

ZHIVAR

ژیوار / گاهنامه انجمن علمی زیست شناسی دانشگاه شاهد / شماره ۸ / بهار ۱۴۰۴

مصاحبه با
دکتر دامون نشتاعلی
مدیرعامل شرکت "هومن ژن پارس"

بیوانفورماتیک دارویی
و فارماکوژنتیک

هوش مصنوعی
در علوم زیستی

بیوانفورماتیک
و زیست شناسی سامانه‌ای

بیوانفورماتیک
محاسباتی

هوش مصنوعی
مرگ یا زندگی!؟

بیوانفورماتیک
ساختاری





صاحب امتیاز:
انجمن علمی زیست شناسی
دانشگاه شاهد

مدیر مسئول:
مهدی ادریسیان

سر دبیر:
محمد صدرا محمدی



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

مشاور علمی:

خانم دکتر زهرا ظفری

دانش آموخته‌ی دکترای ژنتیک پزشکی

عضو هیئت علمی گروه زیست شناسی دانشگاه شاهد

مشاوران موضوعی:

سید هادی پوربخش

دکترای بیوتکنولوژی میکروبی دانشگاه شاهد

مرضیه جرنگیده

کارشناسی ارشد بیوتکنولوژی میکروبی دانشگاه شاهد

فاطمه بهمن پور

کارشناسی ارشد بیوتکنولوژی میکروبی دانشگاه شاهد

طراح و صفحه آرا:

محمد صدرا محمدی

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

هسته‌ی اجرایی و هماهنگی:

مهدی ادریسیان - دبیر انجمن

کارشناسی ارشد بیوتکنولوژی میکروبی دانشگاه شاهد

محمد صالح حاجی نصراله - مدیر علمی

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

سید علی حسینی - مدیر موضوعی

کارشناسی زیست فناوری دانشگاه شاهد

کارگروه ویراستاری:

محمد ابراهیمی آشتیانی

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

محمد صدرا محمدی

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

عوامل مصاحبه:

مریم فلاحی - مصاحبه کننده

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

زهرا درویشیان - دستیار مصاحبه

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

سید مهدی مرتضوی مقدم - فیلمبردار و تدوینگر

کارشناسی زیست فناوری دانشگاه شاهد

امیر محمد بصیر غفوری - نویسنده

کارشناسی زیست شناسی سلولی مولکولی دانشگاه شاهد

هیئت تحریریه:

فاطمه سادات شمسی نژاد

دکترای بیوتکنولوژی میکروبی دانشگاه تهران

ژیلا عبدی

دکترای ریاضی محض دانشگاه شاهد

صبا رحیمی

دکترای داروسازی Charles University-Czechia

محمد یعقوب زاد

دکترای بیوشیمی دانشگاه تبریز

سید علی حسینی

کارشناسی زیست فناوری دانشگاه شاهد

نوید بی یار عارفی

کارشناسی حقوق دانشگاه مراغه

هانیه محمودی هاشمی

دکترای مهندسی بافت دانشگاه شهید بهشتی

حانیه شریفی

کارشناسی ارشد ژنتیک دانشگاه شهید اشرفی اصفهانی

سبحان جره‌ای

کارشناسی زیست فناوری دانشگاه شاهد

فهرست

بخش اول بیوانفورماتیک

- ۶ بیوانفورماتیک سیستمی
- ۱۵ بیوانفورماتیک محاسباتی
- ۱۶ بیوانفورماتیک دارویی
- ۲۲ بیوانفورماتیک ساختاری

بخش دوم مصاحبه

- ۲۶ مصاحبه با دکتر دامون نشتاعلی
مدیرعامل شرکت "هومن ژن پارس"

بخش سوم هوش مصنوعی

- ۳۶ هوش مصنوعی و علوم زیستی
- ۴۶ هوش مصنوعی، مرگ یا زندگی؟!

امروزه نقش فناوری های نوین در توسعه ی علوم زیستی قطعا یکی از مهم ترین موضوعات مورد بحث است. در این نسخه از ژيوار اختصاصا تمام موضوعات و مباحث مطرح شده حول محور های بیوانفورماتیک و هوش مصنوعی می باشند. در بخش اول، حوزه های کلیدی بیوانفورماتیک مورد بررسی قرار گرفته اند؛ در بخش دوم مصاحبه ی جذابی با یکی از شرکت های خوب و تراز اول در حوزه بیوانفورماتیک انجام شده و در نهایت اساس نقش هوش مصنوعی و البته بحث اخلاقیات در این حوزه به طور اجمالی مورد بحث و بررسی قرار گرفته اند. اینجانب و شورای سردبیری این نسخه، از نویسندگان و عوامل عزیزی که برای تهیه ی این نسخه وقت گذاشتند صمیمانه تشکر کرده و امیدواریم در جهت نشر و تبیین علم و گسترش فضای مربوط به این روند، گامی هرچند کوچک برداشته باشیم.



محمد صدرا محمدی
سردبیر نشریه علمی ژيوار



بیوانفورماتیک سیستمی و زیست‌شناسی سامانه‌ای

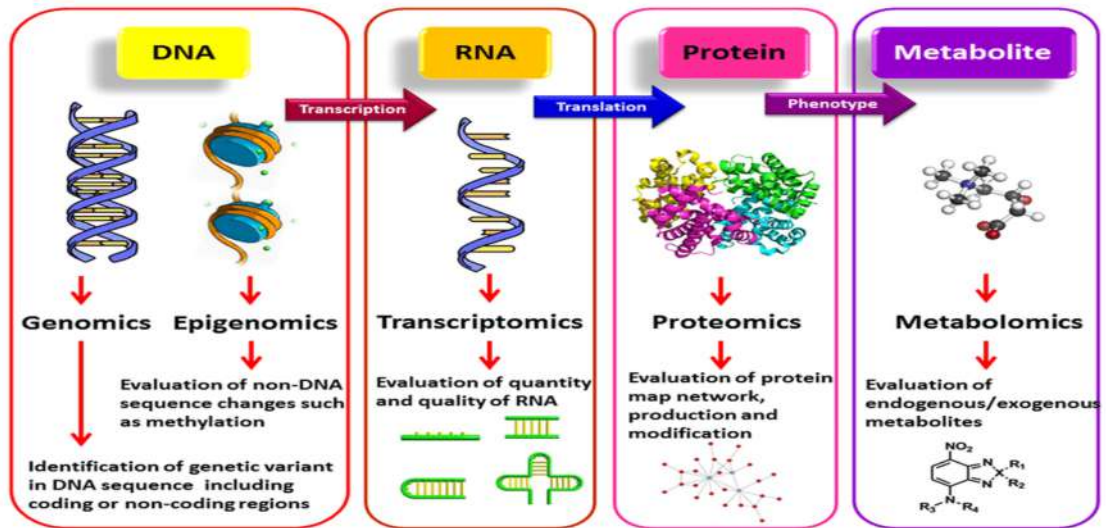
به قلم فاطمه سادات شمسی نژاد - دانشجوی دکترای بیوتکنولوژی میکروبی دانشگاه تهران

مقدمه

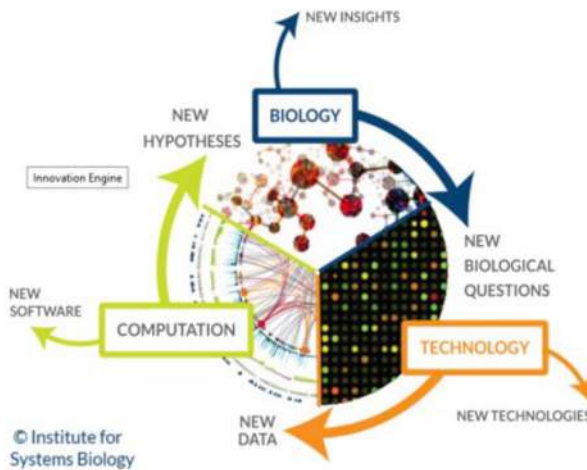
مقاله را با یک پرسش آغاز می‌کنیم؛ آیا یک زیست‌شناس می‌تواند یک رادیو را تعمیر کند؟ یک زیست‌شناس برای تعمیر رادیو، چند رادیوی سالم را بررسی می‌کند تا بتواند رادیوی خراب را تعمیر کند. زمان زیادی می‌برد تا یک زیست‌شناس، یک رادیو را تعمیر کند و هزینه‌ی زیادی برای این کار می‌پردازد. همین نگاه را یک زیست‌شناس در سطح سلول دارد و برای بررسی یک بیماری، توجه خود را روی یک ژن معطوف می‌کند. زیست‌شناسی سامانه‌ای^۱ می‌گوید که یک زیست‌شناس به جای نگاه جزء نگر، نگاه کل نگر داشته باشد؛ یعنی برای بررسی یک بیماری ارتباط اجزای سلول را در نظر گرفته و توجه خود را روی یک ژن معطوف نکند و نگاه سیستماتیک و شبکه‌ای داشته باشد. در مثال تعمیر رادیو، یک زیست‌شناس باید مثل یک مهندس برای تعمیر رادیو، نگاه سیستماتیک داشته باشد و ارتباط اجزای رادیو را در نظر بگیرد تا رادیو را سریع‌تر و با هزینه‌ای کمتر تعمیر کند [۱].

دو دلیل اصلی برای استفاده از رویکرد زیست‌شناسی سامانه‌ای وجود دارد: ۱- نگاه کل نگر و بررسی مسیرها و شبکه‌های زیستی، ۲- وجود تعداد زیادی داده که تفسیر آن‌ها نیازمند استفاده از رویکرد زیست‌شناسی سامانه‌ای است. برای مثال، حوزه تحقیقاتی اومیکس^۲ در سال‌های اخیر پیشرفت زیادی داشته است. این حوزه داده‌های زیادی در مورد سیستم‌های زیستی به ما داده است. تفسیر این مجموعه داده‌ها نیازمند استفاده از زیست‌شناسی سامانه‌ای و رویکرد مدل‌سازی و شبکه‌است. بنابراین، یک سیستم زیستی، یک شبکه پیچیده است که چندین جزء از اجزای زیستی مرتبط را به هم متصل می‌کند. سیستم تنفسی، یک مثال از یک سیستم زیستی است [۱-۴]. در شکل ۱ مهم‌ترین حوزه‌های اومیکس نشان داده شده است.

در واقع زیست‌شناسی سامانه‌ای این گونه تعریف شده است: برای بررسی سیستم‌های زیستی و مدل‌سازی این سیستم‌ها، از آنالیزهای محاسباتی و ریاضی استفاده شود. این یک زمینه مطالعاتی میان رشته‌ای مبتنی بر زیست‌شناسی است که بر برهمکنش‌های پیچیده درون سیستم‌های زیستی تمرکز کرده و با استفاده از رویکردی کل‌نگر (به جای رویکرد سنتی جزء نگر) به تحقیقات زیست‌شناسی می‌پردازد [۱، ۲]. در شکل ۲ روش‌های مورد استفاده در زیست‌شناسی سامانه‌ای به طور خلاصه ذکر شده است.



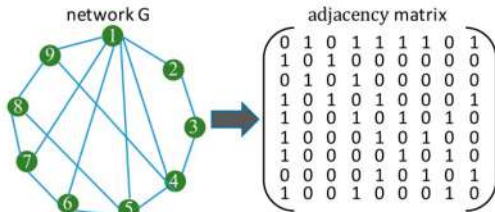
شکل ۱. مهم‌ترین حوزه‌های اومیکس. ژنومیکس، ترانسکریپتومیکس، پروتئومیکس و متابولومیکس مهم‌ترین حوزه‌های اومیکس هستند. تکنیک‌های اومیکس داده‌های زیادی را در اختیار محققان قرار می‌دهد که بررسی و تفسیر این داده‌ها نیازمند مدل‌سازی و رسم شبکه است [۳].



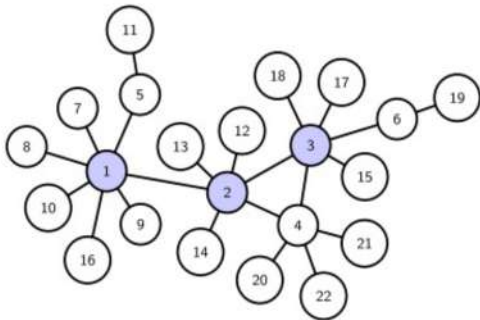
شکل ۲. روش‌های مورد استفاده در زیست‌شناسی سامانه‌ای. زیست‌شناسی سامانه‌ای از داده‌های روش‌های محاسباتی، زیستی، تکنولوژی (داده‌های نتایج تکنیک‌های آزمایشگاهی) برای بررسی یک سیستم زیستی استفاده می‌کند [۵].

شبکه و گراف

شبکه‌ی زیستی^۳: شبکه‌ی زیستی روشی برای نمایش سیستم‌ها به عنوان مجموعه‌های پیچیده‌ای از برهمکنش‌ها و یا روابط بین اجزای درون یک سیستم زیستی است. هر شبکه از گره^۴ و یال^۵ ساخته می‌شود. شبکه مسیره‌های پیام‌رسانی، یک مثال از شبکه زیستی است [۶].



شکل ۳. یک شبکه به همراه ماتریس همسایگی اش. این شبکه ۹ گره دارد. برای رسم ماتریس همسایگی ۹ گره را در سطر نوشته و به همان ترتیب ۹ گره را در ستون ماتریس مینویسیم. عدد 0 یعنی بین دو گره ارتباط وجود ندارد و عدد 1 یعنی بین دو گره ارتباط وجود دارد [۸].



شکل ۴. درجه و hub در یک شبکه. این شبکه ۲۲ گره دارد که گره ۱ دارای بالاترین درجه بوده و بنابراین یک hub محسوب میشود [۹].

نظریه گراف^۶: یک نظریه برای ارائه و نمایش یک شبکه است که از مجموعه گره و یال برای ساخت شبکه استفاده می‌کند. گره نماد مولکول یا پروتئین یا ژن و یال نماد برهمکنش بین مولکول یا پروتئین یا ژن است [۶].

درجه^۷: تعداد برهمکنش‌ها (یال‌ها) که هر جزء (گره) در یک شبکه دارد [۷].

Hub: در آنالیز شبکه، به مولکول یا گره‌ای گفته می‌شود که بالاترین درجه را در شبکه دارد [۷].

در شکل ۳ یک شبکه به همراه ماتریس همسایگی^۸ اش نشان داده شده است. ماتریس همسایگی یک ماتریس دو وجهی است که اجزای گراف و ارتباط آن‌ها را نشان می‌دهد. روی سطر ماتریس اسامی گره‌ها و روی ستون ماتریس به همان ترتیب اسامی گره‌ها نوشته می‌شود. عدد صفر، یعنی بین دو گره ارتباطی وجود ندارد و عدد یک، یعنی بین دو گره ارتباط وجود دارد [۸].

۳- Biological network

۴- Node

۵- Edge

۶- Graph theory

۷- Degree

۸- Adjacency matrix

هوش مصنوعی و زیست‌شناسی سامانه‌ای

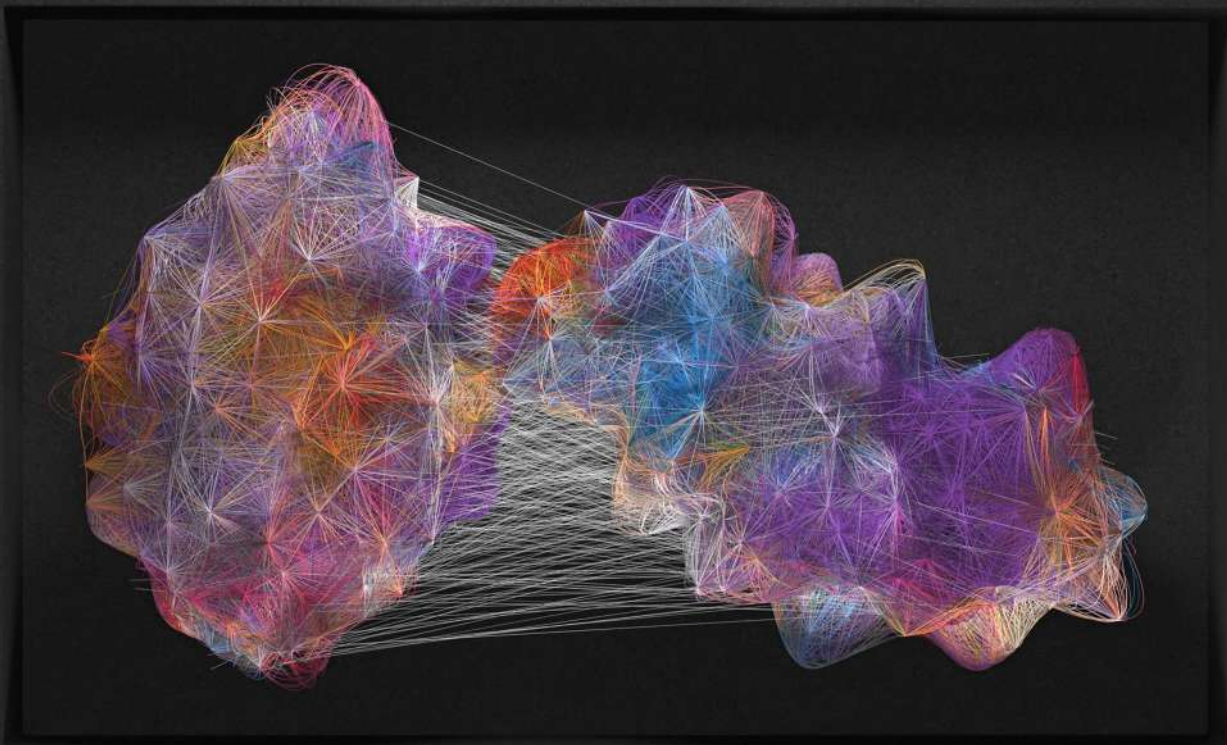
یکی از چالش‌های اصلی در زیست‌شناسی سامانه‌ای تبدیل داده‌های بزرگ به دانش عملی و بهره‌برداری از این داده‌ها است. حوزه تحقیقاتی اومیکس داده‌های زیادی را تولید می‌کند اما بهره‌برداری درست از این داده‌ها برای محققین مشکل ایجاد کرده است. مدلسازی یک بخش ضروری در زیست‌شناسی سامانه‌ای است. ترکیب داده‌های اومیکس برای آنالیز شبکه‌ها و مدل‌سازی رویکرد مناسب برای تحقیق در این حوزه است که استفاده از ابزارهای هوش مصنوعی در این زمینه بسیار کمک‌کننده می‌باشد. رویکردهای هوش مصنوعی شامل یادگیری ماشین، یادگیری عمیق و شبکه‌های عصبی مصنوعی هستند. بنابراین، استفاده از ابزارهای هوش مصنوعی می‌تواند به درک ما از سیستم‌های زیستی کمک کند [۲، ۱۱، ۱۲].

نتیجه‌گیری

در دهه‌های قبل، زیست‌شناسان به صورت رویکرد جزء نگر به سلول نگاه می‌کردند و در مطالعات خود، یک جزء را در نظر می‌گرفتند. زمانی که تکنولوژی اومیکس و تکنیک‌های توالی‌یابی، داده‌های زیادی را در اختیار محققان قرار داد، استفاده از رویکرد زیست‌شناسی سامانه‌ای گسترش روزافزون یافت. بنابراین، محققان رویکرد کل‌نگر را برای بررسی سلول‌ها و سیستم‌های زیستی به‌کار بردند. بررسی و رسم شبکه‌های زیستی به درک ما از مکانیسم‌های زیستی و مسیرهای متابولیک و بیماری کمک کرده است. استفاده از داده‌های حاصل از نتایج آزمایشگاهی همراه با داده‌های حاصل از بررسی شبکه‌های زیستی، می‌تواند به توسعه استراتژی‌های درمانی و تشخیصی جدید کمک کند. هر چند، بررسی شبکه‌های زیستی با چالش‌هایی همراه است، اما امید است در آینده‌ای نزدیک، با پیشرفت هوش مصنوعی و یادگیری ماشین و همچنین، پیشرفت در تکنیک‌های آزمایشگاهی، شرایط برای بررسی شبکه‌های زیستی هموار شود.

