

# کنترل جمعیت سلول‌های سرطانی در مدل غیرخطی سرطان ملانوما با لحاظ عدم قطعیت با استفاده از الگوریتم یادگیری Q تحت سیاست استدلال مبتنی بر مورد (CBR)

امین نوری<sup>۱</sup>    الناز کلهر<sup>۲</sup>    محمد علی صدرنیا<sup>۳</sup>    سارا صبوری راد<sup>۴</sup>

۱- دانشجوی دکتری - دانشکده مهندسی برق و ریاضیات دانشگاه صنعتی شاهرود - شاهرود - ایران

[Amin\\_noori@ieee.org](mailto:Amin_noori@ieee.org)

۲- دانش‌آموخته کارشناسی ارشد - دانشکده مهندسی برق - دانشگاه صنعتی سجاد - مشهد - ایران

[E.kalhor123@sadjad.ac.ir](mailto:E.kalhor123@sadjad.ac.ir)

۳- دانشیار - دانشکده مهندسی برق و ریاضیات دانشگاه صنعتی شاهرود - شاهرود - ایران

[masadrnia@shahroodut.ac.ir](mailto:masadrnia@shahroodut.ac.ir)

۴- عضو هیأت علمی دانشکده پوست - دانشگاه علوم پزشکی مشهد - مشهد - ایران

[Sabourirads@mums.ac.ir](mailto:Sabourirads@mums.ac.ir)

**چکیده:** سرطان پوست یکی از خطرناک‌ترین سرطان‌هایی است که همه ساله افراد زیادی به آن مبتلا می‌شوند. به همین دلیل تشخیص و درمان سریع این سرطان بسیار برای پزشکان حایز اهمیت می‌باشد، در چند دهه اخیر برای بهبود تشخیص و درمان این بیماری استفاده از روش‌های هوشمند بسیار مورد توجه قرار گرفته است. هدف اصلی در این مقاله، تعیین مقدار بهینه دارو برای از بین بردن سلول‌های سرطانی می‌باشد به گونه‌ای که از تاثیر سوء دارو بر روی سلول‌های سالم جلوگیری شود. از الگوریتم یادگیری Q بدین منظور استفاده شده است. برای انتخاب اعمال، از سیاست استدلال مبتنی بر مورد با نام اختصاری CBR که یک نوع سیاست اکتشافی شتاب داده شده می‌باشد، استفاده گشته است که باعث افزایش سرعت یادگیری و کاهش زمان، برای رسیدن به سیاست بهینه می‌شود. مورد دیگری که در این مقاله لحاظ شده است، تاثیر نیمه عمر دارو برای بدست آوردن اثر دارو در هر لحظه در بدن بیمار می‌باشد. برای اینکه عملکرد روش یادگیری تقویتی در کنترل سلول‌های سرطانی و تعیین میزان بهینه دوز دارو بهتر نشان داده شود، این روش با یکی از روش‌های کنترل بهینه به نام روش همیلتونین و روش تزریق دوز داروی ثابت مقایسه شده است. در نهایت نشان داده شده است مجموع دوز داروی تزریقی به بیمار با استفاده از روش یادگیری تقویتی در مقایسه با حالتی که از روش کنترل بهینه و دوز داروی ثابت برای تمام زمان‌ها استفاده شده است، بسیار کاهش پیدا کرده است و در ضمن جمعیت سلول‌های سرطانی نیز کنترل شده است. با اعمال نویز و عدم قطعیت در پارامترهای سیستم و شرایط اولیه باز هم روش انتخابی قادر به کنترل سلول‌های سرطانی می‌باشد.

**واژه‌های کلیدی:** سرطان ملانوما، الگوریتم یادگیری Q، سیاست استدلال مبتنی بر مورد، اثرات سوء دارو، نیمه عمر دارو، کنترل بهینه

تاریخ ارسال مقاله : ۱۳۹۷/۲/۶

تاریخ پذیرش مشروط مقاله : ۱۳۹۷/۰۷/۰۱

تاریخ پذیرش مقاله : ۱۳۹۷/۸/۳۰

نام نویسنده‌ی مسئول : امین نوری

نشانی نویسنده‌ی مسئول : ایران - مشهد - بلوار جلال آل احمد - جلال آل احمد ۶۲ - دانشگاه صنعتی سجاد - دانشکده‌ی برق و

مهندسی پزشکی

سرطان پوست، ایجاد تغییرات غیر عادی بر روی لایه‌ی بیرونی پوست می‌باشد و از متداول‌ترین و شایع‌ترین نوع سرطان‌ها می‌باشد که حدود ۷۵ درصد از کل سرطان‌ها را در جهان تشکیل می‌دهد. در بین انواع سرطان پوست، ملانوما شناخته‌ترین نوع آن می‌باشد که سرطانی شدید و خطرناک بوده و در صورت عدم تشخیص به موقع می‌تواند باعث مرگ بیمار شود [۱]. یکی از مهم‌ترین عوامل ایجاد این نوع سرطان قرارگیری پوست در برابر اشعه فرابنفش خورشید می‌باشد. سرطان ملانوما در دو نوع بدخیم و خوش‌خیم وجود دارد و بر اثر تغییر در ملانوسیت‌ها (سلول‌های رنگدانه پوست) اتفاق می‌افتد [۲]. دارویی به نام اینترفرون (interferon) برای درمان این نوع سرطان و جلوگیری از افزایش سلول‌های سرطانی در نظر گرفته شده است.

در دو دهه‌ی گذشته، تشخیص ملانوما با استفاده از روش‌های هوشمند برای کمک به کار پزشکان و تشخیص سریع‌تر توسط آن‌ها، بسیار مورد توجه قرار گرفت تا جایی که کارهای بسیار زیادی در این زمینه انجام شده است که جدیدترین آن‌ها به شرح زیر می‌باشند: در مراجع [۳، ۴]، ابتدا پردازش اولیه بر روی تصاویر سرطانی انجام شده است، سپس با استفاده از روش ماشین بردار پشتیبان به تشخیص سرطان پرداخته‌اند. در [۵، ۶] نیز با دقت ۷۶٪ و با استفاده از منطق فازی کار تشخیص را انجام داده‌اند. در [۷]، با ترکیب روش K- نزدیک‌ترین همسایه و شبکه عصبی و در [۸] هم با ترکیب الگوریتم ژنتیک و شبکه عصبی کار تشخیص ملانوما را انجام داده‌اند. در [۹]، تشخیص سرطان پستان با استفاده از از انواع روش‌های هوش مصنوعی انجام گرفته است. با دقت ۹۲٪ توانسته‌اند تصاویر سرطانی و غیر سرطانی را از یکدیگر تفکیک کنند.

در زمینه درمان سرطان ملانوما نیز می‌توان به موارد محدودی اشاره کرد، از جمله: در [۱۰]، کنترل سلول‌های سرطانی در بیماران مبتلا به ملانوما از طریق تنظیم ژن‌ها، توسط روش یادگیری تقویتی دسته‌ای (Batch Reinforcement Learning) انجام گرفته است. در این مقاله تزریق دارو به گونه‌ای انجام گرفته است که باعث تغییر در ساختار ژن‌ها و همچنین نحوه قرارگیری آن‌ها کنار یکدیگر شده است که باعث کاهش سلول‌های سرطانی شده است. نحوه قرار گرفتن ژن‌ها به عنوان حالت در مسئله یادگیری تقویتی در نظر گرفته شده است. در تابع پاداش نیز، یک مقدار ثابت به عنوان جرمه برای حالت‌های نامطلوب در نظر گرفته شده است. این امر موجب قرار گرفتن ژن‌ها در حالت مطلوب خواهد شد. در این مقاله، برای بهتر نشان دادن عملکرد روش یادگیری تقویتی، این روش با روش احتمالاتی بولین مقایسه شده است. در نهایت نشان داده شده است که در حالت استفاده از روش یادگیری تقویتی دسته‌ای درصد قرار گرفتن ژن‌ها در حالت مطلوب افزایش و در حالت نامطلوب

کاهش پیدا کرده است. در [۱۱]، با استفاده از الگوریتم ژنتیک، بهینه‌ترین میزان دوز دارو برای بیمار مبتلا به سرطان ملانوما B16F10 تعیین شده است و سعی بر این شده است که تا حد زیادی از اثرات زیان‌بار دارو بر روی سلول‌های سالم جلوگیری شود. دلیل استفاده از این روش وجود عدم قطعیت و نویز خیلی زیاد در محیط می‌باشد که این روش قادر به پیش‌بینی رفتار مدل و کنترل سلول‌های سرطانی با دقت ۷۵٪ بوده است. از معایب روش الگوریتم ژنتیک در تعیین میزان دوز دارو می‌توان به غیرقابل استفاده بودن آن در حالت بر خط (online) اشاره کرد و همچنین این الگوریتم در زمان یادگیری بسیار زمان‌بر می‌باشد. در [۱۲]، از انواع بازدارنده‌ها برای کنترل و درمان سرطان ملانوما استفاده شده است. در نهایت نشان داده شده است که استفاده از مهارکننده‌ی mTOR یک درمان بالقوه می‌باشد و باعث کاهش سلول‌های سرطانی می‌شود.

در زیر به بررسی مقالاتی که درمان و کنترل دیگر سرطان‌ها را با استفاده از روش یادگیری تقویتی انجام داده‌اند، پرداخته شده است. در روش یادگیری تقویتی، عامل و محیط نقش اصلی را بازی می‌کنند و یادگیری بر اساس برهم‌کنش بین آن‌ها انجام می‌گیرد. عامل در هر لحظه با استفاده از اطلاعاتی که از محیط می‌گیرد، خود را به‌روز می‌کند. از جمله مقالاتی که به این طریق کنترل سلول‌های سرطانی را انجام داده‌اند، می‌توان به موارد ذیل اشاره نمود:

در مراجع [۱۳-۲۰]، کنترل سلول‌های سرطانی و بیماری‌های مختلف با استفاده از روش یادگیری تقویتی انجام گرفته است. در مراجع [۱۳، ۱۵، ۱۷، ۱۸، ۱۹]، از روش یادگیری Q که یکی از روش‌های حل مسئله یادگیری تقویتی می‌باشد، استفاده شده است. در مرجع [۱۵]، کنترل میزان دوز دارو در بیماران مبتلا به هیپاتیت انجام گرفته است. در این مقاله برورسانی ارزش جفت حالت و عمل توسط معادله بلمن انجام گرفته است. برای کاهش هر چه بیشتر ویروس‌های آزاد پادشده‌ی بر اساس میزان کاهش این ویروس‌ها تعریف شده است. در مراجع [۱۳، ۱۷، ۱۹]، کنترل سلول‌های سرطانی در بیماران مبتلا به سرطان انجام گرفته است که در [۱۳، ۱۹]، برای نشان دادن خاصیت تطبیق‌پذیری روش یادگیری تقویتی با محیط و عدم وابستگی آن به مدل ریاضی، عدم قطعیت در شرایط اولیه و پارامترهای سیستم لحاظ شده است و کنترل سلول‌های سرطانی در دیگر بیماران مبتلا به سرطان انجام گرفته است. تزریق دوز داروی ثابت در تمام گام‌های زمانی از دیگر کارهایی می‌باشد که در این مقالات انجام گرفته است. در نهایت نشان داده شده است که تزریق دوز دارو توسط روش یادگیری Q نسبت به تزریق دوز داروی ثابت در کل مدت درمان کمتر بوده که این امر موجب کاهش اثرات زیان‌بار دارو خواهد شد. در مرجع [۱۷] نیز، با استفاده از داده عددی و بدون نیاز به مدل ریاضی کنترل سلول‌های سرطانی انجام گرفته است و نشان داده شده است که روش یادگیری تقویتی قابلیت تطبیق‌پذیری با محیط را دارا

