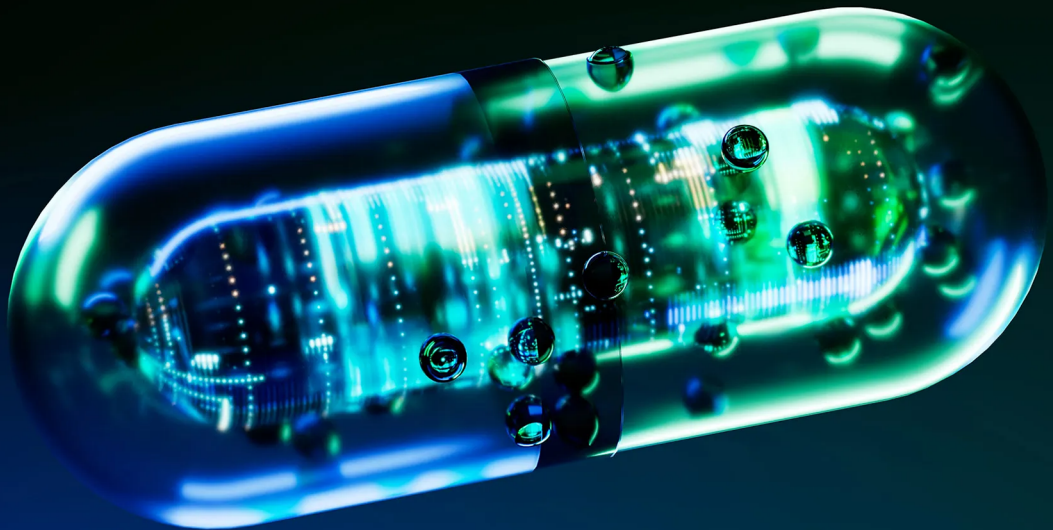


# ZHIVAR

ژیوار / هفته نامه انجمن علمی زیست شناسی دانشگاه شاهد / شماره ۳۱ / هفته چهارم بهمن ۱۴۰۴

اخبار و تازه های زیست شناسی

ماشین لرزینگ و هوش مصنوعی:  
جادوی تحول در ساخت دارو و پروتئین ها









ژیوار، واژه ای ایرانی به معنای زندگی و حیات است...

## صاحب امتیاز: انجمن علمی زیست شناسی دانشگاه شاهد

مدیر مسئول و سردبیر:

سید علی حسینی

دانشجوی کارشناسی بیوتکنولوژی دانشگاه شاهد



ناظر ارشد علمی نشریه:

خانم دکتر طوبی السادات احمدی

عضو هیئت علمی گروه زیست شناسی دانشگاه شاهد



مشاور علمی:

مهدی ادریسیان

دانش آموخته ی کارشناسی ارشد بیوتکنولوژی میکروبی دانشگاه شاهد



مدیر فنی و صفحه آرا:

محمد صدرا محمدی

دانشجوی کارشناسی ارشد ژنتیک دانشگاه آزاد تهران مرکز



سرپرست کارگروه ویراستاری:

محمد ابراهیمی آشتیانی

دانشجوی کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد



## شورای سردبیری:

محمد صدرا محمدی – ناظر فنی و ویراستار

دانشجوی کارشناسی ارشد ژنتیک دانشگاه آزاد تهران مرکز



مهديه حبيبي – سرپرست بخش مطالعات میان رشته‌ای

دانشجوی دکتری نانوشیمی دانشگاه تبریز



الهام کندی – سرپرست خبرنامه‌ی زیوار

دانشجوی کارشناسی زیست‌شناسی سلولی مولکولی دانشگاه شاهد



## هیئت تحریریه:

حنانه دومهری – تحریریه‌ی خبرنامه

دانشجوی کارشناسی زیست‌شناسی سلولی مولکولی دانشگاه شاهد



فاطمه علی‌رضایی – تحریریه‌ی خبرنامه

دانشجوی کارشناسی زیست‌شناسی سلولی مولکولی دانشگاه شاهد



ایلیانکوئی – تحریریه‌ی خبرنامه

دانشجوی کارشناسی بیوتکنولوژی دانشگاه شاهد



فاطمه سبزی

دانشجوی کارشناسی بیوتکنولوژی دانشگاه شاهد



# فهرست

---

۶ اخبار و تازه ها

۱۱ میان رشته‌ای

ماشین لرنینگ و هوش مصنوعی:  
جادوی تحول در ساخت دارو و پروتئین ها





## پایش همزمان فعالیت نورونی و بیان ژن در انسان

نوشته شده توسط خانه دومه‌ری



بالینی و تحقیقاتی، امکان دسترسی به بافت زنده مغز را فراهم کرد. پس از این موفقیت، گروه علمی کوشید طراحی پژوهش را تا حد ممکن دقیق کند؛ از انتخاب ناحیه قشر پیش‌پیشانی گرفته تا هماهنگ‌سازی سن، جنس و وضعیت عصبی نمونه‌ها و همچنین افزایش تعداد نمونه‌ها به صدها مورد برای افزایش اعتبار داده‌ها. نتایج نخستین مطالعه شگفت‌انگیز بود: تفاوت‌های چشمگیری در بیان ژن میان بافت زنده و پسامرگ دیده شد؛ تفاوتی که از نگاه چارنی طبیعی است، زیرا «مغز زنده در حال فعالیت است و مغز مرده نه». پژوهش‌ها سپس به بررسی تفاوت‌ها در پردازش RNA، بیان پروتئین‌ها و تنوع سلولی با ابزارهای آمیگی پرداختند. این رویکرد امکان مشاهده پدیده‌هایی را فراهم می‌کند که در نمونه‌های پسامرگ اساساً قابل بررسی نیستند. حتی می‌توان فعالیت الکتریکی مغز را هنگام انجام یک تکلیف ثبت کرد و سپس این الگوها را مستقیماً با بیان ژن مقایسه نمود. پروژه مغز زنده دریچه‌ای پایدار به سوی نسل آینده علوم اعصاب می‌گشاید. امروزه ایمپلنت‌های مغزی در حال توسعه‌اند تا بسیاری از اختلالات عصبی و روان‌پزشکی را درمان کنند و الگوریتم‌های هوش مصنوعی برای کارایی بیشتر، به داده‌هایی واقعی از مغز زنده نیاز دارند. چنین مجموعه داده‌ای می‌تواند نقشی اساسی در آموزش این سامانه‌ها ایفا کند. با اینکه پرسش اولیه پروژه اکنون پاسخ گرفته است، چارنی می‌گوید: «ما تازه اول راهیم؛ آنچه می‌توان از این روش آموخت، بسیار فراتر از چیزی است که تاکنون دیده‌ایم».

