مجله ایمنی زیستی

دوره ۱۸، شماره ۱، بهار ۱۴۰۴

شاپای چاپی: ۰۶۳۲ – ۲۷۱۷، شاپای الکترونیکی: ۹۸۰۴ - ۲۷۱۶

به‌کارگیری الگوریتم‌های یادگیری ماشین در شناسایی ژن‌ها و مسیرهای مولکولی مرتبط با مقاومت تنش‌ها در گیاهان تراریخته

نوع مقاله: مروری 20.1001.1.27170632.1404.18.1.2.7

نرجس سادات موسوی مدنی[۱،۲]، اکرم صادقی[۱*]، رضا شرفی[۱]

۱. بخش بیوتکنولوژی میکروبی، پژوهشگاه بیوتکنولوژی کشاورزی ایران (ABRII)، سازمان تحقیقات، آموزش و ترویج کشاورزی (AREEO)، کرج، ایران

۲. گروه زیست‌شناسی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران

aksadeghi@abrii.ac.ir

تاریخ دریافت: ۱۴۰۴/۰۱۰/۲۱، تاریخ پذیرش: ۱۴۰۵/۱/۲۰

صفحه ۳۳-۱۵

چکیده

گیاهان در طول چرخهٔ حیات خود با طیف گسترده‌ای از تنش‌های غیرزیستی شامل شوری، خشکی، کم‌آبی، دمایی، اکسیداتیو، فلزات سنگین و تنش های زیستی شامل انواع آفات و بیماری‌ها مواجه می‌شوند. این تنش‌ها یکی از مهم‌ترین عوامل ایجاد شکاف میان پتانسیل ژنتیکی و عملکرد واقعی محصولات زراعی محسوب می‌شوند. در این مقاله، ابتدا نقش تنش‌ها در محدودسازی عملکرد گیاهان و راهبردهای زیست‌فناورانهٔ مقابله با آن‌ها بررسی می‌شود. سپس به‌کارگیری الگوریتم‌های یادگیری ماشین (ML) در شناسایی و اولویت‌بندی ژن‌ها و مسیرهای مولکولی مرتبط با مقاومت به تنش ها، در کنار ابزارهای کلاسیک و نوین زیست‌فناوری گیاهی مورد بحث قرار می‌گیرد. هوش مصنوعی و ML با تحلیل داده‌های چندلایه‌ی ژنومی، ترنسکریپتومی و اُمیکس، امکان مدل‌سازی شبکه‌های سیگنال‌دهی و متابولیکی، پیش‌بینی عملکرد گیاه تحت شرایط تنش و تفسیر برهم‌کنش‌های ژنتیکی و اپی‌ژنتیکی را فراهم می‌کنند و بدین‌ترتیب به افزایش کارایی پژوهش‌ها و کاهش هزینه و زمان منجر می‌شوند. در ادامه، روش‌هایی نظیر درخت تصادفی، ماشین بردار پشتیبان و مدل‌های یادگیری عمیق برای شناسایی عوامل کلیدی مقاومت و پشتیبانی از تصمیم‌گیری‌های اصلاحی ارائه شده است. در انتها، مطالعات موردی انجام‌شده بر روی برنج، ذرت، آرابیدوپسیس و چالش‌ها و محدودیت‌های موجود و ملاحظات ایمنی‌زیستی، قانونی و اقتصادی مورد بحث قرار می‌گیرند.

واژه‌های کلیدی: اُمیکس، بیوتکنولوژی، تراریخته، تنش‌های گیاهان، یادگیری‌ماشین

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

مقدمه

گیاهان زراعی در طول چرخه‌ی رشد و نمو خود به‌طور مداوم در معرض طیف گسترده‌ای از تنش‌های زیستی و غیرزیستی از جمله خشکی، شوری، کم‌آبی، تنش‌های دمایی، تنش اکسیداتیو، فلزات سنگین، آفات و عوامل بیماری‌زا قرار دارند. این تنش‌ها به‌صورت مستقیم و غیرمستقیم رشد، فیزیولوژی و عملکرد نهایی گیاهان را محدود می‌کنند و یکی از مهم‌ترین عوامل ایجاد شکاف میان پتانسیل ژنتیکی و عملکرد واقعی محصولات زراعی محسوب می‌شوند (Raza *et al*., 2024). در دهه‌های اخیر، تشدید تغییرات اقلیمی، افزایش وقوع تنش‌های محیطی و ناپایداری شرایط کشت، این چالش را به یکی از تهدیدهای اصلی امنیت غذایی در سطح جهانی تبدیل کرده است (Zhou *et al*., 2022).

گزارش‌های سازمان خواربار و کشاورزی ملل متحد (FAO) نشان می‌دهد که تنش‌های محیطی می‌توانند تا حدود ۷۰ درصد از فاصله‌ی میان عملکرد بالقوه و عملکرد واقعی محصولات زراعی را در بسیاری از مناطق کشاورزی جهان ایجاد کنند. این موضوع اهمیت توسعه‌ی راهبردهای نوین و کارآمد برای افزایش تحمل تنش در گیاهان زراعی را بیش از پیش برجسته می‌سازد (FAO, 2015). علوم مختلف از جمله مهندسی ژنتیک، مهندسی متابولیک، میکروبیولوژی و زیست‌فناوری سازوکارهای متعددی برای افزایش تحمل به تنش در این گیاهان ارائه داده اند که یکی از این موارد استفاده از میکروارگانیسم ها برای افزایش تحمل تنش به‌صورت زیستی است.

میکروارگانیسم‌های مفید ریزوسفری، به‌ویژه باکتری‌ها و اکتینومیست‌ها، از طریق القای دفاع سیستمیک، بهبود جذب عناصر غذایی و تنظیم پاسخ‌های فیزیولوژیک، نقش مهمی در افزایش تحمل تنش و کیفیت محصولات زراعی ایفا می‌کنند (Abbasi *et al*., 2022). پاسخ گیاهان به تنش‌های زیستی و غیرزیستی فرآیندی پیچیده، چندلایه و وابسته به زمینه است که در آن شبکه‌ای از مسیرهای مولکولی، تنظیم‌های ژنتیکی و اپی‌ژنتیکی، تغییرات متابولیکی و پاسخ‌های فیزیولوژیک به‌صورت هم‌زمان درگیر می‌شوند. برخلاف دیدگاه‌های کلاسیک که پاسخ به تنش را عمدتاً به تنظیم تعداد محدودی ژن یا مسیر منفرد نسبت می‌دادند، شواهد فزاینده نشان می‌دهد که تحمل تنش نتیجه‌ی برهم‌کنش غیرخطی مجموعه‌ای از ژن‌ها، عناصر تنظیمی و مسیرهای سیگنال‌دهی است که شدت و الگوی آن‌ها به نوع تنش، مدت‌زمان مواجهه و مرحلهٔ رشدی گیاه بستگی دارد (Raza *et al*., 2021; Zhou *et al*., 2022). این موضوع اهمیت تحلیل هم‌زمان و سیستم‌محور شبکه‌های تنظیمی را برای کاهش ریسک‌های پیش‌بینی نشده در برنامه‌های اصلاح ژنتیکی برجسته می‌سازد.

این ماهیت شبکه‌ای و چندژنی پاسخ به تنش نشان می‌دهد که شناسایی اهداف مولکولی مناسب برای اصلاح ژنتیکی، به‌ویژه در گیاهان تراریخته، نیازمند رویکردهایی فراتر از تحلیل ژن‌های منفرد است. در چنین شرایطی، انتخاب نادرست ژن هدف می‌تواند نه‌تنها به افزایش تحمل تنش منجر نشود، بلکه پیامدهای ناخواسته‌ای بر رشد، عملکرد و ایمنی زیستی گیاه داشته باشد. این پیچیدگی زیستی سبب شده است که بسیاری از رویکردهای کلاسیک اصلاح نباتات و حتی برخی راهبردهای زیست‌فناوری، علی‌رغم موفقیت‌های آزمایشگاهی، در دستیابی به مقاومت پایدار و چندگانه در شرایط واقعی مزرعه با محدودیت مواجه شوند. به‌ویژه، شناسایی ژن‌ها و مسیرهای کلیدی مؤثر بر تحمل تنش، به دلیل حجم عظیم داده‌های زیستی، تنوع پاسخ‌ها و وجود برهم‌کنش‌های غیرخطی، به

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

فرآیندی زمان‌بر و پرهزینه تبدیل شده است (Van Dijk *et al*., 2021).

در دهه‌ی اخیر، پیشرفت فناوری‌های اُمیکس شامل ژنومیکس، ترنسکریپتومیکس، پروتئومیکس و متابولومیکس، امکان تولید داده‌های چندلایه و پرحجم را در مطالعات تنش‌های گیاهی فراهم کرده است (Jamalirad *et al*., 2024). این داده‌ها دیدی جامع از تغییرات مولکولی و تنظیمی گیاهان در پاسخ به تنش ارائه می‌دهند، اما تحلیل مؤثر آن‌ها و استخراج الگوهای زیستی معنادار، فراتر از توان روش‌های آماری سنتی است و نیازمند رویکردهای داده‌محور پیشرفته است (Libbrecht and Noble, 2015). در چنین بستری، پیچیدگی داده‌های اُمیکس و تعاملات غیرخطی میان اجزای زیستی، عملاً کارایی رویکردهای آماری کلاسیک را در شناسایی روابط علّی محدود می‌کند و زمینه را برای بهره‌گیری از روش‌های داده‌محور پیشرفته فراهم می‌سازد.

