گیاه و زیست فناوری ایران
Iranian Journal of Plant & Biotechnology (IJPB)
واحد گرمسار

نوع مقاله: پژوهشی
DOI:10.71829/BIOLOGY-2025-1205204

بررسی مکانیسم‌های متابولومیکس دفاعی گیاهان مبتنی بر هوش مصنوعی

مژگان زنگنه[۱] و محمدرضا صالحی‌سلمی (نویسنده مسئول)[۲*]

۱- دکتری، گروه علوم و مهندسی باغبانی، دانشکده کشاورزی، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان، ملاثانی، ایران، m.zangeneh@asnrukh.ac.ir

۲- دانشیار، گروه علوم و مهندسی باغبانی، دانشکده کشاورزی، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان، ملاثانی، ایران، Salehi@asnrukh.ac.ir

تاریخ دریافت: اردیبهشت ۱۴۰۴ تاریخ پذیرش: تیر ۱۴۰۴

Investigating Plant Defense Metabolomics Mechanisms Based on Artificial Intelligence

Mozhgan Zangeneh[1] and Mohamadreza Salehi Salmi (Corresponding author)[2*]

1- Ph.D, Department of Horticultural Sciences, Faculty of Agriculture, Agricultural Sciences and Natural Resource University of Khuzestan, Mulathani, Iran, m.zangeneh@asnrukh.ac.ir

2- Associate Professor, Department of Horticultural Sciences, Faculty of Agriculture, Agricultural Sciences and Natural Resource University of Khuzestan, Mulathani, Iran, Salehi@asnrukh.ac.ir

Received: May 2025 Accepted: July 2025

چکیده

متابولومیکس مبتنی بر هوش مصنوعی به‌عنوان یک رویکرد تحول‌ساز، درک بی‌سابقه‌ای از پاسخ‌های متابولیک گیاهان به تنش‌های زیستی و غیرزیستی ارائه می‌دهد. این مقاله، روندهای فعلی و آینده‌ی این فناوری را بررسی کرده و نقش آن را در رمزگشایی شبکه‌های متابولیک، شناسایی زیست‌نشانگرهای مرتبط با تنش و کشف الگوهای پنهان در داده‌های اُمیکس تحلیل می‌کند. در مقایسه با روش‌های سنتی که با چالش‌هایی مانند یکپارچه‌سازی داده‌ها و تفسیر نتایج روبه‌رو هستند، الگوریتم‌های یادگیری ماشین و یادگیری عمیق، امکان تحلیل سریع و دقیق داده‌های پیچیده را فراهم کرده‌اند. این فناوری‌ها نه‌تنها کشف متابولیت‌های کلیدی را که در مکانیسم‌های دفاعی گیاهان نقش دارند تسهیل می‌کنند، بلکه پیش‌بینی تحمل تنش و توسعه‌ی گونه‌های مقاوم را نیز بهبود بخشیده‌اند. یکپارچه‌سازی داده‌های چنداُمیکسی (ژنومیکس، ترانسکریپتومیکس و پروتئومیکس) با متابولومیکس، از طریق پلتفرم‌های هوش مصنوعی، درک جامع‌تری از واکنش گیاهان به تنش‌ها ارائه می‌دهد. با این حال، چالش‌هایی مانند استانداردسازی داده‌ها، تفسیرپذیری مدل‌ها و نیاز به منابع محاسباتی قدرتمند هنوز باقی است. این مقاله همچنین پتانسیل هوش مصنوعی را در بهینه‌سازی کشاورزی، افزایش مقاومت محصولات و تضمین امنیت غذایی در شرایط تغییرات اقلیمی بررسی می‌کند. با رفع محدودیت‌های کنونی و بهره‌گیری از فناوری‌های نوین، این حوزه می‌تواند انقلابی در علوم گیاهی و مقابله با چالش‌های جهانی غذا ایجاد کند.

کلمات کلیدی: تحلیل داده‌ها، تنش، چنداُمیکسی، کشاورزی پایدار، محصول مقاوم

فصلنامه گیاه و زیست فناوری ایران
بهار ۱۴۰۴، دوره ۲۰، شماره ۱، صص ۵۲-۳۸

Abstract

Artificial intelligence-powered metabolomics has emerged as a transformative approach, providing unprecedented insights into plant metabolic responses to biotic and abiotic stresses. This review examines current trends and future directions of AI-based metabolomics techniques, analyzing their role in deciphering metabolic networks, identifying stress-responsive biomarkers, and uncovering hidden patterns within omics datasets. Compared to traditional methods that face challenges in data integration and result interpretation, machine learning and deep learning algorithms enable rapid and accurate analysis of complex datasets. These advanced technologies not only facilitate the discovery of key metabolites involved in plant defense mechanisms but also enhance stress tolerance prediction and the development of resistant crop varieties. The integration of multi-omics data (genomics, transcriptomics, and proteomics) with metabolomics through AI-driven platforms offers a more comprehensive understanding of plant stress responses. However, challenges such as data standardization, model interpretability, and the need for high-performance computational resources remain unresolved. This paper also explores the potential of AI in optimizing agricultural practices, improving crop resilience, and ensuring food security under climate change conditions. By addressing current limitations and leveraging emerging technologies, this field holds revolutionary promise for plant sciences and global food security challenges.

Keywords: Crop resilience, Data analysis, Multi-omics, Stress, Sustainable agriculture

