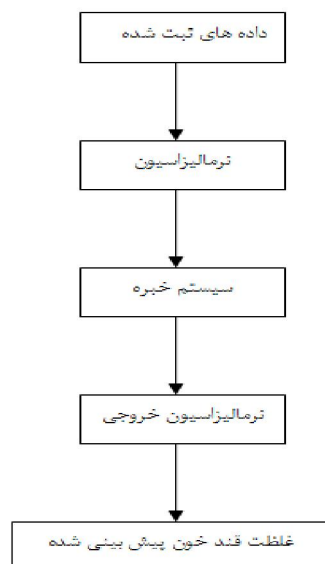
جدول ۱- نمونه‌ای از مجموعه داده

قند خون بعدی	زمان	کربوهیدرات	سطح فعالیت بدنی	سطح استرس	انسولین طولانی اثر	انسولین کوتاه اثر	قند خون اولیه
۲۲۰	۸	۳	۲	۰	۸	۱۰	۱۸۰
۶۵	۹	۶	۰	۱	۵	۹	۲۵۰

روش مورد استفاده

در این روش ابتدا داده‌ها مورد پردازش اولیه قرار می گیرند تا ضمن حذف مولفه‌های غیر ضروری، بتوانند به منظور کاربرد در سیستم خبره مورد استفاده قرار گیرند، به همین

منظور ابتدا داده‌ها نرمالسازی گردیده‌اند در مرحله بعدی داده‌های نرمالسازی شده به عامل یادگیرنده داده می شوند تا فرآیند پیش‌بینی صورت پذیرد، که بلوک دیاگرام سیستم مورد استفاده در این مقاله براساس شکل ۱ است:



شکل ۱- بلوک دیاگرام سیستم خبره طراحی شده برای پیش بینی نوسانات سطح قند خون در افراد مبتلا به دیابت نوع یک



"آپیزود" نامیده می‌شود. پارامتر α پارامتر اندازه مرحله است و بر نرخ یادگیری تاثیر می‌گذارد. پارامتر γ پارامتر تخفیف نرخ است و $0 \leq \gamma \leq 1$ و بر مقدار فعلی پاداش‌های آتی اثر می‌گذارد. مقادیر $Q(s,a)$ می‌توانند به صورت اختیاری آغاز شوند. اگر هیچ عملی برای وضعیت‌های مشخص پیشنهاد نشود زمانی که الگوریتم Fuzzy SARSA آغاز می‌شود همه مقادیر $Q(s,a)$ در یک جدول سیاست می‌توانند با مقدار یکسان آغاز شوند. اگر تعدادی از دانش‌های اولیه درباره منفعت عمل‌های معین موجود باشد عامل ممکن است آن عمل‌ها را در آغاز برای مقدار دهی اولیه $Q(s,a)$ استفاده کند. این عمل‌ها در آغاز انتخاب می‌شوند این امر سبب می‌شود که دوره یادگیری کاهش یابد.

هرچه γ به یک نزدیکتر باشد، الگوریتم یادگیری آینده نگرتر می‌شود و هرچه γ به صفر نزدیکتر باشد الگوریتم یادگیری حریص تر می‌شود، انتخاب $\gamma = 1$ ، غیر مجاز است مگر در سناریوهای با تعداد مراحل محدود.

انتخاب قانون اعزام در سیستم خبره

انتخاب قانون

در این مطالعه از الگوریتم Fuzzy SARSA برای پیش‌بینی نوسانات سطح قند خون در ۳ بیمار زن که داده‌های آن‌ها در سیستم خبره ثبت شده، استفاده می‌کنیم. سیستم خبره شامل یک بافر برای ذخیره‌سازی داده‌های ۳ بیمار است که منتظر پردازش لازم از طرف سیستم خبره هستند. در این تحقیق ما اثر کاربرد الگوریتم Fuzzy SARSA، برای مسئله پیش‌بینی نوسانات سطح قند خون را بررسی می‌کنیم و هدف سیستم خبره در این مطالعه نگه داشتن سطح قند خون در ۳ بیمار زن که براساس داده‌های آن‌ها عمل می‌نمایند، بین ۸۰ تا ۱۲۰ میلی‌گرم در دسی‌لیتر می‌باشد.