می‌باشد. در مرجع [۱۴]، کنترل میزان گلوکز خون در بیماران مبتلا به دیابت مورد بررسی قرار گرفته است. در مرجع [۱۸]، کنترل سلول‌های سرطانی در بیماران مبتلا به سرطان ریه مورد بررسی قرار گرفته است. در هر دو مقاله روش یادگیری تقویتی با روش تزریق دوز داروی ثابت مقایسه شده است. در نهایت نشان داده شده است که میزان تزریق دارو توسط روش یادگیری تقویتی کمتر از میزان تزریق دارو به مقدار ثابت می‌باشد و در این حالت اثرات زیان‌بار دارو نیز کاهش خواهد یافت. از دیگر سیستم‌هایی که از این روش برای کنترل سیستم استفاده کرده‌اند می‌توان به مرجع [۲۰] اشاره کرد که از این روش به منظور بهبود عملکرد پایدار ساز سیستم قدرت سه‌بند استفاده شده است. به طور کلی از جمله مزایای روش یادگیری تقویتی که در مقالات ذکر شده به آن‌ها اشاره شده است، می‌توان به عدم وابستگی این روش به مدل ریاضی اشاره کرد و همچنین با وجود نویز و اغتشاش محیطی قابلیت کنترل سیستم‌های مختلف را دارا می‌باشد. از معایب این روش نیز می‌توان به زمان‌بر بودن فرآیند یادگیری اشاره نمود.

از دیگر روش‌هایی که برای کنترل بیماری‌های مختلف مورد استفاده قرار گرفته است، روش استدلال مبتنی بر مورد (CBR) می‌باشد. در این روش اگر یک مسئله حل شود و حل آن به هر نتیجه‌ای ختم شود به منظور حل مسئله مشابه در آینده نگه داشته می‌شود که اگر این حل با شکست روبه‌رو شود به منظور جلوگیری از اشتباه در آینده به خاطر سپرده می‌شود. در مراجع [۲۱-۲۵] از این روش استفاده شده است. در مرجع [۲۱]، تشخیص بیماری دیابت با استفاده از روش فازی انجام گرفته است. برای بررسی و ارزیابی ویژگی‌ها در تشخیص بیماری از روش استدلال مبتنی بر مورد استفاده شده است. از پنج طبقه‌بند دیگر نیز به منظور تشخیص بیماری دیابت استفاده شده است. در نهایت نشان داده شده است که در حالت استفاده از روش فازی و ترکیب آن با روش استدلال مبتنی بر مورد، تشخیص بیماری دیابت با دقت ۹۷٫۶۷٪ انجام گرفته است. این در حالی می‌باشد که در حالت استفاده از دیگر روش‌ها دقت تشخیص بیماری پایین می‌باشد. در مرجع [۲۲]، به مزایا و عملکرد روش استدلال مبتنی بر مورد و همچنین بررسی چگونگی استفاده از روش‌های هوشمند در موضوعات پزشکی اشاره شده است. چهار سیستم پزشکی در این مقاله مورد بررسی قرار گرفته است. هدف در این سیستم‌ها کنترل سلول‌های بنیادی، کنترل میزان انسولین، کنترل فرآیند دیالیز و کنترل میزان استرس می‌باشد. در کنترل تمام این سیستم‌ها استفاده از روش استدلال مبتنی بر مورد در کنار روش‌های هوشمند باعث افزایش سرعت و دقت شده است. در مرجع [۲۳]، روش استدلال مبتنی بر مورد در جنبه‌های پزشکی از جمله تشخیص، پیش‌آگاهی و درمان سرطان معده مورد بررسی قرار گرفته است. این مقاله یک تحقیق در مورد اطلاعات بالینی ارائه داده است و شبیه‌سازی بر روی روش

پیشنهادی انجام نگرفته است. اما در نهایت با تحقیقاتی که انجام گرفته است نشان داده شده است که استفاده از روش استدلال مبتنی بر مورد در بدست‌آوردن دوز داروی بهینه و درمان سرطان بسیار موثر می‌باشد. در مرجع [۲۴]، تفاوت‌های بین روش استدلال مبتنی بر مورد و الگوریتم‌های کلاسیک در حل مسائل پزشکی مورد بررسی قرار گرفته است. در مرجع [۲۵]، از روش استدلال مبتنی بر مورد برای شناسایی بیماران واجد شرایط در آزمایش‌های بالینی استفاده شده است. داده‌های مربوط به پرونده سلامت الکترونیکی بیماران برای کشف بیماران واجد شرایط مورد بررسی قرار گرفته است. در نهایت نشان داده شده است که روش استدلال مبتنی بر مورد با دقت ۹۵٪ توانسته است افراد واجد شرایط را شناسایی کند. از مزایای این روش می‌توان به افزایش دقت و سرعت در حل مسائل فعلی، به دلیل استفاده از داده‌های گذشته اشاره کرد.

یکی دیگر از روش‌هایی که برای کنترل سرطان مورد استفاده قرار می‌گیرد، روش‌های کنترل کلاسیک (کنترل بهینه) می‌باشد. در این گونه روش‌ها کنترل سلول‌های سرطانی وابسته به مدل ریاضی از بیمار مبتلا به سرطان می‌باشد. در مراجع [۲۶-۳۰]، کنترل سلول‌های سرطانی با استفاده از روش‌های کنترل کلاسیک (کنترل بهینه) انجام گرفته است. در مرجع [۲۶]، از مدل ریاضی دارای تاخیر از بیمار مبتلا به سرطان استفاده شده است. در نهایت نشان داده شده است که روش پیشنهادی این مقاله توانسته است سلول‌های سرطانی را کاهش داده و به صفر برساند. در مرجع [۲۷] نیز، برای ارتقاء شیمی‌درمانی از یکی از روش‌های کنترل بهینه به نام "اصل حداکثری پوینتراگین" استفاده شده است. در نهایت نشان داده شده است که روش پیشنهادی سلول‌های سرطانی را کاهش داده اما قادر به حذف کامل آنها نبوده است. این امر از معایب روش‌های کنترل بهینه محسوب می‌شود. در مرجع [۲۸]، از روش ذکر شده برای کاهش سلول‌های سرطانی در بیماران مبتلا به سرطان خون استفاده شده است. این روش با روش تزریق دوز داروی ثابت مقایسه شده است. در نهایت نشان داده شده است که روش کنترل بهینه با تزریق دوز داروی کمتر توانسته است سلول‌های سرطانی را کاهش دهد، اما هر دو روش قادر به حذف کامل سلول‌های سرطانی نبوده‌اند. در مرجع [۲۹]، از روش کنترل بهینه چند هدفه برای کاهش هم‌زمان سلول‌های سرطانی و میزان دوز داروی تزریقی استفاده شده است. تزریق دوز دارو تا زمانی که سلول‌های سرطانی به صفر نرسیده‌اند انجام می‌گیرد و با حذف کامل سلول‌های سرطانی تزریق دارو متوقف شده است. در مرجع [۳۰]، رژیم دارویی بهینه برای بیماران مبتلا به سرطان پستان مرحله ۳ به‌گونه‌ای تنظیم شده است که اندازه تومور به اندازه مطلوب برای انجام عمل جراحی و برداشتن تومور برسد. هدف اصلی در مقالات ذکر شده تعیین بهینه‌ترین میزان دوز دارو برای کاهش سلول‌های سرطانی می‌باشد که در تابع هزینه تعریف شده، این امر

قابل مشاهده می‌باشد. از جمله معایبی که در تمام روش‌های کنترل کلاسیک وجود دارد، وابستگی آن‌ها به مدل ریاضی می‌باشد و تمامی این روش‌ها بر مبنای خطی‌سازی می‌باشند که رفتار سیستم در این حالت با رفتار سیستم اصلی تفاوت خواهد داشت. بنابراین دوز داروی پیشنهاد شده توسط چنین روش‌هایی از قابلیت اطمینان بالا برخوردار نمی‌باشد. در این مقاله روش‌های یادگیری Q و استدلال مبتنی بر مورد با جزئیات بیشتر توضیح داده شده‌اند. اما نوآوری این مقاله ترکیب این دو روش با یکدیگر می‌باشد که در هیچ یک از کارهای گذشتگان انجام نگرفته است. انتخاب اعمال توسط روش استدلال مبتنی بر مورد باعث افزایش در سرعت یادگیری می‌شود که ترکیب آن با روش یادگیری Q علاوه بر این امر، کنترل سلول‌های سرطانی با تزریق کمترین میزان دوز دارو را به همراه خواهد داشت. برای اینکه مزیت روش پیشنهادی این مقاله در کنترل سلول‌های سرطانی برای بیماران مبتلا به ملانوما بهتر نشان داده شود، این روش با روش مقاله [۳۱] مقایسه شده است. در این مقاله از روش کنترل بهینه که یکی از روش‌های کنترل کلاسیک می‌باشد استفاده شده است. از معایب روش‌های کنترل کلاسیک می‌توان به وابستگی آن‌ها به مدل ریاضی اشاره کرد که اکثر این مدل‌ها غیرخطی می‌باشند. و برای استفاده از چنین روش‌هایی می‌بایست سیستم حتماً خطی گردد. این امر نیز از معایب روش‌های کنترل کلاسیک محسوب می‌شود چراکه تمامی سیستم‌های پزشکی غیرخطی و پویا می‌باشند و با تبدیل آن‌ها به سیستم خطی ممکن است رفتار سیستم تحت تاثیر قرار گیرد. در بعضی مواقع نیز این روش‌ها قادر به کنترل چنین سیستم‌هایی نخواهند بود. بنابراین روش‌های کنترل کلاسیک دارای خطای بالا و دقت پایین می‌باشند. اما روش یادگیری تقویتی قابلیت کنترل سیستم‌های غیرخطی با وجود نویز و اغتشاش محیطی را دارا می‌باشد. با وجود اینکه در این مقاله از مدل ریاضی بیمار مبتلا به سرطان ملانوما استفاده شده است اما این روش کاملاً بی‌نیاز به مدل محیط بوده و قابلیت پیاده‌سازی بر روی مدل واقعی را نیز دارا می‌باشد. روش یادگیری تقویتی دارای قابلیت اطمینان بالاتری نسبت به روش‌های کنترل کلاسیک می‌باشد. استفاده از این روش در کنترل سلول‌های سرطانی با استفاده از مدل ریاضی، به پزشکان کمک می‌کند که با سعی و خطای کمتری کار درمان بیمار را انجام دهند. ادامه مقاله به شرح ذیل می‌باشد. در بخش دوم مقاله در رابطه با مدل ریاضی سیستم و معادلات حاکم بر سیستم صحبت شده است. در بخش سوم در مورد یادگیری تقویتی و کنترل ملانوما با استفاده از الگوریتم یادگیری Q تحت سیاست CBR توضیحاتی ارائه شده است و در بخش چهارم نیز شبیه‌سازی و مقایسه روش پیشنهادی با روش کنترل بهینه و نتایج حاصل از شبیه‌سازی مطرح گشته است.