Iranian Journal of Plant & Biotechnology
Spring 2025, Vol 20, No 1, Pp 38-52

مقدمه و کلیات

دفاع گیاهان در برابر تنش‌های زیستی (عوامل زنده مانند عوامل بیماری‌زا و آفات) و غیرزیستی (عوامل محیطی مانند خشکی، شوری و دماهای شدید) از اهمیت بالایی در تضمین امنیت غذایی جهانی برخوردار است. با رشد مداوم جمعیت جهان، تقاضا برای محصولات کشاورزی در حال افزایش است و این امر لزوم توسعه راهبردهای کارآمد برای محافظت از گیاهان را بیش از پیش آشکار می‌سازد. توانایی محافظت از محصولات در برابر بیماری‌ها، آفات و شرایط نامساعد محیطی برای حفظ بهره‌وری کشاورزی و تأمین امنیت غذایی جهانی ضروری است. تأثیر قابل‌توجه این تنش‌ها با خسارات اقتصادی گسترده، از جمله کاهش چند میلیارد دلاری در عملکرد محصولات تأیید می‌شود (FAO, 2017). تنش‌های زیستی همواره تهدیدی برای سلامت محصولات کشاورزی محسوب می‌شوند. عوامل بیماری‌زا، از جمله قارچ‌ها، باکتری‌ها و ویروس‌ها، می‌توانند جمعیت کامل گیاهان را نابود کنند و منجر به خسارات اقتصادی قابل توجه و کمبود مواد غذایی شوند. به‌طور مشابه، آفات (از حشرات تا نماتدها) توانایی نابودی محصولات را دارند و باعث کاهش عملکرد و کیفیت آن‌ها می‌شوند. محصولات غذایی در سراسر جهان با کاهش عملکرد قابل توجهی به دلیل بیماری‌های میکروبی و آلودگی آفات مواجه هستند. این خسارات قابل توجه است؛ به‌طوری که برنج با کاهش متوسط ۳۰/۳٪، ذرت با ۲۲/٦٪، گندم با ۲۱/۵٪، سویا با ۲۱/٤٪ و سیب‌زمینی با ۱۷/۲٪ مواجه هستند (Savary *et al*., 2019). بیماری‌های گیاهی می‌توانند به‌ویژه ویرانگر باشند و در برخی مناطق منجر به کاهش عملکرد تا ۵۰٪ شوند، که این امر بیشتر کشاورزان خرده‌پا را تحت‌تأثیر قرار داده و چالش‌های اقتصادی قابل‌توجهی ایجاد می‌کند. علاوه بر این، بیماری‌های گیاهی بر تنوع گونه‌ها تأثیر منفی گذاشته، هزینه‌های مرتبط با اقدامات کنترل بیماری را افزایش می‌دهند و حتی پیامدهایی برای سلامت انسان دارند (Ristaino *et al*., 2021). ظهور بیماری‌های جدید گیاهی و شیوع آفات، پیامدهای اقتصادی قابل‌توجهی برای کشاورزی دارد و تهدیدی برای امنیت غذایی، ثبات ملی و سلامت عمومی محسوب می‌شود (Anderson *et al*., 2004). انتظار می‌رود در سال‌های آینده، تغییرات توزیع عوامل بیماری‌زا به دلیل تغییرات اقلیمی و افزایش تجارت جهانی، منجر به شیوع بیشتر و شدت بالاتر بیماری‌های نوظهور گیاهی شود (Bebber *et al*., 2013). نمونه قابل توجه اخیر، شیوع بیماری زنگ قهوه ناشی از قارچ *Hemileia vastatrix* در آمریکای مرکزی است که منجر به خسارات عمده به محصول و بحران‌های اقتصادی شد (Avelino *et al*., 2015). تنش‌های غیرزیستی شامل عوامل غیرزنده‌ای هستند که رشد و بقای گیاهان را به چالش می‌کشند (Zhu, 2016). این عوامل شامل خشکی‌های طولانی‌مدت، دماهای زیاد و کم، شوری خاک و آلودگی به فلزات سنگین هستند (Raza, 2021). اثرات تنش‌های غیرزیستی اغلب ظریف و نامحسوس است و بر عملکرد محصولات، محتوای غذایی و سلامت کلی گیاهان تأثیر می‌گذارد (Zhang, 2020). تلاش‌ها برای مقابله با تنش‌های زیستی و غیرزیستی به‌طور سنتی بر ترکیبی از روش‌ها، از جمله اصلاح نژاد

متعارف، تیمارهای شیمیایی و روش‌های زراعی متکی بوده است (Varshney, 2018). با این حال، این رویکردها اغلب واکنشی هستند و ممکن است محافظت مؤثری به‌ویژه در مواجهه با عوامل بیماری‌زای نوظهور و شرایط محیطی که به سرعت در حال تغییر هستند، ارائه ندهند (Mittler and Blumwald, 2015). اخیراً، فناوری‌های پیشرفته اُمیکس، انقلابی در بررسی مکانیسم‌های مولکولی تنش در گیاهان ایجاد کرده‌اند (Shen *et al*., 2022; Weckwerth, 2020). این روش‌ها اطلاعات گسترده‌ای ارائه می‌دهند و شبکه‌های پیچیده‌ای از ژن‌ها، پروتئین‌ها و متابولیت‌ها را در طول دفاع گیاهان در برابر تنش‌های زیستی آشکار می‌سازند (Pang, 2021). با این حال، حجم عظیم داده‌های تولید شده توسط این فناوری‌های اُمیکس، چالش‌های قابل‌توجهی در زمینه تحلیل و تفسیر ایجاد می‌کند و نیاز به توسعه ابزارهای محاسباتی بسیار مؤثر را ضروری می‌سازد (Libbrecht and Noble, 2015). هوش مصنوعی به عنوان ابزاری قدرتمند برای تحلیل مجموعه‌داده‌های گسترده اُمیکس و درک مکانیسم‌های پیچیده پاسخ گیاهان به تنش کمک کرده است (Singh, 2022). این تکنیک‌ها درک عمیق‌تری از جنبه‌های ژنتیکی، مولکولی، فیزیولوژیکی و فنوتیپی دفاع گیاهان ارائه می‌دهند و امکان توسعه راهبردهای نوآورانه برای افزایش مقاومت محصولات و کاهش خسارات ناشی از تنش را فراهم می‌کنند (Yang, 2021). در تحقیقات سنتی دفاع گیاهی، شبکه‌های پیچیده ژن‌ها، پروتئین‌ها و متابولیت‌های درگیر در پاسخ به تنش، به دلیل حجم و پیچیدگی بالای داده‌های زیستی، چالش‌هایی ایجاد می‌کرد (Van Wallendael *et al*., 2019). ظهور تکنیک‌های اُمیکس با کمک هوش مصنوعی، راه‌حل‌های نوآورانه‌ای برای مواجهه و تفسیر این حجم عظیم داده ارائه کرده است (Kumar., 2022). این تکنیک‌ها شامل روش‌های متنوع هوش مصنوعی هستند که به‌طور کارآمد روابط پیچیده بین اجزای مولکولی و پاسخ‌های تنش را آشکار و مدل‌سازی می‌کنند (Wang, 2023). در این مقاله مطالعات متابولومیکس مختلف شامل عوامل تنش‌زای گوناگون در گیاهان ارائه می‌شود. با استفاده از هوش مصنوعی در متابولومیکس گیاهی، به بررسی پیشرفت‌های نوین در زمینه دفاع گیاهان در برابر تنش‌های زیستی و غیرزیستی و چالش‌های پیش‌رو پرداخته می‌شود. این مقاله به مزایای هوش مصنوعی نسبت به روش‌های سنتی را بررسی کرده، چالش‌های هوش مصنوعی در متابولومیکس گیاهی و جهت‌گیری‌های آینده در تحقیقات دفاع گیاهی را مورد بحث قرار می‌دهد. همچنین، پتانسیل این فناوری‌ها برای توسعه روش‌های کشاورزی پایدار که محافظت از محصولات، تحمل تنش و امنیت غذایی جهانی را بهبود می‌بخشند، برجسته می‌شود.

فرآیند پژوهش

این مقاله، وضعیت کنونی دانش کاربرد تکنیک‌های متابولومیکس با کمک هوش مصنوعی در درک مکانیسم‌های دفاعی گیاهان را بررسی و تحلیل می‌کند. روش‌شناسی این مرور شامل یک رویکرد سیستماتیک و جامع برای شناسایی، تحلیل و جمع‌بندی پژوهش‌های مرتبط، روندها و جهت‌گیری‌های آینده در این حوزه است. در این بررسی یک جستجوی

سیستماتیک در مقالات علمی پژهشی، مقالات مروری و کتاب منتشرشده بین سال‌های ۲۰۰۰ تا ۲۰۲٤ انجام شد. این جستجو در پایگاه‌های داده الکترونیکی PubMed، Web of Science، Scopus، Google Scholar و IEEE Xplore انجام شد. کلمات کلیدی استفاده شده در پایگاه‌ها شامل هوش مصنوعی، یادگیری ماشین، یادگیری عمیق، اُمیکس، متابولومیکس، دفاع گیاهی، برهمکنش گیاه –پاتوژن، بیولوژی، بیوانفورماتیک و کشاورزی دقیق بود. معیارهای ورود شامل پژوهش‌هایی است که بر تکنیک‌های متابولومیکس با کمک هوش مصنوعی در مکانیسم‌های دفاعی گیاهان تمرکز دارند، همچنین پژوهش‌هایی که به ادغام داده‌های چنداُمیکسی برای درک برهم‌کنش گیاه و پاتوژن پرداخته‌اند. علاوه بر این، مقالاتی که روندهای فعلی، چالش‌ها و جهت‌گیری‌های آینده در این حوزه را بررسی می‌کنند نیز در این مجموعه قرار می‌گیرند. در مقابل، معیارهای خروج شامل پژوهش‌هایی است که ارتباطی با دفاع گیاهی یا فناوری‌های متابولومیکس ندارند، همچنین مقالاتی که فاقد مؤلفه‌های هوش مصنوعی یا یادگیری ماشین بوده‌اند. علاوه بر این، انتشارات غیرانگلیسی و مقالاتی که فرآیند داوری را طی نکرده‌اند نیز از دامنه بررسی خارج می‌شوند. داده‌های استخراج‌شده به‌منظور شناسایی روندها، شکاف‌ها و فرصت‌ها در کاربرد تکنیک‌های اُمیکس با کمک هوش مصنوعی در دفاع گیاهی تحلیل شدند. یافته‌های کلیدی به‌منظور ارائه یک نمای جامع از وضعیت فعلی حوزه و پیشنهاد جهت‌گیری‌های پژوهشی آینده واکاوی شدند.