منبع خبر:

Bradford, S., PhD. (2025, November). How brain gene expression changes between life and death. The Scientist. Retrieved from <https://www.the-scientist.com/how-brain-gene-expression-changes-between-life-and-death-73728>

پس از انتشار «پروژه ژنوم انسان»، دانشمندان امکان یافتن لایه‌های ژنتیکی و مولکولی بدن انسان را با دقتی بی‌سابقه بررسی کردند. پژوهش‌های آمیگی - از ترنسکرپتوم تا پروتئوم - بینش‌های تازه‌ای درباره سلامت و بیماری فراهم کرده‌اند. اما در علوم اعصاب، بیشتر این مطالعات بر نمونه‌های مغزی پس از مرگ تکیه داشته و این پرسش را ایجاد کرده است که آیا یافته‌ها واقعاً بازتابی درست از کارکرد مغز زنده هستند یا نه. برای یافتن پاسخی روشن، الکساندر چارنی، روان‌پزشک و ژنتیک‌دان، و برایان کوپل، جراح مغز و اعصاب از مدرسه پزشکی آیکان، پروژه‌ای مشترک را آغاز کردند: «پروژه مغز زنده». چارنی که در سال 2012 سرگرم مطالعه عوامل ژنتیکی اسکیزوفرنی بود، متوجه شد تقریباً همه پژوهش‌ها روی نمونه‌های پسامرگ انجام می‌شود؛ موضوعی که می‌تواند مسیر توسعه درمان‌های روان‌پزشکی را تحت تأثیر قرار دهد. او به این فکر افتاد که شاید بتوان از بافت مغز بیمارانی که جراحی تحریک عمقی مغز (DBS) انجام می‌دهند، نمونه زنده به دست آورد. هرچند تلاش‌های نخستین آن‌ها برای استفاده از سلول‌های باقی‌مانده روی ابزارهای جراحی نتیجه نداد، کوپل روش ساده‌تری پیشنهاد کرد: در جراحی DBS، پیش از قرار دادن الکتروده، بخشی از بافت سوزانده می‌شود و می‌توان درست پیش از آن، نمونه کوچکی برداشت بی‌آنکه خطری متوجه بیمار شود. همین هم‌افزایی میان نگاه

## کشف سرخ‌های ام‌اس در خون، سال‌ها پیش از ظهور علائم بالینی

نوشته شده توسط فاطمه علیرضایی

حمله می‌کند. یافته‌های آنها که در مجله Nature Medicine منتشر شده است، بینش‌هایی در مورد توالی وقایع در طول توسعه ام‌اس و نشانگرهای زیستی بالقوه برای تشخیص زودهنگام بیماری ارائه می‌دهد. هلن تزلت، محقق ام‌اس در دانشگاه بریتیش کلمبیا که در این مطالعه مشارکتی نداشته است، گفت: این یافته‌ها بسیار هیجان‌انگیز و جذاب هستند. آن‌ها از نظر بیولوژیکی به ما می‌گویند که قبل از شروع علائم ام‌اس چه اتفاقی ممکن است رخ دهد و نشان می‌دهند که ما فرصتی داریم تا ام‌اس را زودتر تشخیص دهیم. برای رسیدن به تصویری واضح‌تر از نشانه‌های مولکولی و سلولی در بیماری ام‌اس، عبدالحق و همکارانش بیان پروتئین‌ها را در مراحل مختلف بیماری بررسی کردند. آن‌ها از نمونه‌هایی از مخزن سرم وزارت دفاع که طی آزمایش اجباری اچ‌آی‌وی در میان سربازان تازه استخدام‌شده ارتش ایالات متحده از سال 1989 به دست آمده بود، استفاده کردند. آری گرین، یکی از نویسندگان این مطالعه و دانشمند پزشکی نوروایمونولوژی در UCSF، گفت: چند سال پیش، برخی از محققان متوجه شدند که این یک مجموعه داده واقعاً شگفت‌انگیز است. او توضیح داد که جانبازان تحت مراقبت‌های بهداشتی مداوم قرار می‌گیرند، بنابراین نمونه‌های آنها را می‌توان در طول سال‌ها ردیابی کرد. از این گروه، عبدالحق، گرین و همکارانشان نمونه‌های سرم 134 فرد مبتلا به ام‌اس و نزدیک به 80 فرد بدون این بیماری را تجزیه و تحلیل کردند. بررسی بیش از 5000 پروتئین در این نمونه‌ها نشان داد که سرم افراد این دو گروه پروتئین‌های متمایزی را بیان می‌کند. محققان تقریباً هفت سال قبل از تشخیص بیماری در فرد، افزایش ناگهانی سطح پروتئینی به نام گلیکوپروتئین الیگودندروسیت میلین، که شاخصی از آسیب میلین است، را مشاهده کردند. حدود یک سال بعد، آنها افزایش

علائم اولیه بیماری ام‌اس سال‌ها قبل از بروز علائم در خون ظاهر می‌شوند. پروتئومیکس به کشف امضاهای مولکولی بیماری ام‌اس تا هفت سال قبل از شروع علائم کمک کرد و راه را برای تشخیص و درمان زودهنگام هموار ساخت. مولتیپل اسکلروزیس (MS)، یک اختلال عصبی خودایمنی که در آن میلین پوشاننده نورون‌ها آسیب می‌بیند، بیش از 1.8 میلیون نفر را در سراسر جهان تحت تأثیر قرار می‌دهد. دمیلینه شدن، انتقال سیگنال در سیستم عصبی را مختل می‌کند و در نهایت می‌تواند باعث تجزیه فیبرهای عصبی یا آکسون‌ها شود و در نتیجه طیف وسیعی از علائم جسمی و روانی را ایجاد کند. با این حال، شواهد فزاینده‌ای نشان می‌دهد که تغییرات محرک MS مدت‌ها قبل از بروز علائم شروع می‌شوند. در واقع، هنگامی که احمد عبدالحق، متخصص مغز و اعصاب در دانشگاه کالیفرنیا، سانفرانسیسکو (UCSF) و همکارانش سرم افراد مبتلا به MS را مطالعه کردند، سال‌ها قبل از شروع علائم، اتوانتی‌بادی‌های متمایزی را مشاهده کردند. عبدالحق گفت: کنار هم قرار دادن این قطعات پازل در کنار بسیاری از موارد دیگر در مقالات علمی به ما نشان داد که در واقع زمان قبل از شروع بیماری ام‌اس از نظر ایمونولوژیکی بسیار پیچیده است و اتفاقات زیادی در آنجا رخ می‌دهد. این موضوع محققان را بر آن داشت تا وقایع قبل از شروع علائم ام‌اس را بیشتر بررسی کنند. اکنون، عبدالحق و همکارانش با بهره‌گیری از پروتئومیکس با توان عملیاتی بالا، امضاهای پروتئینی مرتبط با ام‌اس را در افراد مبتلا به ام‌اس تقریباً هفت سال قبل از شروع علائم شناسایی کردند، که نشان می‌دهد سیستم ایمنی حتی زودتر از آنچه قبلاً تصور می‌شد به میلین



