



پایان نامہ کارشناسی ارشد

مهندسی برق - کنترل

کنترل جمعیت ویروس‌ها در بیماران مبتلا به HIV با استفاده از یادگیری تقویتی و لحاظ ضریب تاثیر دوز دارو

امین نوری

اساتید راهنما:

دکتر ناصر پریز

دکتر محمد باقر نقیبی سیستانی

استاد مشاور:

دکتر غزاله شجاع رضوی

تابستان ۱۳۹۰

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

اظهارنامه

اینجانب امین نوری دانشجوی دوره کارشناسی ارشد رشته برق - کنترل دانشکده مهندسی دانشگاه فردوسی مشهد نویسنده پایان نامه "کنترل جمعیت ویروس‌ها در بیماران مبتلا به HIV با استفاده از یادگیری تقویتی" تحت راهنمایی آقای دکتر محمد باقرنقیبی سیستانی و آقای دکتر ناصر پریز متعهد می‌شوم:

- تحقیقات در این رساله/پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در رساله/پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه فردوسی مشهد می‌باشد و مقالات مستخرج با نام « دانشگاه فردوسی مشهد » و یا « Ferdowsi University of Mashhad » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی رساله/پایان نامه تأثیرگذار بوده‌اند در مقالات مستخرج از رساله/ پایان نامه رعایت شده است.
- در کلیه مراحل انجام این رساله/پایان نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این رساله/پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است، اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه‌ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه فردوسی مشهد می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در رساله/پایان نامه بدون ذکر مرجع مجاز نمی‌باشد.

پروردگارا

کمک کن

دریابم که وسعت زندگی فراروی من است، اما چنان عزیز که یک لحظه‌اش تباه نباید کرد
کار کنم تا رسیدن به هدف‌هایم و بدانم که بدست خواهد آمد
و به رویاهایم دست برآورم، با توان، به تصمیم و با ایمان
دریابم که جسارت و امید را فرو نگذارم و تردید را نگذارم که مرا دلسرد کند از انجام آنچه در دل دارم
به یاد نگه دارم که جهان نیازمند آفتاب لبخندهای هرچه بیشتر است
یادم باشد که سهم خود را ادا کنم.

زحمات و تلاش‌های همه اساتید گرامی را پاس داشته و مراتب تقدیر و تشکر قلبی را تقدیم به ساحت مقدس
این فرزندگان می‌نمایم.

از اساتید بزرگوار جناب آقای دکتر نقیبی سیستانی و جناب آقای دکتر پریز که در مراحل مختلف پروژه
بادقت و حوصله ارائه طریق نموده و اینجانب را راهنمایی فرمودند خاضعانه سپاسگزاری می‌نمایم.
از مشاوره و رهنمودهای استاد بزرگوار سرکار خانم دکتر شجاع رضوی کمال تشکر و سپاس را دارم.

تقدیم به

مادر مهربانم

گل خوشبوی زندگیم که بی‌منت وجودت را وقف من نمودی به پاس این همه فداکاری من هم چکیده‌ای از تلاش علمی‌ام را به پای خوبی‌هایت تقدیم می‌کنم.

پدر بزرگووارم

که با اعتماد به شانه‌های محکم واستوارش توانستم قدم درراه تحصیل، علم وکمال بگذارم. به یاد خوبی‌هایت، به یاد ایثارگری‌هایت، به یاد آن همه گذشت و فداکاری که درحق من کردی این مجموعه را تقدیم می‌کنم.

چکیده رساله/پایان نامه :

HIV ویروسی است که با مختل کردن عملکرد و از بین بردن نوع خاصی از سلول‌های ایمنی بدن به نام T-CD⁴ منجر به نقص دستگاه ایمنی بدن انسان می‌شود. این ویروس با از بین بردن این گونه از سلول‌ها، باعث ایجاد بیماری ایدز (نقص ایمنی اکتسابی) می‌شود. امروزه پیشرفت‌های قابل توجهی در داروهای مورد استفاده در درمان بیماران مبتلا به عفونت HIV پدیدار گشته است که سبب بهبود کیفیت زندگی و افزایش طول عمر بیماران شده است، ولی مصرف تمامی این داروها دارای اثرات سوء فراوانی می‌باشند. از آنجا که پزشکان بر اساس علایم بالینی و طبق پروتکل‌های مشخص که برای خیلی از بیماران یکسان می‌باشد، اقدام به تجویز دارو می‌کنند. بنابراین دوز داروی تجویز شده برای هر بیمار بهینه نمی‌باشد. در سال‌های اخیر پژوهش‌های زیادی برای تعیین دوز بهینه دارو برای مقابله با ویروس HIV انجام شده است. دلایل بسیار زیادی برای این درمان بهینه وجود دارد که از آن جمله می‌توان به کمینه کردن مسمومیت بوسیله داروها و هزینه‌های درمان اشاره کرد.