## ۲- مدل ریاضی

مدل‌های ریاضی زیادی برای توضیح دینامیک ملانوما وجود دارد که می‌توان به مراجع [۳۲-۳۷] اشاره کرد. مدل ریاضی استفاده شده در این مقاله توسط کاستیگلیون و پیکلی ارائه شده است [۳۸-۳۹]. یکی از مزایای این مدل ریاضی نسبت به سایر مدل‌ها، در نظر گرفتن معادله دیفرانسیل جدا برای رشد تومورهای سرطانی می‌باشد. در این مدل از پنج معادله دیفرانسیل غیرخطی استفاده شده است که در معادلات (۱) تا (۵) نمایش داده شده‌اند [۳۱]:

$$\dot{H} = a_0 + b_0 DH \left(1 - \frac{H}{n_0}\right) - c_0 H \quad (1)$$

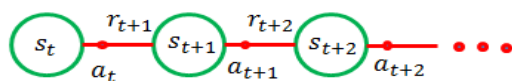
$$\dot{C} = a_1 + b_1 I(M + D)C \left(1 - \frac{C}{n_1}\right) - c_1 C \quad (2)$$

$$\dot{M} = b_2 M \left(1 - \frac{M}{n_2}\right) - d_2 MC \quad (3)$$

$$\dot{D} = -d_3 DC + u \quad (4)$$

$$\dot{I} = b_4 DH - e_4 IC - c_4 I \quad (5)$$

در معادلات فوق، H نشان دهنده جمعیت سلول‌های ایمنی بدن (یاری‌گر CD4) که به طور ذاتی در بدن هر انسانی وجود دارند، می‌باشد. C نشان‌دهنده جمعیت نوع دیگری از سلول‌های ایمنی بدن مرسوم به سرکوبگر CD8 می‌باشد که با ورود سلول‌های سرطانی به بدن انسان سلول‌های ایمنی تبدیل به CD8 شده و با سلول‌های سرطانی مقابله می‌کنند. M جمعیت سلول‌های سرطانی و D و I (Interleukin-2 (IL-2)) نیز به ترتیب سلول‌های دندریتیک و انتقال‌دهنده پیام‌های بین سلولی را نشان می‌دهند. معادله (۱)، نمایش‌گر تغییرات سلول‌های ایمنی بدن (CD4) می‌باشد. معادله (۲)، مربوط به تغییرات سلول‌های ایمنی کشنده بدن می‌باشد که این سلول‌ها با ورود ویروس و یا سلول سرطانی به بدن توسط CD4ها تولید و تا زمان نابودی کامل سلول‌های سرطانی در بدن وجود دارند که جمله دوم در معادله (۲) نشان‌دهنده این موضوع می‌باشد. معادله (۳)، نمایش تغییرات غلظت تومور می‌باشد. جمله اول در این معادله، تکثیر سلول‌های سرطانی را نشان می‌دهد و در جمله دوم نیز از بین رفتن سلول‌های سرطانی به کمک سلول‌های CD8 قابل مشاهده می‌باشد. معادله (۴)، تغییرات سلول‌های دندریتیک را نمایش می‌دهد که این سلول‌ها مستقیماً با تزریق دارو (متغیر u) در بدن تولید شده و با ترشح IL-2 باعث فعال‌سازی CD4ها به طور مستقیم و تبدیل آن‌ها به CD8 فعال می‌شوند. معادله (۵)، نمایش دهنده تغییرات انتقال‌دهنده پیام‌های بین سلولی می‌باشد که در جمله اول این رابطه با ترشح سلول‌های دندریتیک و با توجه به تعداد سلول‌های CD4، IL-2 در خون ترشح می‌شود. در جمله دوم نیز، تعدادی از IL-2های ترشح‌شده با فعال کردن CD8ها برای مبارزه با سلول‌های سرطانی عمل کرده و از بین



شکل(۱): زنجیره حالت و عمل

تخمین تابع ارزش تحت سیاست  $\pi$  که با  $V^\pi$  نمایش داده می‌شود، از رابطه (۶) بدست می‌آید [41].

$$V^\pi(s) = E_n \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s \right\} \quad (۶)$$

در رابطه فوق،  $\gamma$  نرخ فراموشی و مقداری بین صفر و یک می‌باشد.  $k$  گام زمانی می‌باشد. در این رابطه، ارزش هر حالت تحت سیاست  $\pi$ ، بر اساس محاسبه امید ریاضی برای مجموع حاصل ضرب نرخ فراموشی و مقدار پاداش‌های دریافتی از لحظه فعلی ( $s_t$ ) تا انتهای مسیر می‌باشد که با توجه به رابطه (۶)، تاثیر پاداش‌های دریافتی در ارزش حالت فعلی ( $s_t$ ) در هر گام زمانی با ضرب شدن در  $\gamma^k$  خواهد بود.

### ۲-۲- الگوریتم یادگیری Q

یکی از روش‌های یادگیری تقویتی، یادگیری تفاوت‌گذرا می‌باشد. در این روش به‌روزرسانی ارزش جفت حالت و عمل در همان لحظه انجام می‌شود و نیازی به اتمام فرایند جهت ارزش‌گذاری زوج‌های حالت و عمل نمی‌باشد [۴۰]. یکی از الگوریتم‌های مرسوم در روش یادگیری تفاوت‌گذرا، یادگیری Q می‌باشد که در این الگوریتم به‌روزرسانی ارزش جفت حالت و عمل ( $s_t, a_t$ ) طبق معادله (۷) محاسبه می‌شود.

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \quad (۷)$$

در رابطه (۷)،  $\alpha$  نرخ یادگیری می‌باشد که برابر با ۰٫۳ در نظر گرفته شده است و  $\gamma$  که نرخ فراموشی می‌باشد برابر با ۰٫۵ در نظر گرفته شده است. لازم به ذکر می‌باشد که این پارامترها با سعی و خطا بدست آمده‌اند و سعی شده است که بهترین مقدار برای آن‌ها انتخاب شود. فرایند یادگیری در RL بسیار زمان‌بر می‌باشد و ابعاد مسئله بزرگ می‌باشد. در نظر گرفتن فرایند بهینه‌سازی در هر گام زمانی برای پیدا کردن این پارامترها بسیار وقت‌گیر خواهد بود و حجم محاسباتی بالایی دارد.

### ۳-۳- معرفی سیاست استدلال مبتنی بر

#### مورد (CBR) (Case Based Reasoning)

سیاست استدلال مبتنی بر مورد با نام اختصاری CBR یک سیاست اکتشافی شتاب داده شده می‌باشد که برای انتخاب اعمال در الگوریتم‌های یادگیری از آن استفاده می‌شود. بر اساس این

می‌روند. نکته مهم در رابطه با متغیرها نرمال بودن آن‌ها می‌باشد. در جدول ۱ ثوابت سیستم و مقدار آن‌ها آورده شده است.

جدول(۱): معرفی پارامترها و مقادیر در نظر گرفته شده [38,39] (C=cell, h=hour, mm=millimeter)

پارامتر	تعریف پارامتر	مقدار	واحد
$a_0$	نرخ تولید CD4-Tcell	$10^{-4}$	$ch^{-1}mm^{-3}$
$c_0$	نرخ از بین رفتن CD4-Tcell	0.005	$h^{-1}$
$a_1$	نرخ تولید CD8-Tcell	$10^{-4}$	$ch^{-1}mm^{-3}$
$c_1$	نرخ از بین رفتن CD8-Tcell	0.005	$h^{-1}$
$b_2$	نصف حد ثابت تومور	0.02	$h^{-1}$
$n_2$	ظرفیت انتقال از تومور	1	$c^{-1}h^{-1}mm^{-3}$
$b_4$	تولید IL-2 به وسیله CD8-Tcell	$10^{-2}$	$c^{-1}h^{-1}mm^{-3}$
$c_4$	نرخ نابودی IL-2	$10^{-2}$	$h^{-1}mm^{-3}$
$b_6$	نرخ تکثیر CD4-Tcell	$10^{-4}$	$c^{-1}h^{-1}mm^{-3}$
$n_6$	ظرفیت انتقال از CD4-Tcell	1	$c mm^{-3}$
$b_1$	نرخ تکثیر CD8-Tcell	$10^{-2}$	$c^{-1}h^{-1}mm^{-3}$
$n_1$	ظرفیت انتقال از CD8-Tcell	1	$c mm^{-3}$
$d_2$	تاثیر CD8-Tcell	0.1	$c^{-1}h^{-1}mm^{-3}$
$d_3$	از بین رفتن CD8-Tcell	0.1	$c^{-1}h^{-1}mm^{-3}$
$e_4$	IL-2 جذب شده توسط CD8-Tcell ها	$10^{-2}$	$c^{-1}h^{-1}mm^{-3}$

### ۳-۳- کنترل سرطان ملانوما با استفاده از یادگیری

#### تقویتی (Reinforcement Learning) (RL)

روشی که در این مقاله استفاده شده است، استفاده از الگوریتم یادگیری Q برای از بین بردن سلول‌های سرطانی می‌باشد.