یادگیری ماشین (Machine Learning, ML) و هوش مصنوعی به‌عنوان ابزارهایی توانمند برای تحلیل داده‌های پیچیده، مدل‌سازی روابط غیرخطی و شناسایی الگوهای پنهان، به‌سرعت جایگاه ویژه‌ای در علوم گیاهی و اصلاح نباتات یافته‌اند (Washburn *et al*., 2020). الگوریتم‌های یادگیری ماشین با ادغام داده‌های ژنومی، ترنسکریپتومی و فنوتیپی، امکان شناسایی و اولویت‌بندی ژن‌های پاسخ‌دهنده به تنش، بازسازی شبکه‌های تنظیمی و متابولیکی و پیش‌بینی عملکرد گیاهان تحت شرایط مختلف تنش را فراهم می‌کنند (Ubbens and Stavness, 2017). کاربرد این رویکردها زمانی بیشترین اثربخشی را دارد که در کنار ابزارهای زیست‌فناوری گیاهی نظیر تولید گیاهان تراریخته، ویرایش ژنومی مبتنی بر CRISPR/Cas9 و فناوری RNAi (RNA interference) به‌کار گرفته شوند. چنین همگرایی میان داده‌های چندلایه، مدل‌های یادگیری ماشین و فناوری‌های اصلاح ژنتیکی، امکان گذار از اصلاح ژن‌محور به اصلاح شبکه‌محور را فراهم می‌سازد و می‌تواند فرآیند توسعه‌ی گیاهان مقاوم به تنش را هدفمندتر، سریع‌تر و کم‌هزینه‌تر کند (Van Dijk *et al*., 2021; Hesami *et al*., 2022).

بهره‌برداری مؤثر از یادگیری ماشین در زیست‌فناوری گیاهی علی‌رغم پیشرفت‌های بسیار، همچنان با چالش‌هایی نظیر کمبود داده‌های اُمیکس باکیفیت برای بسیاری از گونه‌های زراعی، خطر بیش‌برازش (Overfitting) مدل‌ها، محدودیت‌های زیرساختی و ملاحظات ایمنی‌زیستی و قانونی مرتبط با گیاهان تراریخته همراه است. بیش‌برازش در یادگیری ماشین زمانی رخ می‌دهد که مدل دیتاهای آموزشی خود را بیش از اندازه خوب یاد گرفته است که شامل نویز ها شده و منجر به کاهش عملکرد آن در مقابل داده‌های پیش‌تر دیده نشده می‌شود. ازاین‌رو، نیاز به چارچوب‌های یکپارچه‌ای که بتوانند داده‌های چندلایه‌ی زیستی، الگوریتم‌های یادگیری ماشین و ابزارهای اصلاح ژنتیکی را به‌صورت هم‌افزا ترکیب کنند، بیش از پیش احساس می‌شود (Van Dijk *et al*., 2021).

هدف از این مقاله‌ی مروری، بررسی نقش الگوریتم‌های یادگیری ماشین در شناسایی و اولویت‌بندی ژن‌ها و مسیرهای مولکولی مرتبط با مقاومت به تنش‌های زیستی و غیرزیستی و تحلیل چگونگی همگرایی این رویکرد با زیست‌فناوری گیاهی برای بهبود عملکرد، پایداری و ایمنی زیستی گیاهان زراعی است.

پاسخ گیاهان به تنش‌های زیستی و غیرزیستی

پاسخ گیاهان به تنش‌های زیستی و غیرزیستی حاصل برهم‌کنش پیچیدهٔ عوامل مولکولی، فیزیولوژیک و محیطی است که به‌صورت شبکه‌ای و غیرخطی عمل می‌کنند. برخلاف دیدگاه‌های

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

کلاسیک که پاسخ به تنش را به تنظیم چند ژن یا مسیر منفرد نسبت می‌دادند، مطالعات اخیر نشان داده‌اند که تحمل تنش نتیجهٔ تنظیم هم‌زمان مجموعه‌ای از ژن‌ها، مسیرهای سیگنال‌دهی، تغییرات اپی‌ژنتیکی و بازآرایی متابولیکی است که به‌شدت وابسته به نوع تنش، شدت آن و مرحلهٔ رشدی گیاه هستند (Raza *et al*., 2024).

تنش‌های غیرزیستی نظیر خشکی، شوری و تنش دمایی معمولاً منجر به اختلال در هموستازی سلولی، افزایش تولید گونه‌های فعال اکسیژن و تغییر الگوهای بیان ژن می‌شوند، در حالی‌که تنش‌های زیستی از طریق فعال‌سازی مسیرهای دفاعی وابسته به هورمون‌ها و شبکه‌های تنظیمی ایمنی گیاه عمل می‌کنند. با این حال، شواهد فزاینده‌ای نشان می‌دهد که این پاسخ‌ها مستقل از یکدیگر نبوده و مسیرهای مشترک و نقاط تلاقی متعددی میان پاسخ به تنش‌های زیستی و غیرزیستی وجود دارد که منجر به پاسخ‌های هم‌افزا یا متضاد می‌شوند (Zhou *et al*., 2022).

گزارش شده است که باکتری‌های محرک رشد گیاه می‌توانند از طریق افزایش فعالیت آنزیم‌های آنتی‌اکسیدانی، تنظیم هورمون‌ها و بهبود وضعیت آبی گیاه، تحمل به تنش خشکی را افزایش دهند (Vurukonda *et al*., 2016). برخی مطالعات موردی نشان می‌دهند که مداخلات زیستی مانند میکروارگانیسم‌های ریزوسفری نیز می‌توانند در کنار راهبردهای داده‌محور به بهبود تحمل تنش کمک کنند. به‌عنوان مثال، کاربرد سویه‌های استرپتومایسس در گیاهان دارویی و زراعی نشان داده است که این میکروارگانیسم‌ها می‌توانند علاوه بر افزایش عملکرد، محتوای متابولیت‌های ثانویه را نیز در شرایط تنش بهبود دهند (Esmaeil Zade *et al*., 2018). در سطح مولکولی، پاسخ گیاه به تنش با تغییرات گسترده در ترنسکریپتوم، پروتئوم و متابولوم همراه است که اغلب به‌صورت وابسته به زمان و با الگوهای غیرخطی بروز می‌یابند. افزون بر این، تنظیم‌های اپی‌ژنتیکی نظیر متیلاسیون دی‌ان‌ا و تغییرات هیستونی نقش مهمی در حافظه‌ی تنش و پاسخ‌های تطبیقی بلندمدت ایفا می‌کنند. این چندلایگی داده‌ها و برهم‌کنش‌های پیچیده، تحلیل پاسخ تنش را به مسئله‌ای فراتر از توان رویکردهای آماری خطی و تحلیل‌های تک‌متغیره تبدیل کرده است (Libbrecht and Noble, 2015).

فناوری‌های اُمیکس و فنوتیپینگ پرظرفیت، حجم عظیمی از داده‌های زیستی را در مطالعات تنش‌های گیاهی تولید کرده‌اند، اما استخراج روابط علّی و شناسایی عوامل کلیدی تعیین‌کننده‌ی تحمل تنش از این داده‌ها همچنان یک چالش اساسی محسوب می‌شود. در بسیاری از موارد، تغییرات کوچک در چند ژن یا مسیر تنظیمی می‌توانند اثرات غیرمنتظره و بزرگی بر عملکرد نهایی گیاه تحت تنش داشته باشند، موضوعی که تحلیل آن مستلزم مدل‌سازی روابط پیچیده و چندبعدی است (Washburn *et al*., 2020). در چنین شرایطی، یادگیری ماشین به‌عنوان رویکردی مناسب برای تحلیل داده‌های پرحجم، چندلایه و غیرخطی مطرح شده است. الگوریتم‌های یادگیری ماشین قادرند الگوهای پنهان در داده‌های اُمیکس و فنوتیپی را شناسایی کرده، ژن‌ها و مسیرهای کلیدی را اولویت‌بندی نموده و تعاملات پیچیده میان عوامل ژنتیکی و محیطی را مدل‌سازی کنند. بدین‌ترتیب، پیچیدگی پاسخ‌های گیاهی به تنش نه به‌عنوان یک مانع، بلکه به‌عنوان بستری مناسب برای بهره‌گیری از رویکردهای داده‌محور و هوشمند در اصلاح گیاهان زراعی در نظر گرفته می‌شود (Ubbens and Stavness, 2017; Van Dijk *et al*., 2021).

بنابراین، درک پاسخ گیاهان به تنش‌های زیستی و غیرزیستی نه‌تنها مستلزم شناخت مکانیسم‌های مولکولی منفرد، بلکه نیازمند تحلیل یکپارچه‌ی شبکه‌های تنظیمی، داده‌های چندلایه و تعاملات

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

پیچیده‌ی ژنتیکی است. این رویکرد سیستم‌محور، مبنای منطقی استفاده از فناوری‌های زیست‌فناوری نوین و الگوریتم‌های یادگیری ماشین در طراحی و توسعه‌ی گیاهان مقاوم به تنش را فراهم می‌کند.

مسیرهای بیوتکنولوژیک برای افزایش تحمل تنش

توسعه‌ی راهبردهای بیوتکنولوژیک برای افزایش تحمل تنش در گیاهان زراعی، پاسخی مستقیم به پیچیدگی شبکه‌ای پاسخ‌های گیاهی به تنش‌های زیستی و غیرزیستی است. با این حال، اثربخشی این راهبردها به‌شدت وابسته به انتخاب دقیق اهداف مولکولی و درک جامع از تعاملات ژنتیکی و تنظیمی درگیر در پاسخ به تنش است؛ امری که بدون بهره‌گیری از رویکردهای سیستم‌محور و داده‌محور به‌سختی قابل تحقق است.

۱. تراریختی

فناوری تولید گیاهان تراریخته برای دهه‌ها یکی از راهبردهای اصلی افزایش مقاومت گیاهان زراعی به تنش‌های زیستی و غیرزیستی بوده است. انتقال ژن‌های کاندید مرتبط با تحمل خشکی، شوری، تنش دمایی یا مقاومت به آفات و بیماری‌ها، امکان دستکاری مستقیم مسیرهای مولکولی کلیدی را فراهم کرده است. با این حال، موفقیت این رویکرد به‌شدت وابسته به انتخاب دقیق ژن هدف و درک صحیح از شبکه‌های تنظیمی درگیر در پاسخ به تنش است (Raza *et al*., 2021). در این چارچوب، چالش اصلی نه صرفاً انتقال ژن، بلکه انتخاب ژن یا مجموعه‌ای از ژن‌هاست که بتوانند بدون ایجاد اختلال در رشد و عملکرد، تحمل تنش را در شرایط واقعی مزرعه افزایش دهند. این مسئله اهمیت استفاده از تحلیل‌های داده‌محور برای کاهش ریسک انتخاب اهداف نامناسب در برنامه‌های تراریختی را برجسته می‌سازد.