هدف از این تحقیق، کنترل جمعیت ویروس‌های HIV در بیماران مبتلا به این عفونت و کنترل بهینه دوز دارو با استفاده از الگوریتم‌های حل مسئله یادگیری تقویتی می‌باشد. استراتژی مذکور بر روی مدل پیشنهاد شده توسط وودارز- نوواک مورد استفاده قرار گرفته است. این مدل مشخصات و ویژگی‌های مراحل عفونت و اثر متقابل میان سلول‌های CD⁴ سالم، لنفوسیت‌های پیشرو و عملگر نهایی، ویروس‌های آزاد و داروی مورد استفاده قرار گرفته شده را نشان می‌دهد. وجود متغیرهای لنفوسیت‌های پیشرو، لنفوسیت‌های عملگر نهایی و همچنین وجود جمله‌ای برای بیان تاثیر داروهای مهار کننده ترانس کریپتاز معکوس (RTI)، از مزیت‌های این مدل نسبت به مدل‌های دیگر می‌باشد. در این پژوهش جمعیت ویروس‌های آزاد بوسیله دو الگوریتم یادگیری Q و سارسا کنترل گشته است. با استفاده از اطلاعات ۲ بیمار، استراتژی زیر بهینه‌ای را بدست آورده‌ایم که بر روی ۱۳۰ بیمار با دقت‌های ۹۰٪ و ۹۲٪، با توجه به دو الگوریتم سارسا و یادگیری Q، قابلیت کنترل جمعیت ویروس‌های HIV را دارا می‌باشد. در این تحقیق زمان و سرعت درمان و کاهش دوز دارو از دیگر پارامترهایی بوده است که مورد توجه قرار گرفته است. استراتژی بدست آمده عملکرد بهتری نسبت به تغییر پارامترهای بیماران و مقادیر اولیه متغیرها در مقایسه با روش‌های قبلی دارا می‌باشد.

۱	فصل اول	۱
۲	۱-۱ پیش‌گفتار	۲
۳	۲-۱ توصیف مسئله: کنترل جمعیت ویروس HIV در بدن	۳
۴	۳-۱ مروری بر کارهای انجام شده توسط الگوریتم یادگیری تقویتی در علم پزشکی	۴
۵	۴-۱ مروری بر کارهای انجام شده در رابطه با کنترل عفونت ویروس HIV	۵
۱۰	۲ ویروس HIV و بیماری ایدز	۱۰
۱۱	۱-۲ مقدمه	۱۱
۱۱	۲-۲ تاریخچه‌ی ویروس و پیدایش ایدز	۱۱
۱۳	۳-۲ ویروس HIV	۱۳
۱۳	۱-۳-۲ معرفی ویروس	۱۳
۱۵	۲-۳-۲ ساختمان ویروس	۱۵
۱۵	۳-۳-۲ نحوه عملکرد ویروس	۱۵
۱۷	۴-۳-۲ مکانیسم رستپوری	۱۷
۱۷	۲-۳-۲ چرخه زندگی ویروس	۱۷
۱۹	۴-۲ پاسخ سیستم دفاعی بدن انسان به ویروس HIV	۱۹
۲۲	۵-۲ مراحل عفونت ویروس HIV	۲۲
۲۳	۱-۵-۲ مرحله حاد (مرحله اول آلودگی به ویروس)	۲۳
۲۴	۲-۵-۲ مرحله بی‌علامت (مرحله دوم آلودگی به ویروس)	۲۴
۲۵	۳-۵-۲ مرحله بزرگی غده لنفاوی منتشره پایدار (مرحله سوم آلودگی به ویروس)	۲۵
۲۵	۴-۵-۲ مرحله سندرم ایدز (مرحله چهارم آلودگی به ویروس)	۲۵
۲۷	۶-۲ تشخیص ایدز	۲۷
۲۷	۷-۲ آزمایش‌های تشخیص ویروس HIV	۲۷
۲۸	۱-۷-۲ آزمون‌های سنجش آنتی بادی ویروس	۲۸