تکرار و اعتبارسنجی یافته‌های آنها در یک گروه بزرگتر و جداگانه می‌تواند کمک کند، حتی اگر یافتن چنین نمونه‌های بدون علامتی دشوار باشد. شاید در آینده بتوانیم افراد مبتلا به ام‌اس را زودتر از آنچه در حال حاضر تشخیص می‌دهیم، تشخیص دهیم. گرین موافق بود که نمونه‌هایی که آنها تجزیه و تحلیل کردند ممکن است منعکس‌کننده ویژگی‌های جمعیتی کسانی که معمولاً از ام‌اس رنج می‌برند، نباشد. این تیم در حال حاضر برای غلبه بر این محدودیت تلاش می‌کند. گرین گفت: ما همچنان به آزمایش و جستجوی جمعیت‌هایی ادامه می‌دهیم که بتوانیم برخی از تکنیک‌های موجود را در آنها بیشتر ارزیابی و اعتبارسنجی کنیم. با وجود این، گرین خاطرنشان کرد که این نتایج واقعاً مهم هستند و مستلزم مطالعات بیشتر می‌باشند. او امیدوار است که یافته‌های آنها راه را برای تشخیص بیماری در مراحل اولیه که موفقیت درمان بیشتر است، هموار کند. هرچه زودتر برای درمان افراد اقدام کنید، در جلوگیری از ناتوانی آنها موفق‌تر خواهید بود.

پلی‌پیتید سبک نوروفیلانت را مشاهده کردند که نشان دهنده آسیب به آکسون‌ها است. در مقابل، پروتئین اسیدی فیبریلاری گلیال (GFAP)، که نشان دهنده درگیری آستروسیت‌ها است، تنها در زمان تشخیص بالینی افزایش یافت. عبدالحق گفت که این موضوع نسبتاً غیرمنتظره بود. عبدالحق گفت: بخش شگفت‌انگیز این بود که آسیب میلیون هفت سال قبل از شروع علائم یا یک سال قبل از آسیب عصبی مشاهده نشد. بخش شگفت‌انگیز این بود که مشاهده کردیم فعال شدن گلیال، چیزی که با GFAP دیدیم، بسیار نزدیک به شروع بیماری اتفاق می‌افتد و قبل از شروع اصلاً مشهود نبود. علاوه بر این، محققان همچنین تغییراتی را در پروتئوم سرم مشاهده کردند که به دخالت اینترلوکین 3 و سایر پروتئین‌های دخیل در واکنش‌های ایمنی اشاره دارد. برای بررسی اینکه آیا این پروتئین‌های با بیان متفاوت می‌توانند به تعریف یک پنل نشانگر زیستی کمک کنند، عبدالحق و تیمش مدل‌های یادگیری ماشینی را در مجموعه داده‌های خود به کار بردند. از این تعداد، دو مدل توانستند افراد مبتلا به ام‌اس بدون علامت را با ویژگی و حساسیت بالا شناسایی کنند. شواهد بیماری سال‌ها قبل از شروع علائم، ترملت را متعجب نکرد، تا حدودی به این دلیل که تیمش اخیراً دریافته است که در مقایسه با افرادی که ام‌اس ندارند، افراد مبتلا به ام‌اس تا 15 سال قبل از تشخیص، بیشتر به پزشک مراجعه می‌کنند. او خاطرنشان کرد که این دو نتیجه مکمل یکدیگر هستند. ترملت گفت: یافته‌های آنها هیجان‌انگیز است. این واقعاً یک اثبات جالب برای یک اصل است. با این حال، او هشدار داد که تعداد نمونه‌ها کم بوده و از آنجایی که نمونه‌ها بیشتر متعلق به مردان ارتش بودند، ممکن است قابل تعمیم به جمعیت معمولی که از ام‌اس رنج می‌برند، که در زنان به طور قابل توجهی شایع‌تر است، نباشند. او افزود،

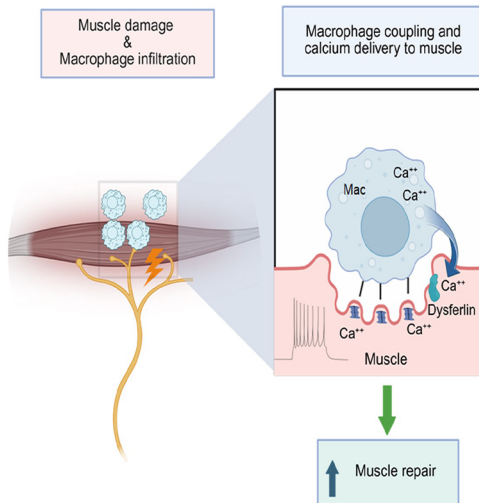
منبع خبر:

<https://www.the-scientist.com/early-signs-of-multi-ple-sclerose-sis-appear-in-blood-years-before-symptoms-73753>

## کشف ترمیم بسیار سریع ماهیچه‌های بدن توسط عاملی باورنکردنی!

نوشته شده توسط ایلینا نکوئی

سیتوکین‌ها و کمیکاین‌ها آزاد می‌کنند که التهاب را هدایت می‌کنند، بر احساس درد تأثیر می‌گذارند و به رشد و بازسازی تارهای عضلانی کمک می‌کنند. اما تیم یانکووسکی کشف کردند که ماکروفاژها فقط سیگنال‌های شیمیایی ارسال نمی‌کنند! بلکه آنها در واقع مانند نورون‌ها رفتار می‌کنند، با فیبرهای عضلانی اتصالات فیزیکی سیناپس‌مانند تشکیل می‌دهند و به‌طور مستقیم یون‌های کلسیم را منتقل می‌کنند.