برای هدف مورد نظر یعنی پایین نگه داشتن سطح قند خون در ۳ بیمار زن ۲ قانون (انسولین کوتاه اثر، انسولین طولانی اثر) را برای استفاده در نظر می‌گیریم. هدفمان این است که ببینیم الگوریتم Fuzzy SARSA برای فرآیند پیش‌بینی نوسانات سطح غلظت قند خون با چه درصد خطایی بهترین عملکرد را برای هر بیمار دارد و سیستم خبره چه تغییری در سیاست و سرعت آموزش عامل یادگیرنده برای کشف جواب پیش‌بینی برای هر

الگوریتم SARSA اولین بار توسط Rummery و [Niranjan ۱۶] در سال ۱۹۹۴ کشف شد، که اصلاح شده الگوریتم Q-Learning نامیده شد. نام SARSA در سال ۱۹۹۶ توسط Sutton معرفی شد. روال زیر الگوریتم SARSA را نشان می‌دهد:

- 1: Initialize $Q(s, a)$ arbitrarily
- 2: for all episode do
- 3: Initialize s
- 4: Choose a from s using policy derived from Q (e.g., ϵ -greedy)
- 5: for all step of episode do
- 6: Take action a , observe r, s'
- 7: Choose a' from s' using policy derived from Q (e.g., ϵ -greedy)
- 8: $Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma Q(s', a') - Q(s, a)]$
- 9: $s \leftarrow s'; a \leftarrow a'$
- 10: end for
- 11: end for

قاعده کلی برای پیاده سازی الگوریتم SARSA با روش Fuzzy در انتخاب هر عمل برای هر قانون عامل یادگیرنده می‌باشد، که نتیجه حاصل از این پیاده‌سازی دسترسی به همه عمل‌ها برای هر قانون می‌باشد، و در فرآیند پیاده‌سازی یک رقابت بین عمل‌ها با این روش به وجود می‌آید، با این هدف که عملیات مطابق هر قانون دارای یک کیفیت می‌باشند، سپس عامل از روی درصد کیفیت هر عمل با احتمال بالایی بهترین عمل را هم برای حال و هم برای آینده انتخاب می‌کند، ضمن اینکه خروجی عمل مجزای از نتیجه استنتاجی بین عملیات مختلف در انتخاب محلی هر کدام می‌باشد. در این روش توانایی پایگاه دانش تنها از روی سیاست قوانین محلی نیست، بلکه به توابع ارزیابی و سیاست بهینه سراسری بستگی دارد.

روال زیر الگوریتم پیشنهادی Fuzzy SARSA را نشان می‌دهد:

- 1: Initialize $Q(s, a)$ arbitrarily
- 2: for all episode do
- 3: Initialize \hat{s}_t
- 4: Choose \hat{a}_t from \hat{s}_t by calculating the centre of mass using all \hat{s}_t that match crisp s and \hat{a}_t following ϵ greedy selection policy.
- 5: for all step of episode do
- 6: Take action \hat{a}_t , observe r, \hat{s}_{t+1}
- 7: Choose \hat{a}_{t+1} from \hat{s}_{t+1} using ϵ greedy selection policy for all \hat{s}_{t+1} match s_{t+1}
- 8: $FQ(\hat{s}_{t+1}^i, \hat{a}_{t+1}^i) = FQ(\hat{s}_{t+1}^i, \hat{a}_{t+1}^i) + \alpha \zeta_{(\hat{s}_{t+1}^i, \hat{a}_{t+1}^i)} [r_t + \gamma \sum FQ(\hat{s}_t^j, \hat{a}_t^j) \zeta_{(\hat{s}_{t+1}^i, \hat{a}_{t+1}^i)} - FQ(\hat{s}_{t+1}^i, \hat{a}_{t+1}^i)]$
- 9: $\hat{s}_t \leftarrow \hat{s}_{t+1}; \hat{a}_t \leftarrow \hat{a}_{t+1}$.
- 10: end for
- 11: end for

در الگوریتم SARSA و Fuzzy SARSA هر تکرار از مراحل ۲ تا ۱۰ یک دوره یادگیری را نمایش می‌دهد که یک