۲۹ آزمون الایزا ۱-۱-۷-۲
۲۹ آزمون وسترن بلات ۲-۱-۷-۲
۳۰ آزمون‌های سنجش آنتی ژن ۲-۷-۲
۳۱ تست آنتی ژن P _{۲۴} ۱-۲-۷-۲
۳۱ HIV-PCR آزمون ۲-۲-۷-۲
۳۲ تست میزان ویروس ۳-۷-۲
۳۲ ۸-۲ داروهای موجود
۳۴ ۹-۲ درمان چند دارویی
۳۶ ۱۰-۲ داروی ایرانی
۳۶ ۱۱-۲ عوارض جانبی داروها
۳۷ ۱۲-۲ درمان ضد رتروویروسی
۳۷ ۱۳-۲ تعاریف و علل شکست درمان ضد رتروویروسی
۳۸ ۱-۱۳-۲ شکست ویرولوژیک
۳۸ ۲-۱۳-۲ شکست ایمونولوژیک
۳۹ ۳-۱۳-۲ ارتباط بین شکست ویرولوژیک، شکست ایمونولوژیک و پسرفت بالینی
۳۹ ۱۴-۲ ارزیابی شکست درمان ضد رتروویروسی
۳۹ ۱۵-۲ وقفه‌های کوتاه مدت درمان
۴۱ ۳ مدل‌های ریاضی بیماری
۴۲ ۱-۳ مقدمه
۴۴ ۲-۳ مدل پایه آلودگی بیماری‌های عفونی
۴۴ ۳-۳ مدل پرلسون
۴۶ ۴-۳ مدل نوواک بنگهام
۴۷ ۵-۳ مدل پرلسون و نلسون
۴۸ ۶-۳ مدل دسوزا
۵۰ ۷-۳ مدل اصلاح شده‌ی پرلسون

۵۰ ۳-۸ مدل اصلاح شده‌ی وودارز- نوک
۵۶ ۴ استفاده از یادگیری تقویتی در کنترل جمعیت ویروس‌ها
۵۷ ۴-۱ مقدمه
۵۷ ۴-۲ تاریخچه الگوریتم یادگیری تقویتی
۵۸ ۴-۳ معرفی یادگیری تقویتی
۶۵ ۴-۳-۱ الگوریتم Q-Learning
۶۷ ۴-۳-۲ الگوریتم‌های سارسا و سارسا λ
۶۸ ۴-۴ تطابق مسئله یادگیری تقویتی و مدل عفونت HIV
۶۹ ۴-۵ دینامیک سیستم
۶۹ ۴-۵-۱ نقطه تعادل اول
۷۰ ۴-۵-۲ نقطه تعادل دوم
۷۰ ۴-۵-۳ نقطه تعادل سوم
۷۱ ۴-۵-۴ نقطه‌ی تعادل چهارم
۷۱ ۴-۶ استفاده از الگوریتم‌های حل مسئله یادگیری تقویتی در کنترل جمعیت ویروس‌ها
۷۴ ۴-۷ نتایج شبیه‌سازی
۸۰ ۴-۸ بررسی رباستنس الگوریتم یادگیری
۸۳ ۴-۹ در نظر گرفتن اثر مقاومت ویروس‌ها در برابر دارو
۸۷ ۴-۱۰ ارائه‌ی مدل جدیدی برای عفونت با در نظر گرفتن ویروس‌های مقاوم در برابر دارو ...
۹۰ ۵ نتایج و پیشنهادات
۹۳ مراجع