مطالعات متعدد نشان داده‌اند که بیش‌بیان (Overexpression) یک ژن منفرد، اگرچه در شرایط کنترل‌شده می‌تواند منجر به افزایش تحمل تنش شود، اما اغلب در شرایط مزرعه‌ای با اثرات ناخواسته‌ای نظیر کاهش رشد یا عملکرد همراه است. این موضوع نشان‌دهنده‌ی ماهیت چندژنی و شبکه‌ای پاسخ به تنش بوده و ضرورت رویکردهایی را برجسته می‌کند که بتوانند تعاملات پیچیده‌ی ژن‌ها و مسیرها را در نظر بگیرند (Zhou *et al*., 2022). از این‌رو، رویکردهای مبتنی بر یادگیری ماشین که قادر به تحلیل داده‌های اُمیکس و شبکه‌ای هستند، می‌توانند به‌عنوان ابزاری مکمل در اولویت‌بندی ژن‌های کاندید برای تراریختی مورد استفاده قرار گیرند و احتمال بروز اثرات ناخواسته‌ی فیزیولوژیک و اکولوژیک را کاهش دهند (Van Dijk *et al*., 2021). درکنار این مسیرهای ژنتیکی، بهره‌گیری از راهبردهای زیستی چندمکانیسمی که بدون دستکاری مستقیم ژنوم گیاه، شبکه‌های دفاعی و فیزیولوژیک را تعدیل می‌کنند نیز حائز اهمیت هستند؛ به‌ویژه میکروارگانیسم‌های مفید خاک‌زی می‌توانند از طریق مهار پاتوژن‌ها، تقویت ایمنی و بهبود وضعیت تغذیه‌ای گیاه، پایداری پاسخ به تنش را افزایش دهند. ضمن اینکه سویه‌های استرپتومایسس با تولید ترکیبات ضدمیکروبی، آنزیم‌های تجزیه‌کننده‌ی دیواره‌ی سلولی و متابولیت‌های ثانویه، نقش مهمی در کنترل عوامل بیماری‌زا و ارتقای سلامت گیاه دارند (Gopalakrishnan *et al*., 2014; Abbasi *et al*., 2020b).

۲. ویرایش ژنومی مبتنی بر کریسپر (CRISPR)

CRISPR/Cas9 یک سامانهٔ قابل برنامه‌ریزی برای ویرایش ژن‌ها است که با استفاده از یک gRNA،

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

آنزیم Cas9 را به نقطه‌ای دقیق از ژنوم هدایت می‌کند تا برش دو رشته‌ای دی.ان.اِ ایجاد کند. ویرایش ژنومی مبتنی بر CRISPR/Cas9 تحول قابل‌توجهی در اصلاح گیاهان زراعی ایجاد کرده است و امکان ایجاد تغییرات دقیق، هدفمند و قابل پیش‌بینی در ژنوم را فراهم نموده است. این فناوری به‌ویژه برای اصلاح صفات پیچیده‌ای مانند تحمل تنش که نیازمند دستکاری چندین ژن یا عناصر تنظیمی هستند، پتانسیل بالایی دارد (Hesami *et al*., 2020).

با وجود دقت بالای کریسپر، چالش اصلی همچنان شناسایی اهداف ژنی مناسب و طراحی راهبردهای ویرایش چندگانه است. پاسخ‌های گیاهی به تنش‌ها اغلب نتیجه‌ی برهم‌کنش شبکه‌ای از ژن‌ها و تنظیم‌کننده‌ها هستند و دستکاری یک هدف منفرد ممکن است پاسخ مطلوب و پایداری ایجاد نکند. در این زمینه، مدل‌های یادگیری ماشین می‌توانند با تحلیل داده‌های ژنومی، ترنسکریپتومی و فنوتیپی، ژن‌های کلیدی، عناصر تنظیمی و ترکیب‌های بهینه‌ی اهداف ویرایشی را پیش‌بینی کنند و بدین‌ترتیب کارایی و دقت ویرایش ژنومی را افزایش دهند (Libbrecht and Noble, 2015; Van Dijk *et al*., 2021).

۳. مهار ژن با RNA interference (RNAi)

فناوری RNAi به‌عنوان ابزاری مؤثر برای خاموش‌سازی هدفمند ژن‌ها، نقش مهمی در مطالعه‌ی عملکرد ژن‌ها و افزایش مقاومت به تنش‌های زیستی، به‌ویژه آفات و عوامل بیماری‌زا، ایفا کرده است. با مهار بیان ژن‌های کلیدی در مسیرهای حساس، RNAi امکان تعدیل پاسخ‌های گیاهی به تنش را بدون ایجاد تغییرات دائمی در توالی ژنومی فراهم می‌کند (Hesami *et al*., 2020).

با این حال، کارایی RNAi به عواملی نظیر انتخاب ژن هدف، طراحی سازه‌های آران‌ا و پایداری خاموش‌سازی وابسته است. علاوه بر این، اثرات خارج از هدف و تغییرات غیرمنتظره در شبکه‌های تنظیمی می‌توانند نتایج این رویکرد را محدود کنند. در این زمینه، رویکردهای داده‌محور و یادگیری ماشین می‌توانند با پیش‌بینی اهداف مناسب، ارزیابی ریسک اثرات خارج از هدف و بهینه‌سازی طراحی توالی‌های RNAi، دقت و اثربخشی این فناوری را افزایش دهند (Ubbens and Stavness, 2017).

در مجموع، فناوری‌های تراریختی، ویرایش ژنومی مبتنی بر کریسپر و RNAi، ابزارهای قدرتمندی برای افزایش تحمل تنش در گیاهان زراعی فراهم کرده‌اند؛ با این حال، اثربخشی و ایمنی این راهبردها به‌شدت به انتخاب هوشمند اهداف مولکولی و درک جامع شبکه‌های زیستی وابسته است. در این میان، یادگیری ماشین به‌عنوان پیونددهنده‌ی داده‌های اُمیکس، تحلیل شبکه‌ای و تصمیم‌گیری اصلاحی، زمینه را برای گذار از اصلاح ژن‌محور به اصلاح سیستم‌محور فراهم می‌کند (Van Dijk *et al*., 2021).

یادگیری ماشین در تحقیقات مربوط به گیاهان تحت تنش

با توجه به ماهیت چندلایه، غیرخطی و وابسته به زمینه‌ی پاسخ گیاهان به تنش، تحلیل داده‌های حاصل از مطالعات تنش گیاهی نیازمند رویکردهایی است که بتوانند روابط پیچیده میان عوامل ژنتیکی، مولکولی و محیطی را به‌صورت یکپارچه مدل‌سازی کنند. در این چارچوب، یادگیری ماشین به‌عنوان ابزاری کلیدی برای استخراج دانش زیستی از داده‌های پرحجم و ناهمگن مطرح شده است (Van Dijk *et al*., 2021).

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

۱. از اُمیکس تا فنومیکس

مطالعات تنش‌های زیستی و غیرزیستی در گیاهان زراعی امروزه به‌شدت متکی بر داده‌های پرحجم و چندلایه‌ی زیستی هستند. فناوری‌های اُمیکس شامل ژنومیکس، ترنسکریپتومیکس، پروتئومیکس و متابولومیکس امکان پایش هم‌زمان هزاران متغیر مولکولی را در پاسخ به تنش فراهم کرده‌اند و دیدی جامع از تنظیم‌های ژنتیکی و متابولیکی ارائه می‌دهند (Raza *et al.*, 2021). افزون بر این، داده‌های اپی‌ژنتیکی نظیر الگوهای متیلاسیون دی‌ان‌ا و تغییرات هیستونی، لایه‌ی دیگری از پیچیدگی را به پاسخ گیاه به تنش اضافه می‌کنند که اغلب به‌صورت وابسته به زمان و شرایط محیطی عمل می‌کند (Libbrecht and Noble, 2015).

در کنار داده‌های مولکولی، توسعه‌ی سامانه‌های فنوتیپینگ پرظرفیت، امکان جمع‌آوری داده‌های دقیق فنوتیپی شامل ویژگی‌های مورفولوژیک، فیزیولوژیک و تصویری گیاهان تحت شرایط تنش را فراهم کرده است. این داده‌ها که اغلب از طریق تصویربرداری چندطیفی، حرارتی و سنسورهای محیطی به‌دست می‌آیند، حجم بالایی داشته و حاوی الگوهای پیچیده‌ای هستند که تفسیر آن‌ها با روش‌های تحلیلی سنتی دشوار است (Washburn *et al.*, 2020). ادغام داده‌های اُمیکس و فنومیکس، اگرچه تصویری جامع از پاسخ گیاه به تنش ارائه می‌دهد، اما به‌دلیل حجم بالا، نویز و هم‌بستگی‌های پیچیده‌ی داده‌ها، تنها در صورت بهره‌گیری از الگوریتم‌های یادگیری ماشین قابل تفسیر مؤثر و کاربردی است (Van Dijk *et al.*, 2021).

۲. الگوریتم‌های یادگیری ماشین به‌کار رفته در تحقیقات مربوط به گیاهان تحت تنش

برای تحلیل داده‌های پیچیده‌ی حاصل از مطالعات تنش گیاهی، طیف متنوعی از الگوریتم‌های یادگیری ماشین مورد استفاده قرار گرفته‌اند. یادگیری ماشین شاخه‌ای از هوش مصنوعی است که به سیستم‌ها توانایی یادگیری خودکار از داده‌ها و بهبود تجربه را بدون برنامه‌نویسی صریح می‌دهد. هسته اصلی این روش‌ها، استخراج الگوهای پیچیده و ساختن مدل‌های پیش‌بینی‌کننده از طریق آموزش بر روی نمونه‌های داده است. انتخاب الگوریتم مناسب بسته به نوع مسئله و ساختار داده‌ها، در چهار گروه اصلی یادگیری نظارت‌شده (برای پیش‌بینی)، بدون نظارت (برای اکتشاف)، تجمعی (برای بهبود دقت) و عمیق (برای داده‌های پیچیده‌ای مانند تصاویر) به کار می‌روند.