منابع:

1. Lazebnik, Y., Can a biologist fix a radio?--Or, what I learned while studying apoptosis. *Cancer Cell*, 2002. 2(3): p. 179-82.
2. Tavassoly, I., J. Goldfarb, and R. Iyengar, *Systems biology primer: the basic methods and approaches*. *Essays Biochem*, 2018. 62(4): p. 487-500.
3. Kohl, M., S. Wiese, and B. Warscheid, *Cytoscape: software for visualization and analysis of biological networks*. *Methods Mol Biol*, 2011. 696: p. 291-303.
4. Kirschner, M.W., *The Meaning of Systems Biology*. *Cell*, 2005. 121(4): p. 503-504.
5. Villa, A. and S.T. Sonis, *2 - System biology*, in *Translational Systems Medicine and Oral Disease*, S.T. Sonis and A. Villa, Editors. 2020, Academic Press. p. 9-16.
6. Koutrouli, M., et al., *A Guide to Conquer the Biological Network Era Using Graph Theory*. *Front Bioeng Biotechnol*, 2020. 8: p. 34.
7. Mitra, K., et al., *Integrative approaches for finding modular structure in biological networks*. *Nat Rev Genet*, 2013. 14(10): p. 719-32.
8. Li, Z., Q. Liu, and L. Chen, *A Novel Edge Rewire Mechanism Based on Multiobjective Optimization for Network Robustness Enhancement*. *Frontiers in Physics*, 2021. 9.
9. Mondragón, R.J., *Estimating degree-degree correlation and network cores from the connectivity of high-degree nodes in complex networks*. *Scientific Reports*, 2020. 10(1): p. 5668.
10. Heraković, N., et al., *Distributed Manufacturing Systems with Digital Agents*. 2019, 2019. 65(11-12): p. 8.
11. Kannan, M., et al., *Leveraging public AI tools to explore systems biology resources in mathematical modeling*. *NPJ Syst Biol Appl*, 2025. 11(1): p. 15.
12. Jamialahmadi, H., et al., *Artificial intelligence and bioinformatics: a journey from traditional techniques to smart approaches*. *Gastroenterol Hepatol Bed Bench*, 2024. 17(3): p. 241-252.



بیوانفورماتیک محاسباتی و یادگیری ماشین

به قلم ژیلا عبدی - دانشجوی دکترای ریاضی محض دانشگاه شاهد

بیوانفورماتیک چیست؟

به نظر می‌رسد اصطلاح بیوانفورماتیک^۱ اولین بار در اواسط دهه 1980 به منظور توصیف کاربرد علم و فناوری اطلاعات در علوم زیستی مورد استفاده قرار گرفت. در آن زمان این تعریف بسیار کلی بود و همه چیز، از رباتیک گرفته تا هوش مصنوعی را در بر می‌گرفت. بعدها بیوانفورماتیک تا حدودی به عنوان «استفاده از رایانه برای بازیابی، پردازش، تجزیه و تحلیل و شبیه‌سازی اطلاعات بیولوژیکی» تعریف شد. یک تعریف محدودتر برای بیوانفورماتیک "کاربرد فناوری اطلاعات^۲ در مدیریت داده های بیولوژیکی" بود. چنین تعاریفی

خلاصه

ابتدا به معرفی اجمالی بیوانفورماتیک می‌پردازیم. بیوانفورماتیک مرکب از زیست‌شناسی و فناوری اطلاعات است که از فناوری اطلاعات برای درک پدیده‌های بیولوژیکی استفاده می‌کند. بیوانفورماتیک را به عنوان بخشی از علم زیست‌محاسباتی نیز در نظر می‌گیرند. زیست‌شناسی محاسباتی از تکنیک‌های تحلیلی کمی برای مدل‌سازی سیستم‌های بیولوژیکی استفاده می‌کند. برای استخراج، تجزیه و تحلیل اطلاعات پنهان در زیست‌شناسی محاسباتی از داده‌کاوی استفاده می‌شود.

بکارگیری ابزارهای تحلیلی داده، روش‌های نظری و مدل سازی ریاضی و تکنیک‌های شبیه سازی نرم افزاری را برای کشف سیستم‌های بیولوژیکی را بر عهده دارند. محاسبات در حال حاضر بخشی ضروری از پروژه‌های تحقیقاتی بیولوژیکی است. به عنوان مثال، بانک‌های داده‌های پروتئین، پایگاه‌های داده ژنومی و تصاویر MRI مغز حاوی مقادیر عظیمی از داده‌های خام هستند

که می‌توانند به اطلاعات

کارآمد و قابل استفاده

درباره همه جوانب

زیست‌شناسی

ترجمه شوند.

اکنون با تقویت

امکانات فناوری

اطلاعات، مطرح

شدن موضوع

بیوانفورماتیک

موضوعیت دارد [۲].

بیوانفورماتیک را

می‌توان از دو جنبه مورد

بررسی قرار داد: یکی برای

زیست‌شناسانی که فناوری اطلاعات را یاد می‌گیرند تا در تخصص خود استفاده کنند، یا برای متخصصان فناوری اطلاعات که زیست‌شناسی را می‌آموزند تا بتوانند مهارت‌های خود را در مسائل بیولوژیکی به کار ببرند. این روند ترکیبی از ریاضیات، علوم کامپیوتر، آمار و زیست‌شناسی است. در مقالات متفاوت فن آوری‌های مختلف علوم کامپیوتر و روش‌های آماری در بیوانفورماتیک مورد توجه قرار گرفته است. رویکرد چند رشته‌ای و میان رشته‌ای اجازه می‌دهد تا درک روابط متقابل بین

می‌توانند محوریت اطلاعات را در زیست‌شناسی به دست آورند. در واقع اگر اطلاعات، بنیادی‌ترین مفهوم زیربنایی زیست‌شناسی باشد و بیوانفورماتیک، کاوش در همه انشعابات و پیامدهای آن مبنا قرار گرفته باشد، پس بیوانفورماتیک در یک موقعیت عالی تعریف شده است تا توجه به مسئله مهم «چستی زندگی» را احیا کند. بنابراین یک

تعریف مناسب از بیوانفورماتیک

این است: «علم چگونگی

تولید، انتقال، دریافت،

ذخیره، پردازش و

تفسیر اطلاعات

در سیستم

بیولوژیکی»

یا به طور خلاصه،

«کاربرد علم

اطلاعات در

زیست‌شناسی».

ظهور نظریه اطلاعات

در اواسط قرن بیستم

امکان ایجاد یک چارچوب

رسمی را فراهم کرد که در آن می‌توان

اطلاعات را کمی کرد [۳]. بیوانفورماتیک

مدرن علمی است که استفاده از روش‌های

کامپیوتری را برای تجزیه و تحلیل انواع داده

های ژنومی^۳ توسعه می‌دهد. با ظهور ابعاد

جدیدی از بیوانفورماتیک، نقش بزرگی در

توسعه این حوزه با توسعه سریع فناوری رایانه

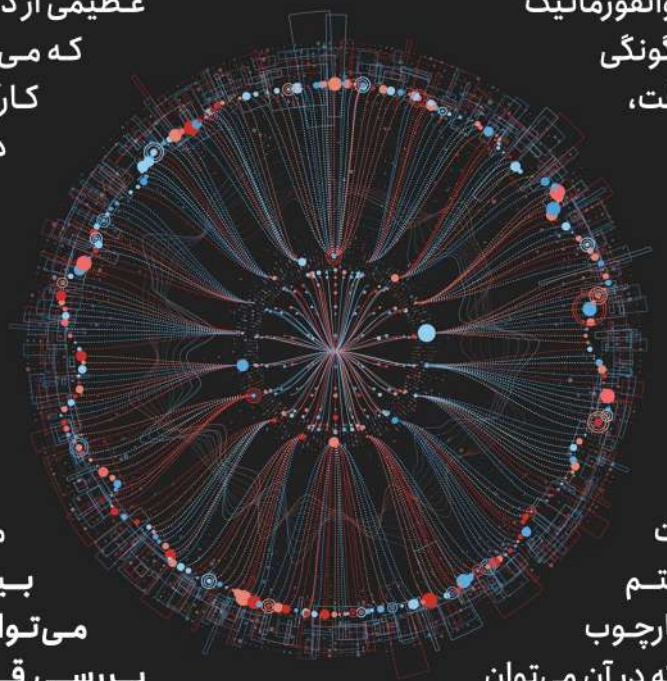
و روش‌های محاسباتی پردازش داده‌ها و ظهور

فناوری‌های مدرن مخابراتی ایفا شد.

بیوانفورماتیک یکی از حوزه‌های علمی است که

غالباً وابسته به اینترنت تعریف می‌شود [۲].

زیست‌شناسان محاسباتی وظیفه توسعه و



فن‌آوری‌های علوم کامپیوتر، روش‌های آماری و کاربردهای بیوانفورماتیک را تسهیل کند. در زمینه زیست‌شناسی محاسباتی، تجربه در زمینه اسکرپیت نویسی، یادگیری ماشین^۴ و کشف و توسعه دارو در اولویت است [۲].

مقایسه تکنیک‌های داده‌کاوی^۵ و روش‌های آماری

تکنیک‌های مختلفی در طول سال‌های متمادی برای کاوش و استخراج اطلاعات از مجموعه داده‌های بزرگ توسعه یافته‌اند. در پایان دهه 1980 رشته جدیدی به نام داده‌کاوی ظهور کرد. تکنیک‌های سنتی تجزیه و تحلیل داده‌ها اغلب در پردازش کارآمد داده‌های حجیم شکست می‌خورند. داده‌کاوی فرآیند کشف اطلاعات معتبر، از پیش ناشناخته و در نهایت قابل درک از ذخایر بزرگ داده‌است. داده‌کاوی فرآیند استخراج دانش پنهان در حجم زیادی از داده‌های خام است. داده‌کاوی فرآیند یافتن روابط و الگوها در داده‌های خام را خودکار می‌کند و نتایجی را ارائه می‌دهد که می‌تواند در یک سیستم پشتیبانی تصمیم‌گیری خودکار استفاده شود یا توسط یک تحلیل‌گر انسانی ارزیابی شود. سیستم‌های مدرن داده‌کاوی کامپیوتری از تاریخچه قبلی سیستم مورد بررسی، خود یاد می‌گیرند و فرضیه‌هایی را در مورد قوانینی که این سیستم از آن‌ها پیروی می‌کند، فرموله و آزمایش می‌کنند. به طور کلی، تکنیک‌های داده‌کاوی را می‌توان به دو دسته کلی تقسیم کرد: داده‌کاوی پیش‌بینی و داده‌کاوی اکتشافی [۲].

کاهش بعد^۶ در داده‌های بیوانفورماتیک

کاهش ابعاد در چارچوب یادگیری ماشین است. برای کشف ساختار و الگوهای پنهان در داده‌ها نیاز به تجزیه و تحلیل آن‌ها وجود دارد. ابعاد داده‌های بیولوژیکی وسیع و فراگیر هستند. تجزیه و تحلیل داده‌ها در مقیاس بزرگ اغلب به دلیل ابعاد بالا و محدودیت‌های مرتبط کار سختی می‌باشد بنابراین، از روش‌های کاهش ابعاد برای این داده‌ها استفاده می‌شود [3]. با توسعه فناوری‌های با توان عملیاتی بالا، مانند ریزآرایه (Microarray) و داده‌های توالی‌یابی نسل بعدی (NGS)، کاوش مجموعه‌های داده با توان بالا برای آشکار کردن اطلاعات مربوطه موجود در داده‌ها به یک ضرورت تبدیل می‌شود. بنابراین، ابزارهای اکتشافی کارآمد، نه تنها برای ارزیابی کیفیت داده‌ها، بلکه برای ارائه یک نمای کلی جامع از سیستم، استخراج اطلاعات قابل توجه و مقابله با ابعاد بالا مورد نیاز است. در واقع، بسیاری از رویکردهای آماری به دو دلیل اصلی شکست می‌خورند یا ضعیف عمل می‌کنند: تعداد نمونه‌ها (یا مشاهدات) بسیار کمتر از تعداد متغیرها (موجودات بیولوژیکی که اندازه‌گیری می‌شوند) و داده‌ها بسیار درهم آمیخته و از نظم گسیخته هستند. یکی از رویکردهای بدون نظارت برای کشف مکانیسم‌های بیولوژیکی جدید و آشکارسازی الگوها و در عین حال کاهش ابعاد در داده‌ها روش تحلیل مولفه اصلی می‌باشد [1]. این رویکردهای اکتشافی، داده‌ها را در یک زیرفضای جدید نشان می‌دهند. این رویکردها امکان کاهش ابعاد را بدون از دست دادن اطلاعات ضروری و تجسم داده‌ها در یک زیرفضای کوچکتر فراهم می‌کنند تجزیه و تحلیل مؤلفه

۴- Machine learning

۵- Data mining

۶- Dimensionality reduction

اصلی یک ابزار کلاسیک برای کاهش ابعاد داده‌ها، تجسم شباهت‌های بین نمونه‌های بیولوژیکی و فیلتر کردن نویز است. این روش اغلب به عنوان یک مرحله پیش پردازش برای تجزیه و تحلیل‌های بعدی استفاده می‌شود. PCA داده‌ها را در فضای جدیدی که توسط مؤلفه‌های اصلی گسترده شده‌است، تصویر می‌کند که غیرهمبسته و متعامد هستند. رایانه‌های شخصی می‌توانند با موفقیت اطلاعات مربوطه را از داده‌ها استخراج کنند[۱].

تحلیل مولفه اصلی و خوشه بندی داده‌ها

گاهی اوقات برای کاهش ابعاد مجموعه داده‌ها قبل از خوشه‌بندی، تحلیل مولفه اصلی اعمال می‌شود. دلیل استفاده از تحلیل مولفه اصلی قبل از تحلیل خوشه‌ای این است که مولفه‌های اصلی بتوانند ساختار خوشه را در مجموعه داده‌ها استخراج کنند[۵].

نتایج آزمایشگاهی

برای بررسی بهتر با ارائه یک مثال کاربرد تحلیل مولفه اصلی در داده‌های بیوانفورماتیکی و در نهایت خوشه‌بندی را بررسی می‌کنیم. ابتدا مجموعه‌ای از داده‌ها را با استفاده از شبیه‌سازی داده‌های بیوانفورماتیکی تولید می‌کنیم. این مجموعه داده مربوط به تحلیل آزمایش خون برای بررسی مبتلا بودن به دیابت یا عدم ابتلا به دیابت است.

```
import pandas as pd
import numpy as np

np.random.seed(42)
data = {
    'Patient_ID': [f'P{i+1}' for i in range(100)],
    'Age': np.random.randint(20, 80, size=100),
    'BMI': np.random.uniform(18, 40, size=100).round(1),
    'Glucose_Level': np.random.randint(80, 200, size=100),
    'Insulin_Level': np.random.uniform(2, 25,
size=100).round(1),
}

df = pd.DataFrame(data)
print(df.head())
```

تصویر ۱. کد‌های مربوط به شبیه‌سازی داده‌های مربوط به آزمایش خون

```

from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA

# Features to include for PCA
features = ['Age', 'BMI', 'Glucose_Level', 'Insulin_Level']
x = df[features].values

# Standardize the features
scaler = StandardScaler()
x_scaled = scaler.fit_transform(x)

# Apply PCA (reduce to 2 principal components for
visualization)
pca = PCA(n_components=2)
principal_components = pca.fit_transform(x_scaled)

# Create a DataFrame for principal components
pc_df = pd.DataFrame(data=principal_components,
columns=['PC1', 'PC2'])
print(pc_df.head())

```

تصویر ۲. کد های مربوط به اجرای الگوریتم تحلیل مولفه اصلی برای کاهش بعد

```

from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Perform K-Means clustering
kmeans = KMeans(n_clusters=2, init=np.array([[1, 0], [0,
1]]), n_init=1, random_state=42)
pc_df = pd.DataFrame(data=principal_components,
columns=['PC1', 'PC2'])
pc_df['Cluster'] = kmeans.fit_predict(principal_components)

# Visualize clusters
plt.figure(figsize=(8, 6))
plt.scatter(pc_df['PC1'], pc_df['PC2'], c=pc_df['Cluster'],
cmap='viridis', s=50)
plt.scatter(kmeans.cluster_centers_[:, 0],
kmeans.cluster_centers_[:, 1], color='red', marker='x',
s=200, label='Cluster Centers')
plt.title('Clusters Centered on PCA Components')
plt.xlabel('Principal Component 1')
plt.ylabel('Principal Component 2')
plt.legend()
plt.grid(True)
plt.show()

# Print PCA cluster results
print("Cluster Centers (PCA Components):")
print(kmeans.cluster_centers_)

```

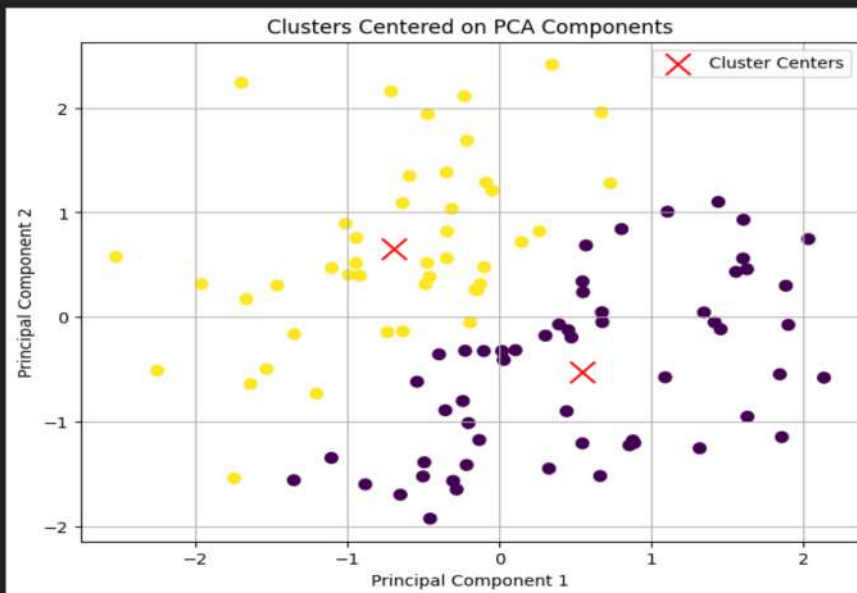
تصویر ۳. کد های مربوط به اجرای خوشه بندی با استفاده از مولفه های اصلی به عنوان مرکز خوشه ها

نتیجه‌گیری

هدف این است با استفاده از کاهش بعد و خوشه‌بندی افراد مبتلا به دیابت را مشخص کنیم. در واقع روش تحلیل مولفه اصلی بعد داده‌ها را با حفظ حداکثر واریانس کاهش می‌دهد. در واقع ساده‌سازی داده‌ها با حفظ ویژگی‌های اصلی. در اینجا $pc1$ و $pc2$ مولفه‌های اصلی هستند که ترکیب خطی از ویژگی‌های اصلی می‌باشند و در شکل با ضربدر مشخص شده‌اند. این مولفه‌ها به عنوان مراکز خوشه‌بندی با استفاده از الگوریتم k -میانگین در نظر می‌گیریم. همانطور که در شکل ۱ مشخص است، افراد به دو گروه با جواب تست دیابت مثبت و منفی تقسیم شده‌است.