شکل ۱-۲. ساختمان ویروس	۱۵
شکل ۲-۲. چرخه زندگی ویروس	۱۹
شکل ۳-۲. رابطه‌ی میان بیماری ایدز و گویچه‌های سفید CD4+T	۲۶
شکل ۱-۳. رابطه‌ی بین متغیرهای حالت مدل وودراز-نوواک	۵۳
شکل ۱-۴. مدل استاندارد یادگیری تقویتی	۵۹
شکل ۲-۴. جمعیت ویروس‌های آزاد طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q، سارسا ..	۷۶	
شکل ۳-۴. جمعیت سلول‌های سالم طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q، سارسا ...	۷۷	
شکل ۴-۴. جمعیت سلول‌های آلوده طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q، سارسا ...	۷۷	
شکل ۵-۴. جمعیت CTL‌های پیشرو طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q، سارسا ..	۷۸	
شکل ۶-۴. جمعیت CTL‌های عملگر نهایی طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q، سارسا	۷۸
شکل ۷-۴. دوز داروی استفاده شده بوسیله‌ی الگوریتم یادگیری-Q در ۳۶۰ روز درمان	۷۹
شکل ۸-۴. دوز داروی استفاده شده بوسیله‌ی الگوریتم سارسا در ۳۶۰ روز درمان	۷۹
شکل ۹-۴. جمعیت ویروس‌های آزاد طی درمان در ۵۰۰ روز با استفاده از الگوریتم یادگیری-Q برای ۳ بیمار با پارامترهای متفاوت	۸۱
شکل ۱۰-۴. جمعیت سلول‌های آلوده طی درمان در ۵۰۰ روز با استفاده از الگوریتم یادگیری-Q برای ۳ بیمار با پارامترهای متفاوت	۸۱
شکل ۱۱-۴. جمعیت سلول‌های سالم طی درمان در ۵۰۰ روز با استفاده از الگوریتم یادگیری-Q برای ۳ بیمار با پارامترهای متفاوت	۸۲
شکل ۱۲-۴. جمعیت CTL‌های پیشرو طی درمان در ۵۰۰ روز با استفاده از الگوریتم یادگیری-Q برای ۳ بیمار با پارامترهای متفاوت	۸۲
شکل ۱۳-۴. جمعیت CTL‌های عملگر نهایی آزاد طی درمان در ۵۰۰ روز با استفاده از الگوریتم یادگیری-Q برای ۳ بیمار با پارامترهای متفاوت	۸۳

- شکل ۴-۱۴. جمعیت ویروس‌های آزاد طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q و سارسا برای بیمار ۲ ۸۵
- شکل ۴-۱۵. جمعیت سلول‌های سالم طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q و سارسا برای بیمار ۲ ۸۵
- شکل ۴-۱۶. جمعیت سلول‌های آلوده طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q و سارسا برای بیمار ۲ ۸۶
- شکل ۴-۱۷. جمعیت CTL‌های پیشرو طی درمان در ۳۶۰ روز با استفاده از الگوریتم‌های یادگیری-Q و سارسا برای بیمار ۲ ۸۶
- شکل ۴-۱۸. دوز داروی استفاده شده بوسیله‌ی الگوریتم یادگیری-Q در ۳۶۰ روز درمان برای بیمار ۲ ۸۷
- شکل ۴-۱۹. دوز داروی استفاده شده بوسیله‌ی الگوریتم سارسا در ۳۶۰ روز درمان برای بیمار ۲ ۸۷

جدول ۱-۳. پارامترهای مدل پرلسون و نلسون	۴۸
جدول ۲-۳. پارامترهای مدل دسوزا	۴۹
جدول ۳-۳. متغیرهای حالت مدل وودارز-نوواک	۵۲
جدول ۱-۳. پارامترهای مدل وودارز-نوواک	۵۵
جدول ۱-۴. مقادیر پارامترهای مدل وودارز-نوواک برای ۴ بیمار	۷۰
جدول ۲-۴. مقادیر نقطه تعادل سوم برای ۴ بیمار	۷۱
جدول ۳-۴. تعریف حالت‌ها	۷۳
جدول ۴-۴. حالت-عمل	۷۴
جدول ۵-۴. مقادیر شرایط اولیه	۷۶
جدول ۶-۴. مقایسه دو الگوریتم یادگیری Q و سارسا در کنترل جمعیت ویروس‌ها	۸۳
جدول ۷-۴. مقایسه دو الگوریتم یادگیری Q و سارسا در مقدار دوز داروی استفاده شده	۸۴

فصل اول

((مقدمه))