برای تحلیل جامع داده‌های مرتبط با تنش گیاهی، از چهار رده اصلی از الگوریتم‌های یادگیری ماشین استفاده شده است: ۱- الگوریتم‌های یادگیری نظارت‌شده که با بهره‌گیری از یک مجموعه داده آموزشی شامل جفت‌های ورودی-خروجی (برچسب‌دار)، یک تابع نگاشت را یادگیری می‌کنند تا بتوانند برچسب (مثلاً مقاوم/حساس) را برای نمونه‌های جدید پیش‌بینی نمایند. ۲- الگوریتم‌های یادگیری بدون نظارت که وظیفه کشف ساختارها، الگوها یا خوشه‌های ذاتی موجود در داده‌های بدون برچسب را بر عهده دارند. ۳- الگوریتم‌های یادگیری تجمعی (Ensemble learning) که با ترکیب خروجی چندین مدل پایه (مانند درختان تصمیم) به منظور کاهش واریانس (Bagging) یا بایاس (Boosting)، به مدل نهایی با کارایی و پایداری بالاتر دست می‌یابند. ۴- الگوریتم‌های یادگیری عمیق که مبتنی بر شبکه‌های عصبی مصنوعی با لایه‌های متعدد (عمیق) هستند و توانایی استخراج خودکار

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

سلسله‌مراتبی ویژگی‌های پیچیده از داده‌های خام (مانند تصاویر یا توالی‌ها) را دارا هستند.

روش‌های یادگیری نظارت‌شده مانند جنگل تصادفی (Random Forest) که یک روش یادگیری تجمعی مبتنی بر مجموعه‌ای از درخت‌های تصمیم است که با رأی‌گیری/میانگین‌گیری خروجی درخت‌ها، دقت و پایداری را افزایش می‌دهد و ماشین بردار پشتیبان (Support Vector Machine) الگوریتمی برای طبقه‌بندی/رگرسیون که با بیشینه‌کردن حاشیه بین کلاس‌ها و (در صورت نیاز) با کرنل، مرزهای غیرخطی را مدل می‌کند؛ به‌طور گسترده برای شناسایی ژن‌های پاسخ‌دهنده به تنش، انتخاب ویژگی‌های کلیدی از داده‌های اُمیکس و طبقه‌بندی فنوتیپ‌های مقاوم و حساس به‌کار رفته‌اند. این الگوریتم‌ها به‌ویژه در شرایطی که تعداد متغیرها بسیار بیشتر از تعداد نمونه‌هاست، عملکرد پایداری از خود نشان داده‌اند (Washburn *et al*., 2020; Gill *et al*., 2022).

علاوه بر این، الگوریتم‌های مبتنی بر یادگیری تجمعی (EL) رویکردی که چند مدل را برای بهبود دقت و کاهش واریانس/بایاس باهم ترکیب می‌کند؛ مانند Gradient Boosting، XGBoost و LightGBM، خانواده‌ای از روش‌های boosting که مدل‌ها را به‌صورت ترتیبی می‌سازند تا خطاهای قبلی را اصلاح کنند؛ XGBoost و LightGBM پیاده‌سازی‌های بهینه و مقیاس‌پذیر این ایده‌اند. این الگوریتم‌ها به‌دلیل توانایی در مدل‌سازی روابط غیرخطی و مدیریت داده‌های پرنویز، در پیش‌بینی عملکرد گیاهان تحت شرایط مختلف تنش و اولویت‌بندی ژن‌های کاندید کاربرد گسترده‌ای یافته‌اند (Shrestha *et al*., 2024; Li *et al*., 2025).

در سال‌های اخیر، زیرمجموعه‌ای از یادگیری ماشین مبتنی بر شبکه‌های عصبی چندلایه برای یادگیری خودکار نمایش/ویژگی از داده‌های پیچیده تحت عنوان مدل‌های یادگیری عمیق (Deep Learning) جایگاه ویژه‌ای در تحلیل داده‌های تصویری و زمانی مرتبط با فنوتیپینگ تنش پیدا کرده‌اند. شبکه‌های عصبی کانولوشنی Convolutional Neural Networks (CNNs) مناسب داده‌های تصویری که با کانولوشن، الگوهای مکانی را استخراج می‌کند؛ به‌طور گسترده برای استخراج خودکار ویژگی‌ها از تصاویر چندطیفی، حرارتی و RGB و تشخیص علائم تنش در مراحل اولیه به‌کار رفته‌اند، در حالی‌که شبکه‌های عصبی بازگشتی Recurrent Neural Networks (RNNs) برای داده‌های ترتیبی/زمانی که وابستگی‌های زمانی را با حالت بازگشتی مدل می‌کند و مدل‌های مبتنی بر حافظهٔ بلند–کوتاه‌مدت Long short-term memory (LSTM) با سازوکار گیت که وابستگی‌های بلندمدت را بهتر حفظ می‌کند، برای تحلیل داده‌های زمانی و پویای پاسخ گیاه به تنش مورد استفاده قرار گرفته‌اند (Ubbens and Stavness, 2017; Singh *et al*., 2018). افزون بر این، رویکردهای یادگیری بدون نظارت و شبکه‌محور، از جمله خوشه‌بندی ژن‌ها، تحلیل شبکه‌های هم‌بیانی و مدل‌های مبتنی بر Graph Neural Networks (GNNs)، امکان بازسازی شبکه‌های تنظیمی ژنی و شناسایی ماژول‌های مولکولی مرتبط با تحمل تنش را فراهم کرده‌اند. این روش‌ها به‌ویژه برای درک ساختار شبکه‌ای پاسخ گیاه به تنش و شناسایی گره‌های تنظیمی کلیدی اهمیت دارند (Chang *et al*., 2024) (جدول ۱).

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

جدول ۱. الگوریتم‌های یادگیری ماشین. خلاصه‌ای از الگوریتم‌های یادگیری ماشین و ویژگی‌های اختصاصی آن‌ها و نمونه داده‌های مورد استفاده در آن‌ها (Hesami *et al.*, 2022)

الگوریتم	توضیحات	نقاط قوت	نقاط ضعف	انواع داده مناسب	کاربرد نمونه
جنگل تصادفی (Random Forest)	روش یادگیری ترکیبی مبتنی بر درخت‌های تصمیم	مقاوم، با داده‌های با ابعاد بالا سازگار است، نسبتاً آسان برای تفسیر	مستعد بیش‌برازش (Overfitting)، از نظر محاسباتی پرهزینه است.	عددی، طبقه‌ای	شناسایی ژن‌های کلیدی مرتبط با تحمل به خشکی
ماشین بردار پشتیبان (SVM)	یافتن صفحه بهینه برای جداسازی کلاس‌های داده	مؤثر در فضاهای با ابعاد بالا، توابع هسته متنوع	حساس به تنظیم پارامترها، از نظر محاسباتی برای مجموعه‌های داده بزرگ پرهزینه است	عددی	شناسایی ژن‌های کلیدی مرتبط با تحمل به خشکی
یادگیری عمیق (CNNها، RNNها)	شبکه‌های عصبی مصنوعی با چندین لایه	می‌تواند به طور خودکار ویژگی‌های پیچیده را یاد بگیرد، عالی برای داده‌های بدون ساختار	نیاز به مجموعه‌های داده بزرگ، تفسیر دشوار، از نظر محاسباتی پرهزینه است	تصویر، سری زمانی، متن	پیش‌بینی مقاومت به بیماری بر اساس تحلیل تصاویر
تقویت گرادیان (Gradient Boosting)	روش یادگیری ترکیبی که به طور متوالی درخت‌های تصمیم را می‌سازد	دقت بالا، با مقادیر گمشده سازگار است، مقاوم در برابر نقاط پرت	مستعد بیش‌برازش در صورت عدم تنظیم مناسب (Regularization)	عددی	پیش‌بینی عملکرد محصول در شرایط تنش مختلف
شبکه‌های عصبی (Networks Neural)	مدل‌هایی الهام گرفته از ساختار و عملکرد مغز انسان	می‌تواند روابط پیچیده و غیرخطی را مدیریت کند	نیاز به مقدار زیادی داده، از نظر محاسباتی پرهزینه است.	عددی، طبقه‌ای	پیش‌بینی تنش

۳. برتری‌های یادگیری ماشین بر روش‌های آماری کلاسیک

یکی از محدودیت‌های اصلی روش‌های آماری کلاسیک در مطالعات تنش گیاهی، اتکای آن‌ها به فرضیات ساده‌ساز مانند خطی بودن روابط و استقلال متغیرها است. در حالی‌که پاسخ گیاه به تنش، حاصل برهم‌کنش‌های غیرخطی و وابسته به زمینه میان تعداد زیادی ژن، مسیر مولکولی و عوامل محیطی است، این فرضیات اغلب نقض می‌شوند (Van Dijk *et al.*, 2021).

یادگیری ماشین در مقایسه با روش‌های آماری سنتی، توانایی بالاتری در مدل‌سازی روابط غیرخطی، مدیریت داده‌های پرابعاد و استخراج الگوهای پیچیده از مجموعه‌داده‌های ناهمگن دارد. این ویژگی‌ها به‌ویژه در تحلیل داده‌های اُمیکس و فنومیکس که با نویز بالا، هم‌بستگی‌های پیچیده و ساختارهای چندسطحی همراه هستند، اهمیت ویژه‌ای پیدا می‌کنند (Libbrecht and Noble, 2015). علاوه بر این، مدل‌های یادگیری ماشین می‌توانند برای اولویت‌بندی ژن‌ها و مسیرهای مولکولی، پیش‌بینی عملکرد تحت شرایط تنش و پشتیبانی از تصمیم‌گیری‌های اصلاحی مورد استفاده قرار گیرند، در حالی‌که روش‌های آماری کلاسیک معمولاً به تحلیل‌های توصیفی یا تک‌متغیره محدود می‌شوند. ضمن اینکه، راهبردهای انتخاب ویژگی و مدل‌های تجمعی در یادگیری ماشین، با اولویت‌بندی ژن‌هایی که به‌طور پایدار در پیش‌بینی عملکرد مدل نقش دارند، احتمال شناسایی کاذب ژن‌ها را کاهش داده و دقت نتایج حاصل از تحلیل داده‌های ژنومی و ترنسکریپتومیک را افزایش می‌دهند (Washburn *et al.*, 2020).

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

در مجموع، یادگیری ماشین نه به‌عنوان جایگزین کامل روش‌های آماری، بلکه به‌عنوان مکملی قدرتمند برای تحلیل داده‌های پیچیده‌ی تنش‌های گیاهی مطرح است که با کاهش عدم قطعیت و افزایش دقت انتخاب می‌تواند شکاف میان داده‌های چندلایه‌ی زیستی و کاربردهای عملی در اصلاح و به‌نژادی گیاهان زراعی را کاهش دهد (Van Dijk *et al*., 2021; Novielli *et al*., 2024; Li *et al*., 2025).