منابع:

- [1] Calabrese, B. (2018). Data Reduction. Reference Module in Life Sciences.
- [2] Raina, K. S., Sofi, M., Dar, N. A., & Zargar, O. (2023). Role of bioinformatics in analyzing big data using statistical computing and computer science. International Journal of Scientific Research in Engineering and Management.
- [3] Ramsden, J. (2023). Bioinformatics: an introduction. Springer Nature.
- [4] Yao, F., Coquery, J., & Lê Cao, K. A. (2012). Independent principal component analysis for biologically meaningful dimension reduction of large biological data sets. BMC bioinformatics, 13, 1-15.
- [5] Yeung, K. Y., & Ruzso, W. L. (2001). Principal component analysis for clustering gene expression data. Bioinformatics, 17(9), 763-774.



شکل ۱. شکل حاصل از خوشه‌بندی با استفاده از مولفه‌های اصلی الگوریتم تحلیل مولفه اصلی به عنوان مراکز خوشه

بیوانفورماتیک دارویی و فارماکوژنتیک

به قلم صبا رحیمی - دانشجوی دکترای داروسازی دانشگاه چارلز (جمهوری چک)

چکیده:

بیوانفورماتیک به عنوان یک علم میان‌رشته‌ای، نقشی اساسی در پیشرفت داروسازی و طراحی دارو ایفا می‌کند. با استفاده از تحلیل داده‌های زیستی، مدل‌سازی مولکولی و یادگیری ماشین، این حوزه به ابزاری کلیدی برای کشف داروهای جدید و بهینه‌سازی فرآیندهای درمانی تبدیل شده است. این مقاله به بررسی کاربردهای بیوانفورماتیک در کشف دارو، بهینه‌سازی ترکیبات دارویی و ارتباط آن با فارماکوژنتیک در توسعه درمان‌های شخصی‌سازی شده می‌پردازد [1].

مقدمه:

الگوریتم‌های یادگیری ماشینی برای پیش‌بینی واکنش‌های دارویی، و سیستم‌های پشتیبانی از تصمیمات بالینی برای هدایت انتخاب دارو و تنظیم دوز آن. همچنین به چالش‌ها و فرصت‌های موجود در این زمینه می‌پردازیم، از جمله ضرورت داشتن داده‌های چندگانه و مقیاس بزرگ، و اهمیت ایجاد مدل‌های ماشین لرنینگ قابل اعتماد و قابل تفسیر. به طور کلی، نرم افزارهای کامپیوتری نقشی اساسی در پیشبرد تحقیقات فارماژنتیک ایفا کرده و این پتانسیل را دارند که با تسهیل توسعه درمان‌های دارویی شخصی‌شده مؤثرتر، انقلاب‌هایی در دنیای پزشکی به وجود آورند.

فرآیند کشف و توسعه دارو یک مسیر پیچیده و پرهزینه است که معمولاً بیش از یک دهه به طول می‌انجامد و نیازمند سرمایه‌گذاری‌های کلان است. با این حال، پیشرفت‌های اخیر در بیوانفورماتیک و کموانفورماتیک، این مسیر را تسهیل کرده و موجب افزایش دقت و کارایی در شناسایی داروهای جدید شده است. بیوانفورماتیک با تحلیل داده‌های ژنومی و پروتئومی، به شناسایی اهداف دارویی بالقوه کمک می‌کند و فرآیند کشف دارو را تسریع می‌کند [2]. ما در این قسمت از نشریه، برخی از نرم افزارهای مهم استفاده شده در تحقیقات فارماژنتیک را بررسی می‌کنیم، از جمله ابزارهای بیوانفورماتیک برای تحلیل داده‌های ژنتیکی،

KGNN34 برای مثال، می‌تواند پیش‌بینی کند که دو دارو ممکن است فشار خون را به میزان خطرناکی کاهش دهند یا با هم ترکیب شوند و اثرات منفی دیگری به وجود آورند. این مدل از داده‌های مختلف دارویی و ژنتیکی برای شبیه‌سازی روابط پیچیده استفاده می‌کند. این به این معناست که سیستم می‌تواند ویژگی‌های مختلف داروها مانند ساختار شیمیایی، نحوه عملکرد آن‌ها در بدن و تأثیرات جانبی را در نظر بگیرد تا تعاملات دقیق‌تری را پیش‌بینی کند.

DPDDI35:

یک پایگاه داده است که به صورت خاص بر روی تعاملات دارویی تمرکز دارد. این پایگاه داده شامل اطلاعاتی است که از مطالعات بالینی و داده‌های دارویی به دست آمده و به شناسایی و پیش‌بینی تعاملات دارویی می‌پردازد. این می‌تواند به پزشکان، داروسازان و محققان کمک کند تا متوجه شوند کدام داروها نباید با هم مصرف شوند. برای مثال، اگر دو دارو با هم ترکیب شوند و یکی از آن‌ها ممکن است تأثیرات منفی جدی مانند کاهش فشار خون یا افزایش خطر خونریزی داشته باشد، این پایگاه داده به راحتی می‌تواند چنین اطلاعاتی را ارائه دهد از مطالعات مختلف در دنیای واقعی و آزمایش‌های بالینی برای جمع‌آوری این داده‌ها استفاده می‌کند، بنابراین اطلاعات موجود در آن به روز و دقیق هستند.

پیش‌بینی‌کننده‌های تعامل و تداخل دارو: تکنیک‌های بیوانفورماتیک که به نام "پیش‌بینی‌کننده‌های تعامل دارو" شناخته می‌شوند، سعی می‌کنند پیش‌بینی کنند که چگونه دو دارو با یکدیگر یا با عواملی مانند غذا یا الکل واکنش نشان می‌دهند. این ابزارها می‌توانند برای توسعه درمان‌های دارویی شخصی‌سازی شده که خطر واکنش‌های ناخواسته دارویی را به حداقل می‌رسانند، استفاده شوند. به عنوان مثال، یک پیش‌بینی‌کننده تعامل دارو می‌تواند برای شناسایی داروهایی که نباید به طور همزمان مصرف شوند، تعاملات دارویی را پیش‌بینی کنند [۳].

KGNN34 و DPDDI35

دو نمونه از روش‌های محاسباتی و پیش‌بینی برای تداخل و تعامل دارو هستند.

KGNN34:

یک مدل محاسباتی پیشرفته است که از تکنیک‌های یادگیری عمیق (Deep Learning) برای پیش‌بینی تعاملات دارویی استفاده می‌کند. یادگیری عمیق به نوعی از یادگیری ماشینی اشاره دارد که به سیستم‌ها کمک می‌کند تا الگوها و روابط پیچیده در داده‌ها را شبیه‌سازی و تحلیل کنند. در مورد تعاملات دارویی، می‌تواند به طور خاص پیش‌بینی کند که وقتی دو دارو با هم مصرف شوند، چه تأثیراتی ممکن است بر روی بدن داشته باشند.

این ابزارها چگونه به کمک می‌آیند؟

هر دو روش KGNN34 و DPDDI35 در واقع به پزشکان و داروسازان کمک می‌کنند تا ترکیب های بهتری انتخاب کنند و از خطرات ناشی از مصرف همزمان داروهای مختلف جلوگیری کنند. به این ترتیب، این ابزارها به ویژه در درمان های شخصی سازی شده اهمیت زیادی دارند، چرا که می‌توانند بر اساس ویژگی های فردی هر بیمار، بهترین دارو و ایمن ترین ترکیب دارویی را پیش بینی کنند. در واقع، بیوانفورماتیک شامل استفاده از پایگاه داده ها، مدل های یادگیری ماشینی و تحلیل های ژنتیکی است که می‌توانند داده های وسیعی از تعاملات دارویی را پردازش کنند. این ابزارها به پژوهشگران کمک می‌کنند تا تعاملات دارویی مختلف را شبیه سازی کنند و از این طریق، درمان های شخصی سازی شده ای طراحی نمایند که به دقت بیشتری نیاز دارند.

مدل سازی:

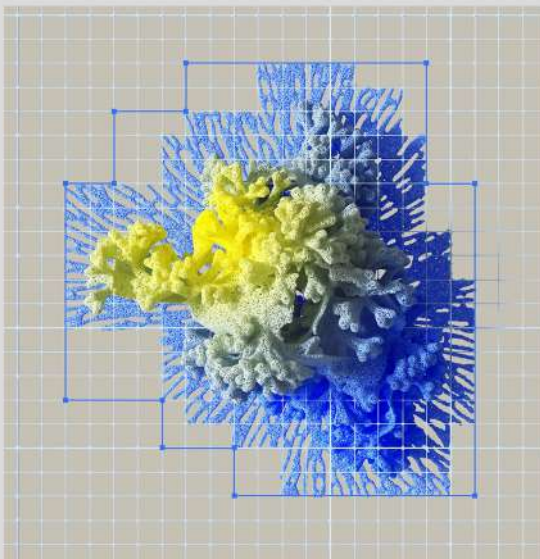
یکی دیگر از کاربردهای مهم فناوری کامپیوتری در فارماژنتیک مدل سازی است. مدل سازی شامل استفاده از مدل های ریاضی و آماری برای پیش بینی پاسخ دارویی بر اساس تفاوت های ژنتیکی است. این مدل ها می‌توانند در توسعه دارو و پزشکی شخصی سازی شده با شناسایی نشانگرهای ژنتیکی که با اثربخشی دارو، سمیت، و فارماکوکینتیک دارو مرتبط هستند، کمک کنند. داکینگ مولکولی فرآیندی است که در آن مولکول های کوچک به ساختارهای ماکرومولکولی متصل می‌شوند تا مقادیر مکمل آن ها در محل های اتصال ارزیابی شود. این یک حوزه تحقیقاتی پویا با کاربردهای گسترده در طراحی دارو مبتنی بر ساختار، بهینه سازی ترکیبات پیشرو، بررسی مسیرهای

بیوشیمیایی و طراحی دارو است که از مهم ترین محسوب می‌شود. مولکولار داکینگ و بیوانفورماتیک هر دو به مطالعه و تحلیل و ساختارهای زیستی مرتبط هستند و در زمینه علوم زیستی و دارویی کاربرد دارند. ارتباط این مشخص می‌شود: طراحی داروهای جدید: بیوانفورماتیک و مولکولار داکینگ هر دو به طراحی داروهای جدید کمک می‌کنند. بیوانفورماتیک با فراهم آوردن پایگاه های داده ساختاری و اطلاعات ژنتیکی، ابزارهایی برای طراحی داروها بر اساس اطلاعات زیستی فراهم می‌کند، در حالی که مولکولار داکینگ با شبیه سازی نحوه اتصال دارو به اهداف مولکولی به انتخاب داروهای مؤثرتر و با کمترین عوارض کمک می‌کند.



های بیان ژن و داده‌های بیومارکر می‌باشد. این داده‌ها از آزمایش‌ها و تحقیقات زیستی استخراج می‌شوند و جمع‌آوری، ذخیره‌سازی، مدیریت و تحلیل آن‌ها ارکان اصلی فناوری بیوانفورماتیک را می‌سازند. در دهه گذشته، ابزارهای بیوانفورماتیک، مانند روش‌های کامپیوتری و تکنیک‌های غربالگری با توان بالا، نقش مهمی در تسریع کشف دارو ایفا کرده‌اند. این روش‌ها به‌طور مؤثر از غربالگری و توسعه ترکیبات طبیعی، سنتزی و نیمه‌سنتزی پشتیبانی کرده و موجب پیشرفت قابل توجهی در تحقیق روی داروهای مؤثر یا مولکول‌های پیش‌گام شده‌اند.

به‌عنوان یکی از بیماری‌های اصلی تهدیدکننده سلامت انسان، درمان دقیق تومورها به دلیل پیشرفت در درمان‌های هدفمند دچار تحول شده است. با توجه به پیچیدگی پاتوژنز سرطان و دشواری توسعه داروهای ضدسرطان، لازم است منابع شیمیایی و زیستی گسترش یابند تا داربست‌های مولکولی بیشتری برای کشف و توسعه داروهای ضدسرطان فراهم شود [۲].



توسعه دارو با کشف یک بیماری و تغییرات آن آغاز می‌شود. زمانی که یک بیماری تهدیدی برای سلامت انسان به شمار می‌آید و کیفیت زندگی را کاهش می‌دهد، داروها ایجاد می‌شوند. داروی ایده‌آل باید نه تنها علائم را کاهش دهد و بیماری را درمان کند، بلکه دارای پروفایل ایمنی بالا، عوارض جانبی کم و هزینه‌های تحقیقاتی پایین باشد. با این حال، این بدون شک یک چالش بزرگ است. فرآیند توسعه دارو هم‌زمان برو پیچیده است و نیاز به سرمایه‌گذاری قابل توجهی از نظر زمان و منابع مالی دارد. بیوانفورماتیک پتانسیل زیادی در کشف داروها نشان می‌دهد. بیوانفورماتیک یک علم میان‌رشته‌ای است که از علوم کامپیوتر، فناوری اطلاعات، ریاضیات کاربردی و روش‌های آماری برای جمع‌آوری، پردازش، ذخیره‌سازی، انتشار، تجزیه و تحلیل و تفسیر اطلاعات زیستی در تحقیقات علوم زیستی استفاده می‌کند. در مه ۱۹۸۵، دانشمند آمریکایی رابرت سینسهایمر اولین بار پروژه ژنوم انسانی را پیشنهاد کرد، که هدف آن تعیین توالی کامل DNA ژنوم انسان در مدت ۱۵ سال، رمزگشایی حدود ۲۵,۰۰۰ ژن در بدن انسان و نقشه‌برداری از ژنوم انسانی بود. پس از تلاش‌های بی‌وقفه، در آوریل ۲۰۰۳، پروژه ژنوم انسانی به‌طور رسمی به پایان رسید و ترکیب ژنوم انسانی تعیین شد: ۴۶ کروموزوم، ۳ میلیارد باز و ۳۰,۰۰۰ ژن کدکننده پروتئین. تکمیل این پروژه آغازگر توسعه سریع بیوانفورماتیک بود [۲].

بیوانفورماتیک در کشف دارو:

داده‌های زیستی اساس فناوری بیوانفورماتیک را تشکیل می‌دهند، که شامل داده‌های ژنوم، داده‌های توالی پروتئین، داده

عملکرد داکینگ مولکولی در بیوانفورماتیک:

عملکرد داکینگ مولکولی شامل دو مرحله اصلی است:

۱- پیش‌بینی ساختار و موقعیت لیگاند در جایگاه فعال پروتئین
این مرحله که به "پوزینگ" معروف است، شامل مدل‌سازی نحوه قرارگیری لیگاند در محل اتصال پروتئین بر اساس قوانین فیزیک شیمیایی و انرژی‌های بین‌مولکولی است. الگوریتم‌های داکینگ مولکولی به کمک روش‌هایی مانند جستجوی تطبیقی، الگوریتم ژنتیک و بهینه‌سازی متا هیوریستیک، ساختارهای ممکن را بررسی و بهترین حالت را انتخاب می‌کنند.

۲- ارزیابی برهم‌کنش و کیفیت اتصال با استفاده از توابع امتیازدهی:
(Scoring Functions) پس از تعیین موقعیت لیگاند، کیفیت اتصال آن با استفاده از توابع امتیازدهی ارزیابی می‌شود این توابع معمولاً بر اساس پارامترهایی مانند نیروهای واندروالس، پیوندهای هیدروژنی، برهم‌کنش‌های الکترواستاتیکی و انرژی اتصال، میزان پایداری کمپلکس را محاسبه می‌کنند. امتیاز بالاتر نشان‌دهنده پتانسیل بالاتر لیگاند برای اتصال و فعالیت زیستی است. [۸]

مزایای استفاده از داکینگ مولکولی در کشف دارو:

کاهش هزینه‌ها و زمان تحقیقاتی روش‌های سنتی کشف دارو معمولاً شامل آزمایش‌های تجربی بر روی هزاران ترکیب هستند که به طور چشمگیری زمان بر و پرهزینه‌اند. با استفاده از داکینگ مولکولی، می‌توان هزاران ترکیب را به طور مجازی و در زمان کوتاه‌تری ارزیابی کرد، که این امر به طور چشمگیری هزینه‌ها را کاهش می‌دهد و سرعت فرآیند کشف دارو را افزایش می‌دهد. افزایش دقت در شناسایی ترکیبات مؤثر با استفاده از پایگاه‌های داده معتبر به محققان این امکان را می‌دهد که به ساختارهای دقیق پروتئینی و ترکیبات دارویی دسترسی پیدا کنند. این اطلاعات، هنگامی که همراه با شبیه‌سازی‌های داکینگ استفاده می‌شوند، دقت انتخاب ترکیبات مؤثر برای بررسی‌های آزمایشگاهی را به طور چشمگیری افزایش می‌دهند [۸].

چالش‌های داکینگ مولکولی و بیوانفورماتیک:

با اینکه داکینگ مولکولی دارای مزایای چشمگیری است، با چالش‌های متعددی مواجه است! دقت پیش‌بینی و صحت داده‌ها یکی از مشکلات عمده در داکینگ مولکولی، دشواری پیش‌بینی دقیق پیکربندی‌های لیگاند-پروتئین است. بسیاری از توابع امتیازدهی قادر به ارزیابی دقیق نیروهای بین‌مولکولی و تعاملات پیچیده در سیستم‌های زیستی نیستند [۸،۹]. انعطاف‌پذیری پروتئین‌ها و لیگاندها در بسیاری از الگوریتم‌های داکینگ، پروتئین‌ها به صورت ساختارهای صلب در نظر گرفته می‌شوند، در حالی که در واقع، پروتئین‌ها در محیط زیستی انعطاف‌پذیر هستند و ممکن است در فرآیند اتصال به لیگاند تغییرات ساختاری داشته باشند. استفاده از داکینگ انعطاف‌پذیر و روش‌های دینامیک مولکولی می‌تواند به حل این مشکل کمک کند [۹].

عوامل محیطی و شرایط فیزیولوژیکی بسیاری از برهم‌کنش‌های زیستی تحت تأثیر شرایط محیطی مانند، pH، دما، غلظت یون‌ها و وجود کوفاکتورها قرار دارند. اما بیشتر مدل‌های داکینگ مولکولی این پارامترها را در نظر نمی‌گیرند، که این موضوع می‌تواند منجر به نتایجی که زیاد دقیق نیستند شود.



نتیجه‌گیری:

به علاوه فارماکوژنومیک اطلاعات سطح ژنوم را درباره پاسخ متغیر دارو فراهم می‌کند، که برای شرکت‌های دارویی خیلی مهم است تا داروی جدید علاوه بر داروی بیماری‌های کمیاب طراحی شود و داروهای موجود پیش از این را ذخیره کند. اگرچه بیوانفورماتیک و فارماکوژنومیک هنوز در فاز اولیه شان و در

حال حاضر با موانعی روبرو هستند. ولی آنها نیز پتانسیل کافی دارند تا به فرایند توسعه دارو در آینده کمک کنند. و در ادامه همینطور محدودیت‌های محاسباتی شبیه‌سازی‌های داکینگ برای مولکول‌های بزرگ یا ترکیبات پیچیده نیازمند قدرت پردازشی بالایی هستند. با این حال، پیشرفت‌های اخیر در پردازش موازی و استفاده از فناوری‌هایی و رایانش ابری، در حال کاهش این محدودیت‌ها هستند و به محققان این امکان را می‌دهند که شبیه‌سازی‌های پیچیده‌تر را با دقت بالاتری انجام دهند.