۱-۱ پیش‌گفتار

سندرم اکتسابی نقص ایمنی (ایدز) که ناشی از ویروس HIV می‌باشد به خاطر انتشار سریع، گستردگی و شدت اثر آن در تاریخ بشر منحصر به فرد است. از زمانی که اولین مورد بیماری ایدز در سال ۱۹۸۱ (۱۳۶۰ شمسی) تشخیص داده شد، جهان تلاش کرده است تا با ابعاد گسترده و همه‌گیر این بیماری مبارزه کند. تلاش‌های اولیه برای پاسخ مناسب به گسترش بیماری بسیار تدریجی، منطقه‌ای، محدود و کاملاً وابسته به امکانات و وضعیت مالی کشورها بوده است. کشورهای کمی خطر پیش‌رو را تشخیص داده بودند و تعداد اندکی از آنها قادر بودند عکس‌العمل موثری نشان دهند. نتیجه آن شد که در طی ۲۹ سال، یعنی تا سال ۲۰۰۹ تقریباً ۳۶ میلیون نفر جان خود را در اثر ابتلا به این بیماری از دست دادند [۱]. اپیدمی HIV به رشد خود در سراسر جهان ادامه می‌دهد، زندگی افراد را نابود می‌کند و در بسیاری از موارد ساختار اصلی جوامع را از بین می‌برد و متلاشی می‌کند. درمان مبتلایان به ویروس HIV یکی از مهمترین مسائلی می‌باشد که بشر به دنبال آن می‌باشد.

وقتی فردی به ویروس HIV آلوده می‌شود، تا تبدیل این آلودگی به بیماری ایدز، بدون درمان دارویی حداکثر ۱۰ سال طول می‌کشد برای درمان این آلودگی تاکنون هیچ درمان قطعی وجود نداشته است و داروهایی که برای این نوع بیماران توسط پزشک تجویز می‌شود، تنها در جهت به تاخیر انداختن مدت زمان ظهور بیماری ایدز می‌باشد؛

ضمن اینکه این داروها خود دارای اثرات سوء بر روی بیماران می‌باشند و از نظر هزینه نیز جزو داروهای گران‌قیمت محسوب می‌شوند. اگر بتوان دوز دارو را به گونه‌ای بهینه برای این بیماران بدست آورد آنگاه هم می‌توان امید داشت که بیماران مدت زمان بیشتری به زندگی خود ادامه دهند و هم از اثرات سوء داروهایی که برای این نوع بیماران تجویز می‌شود کاسته شود. [۲ و ۳]

مطالعه بیماری‌های مسری، از دیدگاه ریاضی، به منظور ارزیابی و کنترل بیماری انجام می‌گیرد. مدل‌های ریاضی متفاوتی برای مدل کردن دینامیک این ویروس در بدن انسان تعریف شده است که مهمترین این مدل‌ها، مدل نوک و همکارنش و مدل پرلسون و همکارنش می‌باشد همچنین مدل‌های ریاضی متفاوتی در رابطه با تاثیر داروهای ضد ویروس HIV تعریف شده است. بدست آوردن یک دوز بهینه دارو برای مهار این ویروس و اثراتش از موضوعات مورد علاقه آکادمیک می‌باشد تا هم بتوانند این ویروس را مهار کنند و هم از اثرات سوء داروهای ضد ویروس بکاهدند و هم اینکه هزینه‌های مادی درمان را کاهش دهند. روش‌هایی که اکنون برای درمان این بیماری در مراکز پزشکی مورد استفاده قرار می‌گیرد روش‌های زیر بهینه می‌باشد.

۱-۲ توصیف مسئله: کنترل جمعیت ویروس HIV در بدن

همانطور که بیان شد بدست آوردن یک دوز بهینه دارو برای مهار ویروس HIV و اثراتش از موضوعات مورد علاقه آکادمیک می‌باشد تا هم بتوانند این ویروس را مهار کنند و هم از اثرات سوء داروهای ضد ویروس بکاهدند و هم اینکه هزینه‌های مادی درمان را کاهش دهند. روش‌هایی که برای درمان این بیماری در مراکز پزشکی اکنون مورد استفاده قرار می‌گیرد روش‌های زیر بهینه می‌باشد. تعیین روشی که هم بتواند این ویروس را کنترل و هم از اثرات سوء داروهای ضد ویروس بکاهد و هم با دقت بالایی با داده‌های واقعی پزشکی مطابقت داشته باشد کاری مشکل می‌باشد.

یادگیری تقویتی در مواردی در علم پزشکی کاربرد داشته است. روش‌های هوشمند ویژگی‌هایی دارند که استفاده از آنها در این دسته از مسائل بسیار مورد توجه است. یکی از این ویژگی‌ها توانایی آنها در حل مسائل مبهم و

در مجموعه داده‌های نویزی است. از آنجا که یادگیری تقویتی به مدل محیط نیاز ندارد در کاربردهای فراوانی مورد توجه قرار گرفته است که از مهمترین این موارد علم پزشکی می‌باشد.