در مقیاس میکروسکوپی، فرایند ترمیم عضله پیچیدگی‌های بسیاری دارد و روند بهبود نیز به انواع آسیب‌های ممکن بستگی دارد؛ برای مثال بدن در برابر پارگی ناگهانی عضله‌ای که ناشی از یک آسیب ورزشی باشد، واکنشی بسیار متفاوتی نشان می‌دهد تا به نسبت آسیبی مانند بیماری تحلیل تدریجی عضله. حال یک گروه پژوهشی در بیمارستان کودکان سینسیناتی سازوکاری مشترک و شگفت‌انگیز شناسایی کرده‌اند که ممکن است روند بهبود را در انواع مختلف این آسیب‌های عضلانی، تقویت کند. یافته‌های شگفت‌انگیز این پژوهش 30 آبان ماه، در ژورنال *Current Biology* انتشار یافته است. این تحقیق توسط دکتر گیانش تریپاتی، نویسنده اول و دکتر مایکل یانکووسکی، نویسنده مسئول انجام شده است. یانکووسکی مدیر بخش پژوهش در دپارتمان بیهوشی بیمارستان کودکان سینسیناتی است و همچنین به عنوان معاون مدیر پژوهش‌های علوم پایه در مرکز تحقیقات درد کودکان فعالیت می‌کند. فرایند جدید شناسایی‌شده بر نوعی یاخته ایمنی به نام ماکروفاژ تمرکز دارد. این یاخته‌ها معمولاً به خاطر پاکسازی باکتری‌ها، سلول‌های مرده و سایر بقایا شناخته می‌شوند. یانکووسکی می‌گوید: «بزرگ‌ترین شگفتی برای ما این بود که فهمیدیم یک ماکروفاژ ویژگی‌ای شبیه سیناپس نورون را دارد و به واسطه آن، یون کلسیم را به فیبر عضلانی منتقل می‌کند تا پس از آسیب، ترمیم را سهولت بخشد. این دقیقاً مانند روشی است که یک نورون عمل می‌کند، با سرعتی بسیار بالا و با ارتباط سیناپسی کار می‌کند تا فرایند ترمیم را تنظیم کند.» دانشمندان مدت‌هاست می‌دانند که ماکروفاژها در واکنش بدن به آسیب عضلانی نقش دارند. این یاخته‌ها، مولکول‌هایی مانند

تیم یانکووسکی در ابتدا قصد داشتند راهی برای کاهش درد در طول دوره بهبودی پس از جراحی پیدا کنند. آنها به دنبال استراتژی‌های جدیدی بودند تا که جایگزین‌های ایمن‌تری نسبت به داروهای مسکن موجود ارائه دهند. اگرچه این پژوهش مسیر مستقیمی برای کاهش درد پیدا نکرد، اما یک مکانیسم ترمیم سریع را کشف کرد که می‌تواند به پیشرفت درمان‌های مرتبط با کاهش حجم عضله و بهبود آسیب‌های عضلانی، کمک کند. نتایج همچنین نشان می‌دهند که ماکروفاژها می‌توانند به عنوان "وسيله‌های انتقال سلولی" در درمان‌های آینده مبتنی بر سلول، برای طیف وسیع‌تری از بیماری‌ها استفاده شوند.



های تحلیل‌برنده ماهیچه هموار کند، هر چند که هنوز فضای بیشتری برای پژوهش درخصوص محدودیت‌های توانایی ماکروفازها وجود دارد.

یانکووسکی همچنین می‌گوید: «اینها ماکروفازهای نفوذی هستند، نوعی خاص از ماکروفازها. آنها یاخته‌هایی نیستند که از قبل در بافت حضور داشته باشند بلکه تنها پس از ایجاد آسیب است که شروع به کار می‌کنند.» در مجموعه‌ای از آزمایش‌ها با استفاده از مدل‌های موش با دو نوع آسیب متفاوت، پژوهشگران بررسی کردند که ماکروفازها چگونه با تارچه‌های عضلانی تعامل دارند. در واقع، آنها توانستند لحظات کلیدی این فعالیت را به‌صورت زنده و در زمان واقعی ثبت کنند. با استفاده از دوزهای کوتاه یک ماده شیمیایی طراحی‌شده برای تحریک فعالیت، تیم محققان مشاهده کردند که ماکروفازها اتصالاتی سیناپس‌مانند با تارچه‌ها تشکیل می‌دهند. سپس این سلول‌های ایمنی، یون‌های کلسیم را مستقیماً به تارچه‌ها منتقل کردند که باعث فعال سازی می‌شود و روند ترمیم آسیب حاد را تسریع می‌کند. در عرض 10 تا 30 ثانیه، پژوهشگران امواج فعالیت الکتریکی را در عضلات آسیب‌دیده شناسایی کردند. یانکووسکی بیان می‌کند: «این اتفاق با سرعت بسیار بالایی رخ می‌دهد. شما می‌توانید ماکروفاز را فعال کنید و تقریباً بلافاصله، انقباض خفیف ماهیچه را مشاهده کنید.» در همین حال، ماکروفازها به موش‌هایی با بیماری‌های عضلانی مانند تحلیل ماهیچه نیز، کمک کردند. پس از شناسایی آسیب، این یاخته‌های ایمنی به ناحیه آسیب‌دیده هجوم می‌آوردند و موج‌های فعالیت را در تارهای عضلانی ایجاد می‌کردند. پس از 10 روز، موش‌های تحت درمان نسبت به گروه کنترل، تعداد بیشتری از فیبرهای عضلانی ترمیم‌یافته و جدید داشتند. این شواهد ما را به این نتیجه‌گیری سوق می‌دهد که پاسخ سیناپس‌مانند ماکروفازها، در هر دو سناریوی عضلات دچار آسیب فیزیکی و دچار بیماری، عمل می‌کند. محققان امیدوارند این کشف بتواند مسیر را برای درمان‌های نوین عضلانی و بیماری

منبع خبر:

<https://scitechdaily.com/researchers-reveal-a-shocking-ly-fast-muscle-healing-mechanism/#comments>