نتایج و بحث

متابولومیکس: مطالعه متابولیت‌ها در سیستم‌های زیستی، برای درک متابولوم گیاهان و آشکارسازی مکانیسم‌های تنظیمی تحت شرایط تنش، از اهمیت بالایی برخوردار است (Saito and Matsuda, 2010). این حوزه، که با توالی‌یابی نسل جدید (NGS) ادغام شده است، بینش‌هایی درباره پاسخ‌های مولکولی در محصولات کشاورزی ارائه می‌دهد و چشم‌اندازی گسترده‌تر از فرآیندهای بیوشیمیایی مؤثر بر عملکرد ژن‌ها فراهم می‌کند (Borrill, 2019). در دفاع گیاهان در برابر تنش‌ها و پاتوژن‌ها، متابولیت‌ها نقش حیاتی ایفا می‌کنند (Piasecka, 2019). به‌عنوان مثال، متابولیت‌هایی که از طریق کروماتوگرافی گازی–طیف‌سنجی جرمی (GC-MS) به‌عنوان نشانگرهای زیستی در ارقام برنج مواجه‌شده با پاتوژن GMB1 شناسایی شده‌اند، اهمیت این متابولیت‌ها را نشان می‌دهند (Chen, 2021). استراتژی‌های مشابه، تجمع متابولیت‌ها را در پاسخ به سایر پاتوژن‌ها در برنج و جو نشان داده است که بیانگر اهمیت متابولومیکس در درک پاسخ گیاهان به تنش‌های زیستی است (Kumar, 2022). در گندم نیز حضور متابولیت‌های فنیل‌پروپانوئید و فنلی در پاسخ به تنش‌های زیستی مشاهده شده است (Hamzehzarghani, 2018). متابولومیکس به‌ویژه در سیستم‌های گیاهی به دلیل تولید فراوان متابولیت‌های ثانویه از اهمیت ویژه‌ای برخوردار است (Fang and Luo, 2019). متابولیت‌های ثانویه مانند پلی‌آمین‌ها که در برنج تحت تنش خشکی شناسایی شده‌اند، اهمیت متابولومیکس محیطی را در درک پاسخ گیاهان به تنش‌های

غیرزیستی برجسته می‌کنند (Do, 2020). تکنیک‌های مختلف متابولومیکس، از جمله LC/GC-MS، GC/EI-TOF-MS، HPLC و NMR، به‌طور گسترده در محصولاتی مانند برنج، گوجه‌فرنگی، ذرت و سویا به‌کار رفته‌اند و بینش‌های ارزشمندی درباره پاسخ این محصولات به شرایط تنش زیستی و غیرزیستی ارائه داده‌اند (Putri, 2021).

مبانی تکنیک‌های هوش مصنوعی: هوش مصنوعی شامل ایجاد سیستم‌های کامپیوتری برای انجام وظایفی است که با هوش انسان مرتبط هستند، مانند یادگیری، حل مسئله و تصمیم‌گیری (Russell and Norvig, 2021). یادگیری ماشین، به‌عنوان زیرمجموعه‌ای از هوش مصنوعی، بر توسعه الگوریتم‌ها و مدل‌های آماری تمرکز دارد که به کامپیوترها امکان انجام وظایف را بدون برنامه‌نویسی صریح می‌دهند (Goodfellow, 2016). یادگیری عمیق، به‌عنوان یک حوزه تخصصی در یادگیری ماشین، شامل آموزش شبکه‌های عصبی مصنوعی برای تقلید از ساختار مغز انسان است و از معماری‌های عمیق با چندین لایه جهت استخراج سلسله‌مراتبی ویژگی‌ها به‌صورت خودکار استفاده می‌کند (LeCun, 2015). مرحله اولیه یادگیری ماشین شامل جمع‌آوری داده‌ها است. انتخاب الگوریتم‌ها نیز بسیار حیاتی است (Bishop, 2006). الگوریتم‌های یادگیری ماشین را می‌توان به سه دسته نظارت‌شده، نیمه‌نظارت‌شده و بدون نظارت تقسیم کرد (Zhou, 2021). روش‌های نظارت‌شده، روابط بین عوامل ورودی و نتایج را بر اساس نمونه‌های آموزشی برقرار می‌کنند (Hastie, 2009). روش‌های بدون نظارت، که عمدتاً خوشه‌بندی هستند، الگوهای داده‌ها را بدون اتکا به نتایج شناخته‌شده شناسایی می‌کنند (Jain, 2010). الگوریتم‌های کاربردی شامل SVM (Cortes and Vapnik, 1995)، DT (Quinlan, 1986)، RF (Breiman, 2001)، ANN (Schmidhuber, 2015) و NB (Rish, 2001) برای یادگیری نظارت‌شده هستند، در حالی که k-means (Lloyd, 1982)، تحلیل مؤلفه‌های مستقل (Hyvärinen and Oja, 2000) و خوشه‌بندی سلسله‌مراتبی (Murtagh and Legendre, 2014) برای یادگیری بدون نظارت استفاده می‌شوند. یادگیری نیمه‌نظارت‌شده، داده‌های ورودی را که هم برچسب‌دار و هم بدون برچسب هستند، مدیریت می‌کند و نمونه‌هایی مانند الگوریتم انتشار برچسب را شامل می‌شود (Zhu and Goldberg, 2009).

استفاده از یادگیری ماشین به جای روش‌های سنتی جهت پردازش داده‌های متابولومیکس: یادگیری ماشین به‌طور فزاینده‌ای در تحلیل داده‌های متابولومیکس گیاهی نسبت به روش‌های سنتی ترجیح داده می‌شود، زیرا توانایی بالایی در مدیریت مجموعه‌های داده بزرگ و پیچیده دارد. فناوری‌های توالی‌یابی با توان عملیاتی بالا، حجم عظیمی از اطلاعات را فراهم کرده‌اند که به زیست‌شناسان امکان می‌دهد ارتباطات پیچیده را بررسی کنند، پاسخ‌های گیاهان به تنش‌ها را رمزگشایی کنند و پیچیدگی‌های پاسخ‌ها را آشکار سازند (Singh *et al*., 2016). با این حال، چالش‌هایی مانند ابعاد بالا، عدم قطعیت و وابستگی متقابل بین متغیرها در داده‌های متابولومیکس گیاهی سبب شده که مدل‌های آماری سنتی در مواجهه با این پیچیدگی‌ها با محدودیت‌هایی روبرو شوند