مشخص می‌کند. از آنجائیکه عامل یادگیرنده سعی در بیشینه‌سازی پاداش مجموع تابع پاداش دارد برای راهنمایی عامل یادگیرنده جهت دستیابی به هدفش استفاده می‌شود. هدف سیستم در این مقاله نگه داشتن سطح قند خون بین ۸۰ تا ۱۲۰ میلی‌گرم در دسی لیتر است. آنچه موجب عوارض بلند مدت بیماری دیابت است در حقیقت بالا بودن قند خون به مدت طولانی و همین‌طور نوسانات زیاد قند خون در بیمار است، بنابراین تابع پاداش در این سیستم باید طوری پیاده‌سازی شود که عامل یادگیرنده با اطلاعات موجود و اطلاعاتی که خود اکتشاف و استخراج می‌کند، بتواند سیاست بهینه‌ای را برای نگه داشتن سطح قند خون در بازه نرمال ۸۰ تا ۱۲۰ نگه دارد و عامل یادگیرنده با عمل‌های موجود (انسولین کوتاه اثر، انسولین طولانی اثر) بتواند تزریق درستی را برای بیمار با استفاده از انسولین اعمال کند، طوری که سطح قند خون بیمار در محدوده اندازه گیری‌ها دچار نوسانات زیادی نشود که برای بیمار خطرناک باشد. در این سیستم وقتی اطلاعات یک بیمار برای پردازش به بافر فرستاده می‌شود، سیستم براساس اطلاعات موجود بیمار تصمیم می‌گیرد در صورتی که قند خون در بازه عادی یعنی ۸۰ تا ۱۲۰ نباشد چه عملی باید انجام شود. اگر پیش‌بینی و عمل انجام شده برخلاف مقدار اندازه گیری عادی باشد، عامل ماشین یادگیرنده یک پاداش -۱ را دریافت می‌کند. در غیر این صورت عامل یادگیرنده یک پاداش +۱ دریافت می‌کند. به عبارت دیگر پیش‌بینی انجام شده و انتخاب عمل نادرست باعث تولید یک جریمه از عمل قبلی روی جدول جفت وضعیت-عمل می‌شود. (جزئیات پاداش در جدول ۲ نشان داده شده است).

بیمار ایجاد می‌کند، و اینکه عامل یادگیرنده با این الگوریتم می‌تواند سطح قند خون را برای هر بیمار بین قند خون عادی یعنی ۸۰ تا ۱۲۰ میلی‌گرم در دسی لیتر نگه دارد و عملکرد سیستم در هنگام تزریق انسولین کوتاه اثر و طولانی اثر چگونه می‌باشد یعنی اینکه عامل با پیش‌بینی‌های لازم توانسته بهترین قانون (انسولین کوتاه اثر و طولانی اثر) را برای هر بیمار اتخاذ کند.

اکتشاف و استخراج

در الگوریتم‌های یادگیری تقویتی، اکتشاف و استخراج موضوعات مهمی هستند. اکتشاف بدین معنی است که عامل باید چیزی را آزمایش کند که تا به حال انجام نشده است، قبل از اینکه پاداش بیشتری بگیرد. از طرف دیگر استخراج بدین معنی است که عامل عمل‌هایی را ترجیح می‌دهد که قبلاً انتخاب شده‌اند و پاداش گرفته‌اند. استخراج ممکن است یک پاداش قابل انتظار را در یک بازه زمانی تضمین کند ولی اکتشاف ممکن است فرصت‌های بیشتری را برای یافتن حداکثر مقدار مجموع پاداش در یک اجرای طولانی فراهم کند. یک رویکرد محبوب برای چنین روشی سیاست ϵ -greedy می‌باشد. روش ϵ -greedy شامل انتخاب عمل با بهترین مقدار و احتمال $(1-\epsilon)$ است در غیر این صورت با احتمال کوچک ϵ یک عمل به صورت تصادفی انتخاب می‌شود.

تابع پاداش

یک تابع پاداش هدف عامل یادگیرنده را تعریف می‌کند و مقدار عمل بلافاصله را براساس وضعیت درک شده از محیط

جدول ۲- تابع پاداش برای پیش‌بینی نوسانات سطح قند خون

If (GH<80 OR GH>120)
Reward = -1
Otherwise
Reward = 1

آموزش دهد، که ببینیم این الگوریتم برای پیش‌بینی نوسانات سطح قند خون چه تاثیری در سرعت آموزش عامل یادگیرنده و سیاست انتخابی عامل برای پیش‌بینی نوسانات سطح قند خون و بهترین قانون و درصد خطا برای عامل ایجاد می‌کند.