# ماشین لرنینگ و هوش مصنوعی: جادوی تحول در ساخت دارو و پروتئین نوشته شده توسط فاطمه سبزی

## مقدمه

یادگیری ماشین (Machine Learning) و کاربرد آن در تولید دارو و پروتئین، یکی از حوزه های پرطرفدار این روزهای دنیای علم است و امروزه پژوهش ها و فعالیت های علمی زیادی در این زمینه انجام می شود. دانشمندان در سال های قبل از شکل گیری ماشین لرنینگ، به تحقیقات در حوزه علوم اعصاب پرداخته بودند و با بهره گیری از نحوه تعامل نرون ها و تقویت ارتباط بین آنها، توانستند پایه های حوزه ماشین لرنینگ را تشکیل دهند و تقریباً در دهه 1950 بود که ماشین لرنینگ به طور رسمی در دنیا شکل گرفت. [1]

پیش از ورود ماشین لرنینگ، تولید داروها و پروتئین ها عمدتاً بر پایه روش های آزمایشگاهی سنتی، استخراج طبیعی، سنتز شیمیایی و مهندسی ژنتیک می گرفت که دارای خطاهای بسیار زیادی بودند. به عنوان مثال، برای تولید انسولین از انسولین های حیوانی یا برای تولید آنتی توکسین ها از سرم های حیوانی استفاده می شد. در حوزه مهندسی ژنتیک از فناوری های ساده ای مثل کلونینگ ساده و ویروس ها برای ورود ژن ها به سلول میزبان استفاده می شد که امروزه و با پیشرفت علوم کامپیوتر از این روش ها، کمتر استفاده می شود.

پس از ورود ماشین لرنینگ در عرصه تولید دارو و پروتئین، داروهای جدیدی با روش های نوین ساخته شدند که از اولین آنها می توان به داروی درمان وسواس فکری (OCD) اشاره کرد که در 12 ماه با استفاده از الگوریتم های هوش مصنوعی کشف و بهینه سازی شد و پس از آن روند داروسازی با استفاده از الگوریتم ها افزایش یافت. کاربرد ماشین لرنینگ و الگوریتم های هوش مصنوعی، تنها به تولید دارو منتهی نشد و این الگوریتم ها درکشف و ساخت پروتئین ها مختلف نیز نقش به سزایی داشتند به طوری که با استفاده از ابزارهای مختلف ماشین لرنینگ مانند AlphaFold، روزانه ساختارهای پروتئینی جدیدی پیش بینی و ساخته می شود. [2]

در این مقاله، به مفهوم ماشین لرنینگ و الگوریتم های آن، معرفی روش های تولید دارو و پروتئین با استفاده از ماشین لرنینگ و جدیدترین تحقیقات در این حوزه می پردازیم.

## ماشین لرنینگ چیست؟

ماشین لرنینگ، زیرمجموعه ای هوش مصنوعی است که با استفاده از الگوریتم های پیشرفته و بدون برنامه ریزی صریح، با بررسی داده ها می توانند آموزش ببینند و با یادگیری روند و الگوهای بین داده ها، داده های جدید را پیش بینی کند. برخلاف الگوریتم ها در برنامه نویسی که برنامه نویس با استفاده از کد، به کامپیوتر دستور مستقیم می دهد؛ الگوریتم ها در ماشین لرنینگ، روش یا فرمول های ریاضی ای هستند که بدون دادن دستور مستقیم، به کامپیوتر اجازه می دهند از داده ها یاد بگیرد و راه حل را پیدا کند.



## انواع یادگیری ماشین:

چهار دسته اصلی از انواع ماشین لرنینگ عبارتند از:

1. یادگیری نظارت‌شده (Supervised Learning) در این نوع یادگیری، الگوریتم‌ها با داده‌های برچسب‌خورده آموزش می‌بینند. در واقع در این تکنیک، ابتدا یک سری ورودی و خروجی‌های حاصل از آن داده می‌شود، و الگوریتم با دیدن و یادگیری الگوی بین ورودی و خروجی‌ها، یاد می‌گیرد که با خروجی‌های حاصل از ورودی‌های آینده را پیش‌بینی کند. وظایف رایج: طبقه‌بندی و رگرسیون.

2. یادگیری بدون نظارت (Unsupervised Learning) الگوریتم‌ها با داده‌های بدون برچسب کار می‌کنند و الگو یا ساختار درونی داده را کشف می‌کنند. در این روش بر خلاف یادگیری نظارت‌شده، ما رابطه بین ورودی‌ها و خروجی‌ها را برای الگوریتم مشخص نمی‌کنیم (داده‌های برچسب‌دار نیستند) بلکه خود الگوریتم موظف به پیدا کردن رابطه و الگو بین ورودی و خروجی‌ها و در نهایت یادگیری آنها است. وظایف رایج: خوشه‌بندی، کشف قوانین انجمنی، کاهش ابعاد، شناسایی ناهنجاری‌ها.

3. یادگیری نیمه‌نظارتی (Semi-supervised Learning) ترکیبی از روش‌های نظارت‌شده و بدون نظارت است. در شرایطی که داده‌های برچسب‌دار محدود و داده‌های بدون برچسب فراوان باشند، کارایی بالایی دارد. کاربردها: ترجمه ماشینی، تشخیص تقلب، دسته‌بندی متن.

4. یادگیری تقویتی (Reinforcement Learning) عاملی نرم‌افزاری در یک محیط پویا عمل کرده و با آزمون و خطا و گرفتن پاداش و تنبیه آموزش می‌بیند. حوزه‌های کاربرد: رباتیک، رانندگی خودکار، بازی‌های هوش مصنوعی. [3]

## الگوریتم‌های یادگیری ماشین و دسته بندی آنها:

### 1. طبقه بندی (Classification)

این روش نوعی یادگیری نظارت‌شده است که در برچسب‌گذاری داده‌های ورودی، کاربرد دارد. در این روش، الگوریتم‌ها داده‌های ورودی را در دسته (Class)‌های مختلف، طبقه‌بندی می‌کنند و از این طریق داده‌های ورودی را برچسب‌گذاری می‌کنند.

### الگوریتم‌های مهم:

- بیز ساده (Naive Bayes)
- رگرسیون لجستیک
- K نزدیک‌ترین همسایه (KNN)
- ماشین بردار پشتیبان (SVM)
- درخت تصمیم و جنگل تصادفی
- روش‌های ترکیبی مثل AdaBoost و XGBoost
- شبکه‌های عصبی [3]



## 2. رگرسیون (Regression):

مدلی برای پیش بینی مقدار عددی با توجه به داده ها (پیش بینی متغیر پیوسته بر اساس متغیرهای وابسته) رگرسیون ها می توانند خطی (Linear Regression) و یا غیرخطی (Polynomial Regression) باشند.