Altman and Krzywinski, 2018; Niazian and) Niedbała, 2020). یادگیری ماشین، به‌ویژه یادگیری عمیق، در غلبه بر این چالش‌ها کارآمد بوده و تحلیل‌های دقیقی از ویژگی‌های گیاهی تحت‌تأثیر برهمکنش متابولومیکس و محیط ارائه می‌دهد (Arsenovic *et al.*, 2019). الگوریتم‌های یادگیری ماشین بدون نظارت و نیمه‌نظارت‌شده در زیست‌شناسی سیستم‌های گیاهی به‌کار گرفته شده‌اند و تحلیل داده‌های بزرگ را بدون نیاز به آموزش تسهیل می‌کنند (Yan and Wang, 2022). کاربرد یادگیری ماشین در بهبود صفات زراعی گیاهان از طریق ادغام داده‌های اُمیکس بزرگ نیز مشهود است (Isewon *et al.*, 2022). مطالعات Farooq و همکاران (۲۰۲۲)، Isewon و همکاران (۲۰۲۲) و Silva و همکاران (۲۰۱۹) برتری روش‌های یادگیری ماشین، به‌ویژه مدل‌های گروهی مبتنی بر درخت تصمیم (Gokalp and Tasci, 2019) را در پیش‌بینی و تحلیل یکپارچه داده‌های اُمیکس گیاهی نشان می‌دهند. پتانسیل یادگیری ماشین در رمزگشایی برهمکنش‌های پیچیده در زیست‌شناسی مولکولی گیاهان مورد تأکید قرار گرفته است (Silva *et al.*, 2019). توانایی یادگیری ماشین در تحلیل چندمتغیره، امکان در نظر گرفتن همزمان تعداد زیادی متغیر را فراهم می‌کند و منجر به کشف نشانگرهای زیستی جدید و توسعه مدل‌های پیش‌بینانه می‌شود (Reel *et al.*, 2021). کاربرد آن در بهبود صفات زراعی در پژوهش‌های متابولومیکس گیاهی آشکار است، اگرچه چالش‌هایی در تحقق کامل پتانسیل ادغام داده‌های چنداُمیکسی وجود دارد و مقیاس‌پذیری یکی از موانع اصلی است (Noor *et al.*, 2019). پیچیدگی داده‌های اُمیکس با ابعاد بالا، نیازمند روش‌های پیشرفته برای انتخاب ویژگی‌ها و استخراج اطلاعات است. مزیت یادگیری ماشین در توانایی آن برای تشخیص و اولویت‌بندی مرتبط‌ترین ویژگی‌ها نهفته است، همان‌طور که Du و همکاران (۲۰۱۹) در تحلیل داده‌های مرتبط با پاسخ به تنش شوری در برنج نشان دادند. استفاده از روش‌های انتخاب ویژگی مبتنی بر یادگیری ماشین، از جمله تحلیل مؤلفه‌های اصلی (PCA) وLASSO، زیرماژول‌های مرتبط با صفات مشاهده‌شده را به‌طور مؤثر آشکار کرد. یادگیری ماشین در وظایف پیش‌بینی و طبقه‌بندی عملکرد برجسته‌ای دارد و به پژوهشگران امکان می‌دهد نتایج فنوتیپی را پیش‌بینی کنند و نشانگرهای زیستی بالقوه را شناسایی کنند. مقیاس‌پذیری آن برای داده‌های اُمیکس در مقیاس بزرگ حیاتی است و آن را از روش‌های سنتی که با چالش‌های محاسباتی مواجه هستند، متمایز می‌کند. در نتیجه، گرایش به سمت یادگیری ماشین در پژوهش‌های متابولومیکس گیاهی ناشی از نقاط قوت منحصربه‌فرد آن در مواجهه با پیچیدگی‌های داده‌ها، امکان مدل‌سازی پیش‌بینانه و تسهیل رویکرد اکتشافی در تحلیل داده‌ها است (Biabi *et al.*, 2019). تشخیص الگو یکی دیگر از نقاط قوت یادگیری ماشین است که امکان کشف الگوها و ارتباطات پیچیده در مجموعه داده‌های پیچیده را فراهم می‌کند. در پژوهش‌های متابولومیکس، که در آن کشف الگوهای عمیق ممکن است بینش‌های جدیدی درباره مکانیسم‌های زیستی ارائه دهد، این قابلیت بسیار ارزشمند است. علاوه بر این، یادگیری ماشین با ناهمگونی‌هایی که اغلب در مجموعه

داده‌های اُمیکس به دلیل تنوع زیستی و تفاوت‌های فنی مشاهده می‌شود، سازگار است. انعطاف‌پذیری و ماهیت آن، همراه با وابستگی کمتر به فرضیات درباره توزیع داده‌ها، آن را برای انواع داده‌ها و طرح‌های آزمایشی مختلف مناسب می‌سازد. ماهیت اکتشافی یادگیری ماشین، که تسهیل کشف الگوها و روابط پنهان را ممکن می‌سازد، در پژوهش‌های متابولومیکس بسیار حیاتی است. این جنبه به پژوهشگران امکان می‌دهد فرضیه‌ها را ایجاد کنند و مسیرهای جدیدی برای بررسی بیشتر شناسایی کنند (Biabi *et al.*, 2019). به‌طورکلی گرایش به سمت یادگیری ماشین در پژوهش‌های متابولومیکس گیاهی ناشی از نقاط قوت منحصربه‌فرد آن در مواجهه با پیچیدگی‌های داده‌های متابولومیکس، تطبیق با چندین متغیر، ادغام مجموعه داده‌های متنوع، ارائه قابلیت‌های مدل‌سازی پیش‌بینانه و طبقه‌بندی، و تسهیل رویکرد اکتشافی در تحلیل داده‌ها است.

کاربرد هوش مصنوعی در متابولومیکس گیاهی برای مقابله با تنش: تکنیک‌های متابولومیکس مبتنی بر هوش مصنوعی در تحقیقات مرتبط با دفاع گیاهی، رویکردی پیشرفته و نوین محسوب می‌شوند که در آنها فناوری‌های مولکولی پیشرفته با هوش مصنوعی ترکیب شده‌اند تا درک عمیق‌تری از مکانیسم‌های پاسخ گیاهان به تنش‌های محیطی فراهم شود. در حالی که روش‌های سنتی با چالش‌های فراوانی در تفسیر شبکه‌های پیچیده متابولیت‌ها مواجه بودند، روش‌های متابولومیکس تقویت‌شده با هوش مصنوعی، راهکارهای نوآورانه‌ای ارائه می‌دهند. این تکنیک‌ها قادرند به سرعت اجزای کلیدی در مسیرهای دفاعی گیاهان را شناسایی کنند، زیست‌نشانگرهای مرتبط را کشف نمایند و الگوهای پنهان را آشکار سازند، که این امر منجر به ارتقای درک ما از فرآیندهای دفاعی گیاهان می‌شود. یکپارچه‌سازی داده‌های حاصل از منابع چنداُمیکسی، امکان دستیابی به درکی جامع‌تر را فراهم می‌کند و با پیشرفت روزافزون فناوری‌های هوش مصنوعی، این روش‌ها پتانسیل بالایی برای توسعه محصولات کشاورزی مقاوم به تنش، بهینه‌سازی شیوه‌های کشاورزی و تضمین تولید پایدار مواد غذایی دارند (Arabnia and Tran, 2011; Yan and Wang, 2022). مطالعه‌ی Liu و همکاران (۲۰۱۷) پیشرفت قابل‌توجهی در زمینه متابولومیکس گیاهی، به‌ویژه در زمینه طبقه‌بندی تاکسونومیک و شناسایی روابط گونه–متابولیت، نشان می‌دهد. محققان با استفاده از یک الگوریتم خوشه‌بندی شبکه‌ای برای تحلیل داده‌های ناقص متابولیتی، نشان دادند که محتوای متابولیتی می‌تواند به عنوان یک نشانگر تاکسونومیک قوی عمل کند. این رویکرد به‌ویژه در مواردی که داده‌های مورفولوژیک یا ژنتیکی سنتی ناکافی یا در دسترس نیستند، بسیار ارزشمند است. استفاده از بردارهای باینری برای نمایش گیاهان، امکان نمایش ساده شده، اما مؤثر داده‌های متابولیتی پیچیده را فراهم کرد و کاربرد تکنیک‌های خوشه‌بندی سلسله‌مراتبی را تسهیل نمود. موفقیت این روش در خوشه‌بندی گیاهان براساس روابط تکاملی شناخته‌شده، اهمیت متابولیت‌های ثانویه در تکامل و سازگاری گیاهان را برجسته می‌سازد. علاوه بر این، یافته‌های این مطالعه در مورد قدرت پیش‌بینی محتوای متابولیتی برای خواص تغذیه‌ای و دارویی، راه‌های