نتایج شبیه‌سازی

شبیه‌سازی سیستم خبره برای پیش‌بینی نوسانات سطح قند خون با الگوریتم یادگیری تقویتی SARSA که با منطق Fuzzy پیاده‌سازی شده است، انجام شده تا یک عامل یادگیرنده را



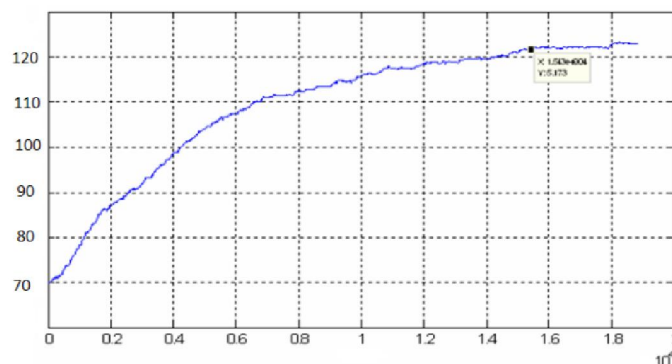
قانونی با بهترین مقدار $Q(s,a)$ انتخاب می شود. الگوریتم Fuzzy SARSA را یک بار با $\alpha = 0.4$ ، $\gamma = 0.9$ و $\epsilon = 0.1$ به کار می گیریم. و یک بار دیگر با $\alpha = 0.4$ ، $\gamma = 0.9$ ، $\epsilon = 0$ به کار می گیریم، تا در ضمن نتایج به دست آمده از این الگوریتم، نقش ϵ را نیز در الگوریتم مورد بررسی قرار دهیم. برنامه پیش بینی نوسانات سطح قند خون که با C# پیاده سازی شده را بر روی کامپیوتر شخصی با اطلاعات موجود برای ۳ بیمار اجرا می کنیم و نتایج شبیه سازی به دست آمده برای عامل یادگیرنده با الگوریتم Fuzzy SARSA را مورد بررسی قرار می دهیم.

نتایج ما برای سیستم خبره با الگوریتم Fuzzy SARSA نشان می دهد که برای پیش بینی نوسانات سطح قند خون با بیش از 10^4 آزمایش انجام شده، عامل در ۸۵ درصد توانسته سطح قند خون را در بازه نرمال یعنی ۸۰ تا ۱۲۰ نگه دارد و در ۱۵ درصد خطایی که عامل داشته مواقعی است که عامل اکتشاف درستی برای تزریق انسولین کوتاه اثر و طولانی اثر نداشته است که این خود برای مطالعات آینده می توان با دادن اطلاعات بیشتر به عامل درصد خطا را کاهش داد (جزئیات کامل نتایج آزمایش های انجام شده برای اجرای سیستم خبره با الگوریتم Fuzzy SARSA در جدول ۳ و شکل ۲ آمده است).

زمانی که الگوریتم آغاز می شود مقادیر جفت های وضعیت-عمل، $Q(s,a)$ می توانند به صورت اختیاری مقداردهی شوند یا مقادیر مرتبط ویژه ای را منتسب کنند. تا اطمینان در انتخاب هر جایگزین را نمایش دهند. در این مقاله همه مقادیر جفت های وضعیت-عمل با اطلاعات اولیه ای که برای هر ۳ بیمار داریم مقداردهی اولیه می شوند. پارامتر اندازه مرحله α که یک تابع مثبت کوچک است بر نرخ یادگیری اثر می گذارد. در این حالت مراحل کوچک تر و کوچک تر می شوند تا زمانی که یادگیری همگرایی مقادیر $Q(s,a)$ را ضمانت کند. مقدار پارامتر نرخ تخفیف γ بین صفر و یک تنظیم می شود. هنگامی که γ به صفر می رسد، عامل بیشتر نزدیک بین است زیرا پاداش فوری دریافت می کند و قوی تر می شود. از طرف دیگر هنگامی که γ به یک می رسد عامل آینده نگرتر می شود و به پاداش هایی که در آینده دریافت خواهد کرد توجه می کند. روش ϵ -greedy برای اکتشاف و استخراج در این مقاله استفاده شده است. اگر $\epsilon = 0.1$ باشد درصد زمانی که استراتژی به صورت تصادفی یکی از دو قانون را مستقل از مقادیر $Q(s,a)$ انتخاب می کند ۱۰ درصد است، این عمل در حقیقت برای اکتشاف در هر گام از دوره یادگیری انجام می شود تا عامل را مجبور کند به اندازه احتمال ϵ یک کشف جدید انجام دهد. در حالی که در ۹۰ درصد دیگر