#### ۱-۳- یادگیری تقویتی

در روش یادگیری تقویتی عامل و محیط نقش اصلی را بازی می‌کنند و در واقع یادگیری، برهم‌کنش بین آنها می‌باشد که در آن عامل با استفاده از اطلاعاتی که هر لحظه از محیط می‌گیرد خود را به‌روز می‌کند. یکی از مزایای یادگیری تقویتی بی‌نیاز بودن به مدل ریاضی محیط می‌باشد. در این روش به عامل گفته نمی‌شود که عمل خوب کدام است بلکه خود عامل با سعی و خطای فراوان و با دریافت پاداش و جریمه از محیط، طی تکرار بالا یاد می‌گیرد بهترین عمل را انتخاب کند و سعی بر این دارد که به سیاست بهینه برسد. سه بخش دیگر یادگیری تقویتی عمل، پاداش و حالت می‌باشد [40]. در هر لحظه عامل حالت محیط ( $s_t$ ) را دریافت می‌کند و عمل ( $a_t$ ) را انجام می‌دهد و به حالت بعدی ( $s_{t+1}$ ) می‌رود و پاداش ( $r_{t+1}$ ) را دریافت می‌کند. یک زنجیره حالت و عمل نوعی در شکل (۱) قابل مشاهده می‌باشد [40].

سیاست، اگر یک مسئله حل شود و حل آن به هر نتیجه‌ای ختم شود به منظور حل مسئله مشابه در آینده نگه داشته می‌شود که اگر این حل با شکست روبرو شود به منظور جلوگیری از اشتباه در این سیاست از یک تابع اکتشافی با نام  $H(s_t, a_t)$  برای انتخاب اعمال استفاده می‌شود که در رابطه‌ی (۸) نمایش داده شده است.

$$H(s, a) = \begin{cases} \max_i \hat{Q}(s, i) - \hat{Q}(s, a) + \eta & a = \pi^H(s) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

در رابطه فوق،  $\eta$  یک مقدار حقیقی کوچک و معمولاً یک می‌باشد و باعث وزن‌دهی به عملی که ارزش بیشتر دارد، می‌شود.  $\pi^H(s)$  سیاستی است که بر اساس آن انتخاب عمل صورت گرفته است که با استفاده از رابطه (۹) محاسبه می‌شود [۴۲].

$$\pi(s) = \begin{cases} \arg_a \max [\hat{Q}(s, a) + \xi H(s, a)^\beta] & q < p \\ a_{\text{random}} & \text{otherwise} \end{cases} \quad (9)$$

در رابطه (۹)،  $\beta$  و  $\xi$  پارامترهای کنترلی هستند که مقدار آن‌ها یک در نظر گرفته شده است.  $q$  یک عدد تصادفی بین صفر و یک می‌باشد و با استفاده از توزیع نرمال بدست می‌آید.  $p$  عددی برای توازن بین جستجو و بهره‌برداری می‌باشد و بین صفر و یک انتخاب می‌شود که هر چه بزرگتر و به یک نزدیک‌تر باشد میزان بهره‌برداری بیشتر خواهد بود و در غیر این صورت هر چه  $p$  کوچکتر باشد بهره‌برداری کمتر و جستجو بیشتر خواهد بود [۴۲].

### ۳-۴- تطبیق الگوریتم یادگیری Q با درمان سرطان ملانوما

هدف این مقاله تنظیم میزان دوز دارو برای کاهش سلول‌های سرطانی در بیماران مبتلا به سرطان ملانوما می‌باشد. محیط در نظر گرفته شده مدل غیرخطی از یک بیمار مبتلا به سرطان ملانوماست. با جستجو در محیط و بهره‌برداری از اطلاعات یادگرفته شده عامل یاد می‌گیرد چه میزان دارویی تزریق کند تا طی تکرار بالا هم سلول‌های سرطانی کاهش پیدا کنند و هم از تاثیر سوء دارو بر روی سلول‌های سالم جلوگیری شود. برای این منظور از الگوریتم یادگیری Q استفاده شده است. روند اجرای الگوریتم یادگیری Q به ترتیب مراحل زیر می‌باشد:

۱. تنظیم مقادیر اولیه:  $H, C, M, D$  و  $I$

۲. تنظیم مقادیر اولیه:  $\gamma, \alpha, p, \xi$

۳. تنظیم جداول  $Q$  و  $H$  بر اساس تعداد حالت‌ها و عمل‌ها

۴. تا زمانی که جمعیت سلول‌های سرطانی بیشتر از  $0.0001$  می‌باشد، مراحل زیر را تکرار کن:

۱،۴. مشاهده وضعیت سلول‌های تومور CD4 ها در لحظه ( $s_t$ )

۲،۴. انتخاب عمل  $a_t$  با استفاده از سیاست CBR

آینده به خاطر سپرده می‌شود. مزیت انتخاب اعمال توسط این سیاست، افزایش سرعت یادگیری و کاهش زمان برای رسیدن به سیاست بهینه و دریافت ماکزیمم مقدار پاداش می‌باشد [۴۲].

۳،۴. اعمال عمل انتخاب شده بر روی مدل غیرخطی از بیمار مبتلا به ملانوما

۴،۴. تاثیر نیمه عمر دارو و مشاهده حالت بعدی بیمار ( $S_{t+1}$ ) بعد از گذشت ۱۰ ساعت از تاثیر دارو و دریافت پاداش ( $r_{t+1}$ )

۵،۴. بروزرسانی جدول  $Q$  با استفاده از رابطه (۱۰)

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \quad (10)$$

۶،۴. در نظر گرفتن حالت جدید برای  $H, C, M, D$  و  $I$  به جای حالت قبلی

### ۳-۴-۱- تعریف پارامترهای یادگیری تقویتی

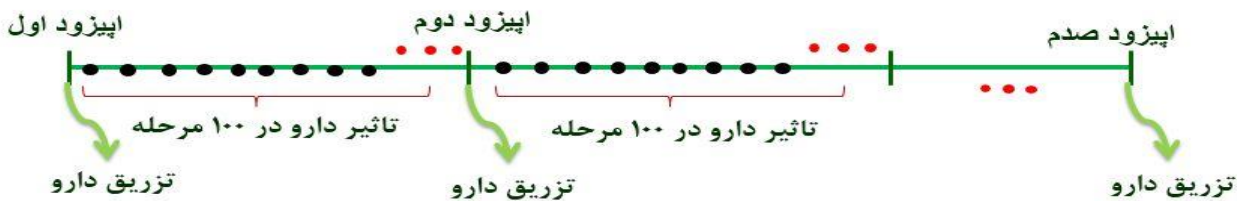
در این مقاله حالت‌ها دو بعدی در نظر گرفته شده‌اند و شامل سلول‌های ایمنی بدن (CD4ها) و سلول‌های تومور می‌باشد. در مدل غیرخطی بیمار مبتلا به سرطان ملانوما  $u$ ، سیگنال کنترلی و عمل می‌باشد و میزان داروی تزریقی در نظر گرفته شده است و مقدار آن بین صفر و یک می‌باشد که با بازه‌های  $0.1$  از هم جدا شده‌اند. تابع پاداش در نظر گرفته شده در این مقاله بر اساس تغییرات میزان غلظت تومور و اختلاف میزان داروی تزریقی در لحظه فعلی و لحظه قبل برای بهینه کردن میزان دارو می‌باشد. تابع پاداش مورد نظر در رابطه (۱۱) نشان داده شده است [۴۱].

$$\text{Reward} = \log\left(\frac{M_{\text{old}}}{M_{\text{new}}}\right) - (dose_{\text{new}} - dose_{\text{old}}) \quad (11)$$

در رابطه فوق،  $M_{\text{old}}$  غلظت تومور در ده ساعت قبل،  $M_{\text{new}}$  غلظت تومور در لحظه فعلی،  $dose_{\text{old}}$  داروی تزریقی در ده ساعت قبل و  $dose_{\text{new}}$  مقدار داروی تزریقی در لحظه فعلی می‌باشند. سیاستی که در این مقاله برای انتخاب عمل استفاده شده است، بر اساس سیاست CBR می‌باشد که در قسمت (۳-۳) عملکرد این سیاست و انتخاب اعمال توسط آن توضیح داده شده است.

### ۴- تاثیر نیمه عمر دارو در بدن

داروهای مختلفی برای درمان و کنترل سلول‌های سرطانی در بیماران مبتلا به ملانوما وجود دارد. در این مقاله برای کنترل سلول‌های سرطانی از داروی اینترفرون استفاده شده است. نیمه عمر این دارو ۱۰ ساعت می‌باشد با توجه به این مورد تزریق دارو در هر ۱۰ ساعت انجام گرفته است. تعداد گام‌های زمانی برای مشاهده تاثیر دارو در بدن بیمار ۱۰۰ گام در نظر گرفته شده است که در واقع در هر شش دقیقه میزان داروی باقیمانده در بدن محاسبه شده است. در شکل (۲)، این گام‌های زمانی قابل مشاهده می‌باشند.



شکل(۲): گام زمانی تأثیر دارو

با توجه به شکل(۴)، هر چقدر تعداد گام‌های مشاهده تأثیر دارو بیشتر شود، سلول‌های سرطانی با سرعت کمتری کاهش پیدا کرده و به صفر می‌رسند. با افزایش تعداد گام‌ها میزان دوز داروی تزریقی به بدن بیمار بین بازه صفر تا ده ساعت کاهش پیدا کرده، بنابراین سلول‌های سرطانی با سرعت کمتری کاهش پیدا می‌کنند. دلیل انتخاب ۱۰۰ گام در این مقاله نیز، بالا رفتن حجم محاسبات، اگر بیشتر انتخاب شود و پایین آمدن دقت، اگر کمتر انتخاب شود، می‌باشد. رابطه(۱۲)، غلظت پلاسمای دارو را نشان می‌دهد که مقدار اولیه آن برابر با صفر خواهد بود و طبق رابطه(۱۳) محاسبه می‌شود [۴۳].

$$C_n(t) = \frac{D}{V} \cdot \frac{(1 - e^{-nk_e \tau})}{(1 - e^{-k_e \tau})} e^{-k_e t} \quad (12)$$

$$C_0 = \frac{D}{V} \quad (13)$$

در این روابط،  $t$  زمان نمونه برداری،  $n$  تعداد گام‌های زمانی برای مشاهده تأثیر دارو (در این مقاله ۱۰۰ گام زمانی در نظر گرفته شده است) و  $\tau$  بازه‌ی تزریق دارو می‌باشد که در این مقاله با توجه به نیمه عمر داروی اینترفرون هر ۱۰ ساعت لحاظ شده است.  $D$  معرف میزان دوز داروی توصیه شده در هر گام است و طبق رابطه(۱۴) این مقدار محاسبه می‌شود، که در این رابطه،  $D_0$  میزان دوز داروی تزریقی در ابتدای هر ده ساعت می‌باشد [۴۳].

$$D = D_0 e^{-k_e \tau} \quad (14)$$

در رابطه فوق،  $k_e$  معرف نرخ حذف دارو می‌باشد که این مقدار طبق رابطه(۱۵) محاسبه می‌شود و مقدار اولیه آن برابر با یک در نظر گرفته شده است [۴۳].