## 3. خوشه بندی (Clustering):

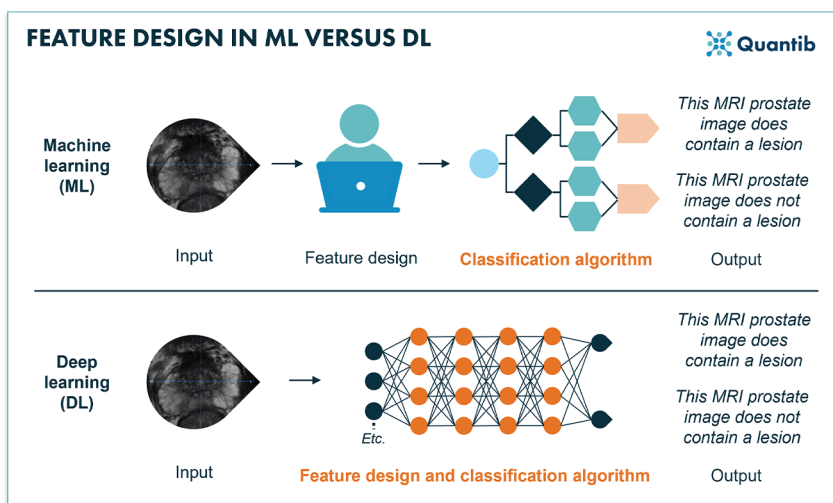
نوعی یادگیری بدون نظارت است که داده های مشابه را در یک گروه قرار می دهد.

### الگوریتم های مهم:

- K-means
- DBSCAN (عموماً برای داده های پراکنده و پرنویز استفاده می شود)
- Hierarchical (خوشه بندی سلسله مراتبی)
- مدل های آماری مثل GMM

## 4. یادگیری عمیق (Deep Learning):

در این نوع یادگیری، ما با مدل های مبتنی بر شبکه های عصبی چندلایه مواجه هستیم. مدل هایی که زیرمجموعه ای از شبکه های عصبی مصنوعی (ANN) هستند. این نوع یادگیری با الگوریتم های خود، به عنوان یکی از پیچیده ترین و پیشرفته ترین بخش های ماشین لرنینگ شناخته می شوند. در حوزه تولید دارو و طراحی پروتئین، به دلیل پیچیده بودن ساختار های دارویی و پروتئینی، بیشتر از این بخش استفاده می شود. [3]

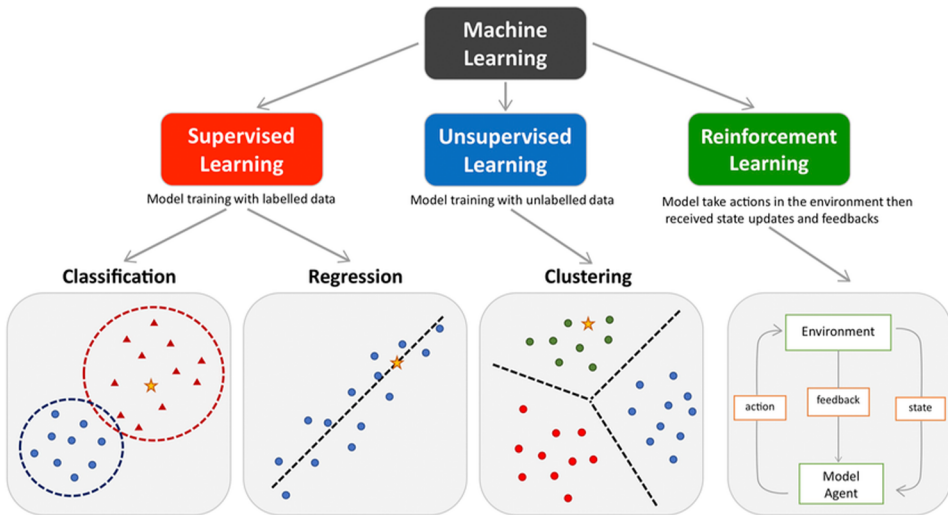


(تصویر 1) تفاوت نحوه یادگیری در یادگیری عمیق با دیگر مدل های یادگیری در ماشین لرنینگ [4]



## مدل های مهم:

- MLP : معماری پایه با لایه های متوالی.
- CNN : معماری تخصصی برای داده های تصویری.
- LSTM-RNN : مناسب برای داده های ترتیبی و سری زمانی.
- مدل های پیشرفته: Autoencoder, RBM, DBN, GAN, Transfer Learning



(تصویر ۲) دسته بندی انواع یادگیری ماشین لرنینگ و مدل های آن [5]

## کاربرد های یادگیری ماشین در طراحی پروتئین ها و تولید دارو:

قبل از معرفی کاربردهای یادگیری ماشین در حوزه طراحی پروتئین و دارو، نیاز است با کمی از اصطلاحات تخصصی این حوزه آشنا شویم: [6]

• **شبکه عصبی (Neural Network):** الگوریتمی الهام گرفته از مغز انسان که با گرفتن داده ها، الگوهای میان آنها را کشف می کند و در زیست شناسی مدل هایی هستند که توانایی یادگیری و پیش بینی ساختار و توالی پروتئین ها را دارا هستند.

• **ترنسفورمر (Transformer):** نوعی معماری خاص از شبکه های عصبی که می تواند با بررسی داده ها، وابستگی های خیلی دور بین قسمت های مختلف داده را کشف کند. عموماً از ترنسفورمرها در پردازش زبان طبیعی مانند ترجمه گوگل، استفاده می شود؛ گرچه در حوزه زیست شناسی، از آن برای کشف ارتباط بین توالی های دور در پروتئین ها استفاده می شود.

• **داکینگ (Docking):** شبیه سازی نحوه قرار گرفتن مولکول های کوچک (مثل دارو) در جایگاه فعال پروتئین و بررسی تطابق با آن جایگاه.



• **نویز (Noise):** نویز، به معنای اختلال و آشفتگی تصادفی در داده ها است. از نویز در مدل یادگیری Diffusion استفاده می شود.

• **مدل انتشار (Diffusion Model):** نوعی الگوریتم که با اضافه کردن نویز به داده ها و پس از آن، یادگیری چگونگی حذف کردن این نویزها از داده ها؛ می تواند داده هایی کاملاً جدید و طبیعی، طراحی کند.