جدیدی برای تحقیقات در زمینه کشف داروهای گیاهی باز می‌کند. شناسایی روابط گونه–متابولیتی که قبلاً ناشناخته بودند، نشان می‌دهد که متابولومیکس می‌تواند به عنوان مکملی برای رویکردهای سنتی در تاکسونومی و فارماکوگنوزی عمل کند و درک جامع‌تری از پایه‌های بیوشیمیایی فعالیت زیستی گیاهان ارائه دهد. تحقیقات آینده می‌توانند به ادغام داده‌های متابولومیک با داده‌های ژنومیک و ترانسکریپتومیک بپردازند تا طبقه‌بندی‌های تاکسونومیک را بهبود بخشند و دقت پیش‌بینی خواص گیاهان را افزایش دهند. همچنین، کاربرد الگوریتم‌های یادگیری ماشین، همان‌طور که در مطالعه Fürtauer (۲۰۱۸) نشان داده شد، می‌تواند به مطالعات متابولومیک گسترش یابد تا مسیرهای متابولیک کلیدی و زیست‌نشانگرهای مرتبط با صفات خاص گیاهی یا پاسخ‌های محیطی شناسایی شوند (Fürtauer, 2018; Alseekh, 2021). Fürtauer (۲۰۱۸) با بررسی تغییرات متابولیک و پروتئومیک ناشی از شرایط تنش غیرزیستی، به منابع موجود در مورد پاسخ‌های گیاهان به تنش کمک کرد. تمرکز آنها روی گیاهان وحشی آرابیدوپسیس و لاین‌های موتانت با نقص در متابولیسم ساکارز یا نشاسته، دیدگاه منحصر به فردی در مورد نقش متابولیسم کربوهیدرات‌ها در سازگاری با تنش ارائه داد (Fürtauer *et al*., 2018; Obata and Fernie, 2012). استفاده از یادگیری ماشین برای طبقه‌بندی لاین‌های موتانت تحت شرایط کنترل و تنش، رویکردی نوآورانه برای درک برهمکنش پیچیده بین تنظیم متابولیک و عوامل تنش‌زای محیطی است (Sweetlove, 2017). شناسایی یک ماژول اصلی متشکل از ۲۳ پروتئین که به طور قابل اعتماد شرایط ترکیبی دما و نور شدید را پیش‌بینی می‌کرد، قابل توجه است، زیرا چارچوبی مکانیکی برای مطالعه پاسخ‌های تنش در سطح مولکولی فراهم می‌کند (Fürtauer *et al*., 2018; Nakabayashi and Saito, 2021). ارتباط ۱۸ پروتئین در این ماژول اصلی با برهمکنش‌های پروتئین–پروتئین، اهمیت تغییرات پس از ترجمه و شبکه‌های پروتئینی در میانجی‌گری پاسخ‌های تنش را برجسته می‌کند. این یافته‌ها با مطالعات قبلی که نقش فسفوریلاسیون پروتئین، یوبیکویتیناسیون و سایر تغییرات را در تنظیم مسیرهای متابولیک تحت شرایط تنش نشان داده‌اند، همسو است (Nakagami *et al*., 2010; Vierstra, 2009). به عنوان مثال، ادغام داده‌های پروتئومیک و متابولومیک در مطالعه Fürtauer و همکاران (۲۰۱۸)، دیدگاهی جامع از پاسخ‌های گیاهان به تنش ارائه می‌دهد و نشان می‌دهد که چگونه تغییرات در سطح پروتئین به تنظیمات متابولیک تبدیل می‌شوند. این رویکرد زیست‌شناسی سیستم‌ها می‌تواند به سایر گونه‌های گیاهی و شرایط تنش گسترش یابد تا مکانیسم‌های سازگاری با تنش که در بین گونه‌ها مشترک یا خاص هستند، شناسایی شوند. دستاورد این مطالعات فراتر از تحقیقات پایه است و کاربردهای بالقوه‌ای در کشاورزی و بیوتکنولوژی دارد. به عنوان مثال، شناسایی متابولیت‌ها و پروتئین‌های کلیدی درگیر در پاسخ‌های تنش می‌تواند به توسعه انواع محصولات مقاوم به تنش از طریق اصلاح نژاد یا مهندسی ژنتیک کمک کند. به طور مشابه، مدل‌های پیش‌بینی توسعه‌یافته توسط Fürtauer و همکاران (۲۰۱۸) می‌توانند برای استفاده در کشاورزی دقیق

تطبیق داده شوند و امکان نظارت بر سطح تنش گیاهان در زمان واقعی و اطلاع‌رسانی تصمیمات مدیریتی برای بهینه‌سازی عملکرد محصول را فراهم کنند. علاوه بر این، ادغام داده‌های متابولومیک و پروتئومیک با سایر مجموعه داده‌های اونیکس، مانند ترانسکریپتومیک و اپی‌ژنومیک، می‌تواند درک جامع‌تری از پاسخ‌های گیاهان به تغییرات محیطی ارائه دهد و توسعه محصولات مقاوم به تغییرات آب‌وهوایی را تسهیل کند.