منابع:

1. Dao NA, Vu T-D, Chu D-T. Bioinformatics in drug discovery. *Advances in Bioinformatics*: Springer; 2024. p. 239-48.
2. Zhang S, Liu K, Liu Y, Hu X, Gu X. The role and application of bioinformatics techniques and tools in drug discovery. *Frontiers in Pharmacology*. 2025;16:1547131.
3. Qiu Y, Zhang Y, Deng Y, Liu S, Zhang W. A comprehensive review of computational methods for drug-drug interaction detection. *IEEE/ACM transactions on computational biology and bioinformatics*. 2021;19(4):1968-85.
4. Ru J, Li P, Wang J, Zhou W, Li B, Huang C, et al. TCMSp: a database of systems pharmacology for drug discovery from herbal medicines. *Journal of cheminformatics*. 2014;6:1-6.
5. Rosita AS, Begum TN. Molecular Docking analysis of the TNiK Receptor protein with a potential Inhibitor from the NPACT databas. *Bioinformation*. 2020;16(5):387.
6. Dunkel M, Fullbeck M, Neumann S, Preissner R. SuperNatural: a searchable database of available natural compounds. *Nucleic acids research*. 2006;34(suppl_1):D678-D83.
7. Pathania S, Ramakrishnan SM, Bagler G. Phytochemica: a platform to explore phytochemicals of medicinal plants. *Database*. 2015;2015:bav075.
8. Saikia S, Bordoloi M. Molecular docking: challenges, advances and its use in drug discovery perspective. *Current drug targets*. 2019;20(5):501-21.
9. Jakhar R, Dangi M, Khichi A, Chhillar AK. Relevance of molecular docking studies in drug designing. *Current Bioinformatics*. 2020;15(4):270-8.

بیوانفورماتیک ساختاری و طراحی پروتئین

به قلم محمد یعقوب زاد - دانشجوی دکترای بیوشیمی دانشگاه تبریز

تحقیقات نشان می‌دهند که طراحی پروتئین به سرعت در حال پیشرفت است و این امکان را فراهم می‌سازد تا در زمینه‌های زیست‌فناوری و پزشکی کاربردهای جدیدی پیدا کند.



طراحی پروتئین یکی از زمینه‌های نوآورانه علمی است که در آن دانشمندان پروتئین‌ها را برای اهداف خاص در زیست‌فناوری و پزشکی ایجاد یا اصلاح می‌کنند. پیشرفت‌های اخیر، با کمک ابزارهای محاسباتی قدرتمند و هوش مصنوعی، این امکان را فراهم کرده است تا پروتئین‌هایی مهندسی شوند که از آنزیم‌های صنعتی گرفته تا داروهای نجات‌بخش زندگی را شامل می‌شوند [۱-۳]. تحقیقات نشان می‌دهد که هوش مصنوعی و یادگیری ماشین — مانند ابزارهایی همچون AlphaFold و RoseTTAFold — طراحی پروتئین را با پیش‌بینی ساختارهای آن با دقت بالا دگرگون کرده‌اند. این موضوع امکان طراحی *de novo* را فراهم می‌سازد، یعنی ساخت پروتئین‌هایی کاملاً جدید از ابتدا، متناسب با عملکردهای خاص. برای مثال، طراحی آنزیم‌هایی برای انجام واکنش‌های شیمیایی و پروتئین‌هایی که بتوانند به سلول‌های سرطانی متصل شوند، در حال انجام است و در توسعه داروهای امیدبخش هستند [۴]. به نظر می‌رسد طراحی پروتئین در حال تحول زیست‌فناوری است؛ از طریق ایجاد آنزیم‌هایی برای فرآیندهای صنعتی مانند تولید سوخت زیستی، و همچنین ساخت نانومواد خودآرا برای انتقال دارو یا حسگرهای زیستی. این پیشرفت‌ها می‌توانند منجر به کاربردهایی پایدارتر و کارآمدتر در صنعت شوند، مانند تولید ترکیبات کایرال برای داروسازی. شواهد نشان می‌دهد طراحی پروتئین نقش کلیدی در پزشکی ایفا می‌کند، به‌ویژه در تولید درمان‌های نوین، واکسن‌ها و سامانه‌های انتقال دارو. به‌عنوان نمونه، آنتی‌بادی‌های مونوکلونال برای درمان سرطان و

از ابتدا — در حال رشد است و در کشف دارو و ساخت واکسن‌ها پتانسیل بالایی دارد. روش‌های محاسباتی در طراحی پروتئین طراحی *de novo* پروتئین

تحول در طراحی پروتئین عمدتاً ناشی از پیشرفت‌های محاسباتی، به‌ویژه در حوزه هوش مصنوعی و یادگیری ماشین بوده است. بر اساس مطالعات موجود، افزایش قدرت محاسباتی و رشد پایگاه‌داده‌های توالی و ساختار پروتئین‌ها منجر به رویکردهای جدیدی شده که مبتنی بر داده‌های حجیم هستند. ابزارهایی مانند AlphaFold و Ro-SETTAfold به دقت‌هایی

دست یافته‌اند که با روش‌های تجربی رقابت می‌کند، و امکان طراحی *de novo* (از ابتدا) را فراهم ساخته‌اند؛ جایی که پروتئین‌ها بدون تکیه بر قالب‌های طبیعی ساخته می‌شوند. طراحی *de novo* به معنای ساخت پروتئین‌ها بر اساس اصول

فیزیکی تاخوردگی (folding) است، به‌گونه‌ای که توالی‌ها برای رسیدن به ساختارهای مورد نظر بهینه‌سازی می‌شوند. طراحی پروتئین *de novo* شامل تولید پروتئین‌هایی با توالی‌هایی است که با پروتئین‌های طبیعی بی‌ارتباط هستند و بر اساس اصول تاخوردگی (fold-*ing*) هدایت می‌شوند. مقاله‌ی *The coming of age of de novo protein design* که در ژورنال *Nature* به چاپ رسیده است بیان می‌کند که روش‌شناسی‌های محاسباتی اکنون این امکان را فراهم کرده‌اند که ساختارهایی با دقت بالا در حد اتم طراحی شوند و فضای

واکسن‌های مبتنی بر پروتئین مانند SKYCoV-19 برای مثال‌هایی از تأثیر مثبت این پروتئین‌های طراحی‌شده بر سلامت انسان هستند [۶-۸]. همچنین پزشکی شخصی، که در آن درمان‌ها متناسب با هر فرد طراحی می‌شود، به‌عنوان یک کاربرد مهم در حال ظهور است. چالش‌هایی همچون حفظ پایداری پروتئین‌ها در بدن و تبدیل طراحی‌ها به استفاده بالینی همچنان پابرجاست، اما پیشرفت در هوش مصنوعی و درک بهتر تاخوردگی پروتئین‌ها می‌تواند راه‌حل‌هایی برای این مشکلات ارائه دهد [۱۰]. این حوزه

پتانسیل بالایی برای حل مسائل مربوط به سلامت و پایداری دارد، هرچند برخی جنبه‌ها مانند اخلاق پزشکی هنوز محل بحث است.

پیشینه طراحی پروتئین طراحی پروتئین شامل ساخت پروتئین‌های جدید یا اصلاح پروتئین‌های موجود به منظور

دستیابی به عملکردهای خاص است که با بهره‌گیری از زیست‌شناسی محاسباتی، هوش مصنوعی و اعتبارسنجی تجربی انجام می‌شود. تا به اکنون (آوریل ۲۰۲۵)، این حوزه شاهد پیشرفت‌های چشمگیری بوده است، به‌ویژه با ادغام ابزارهای هوش مصنوعی که پیش‌بینی و طراحی ساختار پروتئین‌ها را متحول کرده‌اند. این گزارش، یافته‌های اخیر را از مطالعات علمی گردآوری کرده تا نکات کلیدی پیشرفت‌ها و کاربردها را برجسته کند. **شواهد حاکی از آن است که طراحی *de novo* پروتئین — یعنی ساخت آنزیم‌ها و درمان‌های نوین از**

”به نظر می‌رسد ابزارهای هوش مصنوعی و محاسباتی در حال تحول نحوه مهندسی پروتئین‌ها برای عملکردهای خاص هستند.“

وسیع توالی‌ها فراتر از تکامل طبیعی را کاوش کنند.

کاربردهای طراحی پروتئین در پزشکی

در حوزه پزشکی، طراحی پروتئین نوآوری‌هایی در زمینه درمان، واکسن‌سازی و انتقال دارو ایجاد کرده است [۴]. پروتئین‌های طراحی‌شده مسیرهای خاص بیماری را هدف قرار می‌دهند، مانند آنتی‌بادی‌های مونوکلونال، برای درمان سرطان نمونه‌هایی نیز شامل پروتئین‌هایی هستند که تعاملات مرتبط با آلزایمر و HIV را مختل می‌کنند [۴]. در دارو رسانی و انتقال دارو مجتمع‌های نانومقیاس پروتئینی برای رساندن دارو به سلول‌های خاص استفاده می‌شوند که عوارض جانبی را کاهش می‌دهند. این سیستم‌های دارورسانی عمدتاً می‌توانند به محرک‌هایی مانند تغییرات pH و دما پاسخ دهند. یک حوزه دیگر در کاربردهای طراحی پروتئین در پزشکی، تنظیم و طراحی پروتئین‌ها بر اساس پروفایل ژنتیکی فردی (پزشکی فردی) به‌ویژه در انکولوژی (سرطان‌شناسی) بسیار نویدبخش است. این رویکرد امکان درمان‌های سفارشی برای سرطان را فراهم می‌کند و می‌تواند اثربخشی درمان را بهبود بخشیده و عوارض جانبی را کاهش دهد. مهندسی پروتئین‌ها برای درمان، شامل اصلاح ویژگی‌های آن‌ها به منظور افزایش کارایی و ایمنی است که راهبردهای کلیدی هرکدام از داروهای مهندسی و تولید شده در این زمینه را می‌توانید با اسکن کردن کیو آر کد مقابل، در یک جدول ملاحظه بفرمایید:



کاربردهای طراحی پروتئین در زیست‌فناوری

طراحی پروتئین در حال دگرگون کردن زیست‌فناوری است و در بهینه‌سازی آنزیم‌ها، نانومواد و ابزارهای زیستی کاربرد دارد [۶-۸]. در زیست‌فناوری آنزیم‌هایی برای فرایندهای صنعتی مانند تولید سوخت‌های زیستی و سنتز داروها مهندسی شده‌اند [۶]. آنزیم‌های مهندسی‌شده از پایداری و فعالیت بالاتری برخوردارند که برای تولید ترکیبات کایرال دارویی حیاتی است. در حوزه طراحی نانومواد مؤسسه طراحی پروتئین (Institute for Protein Design) نانومواد بر پایه پروتئین خودآرایش‌یافته (self-assembling) را معرفی می‌کند که برای انتقال دارو، بیوسنسورها و مهندسی بافت کاربرد دارند. این ساختارها با دقت در حد اتم طراحی می‌شوند. در حوزه ابزارهای زیستی نیز سویچ‌ها و حسگرهای قابل برنامه‌ریزی مبتنی بر پروتئین در حال توسعه هستند که پتانسیل بالایی در تشخیص بیماری‌ها و درمان دارند.

دو نمونه پروتئین طراحی شده در زیست‌فناوری عبارت‌اند از:

- KumaMax، آنزیمی برای درمان بیماری سلیاک (Coeliac disease) که هم‌اکنون در آزمایش‌های بالینی توسط شرکت داروسازی تاکدا (Takeda Pharmaceutical Company) قرار دارد (طبق گزارش مؤسسه طراحی پروتئین).
- آنزیم‌های طراحی‌شده با هوش مصنوعی که آلودگی‌ها را تجزیه می‌کنند، به‌عنوان مثال PETase برای تجزیه پلاستیک پلی اتیلن ترفتالات (PET).

چشم‌انداز آینده و چالش‌های طراحی پروتئین

با وجود پیشرفت‌های چشمگیر، چالش‌هایی در این حوزه باقی مانده است. این چالش‌ها عبارت‌اند از اطمینان از پایداری پروتئین درون بدن (*in vivo*)، جلوگیری از تجمع و کلوخه شدن پروتئین‌ها (*aggregation*) و تبدیل یافته‌های پژوهشی به کاربردهای بالینی، که با موانع نظارتی و اخلاق پزشکی روبه‌رو است. ادغام گسترده‌تر هوش مصنوعی برای طراحی پروتئین‌هایی با عملکردهای پیچیده و دقیق، به خصوص آنزیم‌های آلوستریک، کشف چینش‌ها و تاخوردگی‌های جدید پروتئینی (*novel folds*) و توسعه کاربردهای زیست‌سنتزی (*Synthetic Biology*) برای ساخت مسیرهای سیگنالینگ سلولی و طراحی سیستم‌های زیستی قابل برنامه‌ریزی مسیرهای آینده شامل طراحی پروتئین را پوشش می‌دهند [۱۰].

جمع بندی

پیشرفت‌های طراحی پروتئین که به واسطه ابزارهای محاسباتی و هوش مصنوعی شکل گرفته‌اند، در حال متحول کردن زیست‌فناوری و پزشکی هستند [۱-۳]. از طراحی آنزیم‌ها برای فرایندهای صنعتی گرفته تا تولید پروتئین‌های درمانی و واکسن‌های نسل جدید، این حوزه پتانسیل عظیمی دارد. تا آوریل ۲۰۲۵، تحقیقات فعال همچنان در حال رفع چالش‌ها و گسترش افق‌های کاربردی این فناوری هستند. با این حال، تبدیل یافته‌های پژوهشی به کاربرد بالینی همچنان یکی از مباحث بحث‌برانگیز باقی مانده و نیازمند مطالعه، نظارت و ارزیابی‌های دقیق‌تر است تا کاربردهای درمانی پروتئین‌های طراحی شده در دنیای واقعی محقق شود [۴].

منابع:

1. Huang, P. S., Boyken, S. E., & Baker, D. (2016). The coming of age of de novo protein design. *Nature*, 537(7620), 320-327. <https://doi.org/10.1038/nature19946>
2. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589. <https://doi.org/10.1038/s41586-021-03819-2>
3. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., ... & Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557), 871-876. <https://doi.org/10.1126/science.abj8754>
4. Scott, A. M., Wolchok, J. D., & Old, L. J. (2012). Antibody therapy of cancer. *Nature Reviews Cancer*, 12(4), 278-287. <https://doi.org/10.1038/nrc3236>
5. Katsnelson, A., Buzdin, A., & Borisov, N. (2014). Personalized medicine in oncology: next generation. *Nature Reviews Drug Discovery*, 13(12), 921-922. <https://doi.org/10.1038/nrd3603>
6. Sheldon, R. A., & Woodley, J. M. (2018). Role of biocatalysis in sustainable chemistry. *Chemical Reviews*, 118(2), 801-838. <https://doi.org/10.1021/acs.chemrev.7b00203>
7. Tarvirdipour, S., Huang, X., Mihali, V., Schoenberger, C. A., & Palivan, C. G. (2020). Peptide and protein-based nanoparticles as drug delivery systems. *Pharmaceutics*, 12(8), 740. <https://doi.org/10.3390/pharmaceutics12080740>
8. Ricci, F., Adornetto, G., & Palleschi, G. (2012). A review of experimental aspects of electrochemical immunosensors. *Electrochimica Acta*, 84, 74-83. <https://doi.org/10.1016/j.electacta.2012.06.033>
9. Austin, H. P., Allen, M. D., Donohoe, B. S., Rorrer, N. A., Kearns, F. L., Silveira, R. L., ... & Beckham, G. T. (2018). Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proceedings of the National Academy of Sciences*, 115(19), E4350-E4357. <https://doi.org/10.1073/pnas.1718804115>
10. Wang, W., & Roberts, C. J. (2018). Protein aggregation



مصاحبه با دکتر دامون نشتاعلی مدیر عامل "شرکت هومن ژن پارس"

مصاحبه‌گر: مریم فلاحی / نویسنده: امیرمحمد بصیر غفوری / بازنگاری و ویراستاری: محمدصدرا محمدی

-در خدمت دکتر نشتاعلی هستیم. خیلی ممنونیم که وقتتون رو به ما اختصاص دادید. لطف کنید در ابتدای کار حوزه کاری خودتون رو معرفی کنید و راجع به فعالیت های شرکتتون توضیح بدید.

بنده دامون نشتاعلی هستم. فارغ التحصیل دکترای مهندسی برق مخابرات از دانشگاه صنعتی شریف. در مقطع دکترا، استاد همکارم که از آمریکا برگشت پیشنهاد داد که روی DNA کار کنیم و چون موضوع جذابی بود، از مخابرات نسل پنجم که کار می‌کردم کلاً سوئیچ کردم به حوزه بیوانفورماتیک، در این حوزه فعالیت کردم و دکترای من در حوزه بیوانفورماتیک بود، در حوزه سخت افزاری بیوانفورماتیک با دستگاه NGS و در حوزه نرم افزار بیوانفورماتیک با تحلیل داده های NGS، و از اواسط مهرماه سال ۹۵ ایده‌ی این استارت‌آپ به ذهنم رسید و شروع به کار کردیم. ایده‌ی ما این بود که با وجود چالش‌های موجود در کشور و افزایش تولید دیتا در جهان، بتوانیم داده‌های ژنتیکی رو در کشور نگه داریم، تحلیل کنیم و خدمت بدیم و با همون ایده شروع به کار کردیم.

-چی شد که از حوزه برق وارد حوزه بیوانفورماتیک شدید؟

من قبل از این در یکی از دبیرستان‌های علامه حلی معاون پژوهشی بودم. هدفم این بود در حوزه دبیرستان با بچه‌ها تعامل داشته باشم و در اون فضا کار کنم و خیلی مایل به کار در صنعت نبودم. دکتر مطهری (استاد همکار من) با من صحبت کردن و منو قانع کردن که ممکنه در حوزه صنعت بیشتر موثر باشم و یک مسئله جذاب رو به من پیشنهاد دادند، توالی یابی DND. اونجا بود که دیگه تصمیم گرفتم به حوزه بیوانفورماتیک سوئیچ کنم که با نظر خود دکتر مطهری و استاد راهنمایی که در دانشکده برق داشتم، به من کمک کردن که وارد این حوزه بشم.

من اونموقع خیلی زیست و شیمی و... بلد نبودم و مجبور شدم دکترا رو بازبینی کنم. برای اینکه دیتا ساینس کار کنیم باید اون دیتا را بشناسیم. در کل خیلی اتفاقی و با نظر دکتر مطهری و کمک‌های خودشون بود که این سوییج اتفاق افتاد.

-ایده اولیه فعالیت هاتون چطور شکل گرفت و چه مسیری رو برای پرورش آن طی کردید؟

ما سه نفر بودیم. دکتر مطهری، من و یکی دیگه از بچه‌ها (دکترطوسی) که البته الان خارج هستن. در صحبت‌های سه نفره‌ای که اتفاق افتاد دکتر طوسی نظرشون این بود که این ایده به درد کشور می‌خوره و می‌تونیم روش کار کنیم و این ایده اولیه "هومن ژن پارس" بود ولی توی مرحله بعد، سال 98 که شرکت شروع به کار کرده بود و به مشتری خدمت می‌داد، ایده‌ها به عنوان یکسری محصولات که ممکنه در حوزه B2C و مشتری عادی بتونن خدمت رو ارائه بدن شکل گرفت. ظهور اون ایده این بود که ما بتونیم اراده کنیم از طریق سازماندهی دیتای ژنتیکی، توی این حوزه وارد بشیم و بعد به دنبالش فضاها دیگه‌ای هم ایجاد شد.