حال به چند نمونه از کاربردهای ماشین لرنینگ در حوزه تولید دارو و ساختارهای پروتئینی، میپردازیم:

### «پیش بینی ساختار پروتئین - نرم افزار AlphaFold»

نرم افزار AlphaFold با استفاده از الگوریتم های یادگیری عمیق (شامل ترنسفورمرها، اجزای Attention و شبکه مخصوص ساختار) با بررسی توالی آمینواسیدی یک پروتئین، اطلاعات تکاملی (MSA) و قالب های ساختاری، می تواند ساختار سه بعدی و حتی نقشه اتمی پروتئین را پیش بینی کند. در نسخه جدیدتر (AlphaFold 3)، بخش های Diffusion هم وارد شده اند تا پیش بینی کمپلکس ها و مولکول های کوچک بهتر شود. گرچه نرم افزار AlphaFold ابزاری عالی برای پیش بینی و طراحی سه بعدی پروتئین ها است اما برای دینامیک های بزرگ، مولکول های چند حالتی یا طراحی مستقیم عملکرد، هنوز کامل نیست و نیازمند به اعتبار سنجی تجربی دارد. [7, 8]

### «پیش بینی ساختار و استخراج ویژگی های پروتئین ها با استفاده از مدل های زبان پروتئینی (PLM) نرم افزار ESM/ESMFold»

Protein Language Models یا به اختصار PML ها، مدل های زبانی مانند ترنسفورمرها (مدل های زبانی طبیعی) هستند که ساختار و معنا را در متن یاد می گیرند. این مدل های زبانی که ESM ها نمونه ای از آنها هستند؛ روی میلیاردها توالی پروتئینی آموزش می بینند و ویژگی های ذاتی بیولوژیک، فضاهای تکاملی و اطلاعات ساختاری در بردارهای آنها رمز می کنند و برای پیش بینی ساختار، عملکرد، اثر جهش ها و طراحی توالی به کار می روند.

در واقع از PML ها هم می توان برای استخراج ویژگی ها برای آموزش مدل های بعدی استفاده کرد؛ هم می توان از مدل هایی مانند (ESMFold) برای پیش بینی مستقیم ساختارهای پروتئینی استفاده کرد. با این حال، باید در نظر گرفت که گرچه این مدل ها برای یادگیری نیازی به برچسب های آزمایشی ندارند و در کارهای اکتشافی ابزارهای قدرتمندی هستند، با این حال به داده های حجیم نیاز دارند و گاه هزینه های محاسباتی آنها زیاد است. [9]



## « طراحی توالی برای یک ساختار مشخص در پروتئین ها با استفاده از شبکه های پیام رسانی گرافی (GNN) \_ الگوریتم ProteinMPNN »

در ریاضیات و علوم کامپیوتر، گراف یعنی یک مجموعه از نقاط (گره یا Node) که با خطوط (یال یا Edge) به یکدیگر وصل شده اند. پروتئین ها و مولکول ها ذاتاً ساختار گرافی دارند که در آن اتم یا آمینواسید معمولاً به عنوان Node و پیوندهای بین آنها به عنوان Edge در نظر گرفته می شوند. GNN ها با استفاده از اطلاعات این ساختارها، می توانند با استفاده از یک ساختار مشخص پروتئینی، توالی های پروتئینی مناسب را پیش بینی کرده و آن را طراحی کنند. برای مثال، الگوریتم ProteinMPNN، از این روش استفاده می کند. GNN ها به دلیل پیچیدگی زیاد، از رگرسیون های خطی به عنوان تکنیک یادگیری استفاده نمی کنند و لازم است از ساختارهای غیرخطی که پیشرفته تر هستند؛ استفاده کنند. این روش گرچه روشی مناسب برای پیش بینی توالی ها پروتئینی است؛ اما نیاز به نمونه های با کیفیت و کنترل آزمایشی دارند. [10]

## « بازسازی داده ها به منظور طراحی ساختار سه بعدی با مدل های انتشار (Diffusion Models) \_ مدل های ProteinGenerator, RFDiffusion »

همانطور که پیشتر اشاره شد، مدل های Diffusion، با اضافه کردن عمدی نویزهای تصادفی، یاد می گیرد که این نویزها را چگونه از داده ها پاک کند و پس از یادگیری، می تواند در مواجهه با داده های مختلف، نویزهای آنها را حذف کند و این داده ها را بازسازی کند. از این فرایند در طراحی پروتئین یا مولکول های سه بعدی به صورت ساختارهای واقع گرایانه و سازگار با هندسه شیمیایی مولکول، استفاده می شود با این حال آموزش و نمونه برداری از نمونه ها می تواند هزینه بر باشد و نیازمند بررسی دقیق برای سازگاری شیمیایی دارد. یکی از نمونه های موفق Diffusion، نمونه RFDiffusion می باشد که با استفاده از شرط گذاری روی خواص مختلف، می تواند پروتئین های نو تولید کند. نمونه دیگر، ProteinGenerator ها هستند که علاوه بر ساختار، می توانند توالی های پروتئینی نیز تولید کنند. [11]

## «تولید و بهینه سازی مولکول ها برای کاربرد های مختلف (مانند تولید دارو) با استفاده از خودرمزگذارها (VAE) \_ نمونه JT-VAE»

VAE ها یک فضای نهان پیوسته می سازند که مولکول ها یا توالی ها را در آن نگاشت می کنند؛ سپس از آن فضا می توان نمونه های جدید (مولوکول های جدید) تولید یا بهینه سازی کرد. JT-VAE یک نسخه ویژه است که مولکول را به زیرساختارها تقسیم می کند (درخت جوشگاه) تا تولید مولکول های شیمیایی معتبر بهتر انجام شود و از آن در تولید مولکول های دارویی برای بهینه سازی خاصیت های آنها (برای مثال حلالیت) استفاده می شود.