چالش‌ها: پیشرفت‌های سریع در تولید داده‌های زیستی و توسعه یادگیری ماشین، امکان‌های قابل توجهی برای رمزگشایی اطلاعات پیچیده زیستی فراهم کرده‌اند. با این حال، ادغام یادگیری ماشین در مطالعات مولکولی گیاهان با چالش‌های قابل‌توجهی مواجه است. رویکردهای یادگیری ماشین، مشابه روش‌های سنتی مولکولی گیاهی، به شدت وابسته به زمینه هستند که این موضوع اهمیت طراحی دقیق آزمایش‌ها را برجسته می‌کند. لازم به ذکر است که اگرچه یادگیری ماشین به دنبال ایجاد مدل‌های پیش‌بینانه است، هر الگوریتم یادگیری ماشین نقاط قوت و ضعف متمایزی دارد که بر کارایی پیش‌بینی تحت شرایط خاص تأثیر می‌گذارد. در نتیجه، یک مدل یادگیری ماشین که برای یک مجموعه داده طراحی شده است، ممکن است در تعمیم به مجموعه داده‌های دیگر به دلیل تفاوت‌های ذاتی زیستی و فنی با مشکل مواجه شود. وجود حجم عظیمی از مجموعه داده‌های اُمیکس، گنجینه‌ای از اطلاعات را فراهم می‌کند. با این حال، بخش قابل‌توجهی از این مجموعه‌های داده با ویژگی‌هایی مانند نویز و پراکندگی مشخص می‌شوند. این موضوع چالش بزرگی در شناسایی دقیق ویژگی‌های زیستی، به‌ویژه در هنگام ادغام منابع مختلف داده‌های اُمیکس ایجاد می‌کند (Joyce and Palsson, 2006). چالش مجموعه‌های داده نامتعادل، که در آن اندازه نمونه‌ها در دسته‌های مختلف متفاوت است، در یادگیری ماشین بسیار رایج است. پژوهشگران این چالش را از طریق استراتژی‌های بازنمونه‌گیری مانند نمونه‌گیری بیش از حد و نمونه‌گیری کمتر برطرف می‌کنند (Maimon and Rokach, 2005). به‌عنوان مثال، Li و همکاران (۲۰۲۱) از تکنیک نمونه‌گیری بیش از حد مصنوعی برای تقویت نمایندگی موارد اقلیت استفاده کردند که گامی کلیدی در شناسایی پروتئین‌های اثرگذار بود. وجود نویز و پراکندگی در این مجموعه‌های داده، عدم قطعیت و پیچیدگی ایجاد می‌کند و ممکن است شناسایی الگوها یا ویژگی‌های معنادار در داده‌های زیستی را با مشکل مواجه کند. پژوهشگران باید با چالش تشخیص سیگنال‌های زیستی واقعی از نویز پس‌زمینه مواجه باشند که این موضوع نیاز به رویکردهای تحلیلی قوی برای اطمینان از یافته‌ها را برجسته می‌سازد. علاوه بر این، مسئله بیش‌برازش (overfitting) به‌ویژه در حوزه یادگیری عمیق بسیار مطرح است. بیش‌برازش زمانی رخ می‌دهد که یک مدل بیش از حد با پیچیدگی‌های یک مجموعه داده خاص سازگار شود و در نتیجه در تعمیم به داده‌های جدید و نادیده با مشکل مواجه شود. این پدیده می‌تواند توانایی‌های پیش‌بینانه مدل را تضعیف کند و کاربردپذیری آن را در مشکلات واقعی محدود سازد. برای مقابله با این مشکل، تکنیک‌هایی مانند dropout به‌کار گرفته شده‌اند (Scholz *et al.*, 2004). Dropout

یک تکنیک تنظیم است که شامل غیرفعال کردن تصادفی زیرمجموعه‌ای از نورون‌ها در طول آموزش یک شبکه عصبی می‌شود (Srivastava, 2014). با این کار، dropout از وابستگی بیش از حد شبکه عصبی به ویژگی‌ها یا روابط خاص موجود در داده‌های آموزشی جلوگیری می‌کند و در نتیجه توانایی آن را برای تعمیم به داده‌های جدید و نادیده افزایش می‌دهد (Baldi and Sadowski, 2013). این تکنیک به‌عنوان یک محافظ در برابر بیش‌برازش عمل می‌کند و مدلی قوی‌تر و انعطاف‌پذیرتر ایجاد می‌کند (Wager, 2013). عوامل مختلفی از جمله پیش‌پردازش داده‌ها، پارامترهای تعریف‌شده توسط کاربر و دانش حوزه‌ای، تأثیر قابل‌توجهی بر اثربخشی مدل‌های یادگیری ماشین دارند (Domingos, 2012). متخصصان یادگیری ماشین نقش محوری در تصمیم‌گیری در طول فرآیند دارند که این موضوع اهمیت ادغام دانش پیشین و تخصص حوزه‌ای برای کشف الگوهای معنادار را برجسته می‌سازد (Provost and Fawcett, 2013). مواجهه با ویژگی‌های داده‌های بزرگ در مطالعات زیست‌شناسی سیستم‌های گیاهی، شامل حجم (volume)، تنوع (variety)، اعتبار (veracity)، ارزش (value) و سرعت (velocity)، چالش‌های خاص خود را دارد (Zhou, 2021). روش‌های یادگیری ماشین باید برای مدیریت داده‌های چندأمیکسی تطبیق یابند و بینش‌های منحصربه‌فرد هر لایه اُمیکس را در نظر بگیرند (Libbrecht and Noble, 2015). چالش‌ها شامل مواجهه با مسائل داده‌های با ابعاد بالا مانند پراکندگی، چندخطی بودن (multicollinearity) و بیش‌برازش است که نیازمند روش‌های سفارشی‌شده و همکاری‌های مشترک در ادغام داده‌ها است (Wold, 2018). تفسیر مدل‌های پیچیده، به‌ویژه در رویکردهای پیشرفته یادگیری ماشین مانند یادگیری عمیق، به دلیل ماهیت «جعبه سیاه» آن‌ها همچنان چالش‌برانگیز است (Samek, 2021). پژوهشگران اغلب درک اهمیت زیستی یک مدل پیش‌بینانه را نسبت به دقت آن در اولویت قرار می‌دهند که نیازمند پردازش دقیق و هم‌بستگی با دانش زیستی موجود است (Greener, 2022). علیرغم این چالش‌ها، مطالعات مورد بحث در این زمینه، نمونه‌های موفقی از کاربرد هوش مصنوعی در اُمیکس گیاهی هستند (Singh, 2022). برای بهره‌برداری کامل از پتانسیل هوش مصنوعی، الگوریتم‌های قوی و مقیاس‌پذیر که بینش‌های زیستی معنادار را آشکار می‌کنند، ضروری هستند (Eraslan, 2019). اطمینان از دقت و قابلیت اطمینان از طریق اعتبارسنجی تجربی، برای تبدیل یافته‌های محاسباتی به کاربردهای عملی ضروری است (Waldrop, 2019). پر کردن شکاف‌های همکاری بین پژوهشگران اُمیکس، دانشمندان داده و متخصصان کشاورزی برای تحقق کامل پتانسیل هوش مصنوعی در دفاع گیاهی و کاربردهای عملی حیاتی است (Rai, 2021).

جهت‌گیری‌های آینده

ادغام هوش مصنوعی و داده‌های متابولومیکس در تشخیص زودهنگام بیماری‌ها: یکی از جهت‌گیری‌های امیدوارکننده آینده، ادغام هوش مصنوعی با انواع داده‌های متابولومیکس برای تشخیص زودهنگام بیماری‌های گیاهی است (Raza, 2021). الگوریتم‌های پیشرفته هوش مصنوعی می‌توانند مجموعه‌های داده چندبُعدی اُمیکس را تحلیل کنند تا تغییرات عمیق در

پروفایل‌های مولکولی گیاهان مرتبط با شروع بیماری (حتی قبل از ظهور علائم قابل مشاهده) را شناسایی کنند (Mahlein, 2019).

مدل‌سازی پیش‌بینانه دینامیک بیماری‌ها: با استفاده از هوش مصنوعی و داده‌های متابولومیکس، می‌توان مدل‌های پیش‌بینانه برای پیش‌بینی دینامیک بیماری‌ها در جمعیت‌های گیاهی توسعه داد (Singh, 2022). مدل‌های یادگیری ماشین و یادگیری عمیق می‌توانند متغیرهای ژنتیکی، مولکولی و محیطی را در نظر بگیرند تا شیوع بیماری‌ها را پیش‌بینی کنند و تأثیر اقدامات پیشگیرانه را ارزیابی کنند (Pudjihartono, 2022).