در اینجا کنترل جمعیت ویروس‌های HIV به کمک داروهای شناخته شده در علم پزشکی با استفاده از روش‌های حل مسئله یادگیری تقویتی مورد بررسی قرار گرفته است. برای کمتر شدن اثرات سوء داروها نیز سعی بر این است تا دوز دارو با توجه به جمعیت ویروس‌ها به حداقل ممکن کاهش یابد.

۱-۳ مروری بر کارهای انجام شده توسط الگوریتم یادگیری تقویتی در علم پزشکی

به کمک یادگیری تقویتی، تعیین دوز بهینه دارو در مواردی انجام شده است که از آن جمله می‌توان به تعیین دوز بهینه دارو برای بیماری کم‌خونی [۴]، تعیین دوز بهینه داروی شیمی درمانی برای بیماران سرطانی [۶ و ۵]، کنترل سطح گلوکز خون در بیماران دیابتی نوع ۱ [۷ و ۸] و کنترل دوز دارو برای تعیین عمق بیهوشی برای بیمارانی که تحت عمل جراحی می‌خواهند قرار بگیرند [۹] اشاره کرد. نکته‌ی جالبی که در تعیین دوز بهینه‌ی دارو برای تمام موارد فوق ذکر شد این می‌باشد که همه‌ی موارد فوق تنها از روش Q-learning در حل مسئله یادگیری تقویتی بهره برده‌اند و از روش‌های دیگر یادگیری تقویتی در تعیین دوز بهینه‌ی دارو استفاده نکرده‌اند.

از یادگیری تقویتی در سایر مسایل پزشکی نیز کم و بیش استفاده شده است. استفاده از یادگیری تقویتی در دستگاه CRT^۱ [۱۰ و ۱۱] و کنترل نرم سیگنال‌های فیزیولوژیکی انسان [۱۲] از جمله این موارد می‌باشد. در مرجع [۱۳] کنترل سیگنال GSR^۲ بوسیله‌ی یادگیری تقویتی با استفاده از تحریک موزیکال و بوسیله‌ی اندازه‌گیری GSC^۳ مطرح شده است. الگوریتم یادگیری تقویتی به عنوان یک تکنیک برای طراحی الگوریتم کنترل برای سیستم‌های واسط مغز-کامپیوتر^۴ (BCI) نیز استفاده شده است [۱۳].

^۱ Cardiac resynchronization therapy

^۲ Galvanic Skin Resistance

^۳ Galvanic Skin Conductance

^۴ Brain Computer Interfaces(BCI)

در [۱۴] یک پروتکل مسیریابی بر پایه یادگیری تقویتی که از QoS^۱ نیز پشتیبانی می‌کند برای شبکه‌های سنسور زیست پزشکی پیشنهاد شده است^۲ (RL-QRP). در RL-QRP خط مشی مسیریابی بهینه از طریق تجربه‌ها و پاداش‌ها بدون داشتن اطلاعات خاصی از شبکه قابل دستیابی می‌باشد. در [۱۵] روشی برای قطعه‌بندی تصاویر پزشکی به کمک یادگیری تقویتی معرفی شده است. در [۱۶] از یادگیری تقویتی برای کنترل حرکات چشم استفاده شده است. در [۱۶] یک مدل مغزی چشم‌ها را برای جستجو و نمونه‌برداری در محیط هدایت می‌کند، و اطلاعاتی که از محیط بدست می‌آید برای بهبود مدل مغزی استفاده می‌شود.

۱-۴ مروری بر کارهای انجام شده در رابطه با کنترل عفونت ویروس HIV

پس از شناسایی ویروس HIV و بیماری ایدز افراد زیادی سعی در بدست آوردن مدل دقیقی از این عفونت و بیماری کردند و مدل‌های ریاضی زیادی نیز در این ارتباط تعریف شده است این مدل‌ها به دو دسته‌ی مدل‌های قطعی و استوکستیک تقسیم می‌شوند که در فصل سه در مورد این مدل‌ها، مزایا و معایب هر کدام به صورت مفصل توضیح داده خواهد شد و پارامترهای این مدل‌ها نیز توضیح داده می‌شود. در ادامه کارهایی که محققین و دانشمندان در رابطه با این عفونت انجام داده‌اند مطرح می‌شود.