جدول ۳- جزئیات آزمایش های انجام شده با الگوریتم Fuzzy SARSA

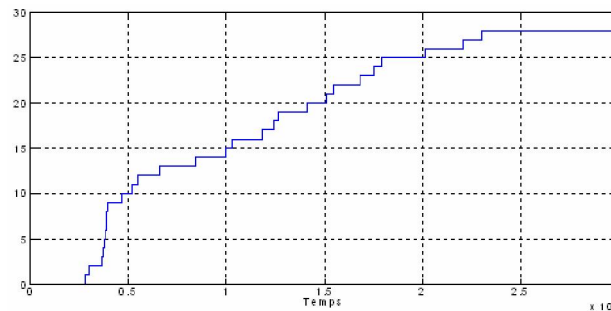
قوانین	درصد انتخاب با $\epsilon = 0.1$	درصد انتخاب با $\epsilon = 0$	پیش بینی نوسانات سطح قند خون
انسولین کوتاه اثر	٪۶۵	٪۵۵	نتایج کلی برای پیش بینی نوسانات
انسولین طولانی اثر	٪۳۵	٪۴۵	
	٪۸۵	٪۸۵	



شکل ۲- درصد کل آزمایش های عامل با الگوریتم Fuzzy SARSA



توانسته سیاستی درست را اتخاذ کند تا جایی که در اواسط دوره یادگیری به بعد نتایج برای عامل تقریباً همگرا شده‌اند که خود نشان دهنده یادگیری درست عامل یادگیرنده می‌باشد. شکل ۳ یکی از آزمایش‌های عامل با الگوریتم Fuzzy SARSA و $\epsilon = 0$ در دوره 10^4 را برای یکی از بیماران نشان می‌دهد. در این آزمایش نوسانات سطح قند خون برای بیمار در ۸۵ درصد دوره یادگیری تقریباً در بازه نرمال سطح قند خون قرار گرفته و سطح قند خون خیلی سریع در محدوده بازه نرمال سطح قند خون یعنی ۸۰ تا ۱۲۰ در حال همگرا شدن است، که خود نشان دهنده یادگیری سریع عامل یادگیرنده در الگوریتم ارائه شده می‌باشد.

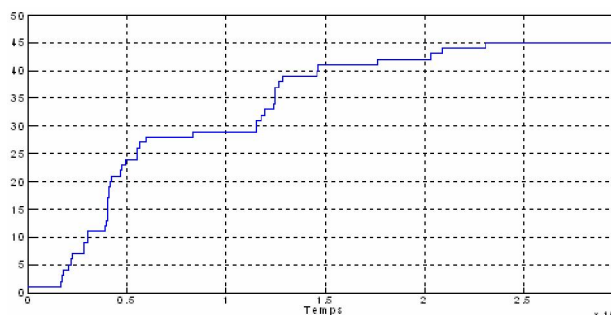


شکل ۳- یکی از آزمایش‌های عامل با الگوریتم Fuzzy SARSA

در اینجا به بررسی آزمایش‌های انجام شده عامل با الگوریتم SARSA Fuzzy می‌پردازیم تا در ادامه نشان دهیم اجرای عامل‌های یادگیری تقویتی با منطق Fuzzy چه تفاوتی و تاثیری در فرآیند یادگیری نسبت به روش‌های دیگر دارند. شکل ۲ درصد کل آزمایش‌های عامل با الگوریتم Fuzzy SARSA را نشان می‌دهد، که عامل در کل فرآیند یادگیری پیش‌بینی نوسانات سطح قند خون توانسته با رویکرد و سیاستی بهینه و کاهش پیچیدگی محاسبات نسبت به روش‌های دیگر، سطح قند خون را در بازه نرمال یعنی ۸۰ تا ۱۲۰ نگه دارد به جز در ۱۵ درصد اوایل دوره یادگیری که عامل هنوز نتوانسته سیاست درستی را اتخاذ کند ولی در ۸۵ درصد دوره یادگیری عامل