اصلاح نژاد سفارشی محصولات: پژوهش‌های آینده ممکن است بر استفاده از تکنیک‌های متابولومیکس با کمک هوش مصنوعی برای سفارشی‌سازی برنامه‌های اصلاح نژاد محصولات به‌منظور افزایش مقاومت به بیماری‌ها متمرکز شوند (Wallace, 2018). با شناسایی نشانگرها و مسیرهای خاص مرتبط با مقاومت، اصلاح‌کنندگان می‌توانند محصولاتی با مکانیسم‌های دفاعی بهبودیافته طراحی کنند (Weckwerth, 2020).

مدیریت پایدار بیماری‌ها با راهنمایی هوش مصنوعی: هوش مصنوعی می‌تواند در بهینه‌سازی استراتژی‌های مدیریت بیماری‌ها کمک کند (Kamilaris, 2018). ادغام توصیه‌های مبتنی بر هوش مصنوعی با داده‌های متابولومیکس، امکان کاربرد دقیق آفت‌کش‌ها را فراهم می‌کند که این امر تأثیرات محیطی را کاهش می‌دهد، هزینه‌ها را کم می‌کند و در عین حال عوامل بیماری‌زای گیاهی را به‌طور مؤثر کنترل می‌کند (Liakos, 2018).

مواجهه با تنش‌های ترکیبی: با تغییرات اقلیمی و تحول در روش‌های کشاورزی، گیاهان اغلب با چالش مواجهه همزمان با چندین تنش روبرو هستند. پژوهش‌های آینده باید بر درک چگونگی پاسخ گیاهان به تنش‌های ترکیبی زیستی و غیرزیستی متمرکز شوند (Zandalinas, 2021)، زیرا این موضوع نمایانگر یک مشکل واقعی مهم است. تکنیک‌های متابولومیکس با کمک هوش مصنوعی می‌توانند نقش محوری در کشف برهمکنش‌های پیچیده بین عوامل تنش مختلف و اثرات تجمعی آن‌ها بر مکانیسم‌های دفاعی گیاهان ایفا کنند (Sewelam, 2020). این دانش برای توسعه استراتژی‌های جامع و مقاوم جهت محافظت از محصولات در محیط‌های تنش‌زا ضروری است و تولید پایدار غذا را تضمین می‌کند.

ادغام داده‌های سنجش از دور با هوش مصنوعی و متابولومیکس: ادغام فناوری سنجش از دور با تکنیک‌های متابولومیکس با کمک هوش مصنوعی، رویکردی قدرتمند برای نظارت و کاهش تنش گیاهان ارائه می‌دهد (Tattaris, 2016). سنجش از دور داده‌های ارزشمند مکانی و زمانی درباره سلامت گیاهان، عوامل تنش و شرایط محیطی فراهم می‌کند (Yang, 2021). با ادغام داده‌های سنجش از دور و اطلاعات متابولومیکس، پژوهشگران می‌توانند درک جامعی از برهمکنش بین پاسخ‌های ژنتیکی و عوامل تنش‌زای محیطی به‌دست آورند. این رویکرد یکپارچه، امکان مداخلات دقیق‌تر و به‌موقع را برای تقویت دفاع گیاهان و کاهش تلفات محصول فراهم می‌کند.

نتیجه‌گیری کلی

این مرور جامع به بررسی چشم‌انداز مطالعات متابولومیکس در دفاع گیاهان در برابر تنش‌های زیستی و غیرزیستی و نقش تحول‌آفرین تکنیک‌های یادگیری ماشین در حوزه‌ی متابولومیکس می‌پردازد. تکنیک‌های مبتنی بر هوش مصنوعی توانایی خود را در استخراج بینش‌های معنادار از مجموعه داده‌های گسترده نشان می‌دهند و از روش‌های سنتی پیشی می‌گیرند. در حالی که بر مزایای یادگیری ماشین تأکید می‌شود، این مرور چالش‌های مرتبط با پیاده‌سازی آن در متابولومیکس گیاهی را نیز مورد توجه قرار می‌دهد و راه را برای توسعه‌های آینده هموار می‌کند. پیشرفت در چارچوب‌های محاسباتی، کاربرد روش‌های مدرن را تسهیل می‌کند. با افزایش حجم داده‌های گیاهی، یادگیری ماشین به‌عنوان یک کاتالیزور در تسریع جنبه‌های مختلف پژوهش‌های گیاهی ظاهر می‌شود. که شامل شناسایی متابولیت‌های مرتبط با مقاومت در برابر تنش‌های زیستی و غیرزیستی و همچنین بهبود درک ما از مکانیسم‌ها است. این پیشرفت‌ها به پژوهشگران کشاورزی کمک می‌کنند تا عملکرد و کیفیت محصولات را بهبود بخشند و مقاومت به تنش‌های زیستی و غیرزیستی را افزایش دهند.

منابع

1) Alseekh, S. 2021. Mass spectrometry-based metabolomics. *Nature Reviews Methods Primers*, 1: 80.
2) Altman, N. and M, Krzywinski. 2018. The curse(s) of dimensionality. *Nature Methods*, 15(6): 399-400.
3) Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R. and P, Daszak. 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, 19(10): 535-544.
4) Arabnia, H. and Q, Tran. 2011. *Software tools and algorithms for biological systems*. Springer.
5) Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A. and D, Stefanovic. 2019. Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry*, 11(7): 939.
6) Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L. and G, Bornemann. 2015. The coffee rust crises in Colombia and Central America, (2008–2013): impacts, plausible causes and proposed solutions. *Food Security*, 7: 303-321.
7) Baldi, P. and P.J, Sadowski. 2013. Understanding dropout. *Advances in Neural Information Processing Systems*, 26: 2814-2822.
8) Bebber, D.P., Ramotowski, M.A. and S.J, Gurr. 2013. Crop pests and pathogens move polewards in a warming world. *Nature Climate Change*, 3(11): 985-988.
9) Biabi H., Abdanan Mehdizadeh, S. and M.R, Salehi. 2019. Design and implementation of a smart system for water management of lilium flower using image processing, *Computers and Electronics in Agriculture.*
10) Bishop, C.M. 2006. Pattern Recognition and Machine Learning. *Journal of Machine Learning Research*, 7: 845-848.
11) Borrill, P. 2019. Applying the latest advances in genomics and phenomics for trait discovery. *Plant Biotechnology Journal*, 17(7): 1267-1286.
12) Breiman, L. 2001. Random forests. *Machine Learning*, 45(1): 5-32.
13) Chen, W. 2021. GC-MS based metabolomics study of rice responses to pathogens. *Plant Physiology*, 186(2): 1239-1255.
14) Cortes, C. and V, Vapnik. 1995. Support-vector networks. *Machine Learning*, 20(3): 273-297.
15) Do, P.T. 2020. Rice drought-responsive metabolites. *Plant, Cell & Environment*, 43(3): 630-644.
16) Domingos, P. 2012. A few useful things to know about machine learning. *Communications of the ACM*, 55(10): 78-87.
17) Du, Q., Campbell, M., Yu, H., Liu, K., Walia, H. and Q, Zhang. 2019. Network-based feature

selection reveals substructures of gene modules responding to salt stress in rice. *Plant Direct, 3*(8): e00154.