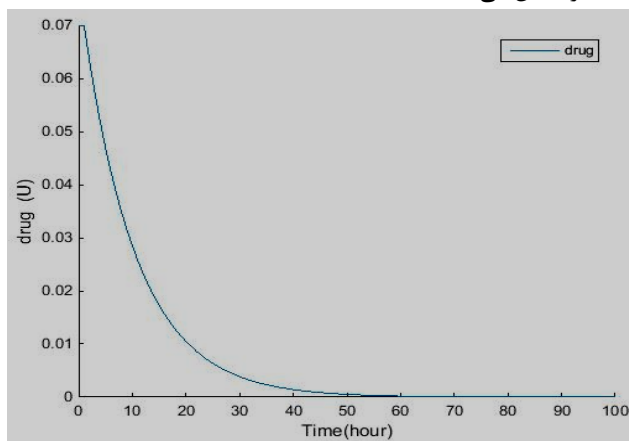
$$k_e = \frac{CL}{V} \quad (15)$$

در رابطه(۱۵)،  $CL$  سرعت حذف دارو می‌باشد که مقدار آن بر اساس رابطه(۱۶) محاسبه می‌شود [۴۳].

$$CL = k_e V \quad (16)$$

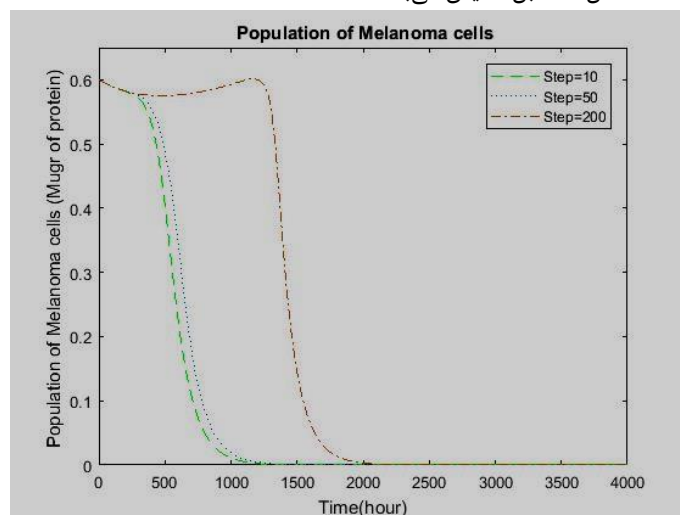
در روابط(۱۵) و(۱۶)،  $V$  حجم توزیع دارو در بدن می‌باشد و طبق رابطه(۱۷) محاسبه می‌شود. در ابتدای کار به دلیل اینکه تمام دارو در بدن وجود دارد، این مقدار برابر با بینهایت خواهد بود که در این مقاله معادل با ۱۰۰۰۰ در نظر گرفته شده است [۴۳].

با توجه به این شکل، تزریق دارو در ابتدای هر اپیزود انجام می‌گیرد و تأثیر آن تا انتهای اپیزود با توجه به نیمه عمر دارو لحاظ شده است. شکل(۳)، نیز روند کاهشی میزان دارو در بدن بیمار در طول ۱۰۰ ساعت را نشان می‌دهد.



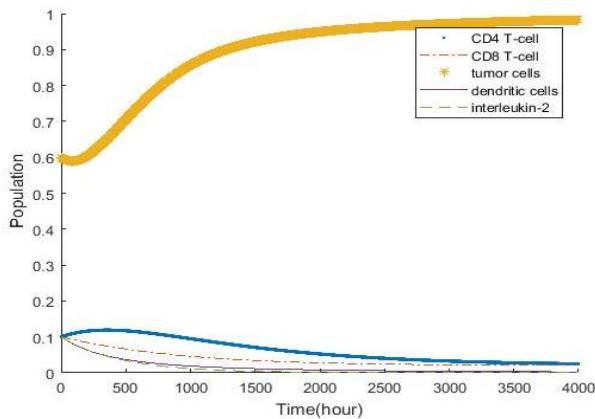
شکل(۳): روند کاهش میزان دارو در بدن در طول ۱۰۰ ساعت (اثر نیمه عمر داروی اینترفرون)

مشاهده تأثیر دارو در بدن بیمار در گام‌های ۱۰، ۵۰ و ۲۰۰ نیز انجام شده است. تأثیر این تعداد گام‌ها در کاهش سلول‌های سرطانی در شکل(۴) قابل نمایش می‌باشد.



شکل(۴): روند کاهش سلول‌های سرطانی در حالت در نظر گرفتن تأثیر دارو در گام‌های ۱۰، ۵۰ و ۲۰۰

نکته مهمی که باید بیان شود این است که تمام متغیرهای سیستم بین صفر و یک نرمالایز شده‌اند و شرایط اولیه به صورت زیر لحاظ شده است  $[31] I_0 = 0.1, D_0 = 0.1, M_0 = 0.6, C_0 = 0.1, H_0 = 0.1$ . برای شبیه‌سازی از نرم افزار متلب استفاده شده است. در ابتدا، رفتار سیستم بدون در نظر گرفتن دارو شبیه‌سازی شده است. شکل (۵)، نمایش گر رفتار سیستم در عدم حضور دارو می‌باشد.



شکل (۵): رفتار سیستم در عدم حضور دارو

همانطور که در شکل (۵) نشان داده شده است، تعداد سلول‌های سرطانی در عدم حضور دارو کم‌کم افزایش پیدا کرده و در نهایت به مقداری تقریباً برابر با یک رسیده است که ممکن است باعث مرگ بیمار شود. سلول‌های ایمنی بدن (CD4 و CD8) هم نیز ابتدا کمی افزایش پیدا کرده و در نهایت به صفر رسیده‌اند که این امر باعث کاهش ایمنی بدن می‌شود. سلول‌های دندریتیک و انتقال‌دهنده‌ی پیام‌های بین سلولی نیز رفته‌رفته کم شده و به صفر می‌رسند. حالت فوق در نهایت بیانگر مرگ بیمار خواهد بود. در قسمت بعدی رفتار سیستم در حضور دارو و تاثیر نیمه عمر دارو به نمایش گذاشته شده است و از الگوریتم یادگیری Q و سیاست CBR برای تعیین میزان بهینه دارو به گونه‌ای که از تاثیر سوء دارو بر روی سلول‌های سالم جلوگیری شود، استفاده گشته است

## ۵-۱- شبیه‌سازی رفتار سیستم در حالت استفاده از الگوریتم یادگیری Q

برای نشان دادن بهتر عملکرد الگوریتم یادگیری Q، شبیه‌سازی ابتدا با دوز داروی صفر و سپس با تزریق مقدار ثابت ۰٫۱ میلی‌گرم و در نهایت با استفاده از الگوریتم یادگیری Q تحت سیاست CBR انجام گرفته است. با توجه به نتایج شبیه‌سازی، استفاده از الگوریتم یادگیری Q در تعیین میزان دوز دارو باعث کاهش میزان داروی مصرفی و نیز کاهش اثرات سوء آن شده است و در عین حال اینکه، بیماری نیز کنترل شده است.

$$V = \frac{D}{k_e T} \cdot \frac{(1 - e^{-k_e T})}{[C_{\max} - (C_{\min} e^{-k_e T})]} \quad (17)$$

در رابطه فوق، T معرف زمان از بین رفتن دارو می‌باشد.  $C_{\max}$  و  $C_{\min}$  به ترتیب ماکزیمم و مینیمم غلظت پلاسما دارو در هر لحظه می‌باشند که طبق رابطه (۱۸) و (۱۹) محاسبه می‌شوند. مقدار مطلوب غلظت پلاسما بین ۱۰ تا ۲۰ میکروگرم می‌باشد [۴۳].

$$C_{\max} = \frac{D}{V \cdot k_e T} \cdot \frac{(1 - e^{-k_e T})}{(1 - e^{-k_e \tau})} \quad (18)$$

$$C_{\min} = C_{\max} e^{-k_e (\tau - T)} \quad (19)$$

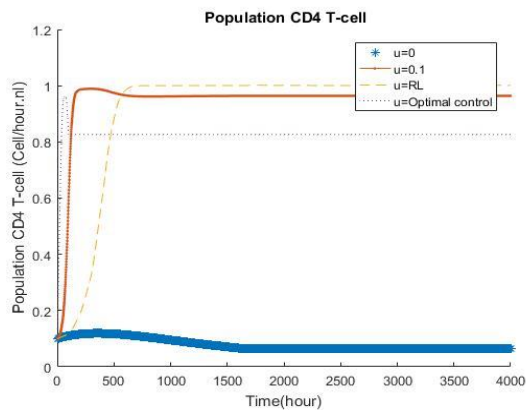
## ۵- شبیه‌سازی

برای اینکه عملکرد روش یادگیری تقویتی در کنترل سلول‌های سرطانی و تعیین میزان بهینه دوز دارو بهتر نشان داده شود، این روش با روش‌های کنترل بهینه و حالت بدون درمان مقایسه شده است. در روش‌های کنترل بهینه، تابعی تحت عنوان تابع هدف تعریف می‌شود و سعی بر این می‌باشد که بهینه‌ترین میزان دوز دارو به گونه‌ای تعیین شود که تابع هدف کمترین مقدار خود را در هر لحظه داشته باشد. جزئیات مربوط به این روش، اثبات روابط و جزئیات شبیه‌سازی در مرجع [۳۱] آورده شده‌اند. در این مقاله روش انتخابی با روش کنترل بهینه (روش همیلتونین) که یکی روش‌های کلاسیک در کنترل سیستم می‌باشد، مقایسه شده است. تابع هدف معرفی شده در [۳۱] بر اساس میزان سلول‌های سرطانی بنیادی و میزان دوز داروهای تزریق شده بیان شده است و هدف مینیمم کردن این تابع در هر لحظه می‌باشد. رابطه مربوط به تابع هزینه در معادله (۲۰) آورده شده است.

$$J = \rho M + \sum_{j=1}^m U_j \quad (20)$$

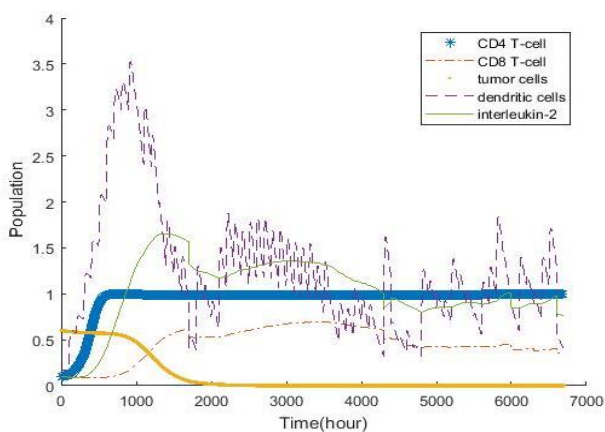
در رابطه فوق،  $\rho$  فاکتور وزن می‌باشد. M جمعیت سلول‌های سرطانی می‌باشد.  $U_j$  نیز میزان دوز داروی تزریق شده در هر لحظه می‌باشد که می‌بایست مینیمم گردد و این امر باعث کاهش اثرات زیانبار دارو خواهد شد.  $m$  تعداد گام در نظر گرفته شده برای تزریق دارو که چند مقدار مختلف و به صورت رندم انتخاب شده است. در نهایت با در نظر گرفتن  $m = 15$  سلول‌های سرطانی بیشترین کاهش را داشتند [31]. از جمله معایب این مدل نسبت به روش پیشنهادی مقاله، می‌توان به نیازمندی این روش به مدل خطی اشاره کرده که مدل خطی شده از بیمار مبتلا به سرطان ممکن است رفتاری مشابه با رفتار سیستم اصلی نداشته باشد. همچنین با وجود تزریق دوز داروی بیشتر قادر به حذف کامل سلول‌های سرطانی نباشد. در حالتی که بیمار هیچ مقدار دارویی دریافت نکرده است نیز، سلول‌های سرطانی شروع به افزایش کرده و سلول‌های سالم نیز کاهش چشم‌گیری پیدا کرده و به صفر می‌رسند که این امر مرگ بیمار را به همراه خواهد داشت.