-از زمان ظهور این ایده تا به این مرحله رسیدن، چه مدت طول کشید؟

ما حدودا از سال 95 شروع کردیم و یکسری فعالیت‌هایی رو انجام می‌دادیم شش ماه برای توسعه‌ی سازمانی تلاش کردیم. با توجه به طولانی شدن زمان، از اوایل 96 وارد بازار شدیم و رفتیم ببینیم بازار چه نیازی داره، گفتیم به جای اینکه خودمون بشینیم و صرفا با ذهنیت خودمون توسعه بدیم، بریم داخل مارکت ببینیم چه خبره. تغییر جدی‌ای که در سیستم و در مجموعه رخ داد، از همون سال 96 بود که رفتیم داخل مارکت و با متخصصین صحبت کردیم، درخواستشون رو فهمیدیم و تلاش کردیم اون درخواست‌ها رو با کمترین هزینه توسعه بدیم و با حداقل هزینه بهشون خدمت ارائه بدیم. از اونجا بود که شرکت یه جورایی تغییر رویکرد داشت، از اون رویکرد اولیه‌اش فاصله گرفت و تلاش کرد که متناسب با خدمتی که داخل بازار نیاز هست، فعالیت کنه و در حوزه بیماری‌های ژنتیکی و سرطان وارد مارکت شدیم.

-چون ایده‌ی شما جدید بود، در این مسیر با چه چالش‌هایی مواجه شدید و چجوری چالش‌ها رو پشت سر گذاشتید؟

تقریبا 2-3 سال قبل، فرهنگ سازی در ذهن متخصصین ژنتیک در حوزه بیماری‌هایی مثل بیماری‌های ژنتیکی و سرطان شروع شده بود و درخواست‌های درمانی جدید بررسی می‌شد. بنابراین ما نیاز به فرهنگ سازی خیلی گسترده‌ای نداشتیم. یه جورایی از اوایل این مارکت داخل کشور بودیم که متخصصین با ان آشنا بودند. بیشترین چالش‌ها که باهاش درگیر بودیم این بود که من یه چیزی متناسب با درخواست متخصصان طراحی بکنم و خدمت بدم تا بتونن ازش استفاده کنن. این چالش اول بود؛ ولی نکته‌ی اینه که اونجا چون دیتای تحلیل ژنتیکی به دیتابیس‌ها خیلی وابستگی دارن، به دلیل اینکه دیتای ایرانی نداشتیم باید همیشه فکر می‌کردیم که این دیتا رو چجوری بسازیم. بنابراین ما با این رویکرد وارد مارکت شدیم که بتونیم دیتابیس بسازیم و بتونیم با توجه به علمی که از دیتابیس استخراج می‌کنیم، در سطح مارکت

خدمت ارائه بدیم. در واقع چالش این بود که من باید خودم رو ثابت می‌کردم. باید با خودم می‌گفتم گرچه برق شریف خوندم ولی NGS بدم و میفهممش. ژنتیک رو درک می‌کنم، آنالیز دیتا رو متوجه میشم. موضوع بعدی جلب اعتماد متخصصین بود که کار سختی بود. ما مذاکره های نفر به نفر داشتیم باهاشون که بتونیم اثبات کنیم که ما متوجه این مسائل هستیم و می‌تونیم بهشون کمک کنیم چون از طرفی اگر دقت کنید، متخصصان در حوزه تشخیص بوده و با خود بیماران در ارتباط هستند و یه اشتباه ممکنه براشون خیلی گرون تموم بشه. ما اینجا با کنترل کیفیت خیلی بالا کار می‌کردیم که بتونیم سرویس‌های درستی بهشون بدیم که به ما اطمینان کنن و کنترل کیفیت ما و اون چیزهایی که بهش پایبند بودیم، زیر پا نداشتیم و همیشه با اون کیفیت مطلوب کار کردیم و بعضا کیفیت رو بیشترش کردیم، آپدیتش کردیم و متناسب با توانمندی‌هایی که داشتیم به طور مرتب به روزش کردیم و خب این خیلی بهمون کمک کرد. درحوزه تست‌هایی که بعدا در B2C شکل گرفت کلا مسئله تغییر کرد. یعنی فرهنگ سازی می‌کردم و نشون می‌دادم این مسائل، مسائلی هستند که بصورت علمی روش کار کردم، داخل مارکت نشون دادم می‌تونن یه سری کاربردهایی داشته باشن که به شما کمک بکنه و اونجا مسئله خیلی پیچیده تر از اول بود چون فرهنگ شکل نگرفته و باید تلاش می‌کردیم واسه تبلیغاتش، متوجه شدیم باید خودمون تلاش بکنیم که به گوش مردم برسونیم و بعد نشون بدیم که این تست‌ها، تست‌های جدیدی هستن و می‌تونن به شما کمک کنن. بنابراین در این حوزه دوم، چالش‌ها چندین برابر حوزه اول بود. بحث فاندینگ و سرمایه گذاری هم چالش داره و پای کار موندن سخته. مثلا ما از سال ۹۶ که به طور جدی شروع کردیم، حدودا ۴-۳ تا تحریم جدی داشتیم که اتفاقا یکی از بنیان‌گذاران اصلی شرکت ما در یکی از تحریم‌ها رفت و ما می‌تونستیم دل‌سرد بشیم و بریم و با این اتفاق ممکن بود دیگه شرکت هم نباشه، ولی خب با این مسائل در کشور کار کردیم، کار سختیه و هر روز چالش‌های خاص خودشو داره و در اثر افزایش قیمت‌ها برای مردم محدودیت‌هایی ایجاد میشه که بعضا توان استفاده از خدمات ندارن. این اتفاق کلاً به عنوان چالش تلقی میشه و مدیریتش در مجموعه‌هایی که استارت‌آپی‌ان، خصوصی‌اند و تلاش می‌کنند که با همین شرایط، ارائه‌ی خدماتشون رو مدیریت کنن کار سختیه.

-درخواست متخصصین از شما چی بوده و با چه ملاک‌هایی شما رو تایید می‌کردن؟

ما در حوزه پزشکی فعالیت می‌کنیم. به نظر می‌رسه که تعامل و کار کردن با پزشکان یک مقدار سخت باشه. از طرفی تعامل با پزشکان ژنتیک (ژنتیک پزشکی) چند لول از پزشکان سخت‌تره. بنابراین ما با جامعه‌ای طرف بودیم که سخت تعامل می‌کرد، سخت قبول می‌کرد که ما می‌تونیم بهشون کمک کنیم و فکر می‌کردن که خودشون از پس کارشون بر میان و ممکنه به ما نیاز نداشته باشن. کاری که ما می‌کردیم در این فضا، درجلسات نشون می‌دادیم که ما می‌فهمیم و از دید علمی کاملاً باهاشون اصولی صحبت می‌کردیم و اونا متوجه می‌شدند که ما از دید علمی مشکلی نداریم و توجیه هستیم. در مرحله انجام تست‌ها و اعتماد سازی، یکی از مدل‌ها این بود

که سمپل هایی که قبلا تست کار کرده بودن رو به ما می دادند که ما تاییدش کنیم و بهشون جواب بدیم. یعنی تست هایی که خودشون دیتا رو داشتن و می خواستن نظر ما رو بدونن ببینن ما هم به نتایج اونا می رسیم یا نه. مثلا در سمپل های جدید ارسالی، خدمات و فرایند ما رو تست می کردن. یکی از عملکرد های ما پیشنهاد و ارائه نتایج سمپل های معروف در دنیا به متخصصین بود. بعد از به اشتراک گذاشتن نتایج باهاشون، علیرغم اینکه خیلی اطمینان نداشتن کم کم با ما شروع کردن به کار کردن. دیدن یه سری مزیت داریم براشون ایجاد می کنیم، کم کم کوتاه اومدن تا دیگه اون اعتماد شکل گرفت. حتی اینجوری بود که ما به اون فرایند قانع نمی شدیم و به طور مرتب بهبود می دادیم و متخصصان هم ما را همراهی می کردن. این فرایند حتی در مرحله ای قرار گرفت که ما در تشخیص بیماری هم کمکشون می کنیم ولی مهم ترین موضوع اینه که اعتماد سازی تو زمان شکل می گیره. اولش خیلی سخته ولی در طول زمان، بر اثر نتیجه ی مثبت این اتفاقاتی که می بینن اعتمادشون شکل می گیره.

-فرایند تحلیل داده های ژنتیکی در شرکت شما به چه صورت انجام می شود؟

ما در شرکت چندین فرایند داریم که صرفا تحلیل دیتا نیست. اینکه سمپل برای ما میاد و ما بررسی روی اون انجام می دیم بعدش میره برای NGS و دیتا برمی گرده. بنابراین یک روند بررسی نهفته اجرایی وجود داره. یک روند پردازش در تحلیل دیتا هست که ما حجم زیادی از دیتای خام رو می گیریم که فرمتش مثلا فست کیو هست. فرض کنید سمپل، نمونه یا مثلا دیتایی که دانلود میشه پردازشش ممکنه 6 ساعت، 8 ساعت روی سرور ها طول بکشه. در پایتون خودکار اتفاق میفته و با علامت گذاری (لیبلینگ) ای که داریم یعنی حریم شخصی وجود داره که سمپل ما، علامت گذاریش، نامش و... محفوظ می مونه و در نهایت نتیجه ای که از این فرایند می گیریم دوباره با پایتون به طور خودکار متناسب با علائمی که متخصصین برای ما می فرستند در کمتر از 5 دقیقه، نهایت 10 دقیقه می تونیم به تشخیص اون بیماری برسیم. مثلا فردی با مشکل ناشنوایی میاد پیش ما، دیتاشون رو میگیریم که این دریافت داده حدودا 6 ساعت طول می کشه. طی یک اتوماسیون مجزا، علائم بالینی رو بهش می دیم مثلا این فرد ناشنوا، 6 سالشه و ازدواج پدر و مادرش خویشاوندی بوده، این رو بررسی می کنیم که چه عامل ژنتیکی ای وجود داشته که باعث ناشنوایی در این بیمار شده. تقریبا این اتوماسیون بعد از 4-5 سال شکل گرفته و تا کنون چندین بار بروز رسانی شده. درحالی که اگر یک متخصص این فرایند رو طی 2-3 ساعت انجام می ده، در اینجا به اندازه 5-10 دقیقه زمان سپری میشه. بنابراین اتوماسیون هم در طول زمان شکل گرفته. این در حوزه تشخیصه ولی در حوزه B2C دیتا میاد و متناسب با دیتابیس و جمعیت ایران و متناسب با فنوتایپی که مد نظرمون هست، ارزیابی میشه که اون ارزیابی از قبل برای ما تفهیم شده و یاد گرفتیم با هم ترکیب کنیم بعدش به یه نتایجی می رسیم و در جامعه ایران عادی نرمالایزش می کنیم. برای مثال می گیم 20% جامعه ایرانی در فنوتایپ های مختلف اینطوری هستن که این فرایند پیچیده تر از فرایند اولیه هست. این تشخیص مهمتره، ریسکش بیشتره و نباید خطا داشته باشیم ولی در اون یکی ریسک کمتره ولی کار پیچیده ترهست.

-چطور از توالی استخراج شده به نتایج کاربردی مثل رژیم غذایی پیشنهادی می‌رسید؟

فضای اول، تشخیص بیماری بود که من اونجا درمورد دلایل صحبت می‌کنم. یعنی در مورد علت این بیماری. یک فردی که ناشنوا هست چرا ناشنوائه و مثلاً ناقل هست یا نه. در مورد 0% و 50% و 100% صحبت می‌کنیم. 100% یعنی من یه مجموعه ای از یافته ها واسه فرد داشتم که علت ناشنواییش را پیدا کردم. 50% یعنی من می‌تونم ناقل بیماری ناشنوایی باشم و گرچه خودم سالمم، ولی به نسل بعد می‌تونم منتقل کنم. 0% یعنی با دانش فعلی من نتونستم علت ناشنوایی را پیدا کنم. توی حوزه B2C که الان بهش میگن ارزیابی ریسک، اینجوریه که شما در مورد یک فنوتایپ کامپلکس صحبت میکنید که لزوماً دلیل نمودش رو نمی‌دونید و نمی‌دونید چرا سرطان رخ میده، نمی‌دونید چرا پارکینسون رخ میده، نمی‌دونید شخصیت فرد چرا داره این سبکی میشه ولی نکته اینه که ژنتیک توی اون دخیله. جایی که ژنتیک با محیط با هم دیگه ارتباط زیادی دارن و فنوتایپ کامپلکس وجود داره، درصدی از این اتفاق وراثتی و درصدی هم محیطیه. مثلاً بخشی از درونگرایی، هوش ریاضی و... ارثی، و یه بخش دیگه اکتسابی به شمار میره. شما با توجه به محیطی که توش زندگی کردین، بعد از تولد در خانواده‌ای که بودین، بازی‌هایی که کردین، مدرسه ای که رفتین، درس‌هایی که خوندین، مطالعه‌ای که داشتین تونستین از اون توانمندی ژنتیکی خودتون استفاده کنید. یا مثلاً ممکنه به یک معنا و نگاه دیگه ای شما ژن خوب داشته باشید و من ژن بد داشته باشم. مثلاً شما با 1 میلیون هزینه کردن در یک سال به مهارت و جایگاهی برسید که من با 100 میلیون هزینه کردن طی 10 سالم نرسم. این ظرفیت در افراد وجود داره و نهفته است. حالا مسئله اینه که من از یه صفت ژنتیکی چقدر می‌تونم این ظرفیت پردازشی را حساب کنم. یعنی مثلاً فرض کنید شما به طور ژنتیکی شخصیت درونگرا دارید که با یک ظرفیتی از پدر و مادرتون به ارث رسیده. من با تست‌های کنونی و علم کنونی ژنتیک چند درصد می‌تونم این رو مشخص کنم؟ و این درصد مشخص شده چقدر صحت داره؟ اینجا معیارهای ارزیابی دقتی که مشخص می‌کنیم، معیارهای سختی هستن. بیشترین دقت این ارزیابی‌ها در سرطان هست و کمترین دقت آن در جاهایی است که اعصاب و مغز درگیر میشن یعنی فنوتایپ‌هایی که مربوط اونا میشه. مثلاً فرض کنید بیماری‌های روانشناختی و اختلالات سایکولوژیکال، بیماری‌های فراگیر عصبی و... اینجاها دقت پایین‌تری دارن. دقتش چند درصده؟ 50% اگر رندوم باشه و 100% اگه بیشترین حالت باشه، اینا حدود 60% هستن. در سرطان تقریباً 85% هست. بقیه فنوتایپ‌ها یا علائمی که در ادامه صحبت می‌کنیم، شخصیت باشه، هوش باشه، هوش ریاضی باشه و... همه اینها در این بازه قرار می‌گیرند که بخشی از اون فاز مثبت داره و بخشی فاز منفی، و برآیند اینها میشه اون امتیازی که داریم در موردش صحبت می‌کنیم ولی نکته مهم ما اینه که ادعای ما بر این اساس بود که ژنتیک به تنهایی نمی‌تونه در این حوزه نظر بده. حتی همین بخش تشخیص بیماری‌ها هم که داریم کار می‌کنیم بعضاً ممکنه چند تا ژن پیدا بشن و باید بره سمت متخصص بالین که یه پزشک مشاور ژنتیکه، که علائم رو دقیق‌تر متوجه میشه، یه سری چیزارو در مریض می‌تونه دوباره چک بکنه و بعد تایید کنه یکیش درسته و دوتاش نادرسته. همیشه ادعای ما این بوده که ژنتیک به متخصص

بالین کمک می‌کنه. این نگرش ما جزو نگرش‌های خاص در دنیا است. اینجوریه که کلی از کمپانی‌های ژنتیک از 2017 ادعاشون این بود که من با موشکافی ژنوم شما، به شما میگم که تغذیه شما کمبود فلان ویتامین را داره و شما بیا از فلان مواد غذایی تغذیه کن. کلی از این کمپانی‌های اینجوری به وجود اومدن و ما چون رویکرد نداشتیم، اکثر این کمپانی‌ها بعد از 2-3 سال شکست خوردند چون نگاه منفرد و تک بعدی داشتن ما می‌گفتیم ژنتیک به تنهایی نمی‌تونه این کار را بکنه و دقت پایینه و حتماً باید یک متخصص بالین اون رو تفسیر کنه. این تست‌ها واقعا یک دقتی داره و یک چیزی را نشون میدن. اینکه من از این تست‌ها چطور استفاده کنم موضوعیت داره. ما از همون ابتدا که این تست‌ها رو راه انداختیم، در کنار هر تست یک متخصص بود، در کنار میزان بروز استعداد یک بیماری روانشناختی بود، بازوی ارزیابی سلامت‌مون یه مشاور ژنتیک بود، در حال حاضر کنار روند نماد گذاریمون یک متخصص تغذیه هست. مسئله اینه که دنبال علیت نیستیم و دنبال انسجام روند هستیم. یعنی مثلاً فرض کنید یکی داره از بیرون میاد داخل، کفشش خیس و جای پای خیسش می‌مونه. به ذهن می‌رسه که احتمالاً این فردی که اومده و جای پاش خیس بوده یه بخشی از یک روند رو داره طی می‌کنه، با اینکه بیرون داره بارون میاد یا اینکه بیرون زمین خیس بوده خب و این قطعاً علیتش نیست. ممکنه این فرد خودش آب گرفته رو پاش و خیس شده باشه. ولی کجا این انسجام روند زیاد میشه؟ وقتی ده نفر میان داخل و مثلاً هر ده نفر پاشون خیس. یا حداقل این ده نفر صحبت‌های مختلفی از رویکرد‌های مختلف در مورد هوای بارانی می‌کنن. اینجا اطمینان من بیشتر میشه. توی مسائل انسجام روند اینجوریه. دیتا زیاده و بررسی میشه و تلاش می‌کنن ببیند که چه شکلی این جامعه مثلاً چشم قهوه‌ای چشمشون آبی شده؟ تفاوتشون چی بوده؟ این تفاوت‌ها را چک می‌کنن. بعد میان تست می‌کنن، تاییدش می‌کنن که مثلاً چند درصد از جامعه رو میتونن تفسیر کنن. و نکته اینه که در اون بستر، چرا مسائل به این سمت میرن؟ بخاطر اینکه در ژنتیک هنوز بخاطر جامعیت غیر قابل نفوذی که وجود داره، مدلی وجود نداره که بفهمم علیت عوض شدن رنگ چشم آبی‌ها چیه. ظاهراً خیلی هم پیچیده است. ممکنه هیچوقت به علیت نرسیم. بنابراین تلاش می‌کنم با زیاد کردن رخدادهای موجود در یک اتفاق، علت‌های اون لایه پایین رو به دست بیارم. در حال حاضر دارن تلاش می‌کنن مالتی اومیکس نگاه کنن به این رویکرد‌ها. یعنی در کنار هم بگذارند و تکی نگاه نکنند. با این کار قطعاً دقت بیشتر میشه. دنیا رو به پیشرفت هست و مدام در حال رشد کردنه، در بحث سرطان خیلی رشد کردن چون دغدغه اول دنیای سلامت بوده و ژنتیک قرار بوده بازوی اصلی کمک‌کننده در حوزه سرطان باشه و اونجا سرمایه بیشتری وجود داره.