ارائه شده می‌باشد، در این آزمایش برخلاف آزمایش قبلی چون عامل یادگیرنده با $\epsilon = 0.1$ یعنی ۱۰ درصد مواقع در حال اکتشاف هست سطح قند خون در مواقع اکتشاف نسبت به آزمایش قبلی کمی تغییر کرده ولی سیاست عامل در دو آزمایش مشابه همدیگر می‌باشد.

شکل ۴ یکی از آزمایش‌های عامل با الگوریتم Fuzzy SARSA و $\epsilon = 0.1$ در دوره 10^4 را برای یکی از بیماران نشان می‌دهد. در این آزمایش نوسانات سطح قند خون برای بیمار در ۸۵ درصد دوره یادگیری تقریباً در بازه نرمال سطح قند خون قرار گرفته و سطح قند خون خیلی سریع در محدوده بازه نرمال سطح قند خون یعنی ۸۰ تا ۱۲۰ در حال همگرا شدن است، که خود نشان دهنده یادگیری سریع عامل یادگیرنده در الگوریتم



شکل ۴- یکی از آزمایش‌های عامل با الگوریتم Fuzzy SARSA



نتیجه گیری

یادگیری تقویتی یک حوزه تحقیقاتی بسیار فعال در یادگیری ماشین است. این مطالعه اثرات به کار گیری الگوریتم معروف یادگیری تقویتی SARSA را که با منطق Fuzzy پیاده سازی شده است، را برای مسئله پیش بینی نوسانات سطح قند خون بررسی می کند. در این مطالعه الگوریتم ارائه شده علاوه بر اینکه با آموزش سریع عامل یادگیرنده تغییرات سریع قند خون را در زمان مورد نیاز پوشش می دهد، در این سیستم سطح قند خون در بیش از ۸۵ درصد زمان تست در محدوده نرمال باقی ماند. از دیگر

مزیت های این سیستم این است که در اینجا انسولین تولید نمی شود بلکه با استفاده از محاسبات ریاضی سیستم مشخص می کنیم چقدر انسولین در زمان های مختلف لازم است و چقدر در روز باید تحویل داده شود، یکی دیگر از مزیت های عامل با الگوریتم ارائه شده کاهش پیچیدگی محاسبات برای پیش بینی نوسانات نسبت به کارهای مرتبط قبلی می باشد. تجربیات حاصل از این مطالعه، پایه خوبی برای مطالعات آینده در به کار گیری تکنیک های یادگیری تقویتی با منطق Fuzzy برای مسائل پیچیده و به خصوص مسائل پیش بینی نوسانات فراهم کند.