شکل(۷): تغییرات سلول‌های ایمنی بدن در حضور دوز داروی صفر، دوز ثابت ۰,۱ میلی‌گرم، تعیین دارو با استفاده از الگوریتم یادگیری Q تحت سیاست CBR و کنترل بهینه(همیلتونین)

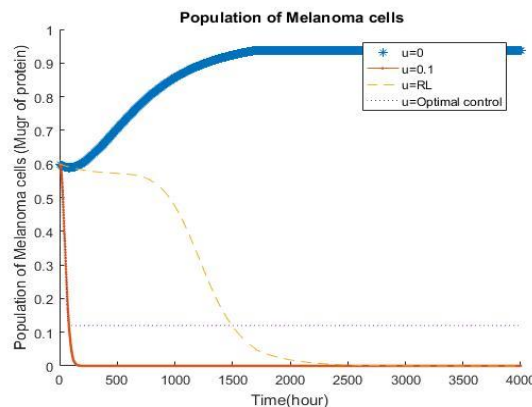
با توجه به شکل(۷)، تعداد سلول‌های ایمنی بدن در حالت تزریق دارو با استفاده از الگوریتم یادگیری Q تحت سیاست CBR بیشتر از حالت تزریق دارو با دوز ثابت ۰,۱ میلی‌گرم می‌باشد که این موضوع نشان‌دهنده افزایش ایمنی بدن در حالت استفاده از این الگوریتم می‌باشد. در حالت استفاده از روش کنترل بهینه نیز سلول‌های ایمنی بدن افزایش پیدا کرده اما به میزانی کمتر از حالت استفاده از الگوریتم یادگیری Q رسیده‌اند. شکل(۸) نیز نمایشگر تاثیر الگوریتم یادگیری Q تحت سیاست CBR بر روی تمام متغیرهای سیستم می‌باشد.



شکل(۸): روند تغییرات پارامترهای سیستم در تعیین دارو با استفاده از الگوریتم یادگیری Q تحت سیاست CBR

همانطور که در شکل(۸) مشاهده می‌شود، با تزریق دارو توسط الگوریتم یادگیری Q تحت سیاست CBR، بعد از گذشت ۲۲۰۰ ساعت، غلظت تومور به مقداری نزدیک به صفر رسیده است و علاوه بر آن CD4ها و CD8ها نیز با از بین رفتن تومور کم کم افزایش پیدا کرده‌اند و این به معنی افزایش ایمنی بدن می‌باشد. سلول‌های دندریتیک نیز به دلیل اینکه ارتباط مستقیم با میزان داروی

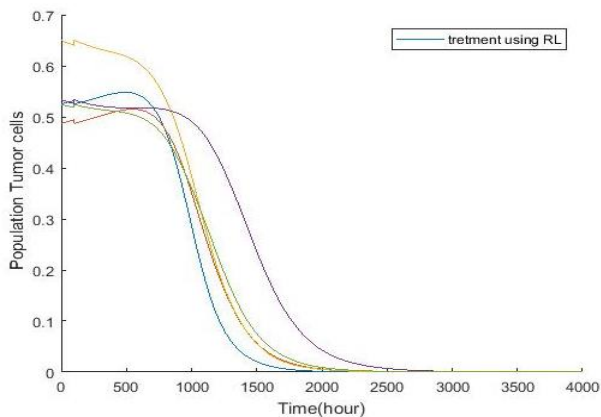
لازم به ذکر می‌باشد که برای جلوگیری از رشد مجدد سلول‌های تومور فرایند یادگیری تا ۶۸۰ نوبت تزریق دارو انجام شده است که در گام‌های انتهایی دوز داروی تزریقی صفر و یا مقداری نزدیک به صفر می‌باشد. در شکل(۶)، تاثیر این الگوریتم در کاهش سلول‌های سرطانی در ۴۰۰۰ ساعت و مقایسه آن با روش کنترل بهینه قابل مشاهده است.



شکل(۶): تغییرات سلول‌های سرطانی در حضور دوز داروی صفر، دوز ثابت ۰,۱ میلی‌گرم، تعیین دارو با استفاده از الگوریتم یادگیری Q تحت سیاست CBR و کنترل بهینه(همیلتونین)

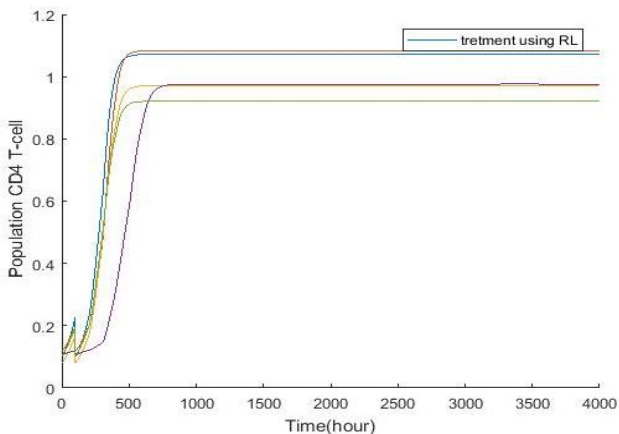
با توجه به شکل(۶)، سرعت کاهش سلول‌های سرطانی در حالت تزریق دوز داروی ثابت به مقدار ۰,۱ میلی‌گرم برای تمام زمان‌ها بیشتر از حالت تزریق دوز دارو با استفاده از الگوریتم یادگیری Q تحت سیاست CBR می‌باشد. طبیعی می‌باشد که اگر دوز داروی زیادی به بیمار تزریق شود تعداد سلول‌های سرطانی به سرعت افت خواهند کرد. اما فقط سرعت نزول آن‌ها مهم نمی‌باشد بلکه باید تاثیرات سوء دارو بر روی بیمار نیز لحاظ شود، بنابراین صرف کاهش سلول‌های سرطانی با یک سرعت بالا ملاک نمی‌باشد. هدف اصلی در این مقاله کاهش میزان داروی مصرفی و به تبع آن کاهش اثرات سوء دارو می‌باشد در عین حالی که جمعیت سلول‌های سرطانی کنترل شده است و سرعت کاهش سلول‌های سرطانی مورد توجه قرار نگرفته است. در حالت استفاده از روش کنترل بهینه نیز سلول‌های سرطانی با سرعت بالایی کاهش پیدا کرده‌اند. با وجود تزریق دوز داروی خیلی زیاد اما این روش قادر به حذف کامل سلول‌های سرطانی نبوده است. در شکل(۷)، تاثیر این الگوریتم بر روی سلول‌های ایمنی بدن در ۴۰۰۰ ساعت قابل مشاهده می‌باشد.

تزریقی دارند در ابتدا که تزریق دارو بیشتر می‌باشد افزایش پیدا کرده‌اند ولی با از بین رفتن تومور سرطانی به دلیل اینکه در بعضی از زمان‌ها برای جلوگیری از افزایش مجدد تومور، دوز داروی کمی به بدن بیمار تزریق می‌شود، نوسانات منظمی خواهند داشت. انتقال‌دهنده پیام‌های بین سلولی نیز به دلیل اینکه ارتباط مستقیم با سلول‌های دندریتیک دارند، ابتدا افزایش پیدا کرده و سپس به یک مقدار تقریباً ثابت رسیده‌اند. برای نمایش بهتر عملکرد تمام متغیرهای سیستم، شکل (۸) برای ۶۸۰۰ ساعت یعنی تا اتمام زمان یادگیری رسم شده است. در شکل (۹) میزان داروی تزریق شده به بدن بیمار مبتلا به سرطان ملانوما در طول درمان مشاهده می‌گردد.



شکل (۱۰): تغییرات سلول‌های سرطانی برای پنج بیمار مبتلا به سرطان ملانوما

همانطور که در شکل (۱۰) نشان داده شده است با استفاده از الگوریتم یادگیری Q تحت سیاست CBR، تعداد سلول‌های سرطانی در پنج بیمار مبتلا به سرطان ملانوما کاهش پیدا کرده‌اند. شکل (۱۱) نمایش سلول‌های ایمنی بدن برای ۵ بیمار مبتلا به سرطان ملانوما می‌باشد.

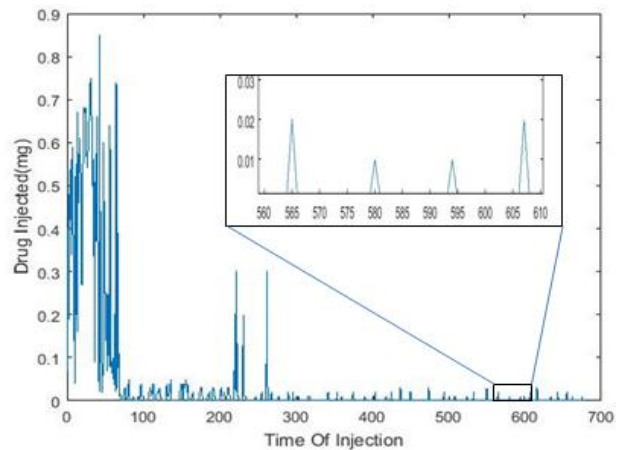


شکل (۱۱): تغییرات سلول‌های ایمنی بدن در پنج بیمار مبتلا به سرطان ملانوما

با توجه به شکل (۱۱)، CD4ها با از بین رفتن سلول‌های تومور در هر پنج بیمار مبتلا به سرطان ملانوما افزایش یافته‌اند که این نشان‌دهنده افزایش ایمنی بدن می‌باشد.