-میزان دقت تحلیل‌های شما چقدر هست؟

یه دقت از بین همون 60% تا 90% و علائم و فنوتایپ‌های مختلف این وسط قرار می‌گیرن و در تعیین درصد‌ها دخیل‌اند. هرچی به سرطان و حس نزدیکتر باشه دقت بالاتره و هرچی به حوزه‌ای که توش ذهن قاطی میشه، بیماری منسجم‌تر میشه و دقت پایین‌تره. برای اسکیزوفرنی یکی از کم‌دقت‌ترین رویکرد‌های اختلالات ژنتیکی ذهنی حدود 60-65% دقت داره

ولی در تشخیص سرطان سینه و دیابت به دلیل موفقیت های اخیر، 85-90% دقت دارد. 50% به طور رندوم و میانگین دقته، و 100% بهترین حالت ممکن.

-آیا تا کنون خطا یا شکایتی در این زمینه براتون پیش اومده ؟

اینجا ما مفهومی از شکایت نداریم بلکه مفهوم از نارضایتی داریم و در مقابل آن مفهوم رضایت داریم که مشتری از این چیزی که بهش گفتیم چقدر استقبال کرد؟ و حالاتوی صفحات فضای مجازی ما می تونید این موارد رو چک کنید. در "مای اسمارت جین" که در واقع تست هایی هستن از مشتری هایی که حتی بهشون گفتیم احتمالاً MS داری و فرد نمیدونست MS داره و 6 ماه بعد MS گرفته. یا مثلاً بهش گفتیم که شما ریسک بالایی برای درگیر شدن با سنگ کلیه دارید بعد رفتن پیش همین پزشک ها و چک کردن و دیده سنگی وجود نداره ولی به مدت بعد سنگ کلیه گرفته. این ها نشون میدن که این تست ها کمک کننده ان ولی اینکه چجوری کمک می کنن مهمه. لزوماً ممکنه همه تشخیص ها درست نباشن و باید یک متخصص در کنارش باشه و به طور صحیح از این تست ها استفاده کنند.

-برخی نظرات منفی در محافل علمی مطرح میشود مبینی بر اینکه نتایج ارائه شده توسط شما مبنای علمی نداشته و حتی به طالع بینی شباهت دارد. پاسخ و دیدگاه شما نسبت به این انتقادات چیست؟

ما طالع بینی می کنیم! ولی به صورت علمی. ژنتیک کلاً فلسفه ش اینه که پیش بینی انجام بده و طالع بینی هم یک نوع پیش گویی هست. مثلاً افراد میان اینجا و تست میدن، اصلاً به یک علت دیگه اومدن ولی بعد از تست متوجه میشن که سرطان هم داره یا دست کم ریسک ابتلای فرد به سرطان رو براش مشخص می کنیم. کسی جهش لایکلی پاتوژن پاتوژن کنسر داشته باشه با احتمال بالای 90% سرطان میگیره و ما داریم پیش گویی می کنیم که این اتفاق می افته. این اصل ژنتیکه که داره یکسری اطلاعاتی میده که هنوز رخ نداده و در طول زمان احتمال رخ دادنش هست. مبنای ما اطلاع رسانی به مخاطب و ارزیابی احتمالات برای فرد هست، بواسطه بررسی اطلاعات ژنتیکی فرد؛ به همین خاطر کارمون یجورایی شبیه به طالع بینی، شبیه جادوگری و اینکه وقتی من دارم یه مفهومی رو میگم که فرد نسبت بهش حس نداره و براش اتفاق نیفتاده، یه گاردی می گیره که این بررسی شما ممکنه اشتباه باشه. این مسئله در واقع ذات ژنتیکه که اینجوریه. مسئله دوم این هست که در علم ژنتیک در سطح دنیا، رویکرد همینه؛ تلاش میشه که ژنتیک به یک تشخیص کمک کنه. اتفاقی که اخیراً داره می افته اینه که پزشکی داره تغییر می کنه و رویکرد های جدید پزشکی دارن وارد دنیا میشن که پیش بینی می کنه و ریسک شما را ارزیابی می کنه و بر اساس اون ریسک، به شما یکسری توصیه و پیشنهاد میده. ما در حال حاضر در این حوزه کار می کنیم. یجورایی ژنتیک در این حوزه، ورژن جدید پزشکی هست که قبل از اینکه شما تب بکنید به شما میگه ریسک تب گرفتن چقدره. و فرقی با تشخیص اینه که نمیکه 0 یا 1. به شما میگه مثلاً 50% ریسک تب گرفتن دارید ولی ممکنه 50% برای شما اطلاعات نداشته باشه، یعنی همون رعایتی که شما تو زندگی می کنید بقیه هم باید بکنن ولی

ولی نکته اینه که در این روش های درمانی جدید که در دنیا داره روش کار میشه اینه که لایف استایل شما و اون اراده شما تو اون زندگی کاملا موثره که شما چه در مسیری حرکت می کنید. نکته خیلی مهم اینه که حتما با متخصص هایی که تو این حوزه کار میکنن ارتباط بگیریم و به یه زبان مشترک برسیم. که ما این کار را انجام دادیم یعنی متخصصین ما هم ژنتیک متوجه میشن و هم ما متوجه میشیم که اون دغدغه متخصص چیه و داریم تلاش میکنیم که یه سرویس واحد بسازیم. ما جواب های که میدیم دست مردم رو QC می کنیم، یعنی قبل از اینکه جواب بدیم به افراد از QC رد شده که اینجا تایید میشه یا ریجکت میشه و هر تستی را وارد مارکت نمی کنیم. -فرایند اعتبارسنجی شما چجوری انجام میشه؟ و نهاد نظارتی یا سازمان بالا دستی شما کدوم سازمانه و از نظر صنفی در چه صنفی تعریف می شید؟

در ایران و در این حوزه مقررات مشخصی نداریم و ما هم کاملا قاچاقی و غیرقانونی کار کردیم! موقعی که کرونا اومده بود کار های مختلفی می کردیم. مثلاً یه کاری که در سطح دنیا انجام شده بود فکر کنم برای MIT هاروارد، سازمان تحقیقاتی شون که درگیر مسائل کرونا بود از کل دنیا متا دیتا (سطح وسیعی از داده های آماری) می گرفت، توالی یابی ژنوتایپ افراد هم می گرفت و بررسی انجام میدادن که این ژنوتایپ چقدر ریسک کرونا داره و اگر کرونا بگیره چقدر سخت می گیره و میزان شدت ابتلای فرد به بیماری چقدر هست، چرا؟ چون کرونا تاثیرش تو افراد بر اساس تایپ ژنتیکی تفاوت هایی داره. یکی کرونا می گیره خیلی خفیف، یکی که خیلی قوی می گیره و درگیر میشه. این مجموعه، دیتای خیلی وسیعی جمع آوری کرده و نتیجه اش رو منتشر کرده بود. ولی هیچ وقت نمی گفت که ماسک نزنید و این کارهای مراقبتی را نکنید. یعنی تلاش میکرد با همون توصیه های راهبردی ای که در دنیا بود (که حتی بخشی از آن هم اشتباه بود) بیاد کمک کنه که شما ریسک خودتون رو بشناسید. و چون این ریسک ها صفر و یک نیستند، شما باید تصمیم بگیرید که در لایف استایل خود چیکار کنید. اگر ریسک ابتلای شما بالا باشه مثلاً باید کمتر جای شلوغ برید و یکسری تشخیص ها مطرح میشه که به خودتون برمی گرده و به پزشکی برنمی گرده. پزشکی نمی تونه بگه چیکار کنید یا نکنید چون این موضوع مستقیماً متوجه لایف استایل خود شما است. اونجا ما دیتا گرفتیم آوردیم تو دیتابیس خودمون آنالیز کردیم و رفتیم جلو بعد در سطح سرویس هامون یک پنل اضافه کردیم که مثلاً اگه میخوای بدونی، من می تونم رایگان بهت بدم. بعد از اون وزارت بهداشت نامه زد که تو کار پزشکی نباید دخالت کنید و ما هم توضیح دادیم که این کار رو نمی کنیم بلکه فقط کمک می کنیم و حتی تشخیص هم نیست. دوباره نامه زد که مجبور میشیم جمع تون کنیم! دیگه ما مجبور شدیم که اون رو برداریم ولی تست های دیگه رو برداشتیم و تا امروز که صحبت می کنیم، هیچ مقررات مصوبی در مورد تست آنالیز دیتای ژنتیکی تو حوزه عیب یابی ژنتیکی افراد موجود نیست؛ یعنی ما که شرکت بیوانفورماتیک هستیم از هیچ نهاد وزارت صنعت، وزارت بهداشت و... نمیتونم مجوزی رو که بر اساس مجموعه ای از مقررات تنظیم شده بگیریم که بگه این کاری که می کنیم درسته. تازه بعد از 10 سال، وزارت بهداشت امسال در مورد مجوز ارسال نمونه به خارج از کشور یک کاری انجام داد. ما چون استارت آپی هستیم، کلاً تو این فضا نمی گنجیم و خیلی بهتر



MSG



می‌تونیم رشد کنیم. مقررات و مجوز و این‌ها معمولاً دست و پا گیره ولی می‌تونه کمک کننده باشه. چون مقررات محدود کننده نبود ما رشد کردیم و به اینجا رسیدیم. در این 7-8 سالی که کار کردیم، از دید حقوقی کاملاً غیرقانونی بودیم. حتی در مواقع عملاً قاچاقچی یکسری کارهای ژنتیکی بودیم! به این معنی که وقتی داشتیم روی دیتاهای ژنتیکی کار میکردیم، یکسری نامه‌ها از مجلس و قوه قضاییه اومده بود که شما دارید کار غیرقانونی می‌کنید...

-برای دانشجویان رشته‌های زیست‌شناسی چه پیشنهادی دارید؟ چه مهارت‌هایی لازم است یاد بگیریم و خودمان را در آن ارتقا دهیم؟

بهترین عملکرد شما اینه که بین رشته‌ای کار کنید. یعنی برای مثال اگر می‌خوان کامپیوتر بفهمین، باید قطعاً با ai بتونید کار کنید و به اینکه فقط در زیست‌شناسی، ژنتیک رو می‌فهمید، قانع نشید. دنیا دنبال این نیست که یه تست‌هایی که 90% یا 100% درسته، داخل مارکت بیاره که ازش استفاده کنند، دنبال اینه که یه سری بسترهایی به وجود بیاره که این بسترها (در تست‌های روتین) دقتشون کمتره، گسترده‌گیش پایین‌تره ولی کار مردم رو راه میندازه. یعنی مثلاً یکسری افراد خاص از این تست‌ها استفاده میکنن. اتفاقی که در دنیا داره میفته در علوم پزشکی و درمانی اینجوریه که داره تلاش میکنه در لایف استایل تغییر ایجاد کنه. اگر می‌خواهیم به این موارد برسیم، اینها همش ترکیب دیتا هست، ترکیب کامپیوتر ساینس هست، ترکیب یکسری الگوریتم هست که باید بین رشته‌ای کار کرد. به خاطر همین بین رشته‌ای کار کردن موضوع مهمی است و دنیا دیگه تک رشته‌ای نیست. سعی کنید محدود به رشته خودتون نشید. سرعت تغییر خیلی زیاده./

با اسکن کردن کد زیر می‌تونید ویدیوی کامل این مصاحبه رو در کانال یوتیوب ژيوار تماشا کنید...



هوش مصنوعی و نقش آن در علوم زیستی

به قلم سید علی حسینی - دانشجوی کارشناسی زیست فناوری دانشگاه شاهد

از آنجایی که حیات چند وجهی بوده و عوامل مختلفی از جمله DNA، RNA و پروتئین ها هستند که از سلولی به سلول دیگر می توانند متفاوت باشند، پیچیدگی خاصی در آن به وجود آمده است. تحلیل و بررسی این پیچیدگی ها از طریق روش های تحلیلی اکتشافی (heuristic) دیگر امکان پذیر نیست. این روش ها معمولاً بر اساس شهود، الگوریتم های ساده یا تجربیات پیشین کار می کنند که برای درک این حجم از تعامل و تغییر کافی نیستند. در دهه گذشته با رویکردهای کل نگر (holistic) مبتنی بر هوش مصنوعی محدود (narrow AI) راه جدیدی برای تحلیل و بررسی پیچیدگی حیات، پیش پای دانشمندان قرار گرفت و اکنون با پیشرفت گسترده ابزارهای عامل محور هوش مصنوعی، که به انفجار کامبریایی (Cambrian Explosion) عوامل هوش مصنوعی تشبیه شده است، دانشمندان علوم زیستی می توانند هوش مصنوعی و زیست

شناسی را گرد هم آورند. این حوزه نوظهور در زیست شناسی با عنوان زیست شناسی مولد (generative biology)

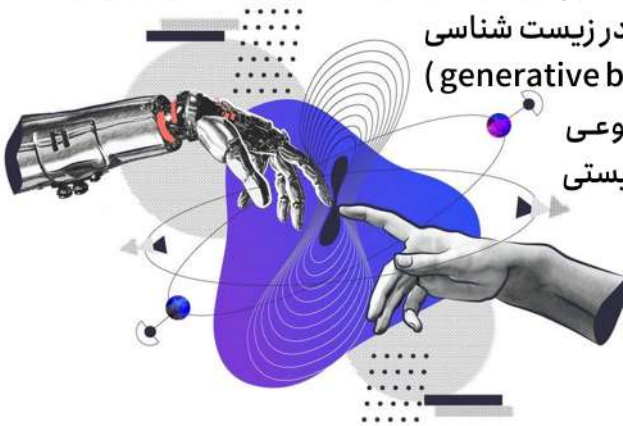
شناخته می شود. ابزارهای هوش مصنوعی می توانند در نقش های مختلفی به علوم زیستی

در فرآیند تحقیقاتی و بالینی کمک کنند و باعث افزایش سرعت فرآیندها شوند. این ابزارها می توانند در ایجاد بارش های فکری یا ارائه مشاوره های تخصصی تا حتی به عنوان کنترل کننده آزمایشگاه ها

در تحقیقات نقش ایفا کنند. این ابزارها بر اساس میزان خودمختاری و میزان اهمیتی کارهایی که انجام می دهند در سطوح مختلفی تقسیم بندی می شوند. البته ظهور ابزارهای هوش مصنوعی و استفاده از آنها در تحقیقات زیستی نگرانی هایی نیز به دنبال داشته است.

هوش مصنوعی چیست؟

در ابتدای بحث درباره نقش هوش مصنوعی در علوم زیستی به بررسی هوش مصنوعی و مفاهیم مربوط به آن میپردازیم. اکنون هوش مصنوعی را به عنوان ماشینی که توانایی انجام وظایف انسانی را بدون دخالت انسان ها دارد، می شناسند. تصمیم گیری، درک بصری، تشخیص گفتاری و یادگیری از داده ها برای پیشبینی بخشی از توانایی های هوش مصنوعی در نظر گرفته می شود. همانگونه که ما انسان ها در ابتدای تولد خود هیچ دانشی نداریم و در طول زندگی خود به دنبال کسب اطلاعات و افزایش آگاهی خود می رویم هوش مصنوعی نیز از ابتدای ایجاد نیاز به آموزش دارد تا بداند هدف آن چیست و باید به کدام سمت حرکت کند. یادگیری ماشین، زیر شاخه ای از هوش مصنوعی است که با استفاده از روش های آماری عملکرد رایانه ها را بهبود می بخشد. یادگیری ماشین را به طور ساده می توان به صورت توانایی یک ماشین در



بروز رفتار هوشمندانه مشابه انسان تعریف کرد. این فناوری ماشین ها را قادر می سازد تا مسائل را حل کرده، وظایف پیچیده ای را به سر انجام برسانند و بر اساس داده های تجربی الگوهایی را شناسایی کنند که کار بسیار سختی برای انسان است. یعنی هوش مصنوعی میتواند یاد بگیرد و متناسب با یادگیری خود عمل کند. نمونه ای از کاربرد این فناوری در زیست شناسی شناسایی جهش های موجود در توالی های DNA یا پروتئین با دقت بالا می باشد که کار محققان را بسیار ساده کرده است [۲].

- این حوزه شامل سه دسته اصلی است:
۱) یادگیری بدون نظارت، که قابلیت کشف الگوها، ساختارها یا روابط پنهان در داده ها را تقویت می کند.

۲) یادگیری نظارت شده، که در آن الگوریتم هایی که بر اساس نمونه های موجود، طبقه بندی و پیش بینی انجام می دهند توسعه پیدا می کنند.

۳) یادگیری تقویتی نیز که روشی برای دستکاری است، در این روش، فضای یک مسئله خاص را از طریق توالی پاداش ها و جریمه ها شکل می دهند.

این سه دسته یادگیری به ماشین ها این امکان را می دهند تا یاد بگیرند و تصمیم گیری های آگاهانه ای را انجام دهند [۷].

یادگیری عمیق نیز زیرمجموعه ای از یادگیری ماشین است که بر اساس اصول شبکه های عصبی مصنوعی عمل می کند. این شبکه ها از لایه های مختلفی شامل لایه های ورودی، خروجی و مخفی تشکیل شده اند که هر لایه شامل واحدهایی است که داده های ورودی را پردازش می کند و اطلاعات را به لایه های بعدی



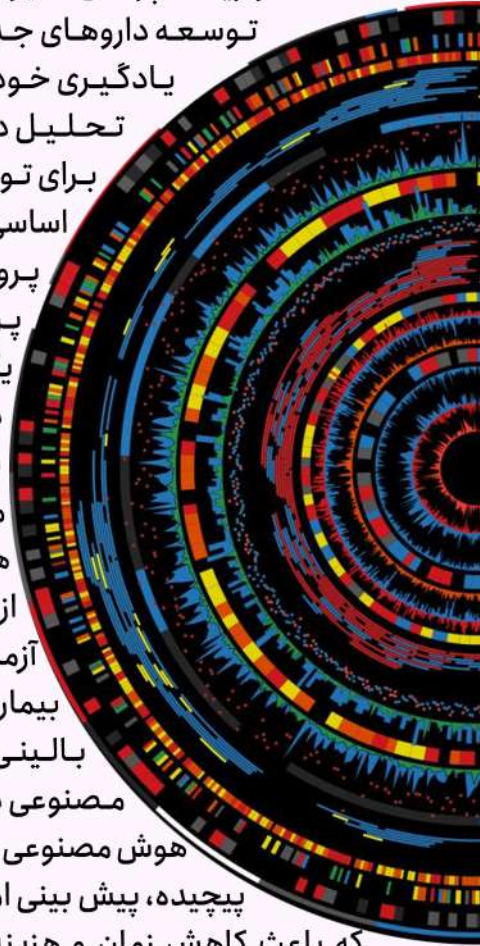
منتقل می‌کند. این فرآیند شبکه را قادر می‌سازد تا وظایف خاص پیش‌بینی را انجام دهد. از طریق این ساختار، ماشین‌ها می‌توانند خودکار از داده‌هایی که با آن‌ها مواجه می‌شوند، یاد بگیرند. با این حال، آموزش مدل‌های یادگیری عمیق منابع زیادی مصرف می‌کند و اغلب نیازمند حجم عظیمی از داده‌ها، سخت‌افزار محاسباتی قدرتمند و دوره‌های آموزشی طولانی است. در شرایطی که این امکانات در دسترس نباشد، یادگیری انتقال به عنوان یک استراتژی ارزشمند مطرح می‌شود. یادگیری انتقال شامل استفاده دوباره از مدل‌های از پیش آموزش دیده با تنظیم لایه‌های نهایی آن‌ها برای حل مسائل جدید است. این رویکرد با بهره‌گیری از دانش به دست آمده از وظایف قبلی، مدت زمان مورد نیاز برای آموزش، حجم داده‌ها و منابع محاسباتی مورد نیاز برای آموزش مدل‌های جدید را به طور قابل توجهی کاهش می‌دهد. این روش فرآیند توسعه را ساده‌تر کرده و امکان استقرار کارآمدتر راه‌حل‌های یادگیری عمیق در حوزه‌های مختلف را فراهم می‌کند [۷].