مآخذ

- Kroll MH. Biological variation of glucose and insulin includes a deterministic chaotic component. *BioSystems* 1999; 50, 189–201.
- Bergman RN. Pathogenesis and prediction of diabetes mellitus: Lessons from integrative physiology, in: Irving L. Schwartz Lecture, Mount Sinai J. *Medicine* 2002; 60, 280–290.
- Leaning MS, and Boroujerdi MA. A system for compartmental modelling and simulation. *Comput. Methods programs biomed*, 1991; vol. 35.
- Salzsieder E, Albrecht G, Fischer U, Rutscher A, and Thierbach U. Computer-aided systems in the management of type I diabetes: the application of a model-based strategy. *Comput. Methods programs biomed* 1990; vol. 32.
- Tresp V, Moody J, and DeLong WR. Neural modelling of physiological processes. *Computational Learning Theory and Natural Learning Systems* 1994; vol. 2, S. Hanson et al., eds., MIT Press.
- قوچانی س، تهامی ا. مقایسه شبکه عصبی خود بازگشتی المن و شبکه عصبی پرسپترون سه لایه به منظور پیش بینی نوسانات سطح قند خون بیماران مبتلا به نوع یک. دوازدهمین کنفرانس بین المللی انجمن کامپیوتر ایران، دانشگاه شهید بهشتی، ۱۳۸۵.
- Emadi M, Bahrami F, Yazdanpanah MJ, Patla A. Movement Prediction Using an MLP without Internal Feedback. *SMCIEEE*, 2004; Netherland.
- عمادی م، بهرامی ف. پیش بینی حرکت برخاستن از صندلی با استفاده از حرکت یادگیری تقویتی دوازدهمین کنفرانس مهندسی پزشکی ایران، ۱۳۸۴.
- Tesauro G. Temporal difference learning and TD-Gammon. *Communications of the ACM* 1995; 38 (3), 58–67.
- Crites RH, Barto AG. Improving elevator performance using reinforcement learning. In: Touretzky DS, Mozer MC, Hasselmo ME. (Eds.), *Advances in Neural Information Processing Systems: Proceedings of the 1995 Conference*, 1996; MIT Press, Cambridge, MA, pp. 1017–1023.
- Samuel AL. Some studies in machine learning using the game of checkers. *IBM Journal on Research and Development* 1959; 3:211–229.
- Sutton RS. Learning to predict by the method of temporal differences. *Machine Learning* 1988; 3:9–44.
- Weiss G. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, 1999; Cambridge, MA.
- Sutton RS, Barto, AG. *Reinforcement Learning: An Introduction*. The MIT Press, 1999; Cambridge, MA.
- Mahadevan S, Kaelbling LP. The NSF Workshop on Reinforcement Learning: Summary and Observations. *AI Magazine Winter* 1996; 89–97.
- Rummery GA, and Niranjan M. *On-line q-learning using connectionist systems. Technical Report CUED/F-INFENG/TR 166*, Cambridge University Engineering Department 1994.

A NEW FUZZY SARSA ALGORITHM FOR BLOOD GLUCOSE LEVEL PREDICTION IN TYPE 1 DIABETICS

Hossein Farzalivand*¹

1. Young Researchers Club, Izeh Branch, Islamic Azad University, Izeh, Iran

ABSTRACT

Background: One of the serious complications of type 1 diabetes is a sudden increase and drop in blood glucose levels causing risks of anesthesia and coma. Thus, an important step towards the optimal control of the disease is to use intelligent methods with low error rate and available information in order to predict and prevent such complications. In this paper, a combined Fuzzy SARSA algorithm was employed to design an expert system to predict fluctuations in blood glucose levels in patients with type 1 diabetes. The aim of this paper is to propose and implement a new algorithm based on reinforcement learning and fuzzy logic called FSA (Fuzzy SARSA Algorithm) to develop an expert system.

Methods: The medical data used in this article were related to 3 Iranian female patients with type 1 diabetes. The data included parameters such as the type and dose of injected insulin and time interval between the hours of recorded blood glucose level and the one at the beginning of the period. A well-known reinforcement learning algorithm called SARSA was implemented with fuzzy logic. The Fuzzy SARSA algorithm was used in an expert system to predict fluctuations in blood glucose levels in patients with type 1 diabetes.

Results: The results of the expert system with the proposed FSA algorithm for a learning factor included the parameters of blood glucose level at the beginning of the period (mg/dlit), short-acting injected insulin dose in to the body (unit), long-acting injected insulin dose (unit), stress level (unit), physical activity level (unit), used carbs (gr), and the interval between the beginning and end of the period (hour). They indicated that the factor managed to keep blood glucose levels at 85% in a normal range of 80 to 120, and it had 15% error. The error was observed in cases in which there was no proper exploratory factor for short-acting and long-acting injected insulin.

Conclusion: According to the results, it was found that the proposed algorithm for intelligent system, with an approximate reduction of 15% in prediction error, can be used as a good model to predict fluctuations in blood glucose levels in patients with type 1 diabetes in comparison with other conventional methods including the ones based on reinforcement learning. In addition, other prediction methods based on artificial intelligence such as neural networks can be used to predict the results accurately.

Keywords: Reinforcement Learning, Fuzzy SARSA Algorithm, Diabetes, Predicting Fluctuations in Blood Sugar Levels

* Khozestan, Izeh, Zagros Street, Izeh Branch, Islamic Azad University, Young Researchers Club Phone: 09384221895
Email: farzalivand@gmail.com