### ۵-۳- افزودن نویز به سیستم در حضور عدم قطعیت در پارامترها و شرایط اولیه

در این بخش تاثیر نویز و اغتشاش در کنترل سلول‌های سرطانی در بیماران مبتلا به سرطان ملانوما بررسی شده است. برای افزودن نویز به دینامیک سیستم، پارامتر  $w$  که در هر گام زمانی مقداری متفاوت و احتمالاتی به خود می‌گیرد به معادلات



شکل (۹): میزان داروی تزریقی توسط الگوریتم یادگیری Q در ۶۸۰ گام زمانی

با توجه به شکل (۹)، در ابتدا که غلظت تومور سرطانی زیاد می‌باشد، میزان داروی پیشنهاد شده توسط الگوریتم یادگیری Q تحت سیاست CBR بیشتر است. کم‌کم که تعداد سلول‌های سرطانی در بدن کاهش می‌یابند داروی پیشنهادی توسط این الگوریتم نیز کاهش یافته و فقط در برخی از زمان‌ها برای جلوگیری از رشد مجدد تومور مقدار بسیار کمی دارو به بدن بیمار تزریق می‌شود که در خیلی از گام‌های زمانی این مقدار تزریقی برابر با صفر می‌باشد.

### ۵-۲- عدم قطعیت در پارامترهای سیستم و شرایط اولیه

در این بخش برای نشان دادن عملکرد بهتر الگوریتم یادگیری Q تحت سیاست CBR، به بررسی کنترل سلول‌های سرطانی با کمترین میزان تزریق دارو در ۵ بیمار مبتلا به سرطان ملانوما پرداخته شده است. برای این کار، ابتدا عدم قطعیت در پارامترها و شرایط اولیه به ترتیب با تغییرات ۱۰ و ۲۵ درصد، به صورت توزیع نرمال لحاظ شده است. شکل (۱۰) کاهش سلول‌های سرطانی برای ۵ بیمار مبتلا به سرطان ملانوما را در ۴۰۰۰ ساعت نشان می‌دهد. همانطور که در این شکل مشاهده می‌گردد، یادگیری که در سیستم اصلی با استفاده از الگوریتم Q تحت سیاست CBR انجام شده است به خوبی توانسته است تعداد

دارو که نتیجه آن کاهش اثرات سوء دارو می باشد سلول های سرطانی را در دیگر بیماران مبتلا به سرطان ملانوما کنترل کند، که این امر افزایش ایمنی بدن را به همراه خواهد داشت.

## ۶- نتیجه گیری

تشخیص و درمان سرطان ملانوما با استفاده از روش های پزشکی امری زمان بر بوده و با ایجاد درد برای بیمار همراه می باشد. عدم تشخیص به موقع و عدم کنترل سریع این نوع سرطان نیز ممکن است مرگ بیمار را به همراه داشته باشد. استفاده از روش های هوشمند در کنار کار پزشکان باعث افزایش سرعت در تشخیص سرطان ملانوما و تزریق میزان بهینه دوز دارو برای کنترل سریع تر سلول های سرطانی خواهد شد که این امر کاهش اثرات زیانبار دارو را به همراه دارد.

در این مقاله، ابتدا با تزریق دوز داروی ثابت، رفتار سیستم بررسی شد. همانطور که مشاهده گردید با تزریق مقدار ثابت دارو در هر گام زمانی غلظت تومور خیلی سریع کاهش پیدا کرده است، اما این روش دو اشکال اساسی دارد: اول اینکه مقدار داروی تزریقی با سعی و خطا بدست آمده است و دوم اینکه تزریق دارو در هر لحظه تاثیرات سوء دارو را بر روی سلول های سالم به همراه خواهد داشت که این مورد در روش تزریق دوز داروی ثابت در نظر گرفته نشده است. ایده اصلی این مقاله استفاده از روش الگوریتم یادگیری Q برای تعیین مقدار بهینه دوز دارو می باشد به گونه ای که علاوه بر کنترل سلول های سرطانی از تاثیرات سوء دارو بر روی سلول های سالم نیز جلوگیری شود. کار دیگری که در این مقاله انجام شده است، برای انتخاب اعمال در الگوریتم یادگیری Q از سیاست اکتشافی شتاب داده شده با نام اختصاری CBR استفاده شده است که باعث افزایش سرعت یادگیری و نیز کاهش زمان برای رسیدن به سیاست بهینه شد. انتخاب اعمال با استفاده از این سیاست و در واقع ترکیب روش یادگیری تقویتی و CBR که نوآوری اصلی این مقاله می باشد، در کارهای پیشین انجام نشده است. با در نظر گرفتن اثر نیمه عمر دارو نیز روند کاهش داروی تزریق شده به بدن بیمار در طول ده ساعت مشاهده شد. با هر بار تزریق دارو و تاثیر دادن نیمه عمر دارو، حالت بیمار برای لحظه بعدی ( $S_{t+1}$ ) محاسبه گردید. با در نظر گرفتن عدم قطعیت در پارامترهای سیستم، مشاهده شد که الگوریتم یادگیری Q به خوبی توانسته است تعداد سلول های سرطانی را در دیگر بیماران مبتلا به ملانوما نیز کاهش دهد و در حضور نویز و اغتشاش نیز قادر به کنترل سلول های سرطانی می باشد.

در انتها برای اینکه عملکرد و مزیت روش پیشنهادی بهتر نشان داده شود، این روش با یکی از روش های کنترل بهینه به نام همیلتونین مقایسه شد. با توجه به نتایج شبیه سازی روش کنترل بهینه با سرعت بالاتر سلول های سرطانی را کاهش داده است اما قادر به

(۱) تا (۵) اضافه شده است و معادلات جدید به فرم روابط (۲۱) تا (۲۵) تغییر خواهند کرد.

$$\dot{H} = a_0 + b_0 DH \left(1 - \frac{H}{n_0}\right) - c_0 H + w \quad (21)$$

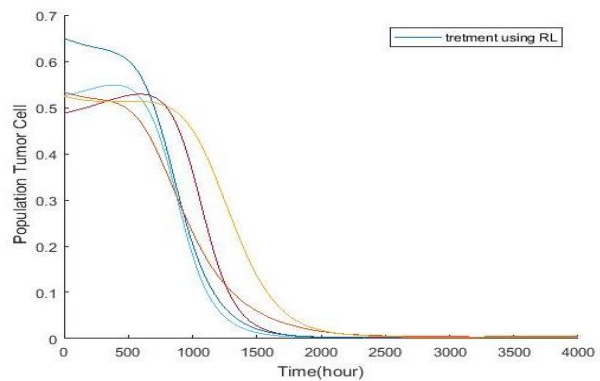
$$\dot{C} = a_1 + b_1 I(M + D)C \left(1 - \frac{C}{n_1}\right) - c_1 C + w \quad (22)$$

$$\dot{M} = b_2 M \left(1 - \frac{M}{n_2}\right) - d_2 MC + w \quad (23)$$

$$\dot{D} = -d_3 DC + u + w \quad (24)$$

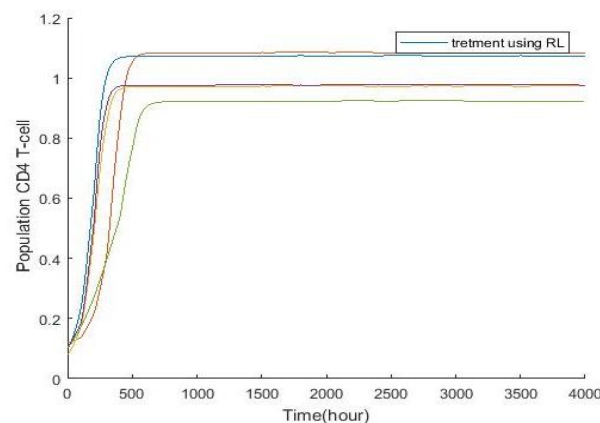
$$\dot{I} = b_4 DH - e_4 IC - c_4 I + w \quad (25)$$

شکل (۱۲)، نمایش کاهش سلول های سرطانی در سیستم دارای عدم قطعیت و نویز برای پنج بیمار مختلف می باشد.



شکل (۱۲): تغییرات سلول های سرطانی برای پنج بیمار مبتلا به سرطان ملانوما در حضور نویز و اغتشاش

شکل (۱۳)، سلول های ایمنی بدن را برای پنج بیمار مبتلا به سرطان ملانوما در حضور نویز و اغتشاش نشان می دهد.



شکل (۱۳): تغییرات سلول های ایمنی بدن برای پنج بیمار مبتلا به سرطان ملانوما در حضور نویز و اغتشاش

با توجه به شکل (۱۲) و (۱۳)، الگوریتم یادگیری Q تحت سیاست CBR به خوبی توانسته است در حضور نویز و با کمترین میزان تزریق

- [1] M. Suryapraha, G. Rajanarayane, P. Kumari, "Analysis of Skin Cancer Classification Using GLCM Based On Feature Extraction in Artificial Neural Network," International Journal of Emerging Technology in Computer Science & Electronics, March 2015.
- [2] E. Fikret. "Neural network diagnosis of malignant melanoma from color images." IEEE Transactions on biomedical engineering, Vol.41, pp. 837-845, 1994.
- [3] W. Chang. "Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study." PloS one, 2013.
- [4] M.Ramezani, A. Karimian, P. Moallem. "Automatic detection of malignant melanoma using macroscopic images." Journal of medical signals and sensors, Vol.4, No.4, 2014.
- [5] S.Ramteke, V. Nilkamal, J. Shweta. "Analysis of Skin Cancer Using Fuzzy and Wavelet Technique-Review & Proposed New Algorithm." International Journal of Engineering Trends and Technology, Vol.4, No.5, pp. 2555-2566, 2013.
- [6] M. Palak, and B. Shah. "Review on Techniques and Steps of Computer Aided Skin Cancer Diagnosis." Procedia Computer Science 85, Vol.85, pp. 309-316, 2016.
- [7] J. Aswin, A. Jaleel, S.Salim. "Hybrid genetic algorithm— Artificial neural network classifier for skin cancer detection." Contrl, Instrumentation, Communication and Computational Technologies, International Conference on.IEEE, 2014.
- [8] محزونیه و محمدی خروشان. "سیستم نرم افزاری تشخیص هوشمند سرطان پوست مبتنی بر اصول پزشکی از راه دور." فصلنامه بیمارستان، ویژه نامه تله مدیسن (۱۳۹۳)، ۲۰۱۴.
- [9] A. Lashkari, F. Pak, M. Firouzmand, "Full intelligent cancer classification of thermal breast images to assist physician in clinical diagnostic applications". Journal of medical signals and sensors, 6(1), 12, 2016.
- [10] U. Sirin, F. Polat, R. Alhaji. "Employing batch reinforcement learning to control gene regulation without explicitly constructing gene regulatory networks." Proceedings of the 23rd International Joint Conference on Artificial Intelligence, 2013.
- [11] S. Mazdeyasna, A. Jafari, J. Hadjati, A. Allahverdy, M. Alavi-Moghaddam. "Modeling the Effect of Chemotherapy on Melanoma B16F10 in Mice Using Cellular Automata and Genetic Algorithm in Tapered Dosage of FBS and Cisplatin." Frontiers in Biomedical Technologies, Vol. 2(2), pp. 103-108, 2015.
- [12] X. Wang, S. Lu, J. Guo. "Treatment algorithm of metastatic mucosal melanoma." Chinese clinical oncology 3.3, Vol.3, NO.3, 2014.
- [13] Y. Zhao, M. R. Kosorok, D. Zeng, (2009). "Reinforcement learning design for cancer clinical trials". Statistics in medicine, Vol. 28(26), pp. 3294-3315, 2009.
- [14] M. K. Bothe, L. Dickens, K. Reichel, A. Tellmann, B. Ellger, M. Westphal, A. A. Faisal. "The use of reinforcement learning algorithms to meet the challenges of an artificial pancreas." Expert review of medical devices, Vol. 10(5), pp. 661-673, 2013.
- [15] A. Noori, M. B. Naghbi Sistani, N. Pariz. "Hepatitis B virus infection control using reinforcement learning", ICEEE 2011, Gonabad Islamic Azad University, 2011.
- [16] A.E. Gaweda, M.K. Muezzinoglu, A.A. Jacobs, G. R. Aronoff, M.E. Brier. "Model predictive control with reinforcement learning for drug delivery in renal anemia