کاربرد ابزارهای هوش مصنوعی در علوم زیستی

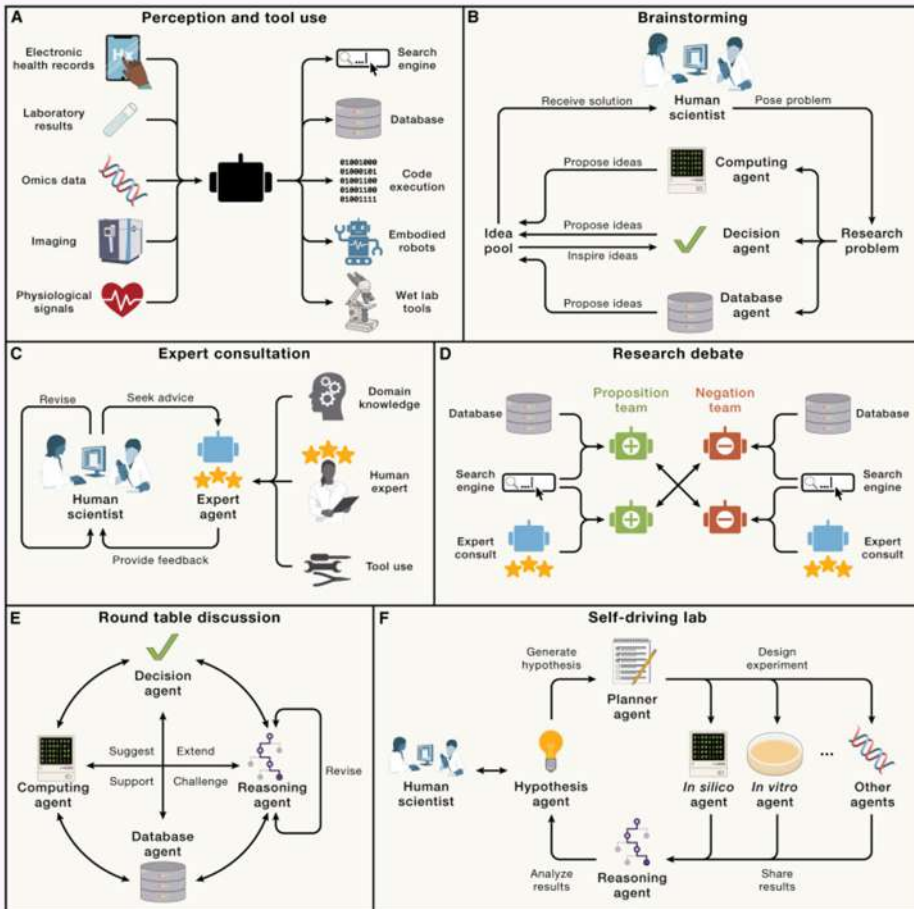
ابزارهای هوش مصنوعی توانسته‌اند با ورود به حوزه‌های مختلف علوم زیستی کمک‌های بسیاری به محققان و متخصصان بکنند و بسیاری از مشکلات را حل کنند و نیز باعث تسهیل فرآیندهای بسیار دشوار و طولانی شوند. در این بخش کاربرد این ابزارها را در دو حوزه علوم زیستی بررسی خواهیم کرد. **حوزه کشاورزی:** ابزارهای هوش مصنوعی با ورود به حوزه‌های مرتبط با گیاهان و کشاورزی توانسته‌اند با توسعه الگوریتم‌های شخصی سازی شده، بر بهبود محصولات کشاورزی اثر بگذارند. این الگوریتم‌ها به تحلیل وضعیت فیزیکی و شیمیایی محصولات کشاورزی مانند داده‌های مرتبط با خاک و شرایط آب و هوایی می‌پردازند. در این رویکرد میتوان با استفاده از ترکیب فنوتیپ برداری با توان بالا (HTP) و هوش مصنوعی به پیشرفت‌های گسترده دست پیدا کرد. فنوتیپ برداری با توان بالا به این معناست که با استفاده از فناوری‌هایی همچون پهپاد، سنسورهای تصویر برداری حرارتی و چند طیفی و ربات تعداد زیادی گیاه را در مزرعه و گلخانه بدون تخریب از نظر ویژگی‌های ظاهری مثل رشد، شکل برگ و مقاومت به بیماری‌ها مشاهده و اندازه‌گیری کنیم. ابزارهای هوش مصنوعی نیز با تجزیه و تحلیل داده‌های تصویری و محیطی، الگوهای پنهان بین ژن‌ها، صفات ظاهری و شرایط محیطی را کشف کرده، مدل‌هایی برای پیش‌بینی آینده رشد گیاه ساخته و به متخصصان در انتخاب سریع‌تر ژن‌های برتر کمک می‌کنند. از مزایای این رویکرد می‌توان به تشخیص بیماری گیاهان با تحلیل تصاویر، پیش‌بینی عملکرد و بهره‌وری محصولات، انتخاب صفات مطلوب ژنتیکی با دقت بالا، بهینه‌سازی آبیاری، کوددهی و کنترل آفات و همچنین برنامه‌ریزی دقیق کاشت و برداشت اشاره کرد. البته همچنان محدودیت‌هایی در این حوزه وجود دارد، از جمله

کمبود نیروی متخصص برای کار با ابزارهای هوش مصنوعی، هزینه های بالای بعضی از تجهیزات و تعیین استاندارد برای جمع آوری و تفسیر داده ها [۴]. از نمونه های جالب این رویکرد این است که تلاش شده است با استفاده از تصویربرداری با نور ساختار یافته (SLI)، کبودی های نامرئی سیب ها را شناسایی کنند. ابزار هوش مصنوعی ای که از طریق الگوریتم پیشرفته تفکیک (segmentation) آموزش دیده است با تشخیص کنتراست نواحی سالم و کبود، کبودی های نامرئی میوه را شناسایی می کند که باعث کاهش ضایعات و خسارات اقتصادی می شود [۵].

زیست پزشکی و سلامت: رشد چشمگیر هوش مصنوعی در حوزه سلامت و زیست پزشکی تاثیر قابل توجهی بر بخش های مختلف از جمله تشخیص دقیق، توسعه داروهای جدید، پزشکی دقیق و کارآزمایی بالینی گذاشته اند. برای مثال یادگیری خودکار و قدرتمند هوش مصنوعی، همراه با توان بالای پردازش و تحلیل داده های عظیم، قابلیت ویژه ای را برای بهره برداری از داده ها برای توسعه دارو فراهم کرده است. از آنجایی که پروتئین ها نقش اساسی در توسعه دارو ها دارند ابزارهای هوش مصنوعی با پیش بینی پروتئین ها، تحلیل عملکرد آن ها و طراحی نوآورانه و از صفر پروتئین ها باعث ایجاد جهش هایی در توسعه دارو ها شده اند. یکی از مراحل مهم در توسعه دارو ها نیز درک فعالیت بیولوژیکی دارو و مکانیسم های خاص آن است. هوش مصنوعی قادر است با تجزیه و تحلیل مجموعه بسیار بزرگی از داده ها کاندیداهای ممکن را برای دارویی با ویژگی خاص شناسایی کند. علاوه بر این هوش مصنوعی با استفاده از الگوریتم هایی برای تطبیق بیماران از طریق تحلیل داده های حجیم بالینی چرخه های طولانی آزمایش های بالینی را کاهش داده و فرآیند طاقت فرسای انتخاب بیماران را تسهیل می کند. همچنین باعث بهینه سازی آزمایش های بالینی و بهبود فرآیند های تصمیم گیری می شود. ابزارهای هوش مصنوعی در حوزه ویرایش ژنوم نیز تاثیر زیادی گذاشته اند. در این حوزه هوش مصنوعی با تحلیل حجم عظیمی از داده های ویرایش ژن، کشف الگوهای پیچیده، پیش بینی امکان پذیری و ایمنی نتایج ویرایش ژن نقش مهمی ایفا می کند که باعث کاهش زمان و هزینه های آزمایش ها و اعتبارسنجی ها می شود. در این زمینه الگوریتم های یادگیری عمیق در ارزیابی کارایی ویرایش باز (Base Editing) و ویرایش اولیه (Primary Editing) عملکرد بسیار موثری از خود نشان داده اند. در زمینه ویرایش بازها، ابزارهای شبیه سازی کامپیوتری متعددی برای کاهش اثرات خارج از هدف در ویرایش ژن طراحی شده اند. برای نمونه ابزار نوینی با نام "BEdeepoff" توسعه یافته است تا مکان های خارج از هدف ویرایشگرهای باز را پیش بینی کند. در این روش، محققان مجموعه از جفت های gRNA و نواحی خارج از هدف برای ویرایشگر آدنوزین (ABE) و ویرایشگر ستوزین (CBE) طراحی کردند و آن ها را به طور پایدار وارد سلول های انسانی نمودند.



پس از انجام آزمایش، داده های مربوط به کارایی این ویرایشگرها جمع آوری شده و برای آزمایش مدل های یادگیری عمیق مورد استفاده قرار گرفتند که به توسعه مدل های "ABE-deepoff" و "CBEdeepoff" منجر شد. این مدل ها این قابلیت را دارند که مکان خارج از هدف را با دقت بالا پیش بینی کنند و از احتمال ایجاد تغییرات ژنتیکی کم کنند. همچنین مشابه کاربرد هوش مصنوعی در سیستم CRISPR هدفمند DNA-cas9، از هوش مصنوعی در سیستم CRISPR هدفمند RNA-cas13 استفاده شده است. به عنوان مثال، محققان با ترکیب هوش مصنوعی و غربالگری CRISPR هدفمند RNA، مدلی بر پایه یادگیری عمیق با نام TIGER توسعه دادند تا از طریق طراحی gRNA بیان ژن را مهار کند. آن ها فعالیت 200 هزار gRNA را که ژن های ضروری را در رده های مختلف سلول انسانی هدف قرار داده بودند، آزمایش کردند تا مجموعه عظیمی از داده برای cas13 تولید کنند. سپس با استفاده از این داده ها مدل TIGER را آموزش دادند تا ارزیابی جامعی از فعالیت هدف گیری و فعالیت خارج از هدف gRNA های cas13 ارائه کنند. نویسندگان با بهره گیری از TIGER توسعه درمان های هدفمند RNA نسل بعدی را پیش بردند [7].



سطوح خودمختاری ابزارهای هوش مصنوعی

ابزارهای هوش مصنوعی را می‌توان بر اساس قابلیت آنها در تولید فرضیه‌ها، طراحی و انجام آزمایش‌ها و استدلال در چهار سطح مختلف از خودمختاری تعریف کرد. سطح بندی یک ابزار هوش مصنوعی در این سیستم به این صورت است که پایین‌ترین سطحی را که یک ابزار هوش مصنوعی در هر یک از این تخصص‌ها کسب می‌کند به عنوان سطح آن هوش مصنوعی به صورت کلی در نظر گرفته می‌شود. به عنوان مثال اگر یک ابزار هوش مصنوعی از نظر توانایی استدلال در سطح ۳ قرار گرفته باشد ولی از نظر طراحی و انجام آزمایش‌ها در سطح ۲ باشد، این هوش مصنوعی را به عنوان ابزار سطح ۲ در نظر می‌گیرند.

سطح ۰: بدون ابزارهای هوش مصنوعی

در این سطح از ابزارهای هوش مصنوعی استفاده می‌شود که به طور مستقل توانایی تولید فرضیه‌های قابل آزمایش نیستند و صرفاً به دانشمندان در طرح فرضیه‌ها کمک می‌کنند تا آنها بتوانند فرضیه‌های دقیقی تولید کنند.

سطح ۱: ابزارهای هوش مصنوعی در نقش دستیار

در این سطح دانشمندان فرضیه‌ها را طراحی می‌کنند و بر اساس فرضیه‌ها وظایفی را برای ابزارهای هوش مصنوعی تعریف می‌کنند و کارهای خاصی را به عهده آن‌ها قرار می‌دهند. در این سطح ابزارهای هوش مصنوعی می‌توانند فرضیه‌ها را فرموله کنند و از مجموعه محدودی از ابزارها برای انجام وظایف خود استفاده می‌کنند.

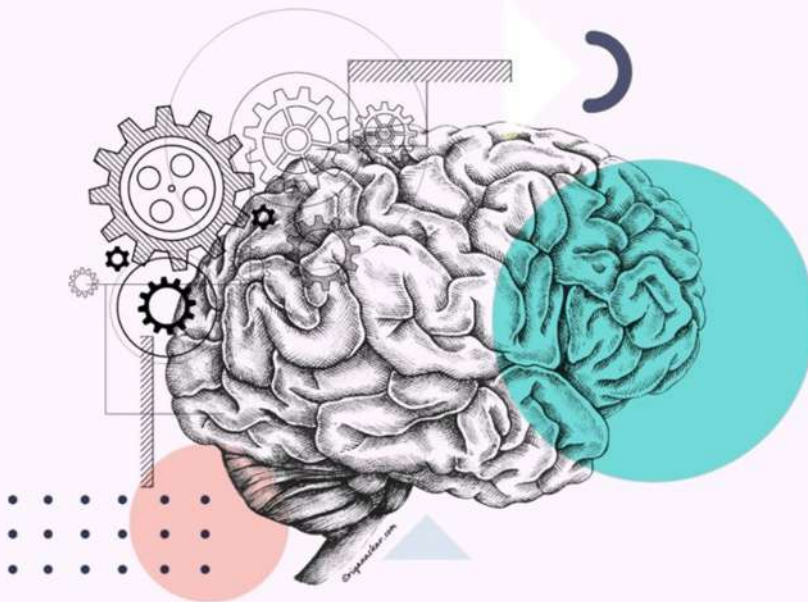
سطح ۲: ابزارهای هوش مصنوعی در نقش همکار

در سطح ۲ نقش ابزارهای هوش مصنوعی گسترده‌تر می‌شود. در این سطح ابزارهای هوش مصنوعی در نقش همکار برای دانشمندان ظاهر می‌شوند و به طور مشترک فرضیه‌ها را اصلاح می‌کنند. در این سطح ابزارهای هوش مصنوعی وظایف حیاتی را برای آزمایش فرضیه‌ها به دوش می‌کشند و در کنار آن مجموعه گسترده‌تری از ابزارها را در اختیار دارند ولی همچنان توانایی درک و فهم پدیده‌های پیچیده علمی را ندارند و در تولید فرضیه‌های علمی ضعف دارند.

سطح ۳: ابزارهای هوش مصنوعی در نقش دانشمند

ابزارهای هوش مصنوعی در این سطح قادر به درک مفاهیم پیچیده علمی هستند و فرضیه‌هایی را فراتر از دامنه تحقیقاتی قبلی تولید می‌کنند. ابزارها در این سطح مفاهیم را فراتر از خلاصه سازی‌های ساده تولید می‌کنند و پیوندهای مفهومی واضح بین یافته‌ها ایجاد می‌کنند [۶].

همکاری انسان و هوش مصنوعی	استدلال	طراحی آزمایش	تولید فرضیه	سطوح خودمختاری
دانشمند فرضیه را تعریف میکند و گاهی از خروجی مدل های هوش مصنوعی برای کمک به تولید فرضیه استفاده میکند؛ دانشمند وظیفه ای برای آزمودن فرضیه تعریف میکند و خودش وظایف را به سرانجام می رساند	ندارند	مدل های یادگیری ماشین وظایف از پیش تعیین شده ای را اجرا میکنند و هیچ تغییری در پروتکل ها بر اساس داده های جدید ایجاد نمی کنند.	ندارند	سطح ۰
دانشمند فرضیه را تعریف میکند و مجموعه ای از وظایف را برای آزمودن فرضیه تعریف میکند؛ عامل هوش مصنوعی وظایف را انجام میدهد	استدلال قوی در یک وظیفه انتخابی؛ خلاصه سازی چندوجهی یافته ها؛ استفاده از داده های آزمایشی و دانش موجود	طراحی محدود پروتکل های آزمایشی و بهره گیری از ابزارهای محاسباتی (silico in) و تجربی	عامل هوش مصنوعی فرضیه های ساده و محدود را تدوین میکند که ترکیبی مستقیم از دانش موجود، داده های اولیه یا مشاهدات هستند.	سطح ۱
دانشمند فرضیه اولیه را پیشنهاد میدهد و فرضیه را همراه با عامل هوش مصنوعی اصلح میکند؛ عامل هوش مصنوعی مجموعه ای از وظایف را برای آزمودن فرضیه تعریف میکند و عامل هوش مصنوعی وظایف را تکمیل میکند.	2: تفسیر یافته ها در چارچوب دانش موجود، در نظر گرفتن توضیحات جایگزین، و ارزیابی میزان اطمینان و اعتبار یافته ها؛ ترکیب مفاهیم فراتر از خلاصه سازی یافته ها؛ همکاری با سایر پژوهشگران تا متخصصان آن حوزه اعتبار یافته ها را بررسی کنند و از اینکه نتایج معتبر و قابل اعتماد هستند اطمینان حاصل کنند.	طراحی پروتکل های آزمایشی دقیق و بهره گیری ماهرانه از طیف گسترده ای از ابزارهای خارج از محیط محاسباتی (silico ex)؛ پس از جمع آوری داده ها، به کارگیری روشهای آماری و محاسباتی برای تحلیل نتایج و تفسیر داده ها برای مشخص کردن اینکه آیا از فرضیه پشتیبانی میکند یا آن را رد میکند	عامل هوش مصنوعی فرضیه هایی تولید میکند که در امتداد داده ها و منابع علمی شناخته شده هستند.	سطح ۲
دانشمند و عامل هوش مصنوعی بهطور مشترک فرضیه را تولید می کنند؛ عامل هوش مصنوعی مجموعه ای از وظایف را برای آزمودن فرضیه تعریف میکند و خود عامل هوش مصنوعی وظایف را تکمیل میکند	با توجه به نتایج و تفسیرها، روش های آزمایشی را به منظور یادگیری مداوم و انطباق اصلاح می شود تا دقت و عمق درک بهبود یابد؛ روابط مفهومی واضح بین یافته ها ایجاد می شود	توسعه روش های آزمایشی برای یافتن قابلیت های جدید؛ گردآوری فعال داده ها از طریق آزمایش ها یا شبیه سازی ها با استفاده از تکنیک ها و ابزارهای گوناگون به منظور اندازه گیری و ثبت پدیده های زیستی	عامل هوش مصنوعی فرضیه های خلاقانه و جدیدی تولید میکند که نتیجه تجزیه و تحلیل دانش موجود هستند.	سطح ۳



سطوح مختلف ابزارهای های هوش مصنوعی در مطالعات زیست شناسی سلولی

سلول‌ها واحدهای اساسی در زیست شناسی سلولی هستند و پیشرفت‌های اخیر در زمینه‌های مختلف مانند omics تک‌سلولی، میکروسکوپ نوری با وضوح بالا و ویرایش ژنی داده‌های زیادی تولید کرده‌اند که باعث افزایش اهمیت مدل سازی سلولی را در بستر کامپیوتری شده است. در این مطالعات از ابزارهای هوش مصنوعی در سطوح مختلفی استفاده می شود. ابزارهای های هوش مصنوعی در تجزیه و تحلیل داده‌ها نقش دارند، اما مدل‌های سطح صفر فاقد قابلیت‌های تحقیقاتی خودمختار هستند. در سطح ۱، ابزارهای هوش مصنوعی، مدل‌های سطح صفر را یکپارچه کرده و به دانشمندان در آزمودن فرضیه‌ها کمک می‌کنند. به عنوان مثال ابزارهای این سطح برای بررسی مقاومت یک ترکیب، پیش‌بینی‌هایی درباره پاسخ‌های سلولی انجام داده و سپس بر اساس آن‌ها پروتکل‌های آزمایشی را طراحی می‌کنند. ابزارهای هوش مصنوعی سطح ۲، وظایف از پیش تعیین شده را انجام داده و فرضیه‌هایی درباره عملکرد سلولی تولید می‌کنند. آنها به طور خودکار وظایف را تنظیم کرده و آزمایش‌ها را برای بررسی پدیده‌هایی مانند مقاومت دارویی بهینه سازی می‌کنند. این عامل‌ها با ترکیب داده‌های پیش‌بینی شده و جدید، به طور مداوم ابزارهای خود را به روزرسانی کرده و مدل‌های سلولی مجازی ایجاد می‌کنند که قادر به تولید فرضیه‌های جدید هستند. ابزارهای سطح ۳ به پیش‌بینی چالش‌ها و جهت‌گیری‌های آینده در تحقیقات زیست شناسی سلولی می‌پردازند. آنها مدل‌های هیبریدی ایجاد کرده و از ترکیب ابزارهای هوش مصنوعی و پلتفرم‌های آزمایشی استفاده می‌کنند. این ابزارهای هوش مصنوعی نه تنها پروتکل‌ها را بهینه سازی می‌کنند، بلکه فناوری‌های نوآورانه‌ای را برای شبیه سازی دقیق‌تر در مدل‌های مختلف توسعه می‌دهند. توانایی این ابزارهای هوش مصنوعی در کشف مکانیسم‌های زیستی و استراتژی‌های درمانی، نقش دانشمندان را از انجام وظایف عملیاتی به تمرکز بر ایده‌پردازی و مدیریت مدل‌ها تغییر می‌دهد [۶].