مکلین و نوواک^۳ [۱۷] از اولین افرادی بودند که پس از مطرح شدن مدل‌های اولیه‌ای در رابطه با این عفونت، شروع به تحلیل ریاضی درمان‌ها و داروهای موجود پرداختند این دو پس از بررسی که بر روی داروها انجام دادند مدلی را مطرح کردند که در آن نقش یک نوع ویروس جدید در بدن انسان را که در برابر درمان از خود مقاومت نشان می‌دهد را نیز در نظر گرفته بودند. اگر^۴ [۱۸] و کوجوکارو^۵ و اگر [۱۹] تاثیر شیمی درمانی را بر روی سلول‌های سالم دستگاه ایمنی بدن بررسی کردند و سپس روشی برای کمینه کردن اثر جانبی داروهای شیمی درمانی

^۱ Quality of Service

^۲ Reinforcement learning based routing protocol with QoS-support

^۳ McLean and Nowak

^۴ Agur

^۵ Cojocararu

در بیماران مبتلا به ویروس HIV ارائه کردند. کریشنر و وب^۱ [۲۰] نیز دوز دارو در مراحل اولیه عفونت را به گونه‌ای بهینه کردند تا از اثرات سوء داروهای شیمی درمانی کاسته شود. سپس کریشنر و همکارانش [۲۱] با استفاده از مدل مطرح شده توسط پرلسون^۲ [۲۲] و مدل پرلسون و کریشنر [۲۳] که روابط بین سیستم ایمنی بدن، ویروس HIV و شیمی درمانی در مراحل اولیه عفونت را توضیح می‌دهد دوز دارو را برای این بیماران با استفاده از تئوری کنترل بهینه، بهینه کردند.

دسوزا و همکارانش^۳ [۲۴] با استفاده از مدل دینامیکی مطرح شده توسط نوواک و همکارانش^۴ [۲۵] مقدار بار ویروسی و اثرات سوء دارو را به کمک تئوری کنترل بهینه به صورت بهینه کنترل کرده‌اند. تابع هزینه‌ای که دسوزا در نظر گرفت به صورت رابطه‌ی (۱-۱) می‌باشد:

$$J[m(\cdot)] = \Psi \frac{V^{\lambda}(t_f)}{\gamma} + \int_{t_0}^{t_f} \frac{m^{\lambda}(t)}{\gamma} dt \quad (1-1)$$

در معادله فوق $m(t)$ نشان دهنده دوز دارو می‌باشد و $v(t)$ نشان دهنده‌ی بار ویروسی می‌باشد. تابع هزینه پیشنهاد شده توسط دسوزا سعی در برقرار کردن تعادل مابین اثرات سوء دارو (جمله‌ی انتگرالی) و بار ویروسی دارد.

برندت و چن^۵ [۲۶] با استفاده از مدل مطرح شده توسط دسوزا [۲۷] یک برنامه دارویی را با توجه به کنترل فیدبک خارجی ارائه داده‌اند هدف از درمان دارویی مطرح شده هم کمینه کردن مقدار بار ویروسی و هم دوز دارو می‌باشد تا از اثرات سوء دارو نیز کاسته شود.

کوبیاک و همکارانش^۶ [۲۸] بر روی مدل مطرح شده توسط وودارز^۷ و نوواک [۲۹] اصلاحاتی انجام دادند و پارامتر v که نشان دهنده بار ویروس می‌باشد را وارد مدل کردند. آن‌ها برای وارد کردن بار ویروس در مدل، بار

^۱ Kirschner and Webb

^۲ Perelson

^۳ De Souza et al.

^۴ Nowak et al.

^۵ Brandt and Chen

^۶ Kubiak et al.

^۷ Woodarz

ویروس را در یک بازه‌ی زمانی به صورت اختلاف ویروس‌های تازه تولید شده، که از مرگ سلول‌های آلوده‌ی دستگاه ایمنی بدن حاصل می‌شوند و نرخ مرگ ویروس‌ها (ویروس‌های با عمر کوتاه) تعریف کردند. سپس در این مدل حساسیت پارامترها را بدست آوردند و بیان کردند که کدام یک از پارامترها در این مدل بیشترین حساسیت را دارا می‌باشند.

استفورد و همکارانش^۱ [۳۰] نشان دادند که مدل پرلسون [۲۲] بیشتر برای مراحل اولیه بیماری مناسب می‌باشد و برای بیان مراحل بعدی عفونت مناسب نمی‌باشد دی بوئر و همکارانش^۲ [۳۱] مدل‌های مختلفی را از دیدگاه‌های مختلفی بررسی کردند و بیان کردند از دیدگاه تئوری تقریباً این مدل‌ها به صورت یکسان رفتار می‌کنند.