از داده تا بهبود محصولات کشاورزی

۱. از داده‌های اُمیکس تا استنتاج مبتنی بر یادگیری ماشین

پیشرفت فناوری‌های اُمیکس امکان تولید داده‌های گسترده‌ای از سطوح مختلف سازمان‌یافتگی زیستی، از ژنوم و ترنسکریپتوم تا پروتئوم و متابولوم را در مطالعات تنش‌های گیاهی فراهم کرده است. این داده‌ها اطلاعات ارزشمندی درباره‌ی پاسخ‌های مولکولی گیاه به تنش‌های زیستی و غیرزیستی ارائه می‌دهند، اما به‌دلیل حجم بالا، ناهمگنی و هم‌بستگی‌های پیچیده، تفسیر مستقیم آن‌ها با رویکردهای سنتی دشوار است (Raza *et al*., 2021; Zhou *et al*., 2022).

یادگیری ماشین نقش کلیدی در گذار از داده‌های توصیفی به استنتاج‌های پیش‌بینانه ایفا می‌کند و با کاهش ابعاد داده، شناسایی الگوهای پنهان و مدل‌سازی روابط غیرخطی، امکان استخراج دانش زیستی قابل استفاده در اصلاح گیاهان زراعی را فراهم می‌سازد (Libbrecht and Noble, 2015; Tsega and Mullualem, 2026).

الگوریتم‌های یادگیری ماشین با ادغام داده‌های چندلایه‌ی اُمیکس قادرند روابط غیرخطی میان ژن‌ها، مسیرهای مولکولی و فنوتیپ‌های مرتبط با تنش را مدل‌سازی کرده و سیگنال‌های زیستی ضعیف اما معنادار را که در تحلیل‌های آماری کلاسیک نادیده گرفته می‌شوند، شناسایی کنند. این رویکرد داده‌محور، امکان گذار از تحلیل‌های توصیفی به پیش‌بینی عملکرد گیاه تحت شرایط مختلف تنش را فراهم می‌سازد (Van Dijk *et al*., 2021; Tsega and Mullualem, 2026).

۲. شناسایی و اولویت‌بندی ژن‌های منتخب

یکی از مراحل کلیدی در اصلاح گیاهان مقاوم به تنش، شناسایی و اولویت‌بندی ژن‌های کاندید است که بیشترین نقش را در پاسخ گیاه به شرایط تنش ایفا می‌کنند. مدل‌های یادگیری ماشین می‌توانند با تحلیل هم‌زمان داده‌های بیان ژن، شبکه‌های تنظیمی و اطلاعات فنوتیپی، ژن‌ها و مسیرهایی را که سهم بیشتری در تحمل تنش دارند، رتبه‌بندی کنند (Washburn *et al*., 2020; Gill *et al*., 2022). این اولویت‌بندی داده‌محور، به‌ویژه برای صفات پیچیده و چندژنی، جایگزینی کارآمد برای رویکردهای مبتنی بر ژن‌های منفرد محسوب می‌شود و می‌تواند ریسک انتخاب اهداف نامناسب در برنامه‌های تراریختی و ویرایش ژنومی را کاهش دهد (Van Dijk *et al*., 2021; Shrestha *et al*., 2024).

تحلیل‌های مبتنی بر یادگیری ماشین اغلب خانواده‌های ژنی تنظیم‌کننده‌ای مانند DREB، NAC و WRKY و همچنین اجزای کلیدی مسیرهای سیگنال‌دهی از جمله MAPK و SnRK2 را به‌عنوان گره‌های تنظیمی مرکزی در پاسخ گیاهان به تنش معرفی می‌کنند؛ ژن‌هایی که نقش آن‌ها در کنترل هم‌زمان چندین مسیر پاسخ به تنش در گونه‌های مختلف گیاهی گزارش شده است (Raza *et al*., 2021; Zhou *et al*., 2022; Abassi *et al*., 2022). ادغام داده‌های بیان ژن، فنوتیپ و شرایط محیطی در مدل‌های یادگیری ماشین، امکان شناسایی سیستم‌محور ژن‌های کاندید را فراهم می‌کند؛ به‌گونه‌ای که اهداف

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

اصلاحی منتخب، در مقایسه با رویکردهای هم‌بستگی تک‌ژنی، پایداری و کارایی بالاتری در ایجاد تحمل تنش نشان می‌دهند (Van Dijk *et al.*, 2021).

۳. اعتبارسنجی و بازخورد به مدل‌های یادگیری ماشین

با وجود توان بالای مدل‌های یادگیری ماشین در پیش‌بینی ژن‌ها و مسیرهای مرتبط با تحمل تنش، اعتبارسنجی تجربی همچنان مرحله‌ای ضروری در پایپ‌لاین‌های (pipeline) اصلاح گیاهان زراعی محسوب می‌شود. پایپ‌لاین یک گردشکار خودکار و مرحله‌به‌مرحله برای پردازش و تحلیل داده‌ها است، که قابلیت تکرارپذیری و کارایی را افزایش میدهد. پایپ‌لاین ژن‌های کاندید شناسایی‌شده باید از طریق آزمایش‌های مولکولی، فیزیولوژیک و فنوتیپی مورد ارزیابی قرار گیرند تا نقش آن‌ها در تحمل تنش به‌طور قطعی تأیید شود (Hesami *et al.*, 2020; Van Dijk *et al.*, 2021). اعتبارسنجی موفق ژن‌های منتخب، علاوه بر ارتقای درک مکانیسم‌های پاسخ به تنش، امکان بازخورد داده‌های تجربی به مدل‌های یادگیری ماشین و بهبود تدریجی دقت پیش‌بینی‌ها را نیز فراهم می‌کند؛ رویکردی چرخه‌ای که هسته‌ی پایپ‌لاین‌های داده‌محور اصلاح گیاهان را تشکیل می‌دهد (Van Dijk *et al.*, 2021; Li *et al.*, 2025).

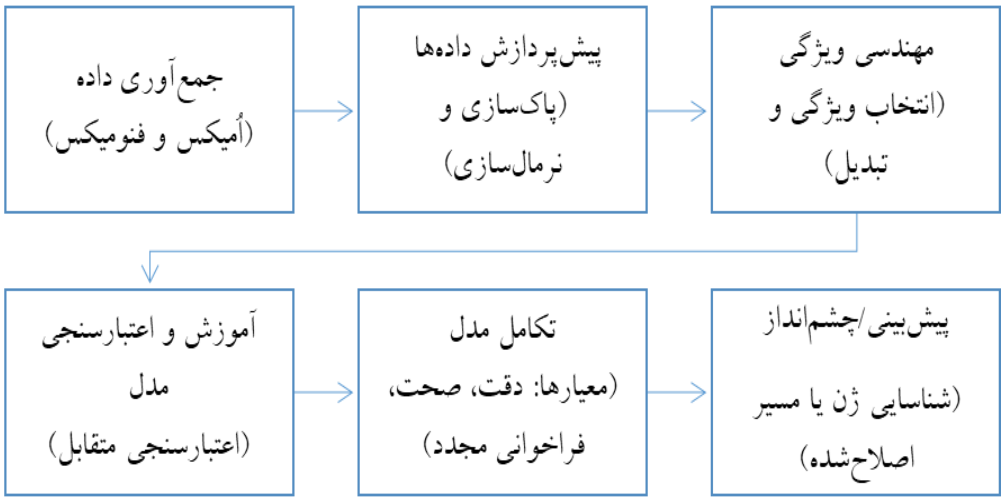
۴. از فناوری‌های ویرایش ژن تا توسعهٔ محصولات تراریخته

پس از شناسایی و اعتبارسنجی ژن‌های کلیدی، ابزارهای زیست‌فناوری نظیر تراریختی، ویرایش ژنومی مبتنی بر CRISPR/Cas9 و فناوری RNAi برای ایجاد تغییرات هدفمند در ژنوم گیاهان زراعی به‌کار گرفته می‌شوند. ترکیب خروجی مدل‌های یادگیری ماشین با این فناوری‌ها، امکان طراحی راهبردهای اصلاحی دقیق‌تر، از جمله ویرایش هم‌زمان چند ژن یا عناصر تنظیمی، را فراهم می‌کند (Hesami *et al.*, 2020; Van Dijk *et al.*, 2021).

در این چارچوب، پایپ‌لاین یکپارچه‌ای که از داده به مدل سپس ژن کاندید، اعتبار سنجی می‌شود و نهایتا به گیاه اصلاح‌شده یا تراریخته می‌رسد؛ به‌عنوان رویکردی کارآمد برای توسعه‌ی گیاهان مقاوم به تنش مطرح می‌شود؛ رویکردی که می‌تواند احتمال موفقیت اصلاح ژنتیکی را افزایش داده و ریسک بروز اثرات ناخواسته‌ی زیستی و اکولوژیک را کاهش دهد (Zhou *et al.*, 2022; Kumar *et al.*, 2021) (شکل ۱).

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]



شکل ۱. یک پایپ‌لاین یادگیری ماشین: حالت استاندارد یک چارچوب داده‌محور در اکتشافات

کارایی رویکردهای مبتنی بر یادگیری ماشین و زیست‌فناوری گیاهی زمانی به‌طور ملموس قابل ارزیابی است که در بستر محصولات زراعی بااهمیت مورد بررسی قرار گیرد. ازاین‌رو، در بخش بعدی، مطالعات موردی انجام‌شده بر روی برنج، ذرت و آرابیدوپسیس به‌عنوان نمونه‌هایی از کاربرد عملی این رویکردها مرور می‌شوند.

مطالعات موردی در محصولات زراعی بااهمیت

شواهد نشان می‌دهد که باکتری‌های محرک رشد گیاه از طریق تنظیم تعادل یونی، افزایش دسترسی عناصر غذایی و کاهش تنش اکسیداتیو، نقش مؤثری در افزایش تحمل گیاهان به تنش شوری ایفا می‌کنند (Kumar *et al.*, 2021). علاوه بر این، رویکردهای نوین مبتنی بر تحلیل داده‌های اُمیکس و الگوریتم‌های یادگیری ماشین نیز به‌عنوان ابزارهای مکمل، امکان شناسایی ژن‌ها، مسیرهای سیگنال‌دهی و شبکه‌های تنظیمی مرتبط با تحمل تنش را فراهم کرده‌اند. در همین راستا، مطالعات موردی انجام‌شده بر روی محصولات زراعی مهم نظیر برنج، ذرت و آرابیدوپسیس نشان می‌دهد که ادغام فناوری‌های زیستی با مدل‌های داده‌محور می‌تواند کارایی برنامه‌های اصلاح گیاهان مقاوم به تنش را به‌طور چشمگیری افزایش دهد.