حذف کامل این سلول‌ها نبوده است. در روش تزریق دوز داروی ثابت نیز سلول‌های سرطانی با سرعت بالا کاهش پیدا کرده‌اند اما مجموع داروی تزریقی در کل مدت درمان بسیار زیاد می‌باشد که این امر افزایش اثرات زینبار دارو را به همراه خواهد داشت. این درحالی می‌باشد که روش پیشنهادی این مقاله که ترکیب دو روش الگوریتم یادگیری Q و سیاست استدلال مبتنی بر مورد می‌باشد، توانسته است علاوه بر تزریق دوز داروی کمتر سلول‌های سرطانی را به طور کامل از بین برده و به صفر برساند. علاوه بر این، روش‌های کنترل بهینه بر مبنای مدل خطی شده سیستم کار می‌کنند درحالی‌که روش یادگیری تقویتی نه تنها قابلیت پیاده‌سازی بر روی مدل غیر خطی را دارد بلکه بدون استفاده از مدل ریاضی نیز قابلیت کنترل سیستم و خاصیت تطبیق‌پذیری با محیط را دارا می‌باشد. این روش سعی و خطای پزشک را در بدست آوردن دوز داروی بهینه برای کنترل سلول‌های سرطانی کاهش خواهد داد. همانطور که در بخش مقدمه بیان شد، با استفاده از الگوریتم ژنتیک نیز کنترل سلول‌های سرطانی در بیماران مبتلا به سرطان ملانوما انجام گرفته است. اما این روش به صورت آبی و در لحظه قابلیت تعیین بهینه‌ترین میزان دوز دارو را ندارد که این امر از معایب این روش محسوب می‌شود. این در حالی می‌باشد که روش پیشنهادی در این مقاله بعد از یادگیری قابلیت تعیین میزان بهینه دوز دارو را برای بیماران واقعی دارا می‌باشد. تعیین پارامترها در بخش یادگیری تقویتی به صورت سعی و خطا تعیین گشته است. تعیین این پارامترها با استفاده از روش‌های بهینه‌سازی باعث افزایش دقت در کنترل سلول‌های سرطانی می‌شود. در این مقاله به دلیل زمان‌بر بودن روش یادگیری، تعیین پارامترها به این روش انجام نگرفته است.

جدول (۲): مقایسه مجموع داروهای تزریقی به بیمار در دو حالت ثابت

و با استفاده از الگوریتم یادگیری Q

روش تزریق دارو	مجموع داروی تزریقی به بیمار (۶۸۰ نوبت)
تزریق داروی ثابت به میزان ۰.۱	۶۸ میلی گرم
تزریق دارو با استفاده از الگوریتم یادگیری Q	۳۰.۱۶ میلی گرم
تزریق دارو با استفاده از روش همیلتونین	۳۱۳.۵۱۵ میلی گرم

همان‌طور که در جدول فوق نشان داده شده است، در حالت استفاده از روش الگوریتم یادگیری Q علاوه بر حذف کامل سلول‌های سرطانی، میزان دوز داروی تزریقی نیز بسیار کمتر از حالت استفاده از روش کنترل بهینه و تزریق دوز داروی ثابت می‌باشد که کاهش اثرات زینبار دارو برای بیمار را به همراه خواهد داشت.

مراجع

- [30] M. Andrea, F. Topputo, F. Bernelli-Zazzera. "Controlled drug delivery in cancer immunotherapy: Stability, optimization, and Monte Carlo analysis." *SIAM Journal on Applied Mathematics* 71.6, pp. 2229-2245, 2011.
- [31] S. Eikenberry, T. Craig, K. Yang. "Tumor-immune interaction, surgical treatment, and cancer recurrence in a mathematical model of melanoma." *PLoS computational biology* 5.4, Vol.5, No.4, 2009.
- [32] M. Pennisi. "A mathematical model of immune-system-melanoma competition." *Computational and mathematical methods in medicine*, Vol.2.12., 2012.
- [33] Y. Kogan, A. Zvia, E. Moran. "A mathematical model for the immunotherapeutic control of the Th1/Th2 imbalance in melanoma." *Discrete And Continuous Dynamical Systems Series B*, Vol.18, pp.1017-130, 2013.
- [34] X. Sun, J. Bao, Y. Shao. "Mathematical modeling of therapy-induced cancer drug resistance: connecting cancer mechanisms to population survival rates." *Scientific reports* 6, 2016.
- [35] A. Kłusek, W. Dzwinel, V. Vasilyev. "Super modeling in simulation of melanoma progression." *Procedia Computer Science*, Vol.80, pp.999-1010, 2016.
- [36] A. Isabel. "On the geometric modulation of skin lesion growth: a mathematical model for melanoma." *Research on Biomedical Engineering AHEAD*, Vol.32, pp.44-54, 2016.
- [37] F. Castiglione, B. Piccoli, "Optimal Control in a Model of Dendritic Cell Transfection Cancer Immunotherapy" *Bulletin of Mathematical Biology*, Vol.68, pp. 255-274, 2006.
- [38] F. Castiglione, B. Piccoli, "Cancer Immunotherapy, Mathematical Modeling and Optimal Control." *Journal of Theoretical Biology*, Vol.247, pp. 723-732, 2007.
- [39] R. Sutton and A. Barto. *Reinforcement learning: An introduction*. The MIT Press Cambridge, Massachusetts London, England. 2011.
- [40] H. Gholizade-Narm, A. Noori. "Control the population of free viruses in nonlinear uncertain HIV system using Q-learning." *International Journal of Machine Learning and Cybernetics*, Vol.9, pp. 1169-1179, July 2018.
- [41] R. Bianchi, R. Mantaras. "Case-Based Multi agent Reinforcement Learning: Cases as Heuristics for Selection of Actions." *ECAI*. Vol.215, pp.355-360, 2010.
- [42] S. Dhillon, A. Kostrzewski, "Clinical pharmacokinetics." *Pharmaceutical Press*, 2006.
- management." *Engineering in Medicine and Biology Society*, 2006. EMBS'06. 28th Annual International Conference of the IEEE, New York, pp. 5177-5180, 2006.
- [17] G. Czibula, I. M. Bocicor, I. G. Czibula. "Temporal ordering of cancer microarray data through a reinforcement learning based approach". *PloS one*, Vol. 8(4), 2013.
- [18] Y. Zhao, D. Zeng, M. A. Socinski, M. R. Kosorok. "Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer". *Biometrics*, Vol. 67(4), pp. 1422-1433, 2011.
- [19] R. Padmanabhan, N. Meskin, W. M. Haddad. "Reinforcement learning-based control of drug dosing for cancer chemotherapy treatment". *Mathematical biosciences*, Vol. 293, pp. 11-20, 2017.
- [۲۰] یونسسی، شایقی، اکبری مجد و یاشار هاشمی. "طراحی پایدار ساز PSS3B بر اساس الگوریتم KH و Q-learning برای میراسازی نوسانات فرکانس پایین سیستم قدرت تک ماشینه". *نشریه مهندسی برق و الکترونیک ایران*. دوره ۱۴، شماره ۳، ۱۳۹۶.
- [20] S. El-Sappagh, M. Elmogy, A. M. Riad. "A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis". *Artificial intelligence in medicine*, Vol. 65(3), pp. 179-208, 2015.
- [21] J. M. Malof, A. E. Gaweda. "Optimizing drug therapy with reinforcement learning: The case of anemia management." In *Neural Networks (IJCNN)*, The 2011 International Joint Conference on, pp. 2088-2092. IEEE, 2011, July.
- [22] Y. Shen, J. Colloc, A. Jacquet-Andrieu, K. Lei. "Emerging medical informatics with case-based reasoning for aiding clinical decision in multi-agent system". *Journal of biomedical informatics*, Vol. 56, pp. 307-317, 2015.
- [23] M. G. Voskoglou, A. B. M. Salem. "Analogy-based and case-based reasoning: Two sides of the same coin". *arXiv preprint arXiv:1405.7567*, 2014.
- [24] R. Miotto, C. Weng. "Case-based reasoning using electronic health records efficiently identifies eligible patients for clinical trials". *Journal of the American Medical Informatics Association*, Vol. 22(e1), 2015.
- [25] L. Göllmann, H. Maurer. "Optimal control problems with time delays: Two case studies in biomedicine". *Mathematical Biosciences & Engineering*, Vol. 15(5), pp. 1137-1154, 2018.
- [26] J. Malinzi, R. Ouifki, A. Eladdadi, D. F. Torres, K. A. White. "Enhancement of chemotherapy using oncolytic virotherapy: Mathematical and optimal control analysis". *arXiv preprint arXiv:1807.04329*, 2018.
- [27] H. Moore. "How to mathematically optimize drug regimens using optimal control". *Journal of pharmacokinetics and pharmacodynamics*, Vol. 45(1), pp. 127-137, 2018.
- [28] A. M. A. Rocha, M. F. P. Costa, E. M. Fernandes. "On a multiobjective optimal control of a tumor growth model with immune response and drug therapies". *International Transactions in Operational Research*, Vol. 25(1), pp. 269-294, 2018.
- [29] H. Khaloozadeh, P. Yazdanbakhsh, F. Homaei-Shandiz. "The Optimal Dose of Drug in Neoadjuvant Chemotherapy before Surgery for the Patients Suffering from Breast Cancer Stage III". *Iranian Journal of Biomedical Engineering*, Vol. 1, pp. 319-334, 2008.