کاتچ و گورفیل [۳۲] نحوه تاثیر دارو بر روی مدل پرلسون و کریشنر را بررسی کردند و تغییراتی در مدل ایجاد کردند. همانطور که در فصل بعد مفصل توضیح داده خواهد شد داروهایی که برای درمان و جلوگیری از پیشرفت ویروس‌ها در بدن مورد استفاده قرار می‌گیرند پس از مدتی اثر خود را از دست می‌دهند. به بیان دیگر بعضی از ویروس‌ها در برابر داروها از خود مقاومت نشان می‌دهند کاتچ و گورفیل این ویروس‌ها را نیز در مدل دینامیک در نظر گرفتند و به این ترتیب مدل جدیدی را مطرح کردند. لدزویچ و شاتلر^۳ [۳۳] کنترل بهینه شیمی درمانی برای مدل ریاضی HIV را به کمک L_1 و L_2 انجام داده‌اند.

آدامز و همکارانش [۳۴] ابتدا یک مدل ریاضی که برهم‌کنش سیستم ایمنی بدن و ویروس HIV را نشان می‌دهد فرموله کردند سپس یک استراتژی درمانی برای بیماران مبتلا به HIV به کمک کنترل بهینه و با استفاده از دو داروی RTI و PI بدست آوردند و همچنین با استفاده از ایده‌های برنامه‌ریزی پویا ساختار درمانی جدیدی تحت عنوان وقفه‌های درمانی ساختار یافته زیر بهینه^۴ (STI) را تعریف کردند. باجاریا و همکارانش [۳۵] شبیه‌سازی عددی با توجه به این ساختار درمانی (STI) ارائه کردند.

^۱ Stafford et al.

^۲ De Boer et al.

^۳ Ledzewicz and Schattler

^۴ Structured Treatment Interruptions

اوآتارا^۱ [۳۶ و ۳۷] روشی برای تقریب پارامترهایی که در مدل ریاضی عفونت HIV-۱ می‌باشد را ارائه کرده است. او در یک بازه‌ی زمانی طولانی (حدود ۶ سال) داده‌های کلینیکی موجود برای ۲ بیمار را مورد بررسی قرار داده است و پارامترهای موجود در مدلی که برای بیماری‌های عفونی تعریف می‌شود را بدست آورده است البته افراد دیگری نیز پارامترهای مدل‌های متفاوت دیگری از عفونت HIV را نیز بدست آورده‌اند که اکثر آن‌ها برای بدست آوردن پارامترها داده‌های کلینیکی را در مدت کوتاهی مورد مشاهده قرار داده‌اند. ژیا^۲ [۳۸] با اندازه‌گیری خروجی، پارامترهای موجود در مدل پایه عفونت HIV را که توسط نوواک و می [۳۹] ارائه شده بود را تخمین زد و سپس عملکرد ترکیبی هر کدام از داروها را یافته است و سپس در یک مطالعه با بدست آوردن ضرایب مختلف برای بیمار با توجه به داده‌های کلینیکی فهمیده است که کدام یک بهتر است.

شوژی و همکارانش^۳ [۴۰] با بررسی خصوصیات دینامیکی سیستم، سیستم اصلی را ابتدا به دو زیر سیستم تقسیم کردند سپس یک کنترلر غیر خطی سراسری بوسیله‌ی طراحی کنترلر بر روی دو ناحیه مکمل هم طراحی کرده‌اند. این دو کنترلر شامل یک کنترلر محلی بر روی یک محدوده‌ی مشخص و یک کنترلر سراسری بر روی ناحیه‌ی مکمل آن می‌باشد. بیافور و آتلیس^۴ [۴۱] یک استراتژی کنترل برای بیماران مبتلا به HIV بر اساس تئوری کنترلر غیر خطی مطرح کردند. این استراتژی بر روی مدل دینامیکی بیماری HIV که توسط نوواک و می [۳۹] مطرح شده بود به کار برده شد. نتیجه‌ای که بیافور و همکارش گرفتند نشان دهنده‌ی این بود که کنترلر غیر خطی توسعه داده شده، در کمینه کردن بار ویروسی در مراحل اولیه‌ی عفونت و دوره‌ی بدون نشانه و عوارض^۵ عفونت موثر می‌باشد.

^۱ Ouattara

^۲ Xia

^۳ Shuzhi

^۴ Biafore and Attellis

^۵ Asymptomatic period