چالش‌ها و نگرانی

همانطور که ورود و توسعه هوش مصنوعی به عرصه‌های تحقیقاتی و بالینی باعث افزایش بهره‌وری در وظایف و خودکارسازی فرآیندهای تکراری می‌شود، امکان تجزیه و تحلیل حجم عظیمی از داده‌ها را فراهم کرده و انجام پژوهش‌هایی که قبلاً تقریباً غیرممکن بود را امکان‌پذیر می‌کند اما همچنان نگرانی‌ها و چالش‌هایی را به همراه خود دارد که بی‌توجهی به آنها مشکلاتی را در پی خواهد داشت. این چالش‌ها و نگرانی‌ها موانعی را ایجاد کرده است. در مطالعه‌ای که توسط پنل علم و فناوری پارلمان اروپا درباره خطرات استفاده از هوش مصنوعی در علوم زیستی انجام داده است، برای نمونه به هفت مورد از این خطرات اشاره شده است؛

- وجود داده‌های مصنوعی در ورودی و اندازه‌گیری‌های بالینی، تفاوت بین داده‌های آموزشی و داده‌های واقعی و تغییرات غیرمنتظره در محیط‌های بالینی می‌تواند باعث ایجاد خطای ابزارهای هوش مصنوعی در محیط‌های پزشکی و آسیب به بیماران شود. این خطاها می‌توانند در تشخیص نادرست، درمان ناکافی و اولویت‌بندی نادرست خود را نشان دهند.

- هر چند که ابزارهای هوش مصنوعی بسیار دقیق باشند، اما تاثیر آن ابزارها وابسته به استفاده انسان‌ها از نتایج آن‌هاست. استفاده نادرست متخصصان از آن‌ها می‌تواند مشکلاتی را به وجود آورد. استفاده نادرست از این ابزارها توسط متخصصان می‌تواند ناشی از مشارکت محدود متخصصان و محققان علوم زیستی در توسعه هوش مصنوعی، کمبود آموزش در زمینه هوش مصنوعی و... باشد.

- سوگیری‌های سیستماتیک همانگونه که بین انسان‌ها وجود دارد می‌تواند به مدل‌های هوش مصنوعی نیز راه پیدا کند. این سوگیری‌ها می‌تواند مبتنی بر جنسیت، نژاد و قومیت، سن، وضعیت اجتماعی-اقتصادی و... باشد.

- یکی از خطرات قابل‌تأمل استفاده از هوش مصنوعی عدم شفافیت در مورد طراحی، توسعه، ارزیابی و استقرار ابزارهای آن است. شفافیت هوش مصنوعی با دو مفهوم ردیابی پذیری و توضیح‌پذیری مرتبط است که باعث تقسیم بحث به دو بخش می‌شود؛ شفافیت در فرآیندهای توسعه و استفاده از هوش مصنوعی با ردیابی‌پذیری مرتبط است و شفافیت در تصمیم‌های گرفته شده توسط هوش مصنوعی با توضیح‌پذیری ارتباط دارد. از خطرات عدم شفافیت مرتبط با هوش مصنوعی می‌توان به عدم درک و اعتماد به پیش‌بینی‌ها و تصمیم‌گیری‌های تولید شده توسط ابزارهای هوش مصنوعی، دشواری در بازتولید و ارزیابی مستقل الگوریتم‌های هوش مصنوعی، دشواری در شناسایی منشأ خطاهای هوش مصنوعی و تعیین مسئولیت آن‌ها و... اشاره کرد.

- یکی از خطراتی که در بخش زیست‌پزشکی و تحقیقات انسانی با توسعه راهکارها و فناوری‌های هوش مصنوعی بروز پیدا می‌کند و اخیراً با اتکا به داده‌های کلان در طول پاندمی کووید-19 برجسته‌تر شده است، فقدان حریم خصوصی، رازداری و حفاظت از داده‌های بیماران است. این خطرات شامل اشتراک‌گذاری داده‌های شهروندان بدون رضایت آگاهانه و کامل آنها، استفاده مجدد از داده‌های بیماران بدون اطلاع آن‌ها، نقض داده‌ها که می‌تواند باعث افشای اطلاعات حساس و شخصی شهروندان شود و خطر حملات سایبری می‌باشد.

- از دیگر مشکلات استفاده از هوش مصنوعی در علوم زیستی و پزشکی نامشخص بودن فرد پاسخگو است. خلأهای قانونی بسیاری در مورد اینکه چه کسی باید برای خطاها و شکست های این سیستم پاسخگو یا مسئول باشد نیز وجود دارد. این مشکل ناشی از تعدد دست اندرکاران در فرآیند هوش مصنوعی های پزشکی و زیستی، از طراحی تا استقرار است که باعث آسیب پذیری پزشکان، محققان و متخصصان می شود [۳].

با وجود تمام پیشرفت هایی که در حوزه عامل های هوش مصنوعی وجود داشته اما چالش ها و نگرانی های مهمی نیز در این حوزه وجود دارد. همه این ها در کنار هم به ما امید می دهد تا شاید بتوان به مرحله ای برسیم که عامل های هوش مصنوعی تبدیل به دانشمندی شوند که تمام مراحل یک فرآیند تحقیقاتی از تولید فرضیه تا انجام آزمایش ها و همچنین تجزیه و تحلیل داده ها و به دست آوردن نتایج قابل اطمینان و راه حل هایی برای مشکلات تحقیقاتی را انجام دهند ولی همچنان راه زیادی تا آن مرحله مانده که نباید شتاب زده عمل کرد تا باعث مشکلات غیر قابل جبران شد.

منابع:

1. Lazebnik, Y., Can a biologist fix a radio?--Or, what I learned while studying apoptosis. *Cancer Cell*, 2002. 2(3): p. 179-82.
2. Tavassoly, I., J. Goldfarb, and R. Iyengar, Systems biology primer: the basic methods and approaches. *Essays Biochem*, 2018. 62(4): p. 487-500.
3. Kohl, M., S. Wiese, and B. Warscheid, Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol*, 2011. 696: p. 291-303.
4. Kirschner, M.W., The Meaning of Systems Biology. *Cell*, 2005. 121(4): p. 503-504.
5. Villa, A. and S.T. Sonis, 2 - System biology, in *Translational Systems Medicine and Oral Disease*, S.T. Sonis and A. Villa, Editors. 2020, Academic Press. p. 9-16.
6. Koutrouli, M., et al., A Guide to Conquer the Biological Network Era Using Graph Theory. *Front Bioeng Biotechnol*, 2020. 8: p. 34.
7. Mitra, K., et al., Integrative approaches for finding modular structure in biological networks. *Nat Rev Genet*, 2013. 14(10): p. 719-32.
8. Li, Z., Q. Liu, and L. Chen, A Novel Edge Rewire Mechanism Based on Multiobjective Optimization for Network Robustness Enhancement. *Frontiers in Physics*, 2021. 9.
9. Mondragón, R.J., Estimating degree-degree correlation and network cores from the connectivity of high-degree nodes in complex networks. *Scientific Reports*, 2020. 10(1): p. 5668.
10. Herakovič, N., et al., Distributed Manufacturing Systems with Digital Agents. 2019, 2019. 65(11-12): p. 8.
11. Kannan, M., et al., Leveraging public AI tools to explore systems biology resources in mathematical modeling. *NPJ Syst Biol Appl*, 2025. 11(1): p. 15.
12. Jamialahmadi, H., et al., Artificial intelligence and bioinformatics: a journey from traditional techniques to smart approaches. *Gastroenterol Hepatol Bed Bench*, 2024. 17(3): p. 241-252.

هوش مصنوعی، مرگ یا زندگی؟!

به قلم نوید بی یار عارفی - دانشجوی کارشناسی حقوق دانشگاه مراغه

مقدمه

از نخستین دفعه‌ای که واژه هوش مصنوعی در مکانی مطرح شد تقریباً ۷۰ سال می‌گذرد. حالا اما این واژه نوظهور به منظومه‌ای شگفت‌انگیز بدل شده که در تار و پود زندگی انسانِ امروزی رسوخ کرده و هرروز پرنرنگ تر از دیروز می‌شود. تا آنجا که می‌توان گفت حتی مفاهیم حیات و ممات نیز در آینده‌ای نه چندان دور، متأثر از این پدیده می‌گردند. تعریف ارائه شده از مرگ به صورت مختصر و مفید چنین است: توقف چرخه‌ی زیستی جاندار زنده [۱]. حالاً یک چوب‌دستی از آستین بشر بیرون آمده که گویی توان زیر میز زدن را دارد، آن هم پدیده هوش مصنوعی است. امروزه با ظهور هوشمان‌های محاوره‌ای (ChatBots) و آواتارهای دیجیتال امکان ساخت شخصیت‌هایی فراهم شده که می‌توانند رفتارها و گفتارها و احساساتی مشابه آنچه که متوفی داشته است را داشته باشند. این وضعیت ما را ملزم به آن می‌کند تا درباره‌ی هوش مصنوعی و آثارش بیندیشیم. آیا اساساً می‌توان گفت ماهیت مرگ تغییر می‌کند؟ آیا می‌توان گفت چنین عملی به طور اخلاقی رواست؟ مسئولیت اخلاقی این شخصیت‌ها با چه کسی خواهد بود؟ مسئولیت مدنی این روند به چه صورت است؟ در این نوشتار نهایت کوشش را به عمل آورده‌ام تا بلکه بتوان این مهم پیش رو را در ابعاد مختلف به صورت مختصر و مفید و با زبانی شیوا بررسی و تحلیل کرد. امید است که مورد توجه علاقه‌مندان قرار گیرد.

تغییر مفهوم مرگ زیر سایه‌ی هوش مصنوعی

در دنیای امروز به ویژه با پیشرفت‌های قابل توجهی که صورت گرفته است، امکان بازسازی و احیای افرادی که در قید حیات نیستند به صورت شخصیت‌های دیجیتال وجود دارد. از مهم‌ترین نمونه‌های واقعی و معروف ماجرای |چت‌بات جسیکا پریرا| است. (2) نامزد جسیکا با دادن اطلاعات او به یک چت‌بات موفق شد یک نسخه‌ی دیجیتالی از وی بسازد که به طرز عجیبی شبیه به جسیکا بود. نمونه‌های دیگری از این قبیل نیز وجود دارد که ذکر آن در این نوشتار نمی‌گنجد. حال مسئله‌ای که مطرح می‌شود این است که آیا می‌توان گفت تعریف مرگ دستخوش تغییر شده است؟ پاسخ به این پرسش به سختی امکان‌پذیر است؛ چرا که از منظر بیولوژیکی نمی‌توان چنین موجودی را زنده دانست. این موجود صرفاً یک ماشین دیجیتال است که بر اساس احتمالات و تقلید پیش می‌رود. اما از آن طرف، دارای حضوری معنادار و قابل تعامل است که توانایی ایجاد احساس را در طرف مقابل را دارد. شاید سخت‌ترین قسمت ارتباط انسان و هوش مصنوعی همین جا باشد؛ اینکه این موجود به لحاظ فیزیکی وجود



ندارد اما حضور روانی و معناداری دارد، ایجاد احساس می‌کند و در ذهن فرد نقش می‌بندد، تاثیر می‌گذارد و تاثیر می‌پذیرد. همه‌ی این موضوعات باعث می‌شوند که پاسخ قاطع به این پرسش، به دشواری ممکن شود.

چالش‌های اخلاقی مرتبط با شخصیت‌های دیجیتال
مهم‌ترین پرسش اخلاقی در خصوص شخصیت دیجیتال، امکان وجود مسئولیت اخلاقی است. نظر بسیاری از فلاسفه و متکلمان بر این است که هنگامی می‌توان مسئولیت اخلاقی را متصور دانست، که عمل فاعل مبتنی بر اراده‌ی آزاد، آگاهی و اختیار باشد [۳]. پس، از این حیث فعلاً نمی‌توان شخصیت دیجیتال را دارای مسئولیت دانست، لذا می‌توان مسئولیت چنین شخصیتی را برعهده‌ی خالقان آن یا در صورت لزوم برعهده‌ی مالکان آن گذاشت و در صورت وقوع خطا، ایشان موظف به رسیدگی هستند.

چالش‌های حقوقی مربوط به شخصیت دیجیتال
اولین نکته‌ی ای که باید لحاظ کرد این است که شخصیت دیجیتال متوفی مبتنی بر رضایت فرد شکل بگیرد. چرا که حق هر فرد بر اطلاعات شخصی خودش تعریف شده است [۴] و در این زمینه قوانین حفاظتی زیادی به تصویب رسیده‌اند. نظیر مقررات عمومی حفاظت از داده اتحادیه اروپا که به منظور حفظ هرچه بهتر حقوق افراد وضع شده است. پس استفاده از این اطلاعات بدون رضایت مالک اخلاقاً و بعضاً قانوناً درست نیست. دومین نکته این است که مسئولیت حقوقی این شخصیت با چه کسی است؟ پیش‌تر در مبحث چالش‌های اخلاقی گفتیم که چرا فرض مسئولیت داشتن این شخصیت منتفی است. اما به بهانه فقدان جاری شدن مسئولیت

نمی‌توان روی این مسئله سرپوش گذاشت چرا که سکوت قانون در این خصوص زمینه‌ساز سوءاستفاده سودجویان می‌گردد. با توجه به ماهیت شخصیت‌های دیجیتال که برپایه‌ی داده‌های ورودی خانواده متوفی شکل می‌گیرند می‌توان رفتار این شخصیت‌ها را بازتابی از داده‌های ایشان دانست، پس در صورت بروز خطا یا ارتکاب رفتار نادرست امکان رجوع به کسی که داده‌ها را ارائه داده، فراهم است ولی این امر موجب عدم امکان رجوع به شرکت فروشنده و سازندگان نرم‌افزاری این شخصیت‌ها نمی‌شود. آن هم در شرایطی که مشخص گردد وقوع رفتار نه ناشی از داده‌های ارائه شده که نتیجه‌ی وجود نقص در طراحی محصول بوده است. این راه به عدالت و انصاف نیز نزدیک‌تر است، چرا که مصرف‌کننده در واقع همان سازنده‌ی اخلاقی شخصیت دیجیتال است و باید در مرحله‌ی اول در قبال آن پاسخگو باشد. پس چنانچه وقوع رفتار، ناشی از تقصیر مصرف‌کننده نباشد می‌توان به سازنده و طراح رجوع کرد. در نهایت باید گفت توسل به تنها یک نظر در خصوص مسئولیت مدنی شخصیت دیجیتال به سختی ممکن است و آنچه که هویداست، ضرورت تدوین قوانین و مقررات دقیق و کاربردی، ضمن هم‌فکری حقوقدانان است.

نتیجه گیری

در پایان باید گفت با وجود پیشرفت‌های موجود در حوزه فناوری و هوش مصنوعی، داشتن نگاهی تازه به مفاهیم کلاسیک مرگ و زندگی و مسئولیت مدنی لازم بوده و از طرفی نیز شاید بتوان گفت زنگ خطری برای انسان است؛ چرا که چه بسا باعث عدم کنترل سوگ و انقطاع، و دوری فرد از واقعیت زندگی بشود. در هر صورت بهتر است قبل از ایجاد معضلات اساسی در حوزه‌های عمومی و خصوصی راه حل‌ها و تدابیری اندیشید. گریزگاهی از پیشرفت فناوری و دانش نیست. آنانی نیز که به دنبال فرار از سازگاری با روزآمدی جهان و پافشاری بر موقعیت سنتی هستند زیر چرخ‌های آینده له خواهند شد؛ زیرا گونه ای که سازگارتر است بقا می‌یابد.

منابع:



برای مطالعه سایر مطالب این نسخه، کدهای مربوطه را اسکن فرمایید...



دکینگ مولکولی هنر و علم اتصال

به قلم هانیه محمودی هاشمی - دانشجوی دکترای مهندسی بافت دانشگاه شهید بهشتی



اپی ژنتیک پلی میان ژن ها و محیط

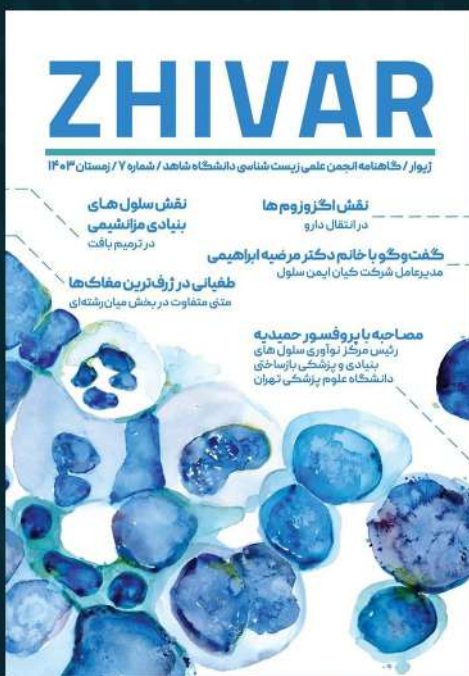
به قلم حانیه شریفی - دانشجوی کارشناسی ارشد ژنتیک دانشگاه شهید اشرفی اصفهانی



میکروبیوم و روش های توالی یابی

به قلم سبحان جره ای - دانشجوی کارشناسی زیست فناوری دانشگاه شاهد

برای مطالعه نسخه پیشین نشریه ژيوار كد زير را اسكن كنيد





دانشگاه شاهرود
دانشکده علوم پایه
موسسه تحقیقات بیولوژی



پل های ارتباطی :

✉ Zhivarpr@gmail.com |  [Shahedbiology](https://www.instagram.com/Shahedbiology)