گرچه این روش، روشی نو و جالب برای تولید مولکول های جدید است اما در حال حاضر، هنوز در تولید مولکول هایی با خاصیت های مطلوب، چالش ها و محدودیت هایی وجود دارد و ممکن است مولکول های تولیدی از لحاظ سنتز شیمیایی یا سمیت مناسب نباشند. [12]

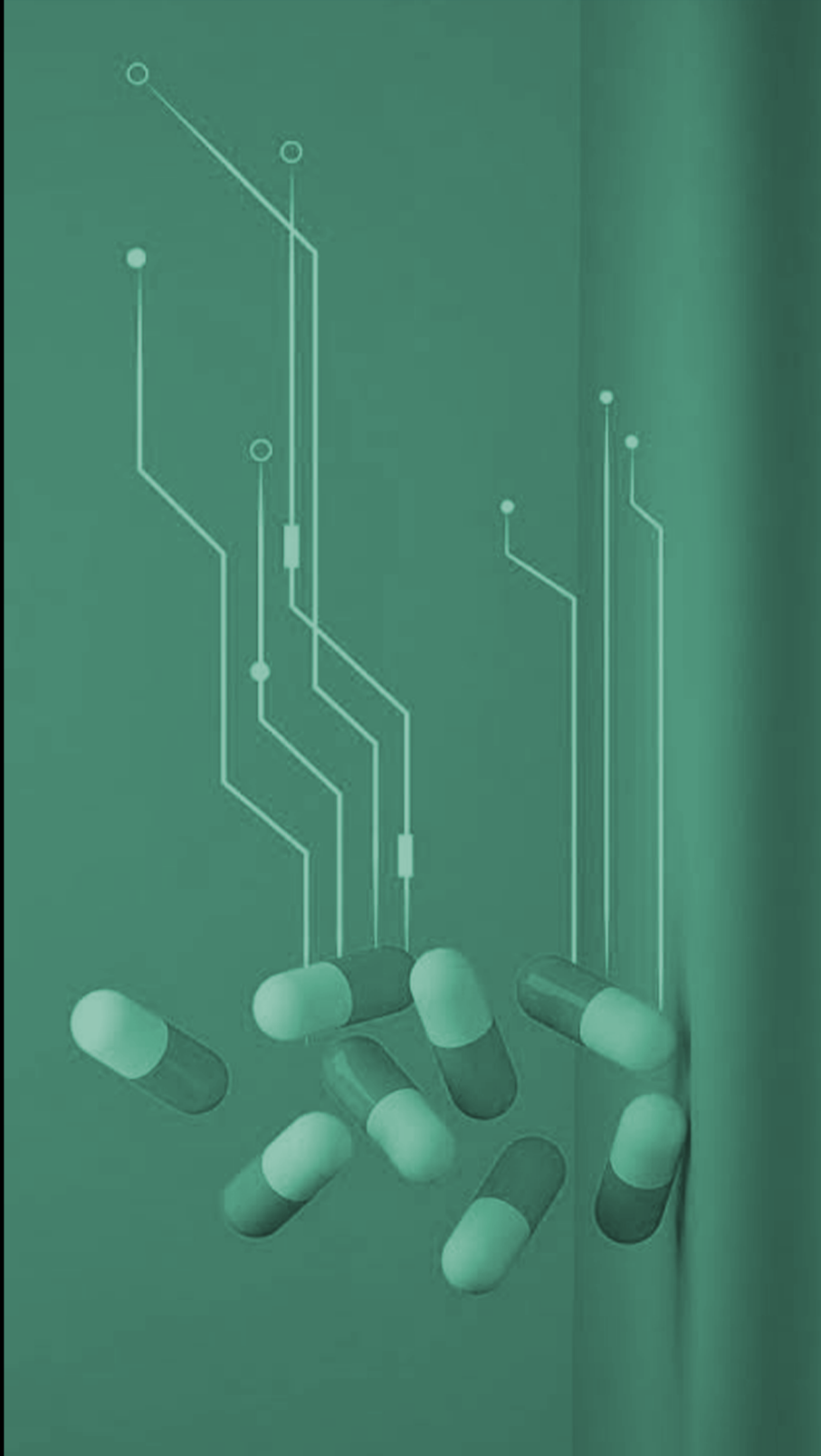


## جمع بندی:

ماشین لرنینگ در سال های اخیر، تبدیل به یکی از مهم ترین و پرفرودارترین مباحث در حوزه هوش مصنوعی تبدیل شده است. مهارتی که یکی از پایه های مهم پژوهش های زیست شناسی است و در تولید پروتئین های جدید و داروها نقش بسزایی ایفا می کند پیشرفت چشمگیر در این حوزه، بسیاری از مشکلاتی که دانشمندان برای تولید این مولکول های اساسی زیستی با آن مواجه بودند را حل کرده است. از تحلیل ساختار مولکول ها و تولید توالی های پروتئینی تا بهینه سازی خواص داروها، همگی با مدل ها و الگوریتم های ماشین لرنینگ امکان پذیر هستند. ترکیب تکنیک های معرفی شده در ماشین لرنینگ، باعث می شود مسیر کشف دارو و طراحی پروتئین سریعتر، کم هزینه تر و دقیق تر از روش های سنتی باشد؛ هر چند هنوز نیاز به آزمایش های آزمایشگاهی و بالینی برای اعتبار سنجی نهایی وجود دارد.

## منابع:

1. Foote, K.D., A Brief History of Machine Learning. Dataversity, December 3, 2021.
2. Review), V.R.D.T., article  
Discovering and designing drugs with artificial intelligence. Drug target review, 4 March 2020.
3. Sarker, I.H., Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci, 2021. 2(3): p. 160.
4. in OBH Energy.
5. Peng, J., et al., Machine Learning Techniques for Personalised Medicine Approaches in Immune-Mediated Chronic Inflammatory Diseases: Applications and Challenges. Frontiers in Pharmacology, 2021. 12.
6. OpenAI. ChatGPT. Available from: <https://chat.openai.com/>.
7. Jumper, J., et al., Highly accurate protein structure prediction with AlphaFold. Nature, 2021. 596(7873): p. 583-589.
8. Abramson, J., et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 2024. 630(8016): p. 493-500.
9. Lin, Z., et al., Evolutionary-scale prediction of atomic-level protein structure with a language model. Science, 2023. 379(6637): p. 1123-1130.
10. Dauparas, J., et al., Robust deep learning-based protein sequence design using ProteinMPNN. Science, 2022. 378(6615): p. 49-56.
11. Watson, J.L., et al., De novo design of protein structure and function with RFdiffusion. Nature, 2023. 620(7976): p. 1089-1100.
12. Wengong Jin, R.B., Tommi Jaakkola. Junction Tree Variational Autoencoder for Molecular Graph Generation. 12 Feb 2018; Available from: <https://arxiv.org/abs/1802.04364v4>.



پل‌های ارتباطی :

✉ [Zhivarpr@gmail.com](mailto:Zhivarpr@gmail.com)

📍 [Shahedbiology](#)