۱. برنج (*Oryza sativa*)

برنج به‌عنوان یکی از مهم‌ترین محصولات غذایی جهان، به‌شدت تحت تأثیر تنش‌های غیرزیستی نظیر خشکی، شوری و تنش دمایی قرار دارد و به همین دلیل به مدل مناسبی برای کاربرد رویکردهای داده‌محور در اصلاح گیاهان مقاوم به تنش تبدیل شده است. مطالعات مبتنی بر ادغام داده‌های ترنسکریپتومیک، ژنومیک و فنوتیپی نشان داده‌اند که الگوریتم‌های یادگیری ماشین می‌توانند ژن‌ها و مسیرهای کلیدی دخیل در تحمل تنش را با دقت بالاتری نسبت به روش‌های کلاسیک شناسایی و اولویت‌بندی کنند (Raza *et al.*, 2021).

در این مطالعات، مدل‌های یادگیری ماشین با تحلیل الگوهای بیان ژن تحت شرایط مختلف تنش، مجموعه‌ای از ژن‌های تنظیمی و شبکه‌های مولکولی را شناسایی کرده‌اند که دستکاری هدفمند (Genetic Manpulation) آن‌ها می‌تواند

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

تحمل تنش را بدون افت قابل‌توجه عملکرد افزایش دهد؛ موضوعی که اهمیت این رویکردها را برای اصلاح دقیق و کم‌ریسک برنج برجسته می‌سازد (Van Dijk *et al*., 2021; Tsega and Mullualem, 2026).

این یافته‌ها زمینه‌ساز توسعه‌ی راهبردهای تراریختی و ویرایش ژنومی هدفمند در برنج شده و نشان داده‌اند که استفاده از یادگیری ماشین می‌تواند احتمال موفقیت اصلاح ژنتیکی در شرایط مزرعه‌ای را افزایش دهد (Kumar *et al*., 2021; Zhou *et al*., 2022).

در برنج، رویکردهای شبکه‌محور و یادگیری ماشین برای استخراج ژن‌های کلیدی پاسخ به تنش از داده‌های چندلایه به‌کار گرفته شده‌اند. برای مثال، Liu و همکاران (۲۰۲۴) با ادغام شبکه‌های پروتئین–پروتئین، هم‌بیانی و مسیرهای زیستی و استفاده از الگوریتم Random Walk روی شبکه‌های چندلایه، ژن‌های کاندید مرتبط با پاسخ به خشکی را رتبه‌بندی کردند و سپس بخشی از ژن‌های برتر را با qRT-PCR به‌عنوان ژن‌های پاسخ‌گو تأیید کردند. این نوع خروجی‌های مبتنی بر رتبه‌بندی ژن، می‌تواند «ورودی» مرحله‌ی بعدی در زیست‌فناوری باشد؛ یعنی انتخاب هدف(ها) برای تراریختی/ویرایش ژن. به‌عنوان نمونه از هدف‌های شناخته‌شده در مسیرهای تنش که قابلیت قرار گرفتن در این پایپ‌لاین را دارند، OsSAPK7 (یک عضو خانواده SnRK2) است که نشان داده شده دستکاری آن در برنج به‌طور معنی‌دار تحمل شوری را تغییر می‌دهد؛ به‌طوری که لاین‌های بیش‌بیان نسبت به گروه شاهد، بقا و شاخص‌های فیزیولوژیک بهتری تحت تنش نمکی دارند و لاین‌های knockout شده، حساس‌تر می‌شوند (Zeng *et al*., 2021). بنابراین، چارچوب عملیِ این پایپ‌لاین می‌تواند به‌صورت شفاف نشان دهد که روش‌های ML چگونه به تصمیم‌گیری ژنتیکی منتهی می‌شوند.

۲. ذرت (*Zea mays*)

ذرت به‌دلیل ژنوم پیچیده، تنوع ژنتیک بالا و حساسیت عملکرد به تنش‌های محیطی، چالشی جدی برای اصلاح کلاسیک محسوب می‌شود. در سال‌های اخیر، ادغام داده‌های فنومیکس پرظرفیت و اُمیکس با مدل‌های یادگیری ماشین، امکان پیش‌بینی عملکرد ذرت تحت شرایط تنش و شناسایی عوامل ژنتیکی مؤثر بر تحمل تنش را فراهم کرده است (Washburn *et al*., 2020; Van Dijk *et al*., 2021).

نتایج این مطالعات نشان می‌دهد که مدل‌های یادگیری ماشین قادرند روابط غیرخطی میان ویژگی‌های فنوتیپی، داده‌های ژنتیکی و شرایط محیطی را با دقت بالاتری نسبت به روش‌های آماری کلاسیک مدل‌سازی کنند؛ قابلیتی که برای صفات پیچیده و چندژنی مانند تحمل تنش در ذرت اهمیت ویژه‌ای دارد (Gill *et al*., 2022; Li *et al*., 2025). این رویکردها نقش مهمی در طراحی راهبردهای اصلاح ژنتیکی مبتنی بر کریسپر و انتخاب ژن‌های کاندید ایفا کرده‌اند و به کاهش عدم قطعیت در تصمیم‌گیری‌های اصلاحی کمک نموده‌اند (Van Dijk *et al*., 2021; Kumar *et al*., 2021).

۳. آرابیدوپسیس (*Arabidopsis thaliana*)

آرابیدوپسیس به‌عنوان گیاه مدل، نقش کلیدی در توسعه و ارزیابی رویکردهای یادگیری ماشین در مطالعات تنش گیاهی داشته است. دسترسی به داده‌های گسترده‌ی اُمیکس، شبکه‌های تنظیمی دقیق و منابع ژنتیکی غنی، این گونه را به بستری مناسب برای آزمون مدل‌های داده‌محور و توسعه‌ی چارچوب‌های مفهومی پاسخ به تنش تبدیل کرده است (Libbrecht and Noble, 2015; Ubbens and Stavness, 2017).

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

مطالعات مبتنی بر یادگیری ماشین در آرابیدوپسیس نشان داده‌اند که این رویکردها می‌توانند تعاملات پیچیده‌ی ژنتیکی و اپی‌ژنتیکی دخیل در پاسخ به تنش را آشکار سازند و مسیرهای مولکولی مشترک میان تنش‌های مختلف را شناسایی کنند؛ یافته‌هایی که مبنای انتقال دانش به گیاهان زراعی را فراهم می‌سازد (Van Dijk *et al.*, 2021). نتایج حاصل از این مطالعات، علاوه بر ارتقای درک بنیادی از زیست‌شناسی تنش، به‌عنوان چارچوبی مفهومی برای طراحی راهبردهای اصلاح ژنتیکی هدفمند در محصولات زراعی مورد استفاده قرار گرفته‌اند (Ubbens and Stavness, 2017; Zhou *et al.*, 2022).

بررسی مطالعات موردی انجام‌شده بر روی برنج، ذرت و آرابیدوپسیس نشان می‌دهد که تحلیل‌های مبتنی بر یادگیری ماشین به‌طور مکرر ژن‌های تنظیمی کلیدی مانند DREB، NAC و WRKY، به‌همراه اجزای مرکزی مسیرهای سیگنال‌دهی و تنظیم یونی را به‌عنوان عوامل اصلی تحمل تنش شناسایی می‌کنند. تکرارپذیری این یافته‌ها در گونه‌های مختلف گیاهی، نشان‌دهنده‌ی پایداری و قابلیت تعمیم رویکردهای داده‌محور در اصلاح گیاهان مقاوم به تنش است (Washburn *et al.*, 2020; Raza *et al.*, 2021; Zhou *et al.*, 2022).

در مجموع، مطالعات موردی تأکید می‌کنند که موفقیت اصلاح ژنتیکی گیاهان زراعی، بیش از آن‌که به یک فناوری منفرد وابسته باشد، به همگرایی داده‌های چندلایه، مدل‌سازی هوشمند، اعتبارسنجی زیستی و در نظر گرفتن ملاحظات ایمنی زیستی متکی است (Van Dijk *et al.*, 2021; Kumar *et al.*, 2021). با وجود نتایج امیدوارکننده‌ی حاصل از مطالعات موردی، کاربرد گسترده‌ی یادگیری ماشین و زیست‌فناوری در اصلاح گیاهان مقاوم به تنش همچنان با چالش‌ها، محدودیت‌ها و ملاحظات اخلاقی و قانونی همراه است که در بخش بعدی مورد بررسی قرار می‌گیرند.

چالش‌ها، محدودیت‌ها و ملاحظات اخلاقی

۱. بیش‌برازش مدل‌ها و چالش‌های تعمیم‌پذیری نتایج

یکی از چالش‌های اساسی در کاربرد یادگیری ماشین در پژوهش‌های تنش گیاهی، خطر بیش‌برازش مدل‌ها است. در این حالت، مدل به‌جای یادگیری الگوهای زیستی تعمیم‌پذیر، ویژگی‌های خاص و نویز موجود در مجموعه‌داده‌ی آموزشی را می‌آموزد؛ به‌طوری که عملکرد آن در داده‌های جدید یا شرایط محیطی متفاوت به‌طور قابل‌توجهی کاهش می‌یابد (Libbrecht and Noble, 2015). این مسئله به‌ویژه در مطالعات تنش گیاهی که اغلب بر داده‌های محدود، ناهمگن یا حاصل از شرایط کنترل‌شده‌ی آزمایشگاهی متکی هستند، اهمیت بیشتری پیدا می‌کند. مدل‌هایی که در چنین شرایطی آموزش می‌بینند، ممکن است توان پیش‌بینی پاسخ گیاه به تنش را در مقیاس مزرعه یا در گونه‌های دیگر به‌درستی منعکس نکنند (Singh *et al.*, 2018; Van Dijk *et al.*, 2021).

برای کاهش خطر بیش‌برازش، استفاده از راهبردهایی نظیر اعتبارسنجی متقابل (cross-validation)، به‌کارگیری مجموعه‌داده‌های مستقل و چندمحیطی، ساده‌سازی مدل‌ها و ترکیب چند الگوریتم در قالب مدل‌های تجمعی توصیه می‌شود. علاوه بر این، توسعه‌ی مدل‌های تفسیرپذیر و ادغام دانش زیستی پیشین می‌تواند به افزایش قابلیت تعمیم و اعتمادپذیری نتایج کمک کند (Van Dijk *et al.*, 2021; Novielli *et al.*, 2024).

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

۲. سوگیری داده‌ها و محدودیت‌های منابع اُمیکس

سوگیری داده یکی دیگر از محدودیت‌های مهم در کاربرد یادگیری ماشین در اصلاح گیاهان مقاوم به تنش است. بخش عمده‌ای از داده‌های اُمیکس و فنومیکس در دسترس، مربوط به گونه‌های مدل یا تعداد محدودی از محصولات زراعی بوده و اغلب تحت شرایط آزمایشگاهی خاص تولید شده‌اند. این تمرکز نامتوازن می‌تواند منجر به توسعه‌ی مدل‌هایی شود که برای گونه‌های کمتر مطالعه‌شده یا شرایط محیطی واقعی کارایی محدودی دارند (Washburn *et al*., 2020; Raza *et al*., 2021).

علاوه بر این، تفاوت در پروتکل‌های نمونه‌برداری، کیفیت متغیر داده‌ها و کمبود داده‌های برچسب‌خورده‌ی دقیق، چالش‌هایی جدی برای آموزش مدل‌های یادگیری ماشین ایجاد می‌کند. این محدودیت‌ها به‌ویژه در مطالعات چنداُمیکسی که نیازمند هم‌ترازی دقیق داده‌ها هستند، برجسته‌تر می‌شوند (Zhou *et al*., 2022; Tsega and Mullualem, 2026).

رفع این چالش‌ها مستلزم توسعه‌ی پایگاه‌های داده‌ی استاندارد، اشتراک‌گذاری داده‌ها در مقیاس بین‌المللی و طراحی آزمایش‌های هماهنگ میان گروه‌های تحقیقاتی است؛ اقداماتی که می‌توانند کیفیت داده‌ها و قابلیت تعمیم مدل‌های یادگیری ماشین را به‌طور قابل‌توجهی بهبود بخشند (Tsega and Mullualem, 2026).

۳. موانع اخلاقی و اجتماعی درمقابل دستاوردهای نوین

با وجود پیشرفت‌های چشمگیر در فناوری‌های تراریختی و ویرایش ژنومی، ملاحظات قانونی، اخلاقی و اجتماعی همچنان مانعی جدی در مسیر کاربرد گسترده‌ی این دستاوردها در کشاورزی محسوب می‌شوند. قوانین سخت‌گیرانه‌ی مرتبط با تولید، ارزیابی ریسک و رهاسازی گیاهان تراریخته، همراه با تفاوت‌های منطقه‌ای در مقررات، می‌توانند انتقال یافته‌های آزمایشگاهی به مزرعه را به تأخیر اندازند (Zhou *et al*., 2022).

در این میان، کاربرد یادگیری ماشین در شناسایی دقیق‌تر اهداف ژنتیکی و پیش‌بینی پیامدهای دستکاری ژنومی، این ظرفیت را دارد که ریسک‌های ایمنی زیستی را کاهش داده و شفافیت علمی در فرآیند تصمیم‌گیری را افزایش دهد. با این حال، خودِ مدل‌های یادگیری ماشین نیز ممکن است به‌دلیل سوگیری داده‌ها یا پیچیدگی ساختاری، تصمیم‌هایی ارائه دهند که نیازمند ارزیابی انتقادی و نظارت انسانی هستند (Van Dijk *et al*., 2021; Lam and Mutwil, 2024).

ازاین‌رو، توسعه و به‌کارگیری مسئولانه‌ی این فناوری‌ها مستلزم ایجاد چارچوب‌های اخلاقی روشن، ارزیابی مستمر ریسک و مشارکت فعال میان دانشمندان، سیاست‌گذاران و جامعه است؛ رویکردی که می‌تواند پذیرش اجتماعی فناوری‌های زیست‌فناورانه را تقویت کند (Singh *et al*., 2018; Kumar *et al*., 2021).

۴. چشم‌اندازهای آتی و جمع‌بندی بخش چالش‌ها

ترکیب داده‌های چندلایه‌ی اُمیکس، مدل‌های یادگیری ماشین و ابزارهای پیشرفته‌ی زیست‌فناوری، چشم‌انداز جدیدی برای اصلاح گیاهان مقاوم به تنش ترسیم کرده است. این همگرایی فناوری‌ها امکان گذار از اصلاح ژنتیکی مبتنی بر آزمون‌وخطا به اصلاح دقیق، پیش‌بینانه و سیستم‌محور را فراهم می‌سازد (Libbrecht and Noble, 2015; Van Dijk *et al*., 2021).

با توجه به پیشرفت‌های اخیر در حوزه‌ی هوش مصنوعی، انتظار می‌رود ادغام ابزارهای هوشمند با زیست‌فناوری گیاهی، منجر به توسعه‌ی راهبردهای

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

ایمن‌تر و هدفمندتر در کشاورزی آینده است (Kazemi Darsanaki, 2024).

با این حال، تحقق کامل این چشم‌انداز به پیشرفت هم‌زمان در سه حوزهٔ کلیدی توسعهٔ مدل‌های یادگیری ماشین تفسیرپذیر و تعمیم‌پذیر، بهبود کیفیت و تنوع داده‌های زیستی و ایجاد چارچوب‌های قانونی و اخلاقی شفاف وابسته است. توجه هم‌زمان به این عوامل می‌تواند نقش یادگیری ماشین را به‌عنوان ابزاری تصمیم‌یار و کم‌ریسک در زیست‌فناوری گیاهی تقویت کرده و مسیر کاربرد عملی آن را در کشاورزی پایدار هموار سازد (Kumar *et al*., 2021; Novielli *et al*., 2024).

نتیجه گیری

تحمل تنش‌های زیستی و غیرزیستی یکی از پیچیده‌ترین و درعین‌حال حیاتی‌ترین چالش‌ها در بهبود پایداری تولیدات کشاورزی و تضمین امنیت غذایی در شرایط تغییرات اقلیمی محسوب می‌شود. شواهد مرورشده در این مقاله نشان می‌دهد که پاسخ گیاهان به تنش‌ها ماهیتی چندلایه، شبکه‌ای و وابسته به زمینه دارد و نمی‌توان آن را صرفاً با تمرکز بر ژن‌ها یا مسیرهای منفرد به‌طور مؤثر مدیریت کرد (Raza *et al*., 2021; Zhou *et al*., 2022).

شواهد تجربی نشان می‌دهد که بهره‌گیری هدفمند از میکروارگانیسم‌های مفید، در کنار رویکردهای داده‌محور، می‌تواند به توسعه‌ی راهبردهای کم‌ریسک و پایدار در بهبود تحمل تنش گیاهان منجر شود (Abbasi *et al*., 2020a).

در این چارچوب، همگرایی فناوری‌های اُمیکس با الگوریتم‌های یادگیری ماشین، امکان تحلیل یکپارچه‌ی داده‌های پرحجم، شناسایی الگوهای زیستی معنادار و اولویت‌بندی ژن‌ها و مسیرهای مولکولی کلیدی را فراهم می‌سازد. این رویکرد داده‌محور، به‌ویژه در ترکیب با ابزارهای زیست‌فناوری گیاهی نظیر تراریختی، ویرایش ژنومی مبتنی بر CRISPR/Cas9 و فناوری RNAi، مسیر گذار از اصلاح ژن‌محور به اصلاح سیستم‌محور را هموار کرده و ظرفیت توسعه‌ی گیاهان مقاوم به تنش را به‌طور قابل‌توجهی افزایش می‌دهد (Van Dijk *et al*., 2021; Hesami *et al*., 2020).

مطالعات موردی بررسی‌شده در محصولات زراعی بااهمیت نشان می‌دهند که کاربرد یادگیری ماشین نه‌تنها دقت شناسایی اهداف اصلاحی را بهبود می‌بخشد، بلکه با کاهش عدم قطعیت در تصمیم‌گیری، احتمال بروز اثرات ناخواسته‌ی فیزیولوژیک و اکولوژیک را نیز کاهش می‌دهد. این ویژگی، یادگیری ماشین را به ابزاری تصمیم‌یار در طراحی راهبردهای اصلاح ژنتیکی کم‌ریسک‌تر و سازگار با اصول ایمنی زیستی تبدیل می‌کند (Washburn *et al*., 2020; Kumar *et al*., 2021).

با وجود این ظرفیت‌ها، بهره‌برداری مؤثر از یادگیری ماشین در زیست‌فناوری گیاهی مستلزم توجه جدی به چالش‌هایی نظیر کیفیت و سوگیری داده‌ها، خطر بیش‌برازش مدل‌ها، محدودیت‌های زیرساختی و ملاحظات قانونی و اخلاقی مرتبط با گیاهان تراریخته است. توسعه‌ی مدل‌های تفسیرپذیر، استفاده از داده‌های چندمحیطی و اعتبارسنجی تجربی نتایج، از الزامات اساسی برای افزایش قابلیت تعمیم و اعتمادپذیری این رویکردها به‌شمار می‌روند (Libbrecht and Noble, 2015; Van Dijk *et al*., 2021; Novielli *et al*., 2024).

در نهایت، این مقاله نشان می‌دهد که یادگیری ماشین، زمانی که در قالب چارچوب‌های یکپارچه و مسئولانه به‌کار گرفته شود، می‌تواند نقشی کلیدی در کاهش ریسک‌های ایمنی زیستی، بهینه‌سازی فرآیند اصلاح گیاهان و تسریع توسعه‌ی گیاهان تراریخته مقاوم به تنش ایفا کند. آینده‌ی اصلاح

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

گیاهان زراعی نه در اتکای به یک فناوری منفرد، بلکه در هم‌افزایی داده‌های چندلایه، مدل‌سازی هوشمند، اعتبارسنجی زیستی و سیاست‌گذاری آگاهانه نهفته است؛ رویکردی که می‌تواند مسیر دستیابی به کشاورزی پایدار و امنیت غذایی در مواجهه با تغییرات اقلیمی را هموار سازد (Kumar *et al.*, 2021; Zhou *et al.*, 2022).

منابع

Abbasi S, Sadeghi A, Safaie N. (2020a) Streptomyces alleviate drought stress in tomato plants and modulate the expression of transcription factors ERF1 and WRKY70 genes. ***Scientia Horticulturae*** 265. *8*(10): 109206. https://doi.org/10.1016/j.scienta.2020.109206.

Abbasi S, Sadeghi A, Safaei N. (2020b) Biological control of cucumber damping-off disease by siderophore- and cellulase-producing Streptomyces strains under unfavorable conditions. ***Journal of Microorganism Biology***. *9*(33): 1–13. https://doi.org/10.22108/BJM.2020.119404.1228.

Abbasi S, Alipour Kafi S, Karimi E, Sadeghi A. (2022) Streptomyces consortium improved quality attributes of bell pepper fruits, induced plant defense priming, and changed microbial communities of rhizosphere under commercial greenhouse conditions. ***Rhizosphere***. *23*: 100570. https://doi.org/10.1016/j.rhisph.2022.100570.

Chang L, Jin X, Rao Y, Zhang X. (2024) Predicting abiotic stress-responsive miRNA in plants based on multi-source features fusion and graph neural network. ***Plant Methods***. *20*: 33. https://doi.org/10.1186/s13007-024-01158-7.

Esmaeil Zadeh NS, Sadeghi A, Moradi P. (2018) Streptomyces strains alleviate water stress and increase peppermint yield and essential oils. ***Plant and Soil***. *431*(1–2): 1–17. https://doi.org/10.1007/s11104-018-3862-8.

Food and Agriculture Organization of the United Nations (FAO) & Water for Food Institute, Daugherty Water for Food Global Institute (DWFI). (2015) Yield gap analysis of field crops: Methods and case studies (FAO Water Reports No. 41). Rome.

Gill M, Anderson R, Hu H, Bennamoun M, Petereit J, Valliyodan B, Nguyen HT, Batley J, Bayer PE. (2022) Machine learning models outperform deep learning models, provide interpretation and facilitate feature selection for soybean trait prediction. ***BMC Plant Biology***. *22*(1): 180. https://doi.org/10.1186/s12870-022-03559-z.

Gopalakrishnan S, Srinivas V, Alekhya G, Prakash B. (2014) Plant growth-promoting traits of Streptomyces with biocontrol potential against plant pathogens. ***Applied Soil Ecology***. *80*: 15–23. https://doi.org/10.1016/j.apsoil.2014.03.008.

Hesami M, Jones AMP. (2020) Application of artificial intelligence models and optimization algorithms in plant cell and tissue culture. ***Applied Microbiology and Biotechnology***. *104*: 9449–9485. https://doi.org/10.1007/s00253-020-10888-2.

Hesami M, Alizadeh M, Jones AMP, Torkamaneh D. (2022) Machine learning: Its challenges and opportunities in plant system biology. ***Applied Microbiology and Biotechnology***. *106*: 3507–3530. https://doi.org/10.1007/s00253-022-11963-6.

Jamalirad S, Azimi MR, Ghaffari MR. (2024) Correlation network analysis of genes responding to salinity stress in *Suaeda salsa*. ***Journal of Biosafety***. *17*(1): 78-92. https://dorl.net/dor/20.1001.1.27170632.1403.17.1.5.3. [In Persian]

Kazemi Darsanaki R. (2024) A review on innovative applications of artificial intelligence in drug discovery, vaccine development, and therapy. ***Journal of Biosafety***. *17*(1): 59–77. https://dorl.net/dor/20.1001.1.27170632.1403.17.1.5.3. [In Persian]

Kumar A, Singh S, Gaurav AK, Srivastava S, Verma JP. (2021) Plant growth-promoting bacteria: Biological tools for the mitigation of salt stress in plants. ***Frontiers in Microbiology***. *12*: 1–17.

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

https://doi.org/10.3389/fmicb.2021.718398.

Lam HM, Mutwil M. (2024) Large language models in plant biology: Opportunities and challenges. ***Trends in Plant Science***. *29*(11): 1254–1266. https://doi.org/10.1016/j.tplants.2024.06.002.

Li Y, Ren S, Li J, Lee J, Wan J, Gan X. (2025) MeNet: A mixed-effect deep neural network for multi-environment genomic prediction of agronomic traits. ***Plant Communications***. 101620. https://doi.org/10.1016/j.xplc.2025.101620.

Libbrecht MW, Noble WS. (2015) Machine learning applications in genetics and genomics. ***Nature Reviews Genetics***. *16*(6): 321–332. https://doi.org/10.1038/nrg3920.

Liu Y, Zhang H, Chen X, Wang L, Li J. (2024) Identification of drought-responsive genes in rice using multilayer network integration and random walk algorithm. ***BMC Plant Biology***. *24*(1): 112. https://doi.org/10.1186/s12870-024-04123-7.

Novielli P, Romano D, Pavan S, Losciale P, Stellacci AM, Diacono D, Bellotti R, Tangaro S. (2024) Explainable artificial intelligence for genotype-to-phenotype prediction in plant breeding: A case study with a dataset from an almond germplasm collection. ***Frontiers in Plant Science***. *15*: 1434229. https://doi.org/10.3389/fpls.2024.1434229.

Raza A, Razzaq A, Mehmood SS, Hussain MA, Wei S, He H, Zaman QU, Xuekun Z, Yong C, Hasanuzzman M. (2021) Omics approaches in abiotic stress tolerance. ***Critical Reviews in Biotechnology***. *44*: 767-779 https://doi.org/10.1080/07388551.2021.1898332.

Raza A, Gangurde SS, Sandhu KS, Lv Y. (2024) Omics-assisted crop improvement under abiotic stress conditions. ***Plant Stress***. *14*: 100626. https://doi.org/10.1016/j.stress.2024.100626.

Singh AK, Ganapathysubramanian B, Sarkar S, Singh A. (2018) Deep learning for plant stress phenotyping: Trends and future perspectives. ***Trends in Plant Science***. *23*(10): 883–898. https://doi.org/10.1016/j.tplants.2018.07.004.

Shrestha S, Ojha G, Sharma G, Mainali R, Galvin L. (2024) Automated grassweed detection in wheat cropping system: Current techniques and future scope. ***Journal of Precision Agriculture***. *1*(1): 19-37. https://doi.org/10.14302/issn.2998-1506.jpa-24- 5058.

Tsega A, Mullualem D. (2026) Machine learning for multi-omics data integration in crop improvement: A systematic review. ***BMC Bioinformatics***. *27*: 81. https://doi.org/10.1186/s12859-026-06438-8.

Ubbens JR, Stavness I. (2017) Deep plant phenomics: A deep learning platform for complex plant phenotyping tasks. ***Frontiers in Plant Science***. *8*: 1190. https://doi.org/10.3389/fpls.2017.01190.

Van Dijk ADJ, Kootstra G, Kruijer W, de Ridder D. (2021) Machine learning in plant science and plant breeding. ***Science***. *24*(1): 101890. https://doi.org/10.1016/j.isci.2020.101890.

Vurukonda SSKP, Vardharajula S, Shrivastava M, Ali S. (2016) Enhancement of drought stress tolerance in crops by plant growth promoting rhizobacteria. ***Microbiological Research***. *184*: 13–24. https://doi.org/10.1016/j.micres.2015.12.003.

Washburn JD, Burch MB, Valdes Franco JA. (2020) Predictive breeding for maize using machine learning and crop simulation models. ***The Plant Genome***. *13*(3): e20042. https://doi.org/10.1002/tpg2.20042.

Yan J, Wang X. (2023) Machine learning bridges omics sciences and plant breeding. ***Trends in Plant Science***. *28*(2): 199–210. https://doi.org/10.1016/j.tplants.2022.08.018.

Zeng D, Li X, Huang J, Chen G. (2021) OSSAPK7 regulates salt stress tolerance in rice through modulation of ABA signaling and ion homeostasis. ***Plant Physiology and Biochemistry***. *166*: 678-689. https://doi.org/10.1016/j.plaphy.2021.06.018.

Zhou R, Jiang F, Niu L, Song X, Yu L, Yang Y, Wu Z. (2022) Increase crop resilience to heat stress using omic strategies. ***Frontiers in Plant Science***. *13*: 814599. https://doi.org/10.3389/fpspl.2022.814599.

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]

Integration of Machine Learning Algorithms in Identifying Genes and Molecular Pathways Related to Stresse Resistance in Transgenic Plants

Narjessadat Mousavimadani[1,2], Akram Sadeghi[1*], Reza Sharafi[1]

1. Department of Microbial Biotechnology, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran
2. Department of Biology, CT.C, Islamic Azad University, Tehran, Iran

aksadeghi@abrii.ac.ir

Abstract

Throughout their life cycle, plants are exposed to a wide range of abiotic stresses, including salinity, drought, water deficit, temperature extremes, oxidative stress, heavy metal toxicity, as well as biotic stresses such as various pests and diseases. These stresses are major factors hindering crops from reaching their genetic potential and achieving optimal performance. This review article first examines the role of these stresses in limiting plant performance and explores biotechnological strategies for managing them. It then discusses the application of machine learning (ML) algorithms in identifying and prioritizing genes and molecular pathways associated with stress resistance, alongside both traditional and modern agricultural biotechnology tools. Artificial intelligence (AI) and ML, through the analysis of multilayered genomic, transcriptomic, and other omics datasets, enable the modeling of signaling and metabolic networks, the simulation of plant performance under stress conditions, and the interpretation of genetic and epigenetic interactions. This approach ultimately enhances research efficiency and reduces costs and time requirements. The article further presents ML algorithms such as Random Forest (RF), Support Vector Machines (SVM), and deep learning (DL) models for identifying key resistance factors and supporting breeding decisions. Finally, it reviews case studies on crops like rice, corn, and Arabidopsis, and discusses associated challenges and limitations, as well as biosafety, legal, and economic considerations.

Keywords: Omics, Biotecchnology, Transgenic, Plant Stresses, Machine Learning

[Downloaded from journalofbiosafety.ir on 2026-11-25]

[DOR: 20.1001.1.27170632.1404.18.1.2.7]