18) Eraslan, G. 2019. Scalable AI for omics data. *Nature Methods*, 16(8): 647-653.
19) Fang, C. and J, Luo. 2019. Metabolic GWAS and plant specialized metabolism. *Trends in Plant Science*, 24(6): 550-561.
20) Farooq, A., Mir, R. A., Sharma, V., Pakhtoon, M.M., Bhat, K.A. and A.A, Shah. 2022. Crop proteomics: Towards systemic analysis of abiotic stress responses. In *Advancements in developing abiotic stress-resilient plants* (pp. 265-285). CRC Press.
21) Food and Agriculture Organization of the United Nations. 2017. The future of food and agriculture – Trends and challenges. Rome: FAO.
22) Fürtauer, L. 2018. Resolving subcellular plant metabolism. *The Plant Journal*, 94(5): 887-890.
23) Fürtauer, L., Pschenitschnigg, A., Scharkosi, H., Weckwerth, W. and T, Nägele. 2018. Combined multivariate analysis and machine learning reveals a predictive module of metabolic stress response in Arabidopsis thaliana. *Molecular Omics,* 14(6): 437-449.
24) Gokalp, O. and E, Tasci. 2019. Weighted voting based ensemble classification with hyper-parameter optimization. In *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-4). IEEE.
25) Goodfellow, I. 2016. Deep Learning. *Nature*, 521(7553): 436-444.
26) Greener, J.G. 2022. Biological interpretability in machine learning. *Nature Machine Intelligence*, 4(6): 510-519.
27) Guyon, I. and A, Elisseeff. 2006. An introduction to feature extraction. In I. Guyon, S. Gunn, M. Nikravesh. and L.A, Zadeh (Eds.), *Feature extraction: foundations and applications* (pp. 1-25). Springer.
28) Guyon, I., Gunn, S., Nikravesh, M. and L.A, Zadeh. 2008. *Feature extraction: foundations and applications* (Vol. 207). Springer.
29) Hamzehzarghani, H. 2018. Metabolite profiling of wheat resistance to Fusarium. *Frontiers in Plant Science*, 9: 667.
30) Hastie, T. 2009. The Elements of Statistical Learning. *Statistical Science*, 24(3): 337-358.
31) Hyvärinen, A. and E, Oja. 2000. Independent component analysis. *Neural Networks*, 13(4-5): 411-430.
32) Isewon, I., Apata, O., Oluwamuyiwa, F., Aromolaran, O. and J, Oyelade. 2022. Machine learning algorithms: their applications in plant omics and agronomic traits' improvement. *F1000Research, 11*: 1256.
33) Jain, A.K. 2010. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8): 651-666.
34) Joyce, A.R. and B.Ø, Palsson. 2006. The model organism as a system: integrating 'omics' data sets. *Nature Reviews Molecular Cell Biology,* 7(3): 198-210.
35) Kamilaris, A. 2018. AI in agriculture. *Agricultural Systems*, 167: 86-96.
36) Khalid, S., Khalil, T. and S, Nasreen. 2014. A survey of feature selection and feature extraction techniques in machine learning. In *2014 science and information conference* (pp. 372-378). IEEE.
37) Kumar, R. 2022. Metabolomics in plant-pathogen interactions. *Journal of Plant Physiology*, 268: 153569.
38) LeCun, Y. 2015. Deep learning. *Nature*, 521(7553): 436-444.
39) Li, J., Wei, L., Guo, F. and Q, Zou. 2021. EP3: an ensemble predictor that accurately identifies type III secreted effectors. *Briefings in Bioinformatics, 22*(2): 1918-1928.
40) Liakos, K.G. 2018. Machine learning in agriculture. *Biosystems Engineering*, 171: 31-43.
41) Libbrecht, M.W. and W.S, Noble. 2015. Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, 16(6): 321-332.
42) Liu, K., Abdullah, A.A., Huang, M., Nishioka, T., Altaf-Ul-Amin, M. and S, Kanaya. 2017. Novel approach to classify plants based on metabolite-content similarity. *BioMed Research International.*
43) Lloyd, S. 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2): 129-137.

44) Mahlein, A.K. 2019. Plant disease detection by imaging sensors. *Annual Review of Phytopathology*, 57: 459-482.
45) Maimon, O. and L, Rokach. 2005. *Data Mining and Knowledge Discovery Handbook* (Vol. 2). Springer.
46) Mittler, R. and E. Blumwald. 2015. The roles of ROS and ABA in response to stress. *Trends in Plant Science*, 20(1): 3-11.
47) Murtagh, F. and P. Legendre. 2014. Ward's hierarchical clustering method. *Journal of Classification*, 31(3): 274-295.
48) Nakabayashi, R. and K, Saito. 2021. Metabolomics for unknown plant metabolites. *Phytochemistry*, 182: 112594.
49) Nakagami, H., Sugiyama, N., Mochida, K., Daudi, A., Yoshida, Y., Toyoda, T., Tomita, M., Ishihama, Y. and K, Shirasu. 2010. Large-scale comparative phosphoproteomics identifies conserved phosphorylation motifs in plants. *Plant Physiology, 152*(2): 711-723.
50) Niazian, M. and G, Niedbała. 2020. Machine learning for plant breeding and biotechnology. *Agriculture*, 10(10): 436.
51) Noor, E., Cherkaoui, S. and U, Sauer. 2019. Biological insights through omics data integration. *Current Opinion in Systems Biology, 15*: 39-47.
52) Obata, T. and A.R, Fernie. 2012. The use of metabolomics to dissect plant responses to abiotic stresses. *Cellular and Molecular Life Sciences*, 69(19): 3225-3243.
53) Pang, Z. 2021. Metabolomics and transcriptomics profiling reveals molecular mechanisms of rice in response to abiotic stresses. *Plant, Cell & Environment*, 44(5): 1486-1503.
54) Piasecka, A. 2019. Secondary metabolites in plant innate immunity. Phytochemistry, 160: 16-28.
55) Provost, F. and T, Fawcett. 2013. Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1): 51-59.
56) Pudjihartono, N. 2022. AI for disease prediction. *Computers and Electronics in Agriculture*, 193: 106668
57) Putri, S.P. 2021. Current metabolomics in plant abiotic stress. *Metabolites*, 11(2): 105.
58) Quinlan, J.R. 1986. Induction of decision trees. Machine Learning, 1(1): 81-106.
59) Rai, A. 2021. Interdisciplinary approaches in agri-tech. *Trends in Biotechnology*, 39(3): 227-239.
60) Raza, A. 2021. AI in plant disease detection. *Nature Plants*, 7(6): 730-741.
61) Reel, P.S., Reel, S., Pearson, E., Trucco, E. and E, Jefferson. 2021. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology Advances, 49*: 107739.
62) Rish, I. 2001. An empirical study of the naive Bayes classifier. *Journal of Machine Learning Research*, 1: 41-46.
63) Ristaino, J.B., Anderson, P.K., Bebber, D.P., Brauman, K.A., Cunniffe, N.J. and N, Fedoroff. 2021. The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences, **118***(23): e2022239118.
64) Russell, S. and P, Norvig. 2021. Artificial Intelligence: A Modern Approach (4th ed.). *AI Magazine*, 42(2): 96-98.
65) Saito, K. and F, Matsuda. 2010. Metabolomics for functional genomics. *Annual Review of Plant Biology*, 61: 463-489.
66) Samek, W. 2021. Explainable AI: Interpreting deep learning models. *Pattern Recognition*, 120: 108194.
67) Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N. and A, Nelson. 2019. The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3(3): 430-439.
68) Schmidhuber, J. 2015. Deep learning in neural networks. *Neural Networks*, 61: 85-117.
69) Scholz, M., Gatzek, S., Sterling, A., Fiehn, O. and J, Selbig. 2004. Metabolite fingerprinting: detecting biological features by independent component analysis. *Bioinformatics*, 20(15): 2447-2454.
70) Sewelam, N. 2020. Molecular plant stress responses. *Trends in Plant Science*, 25(10): 1040-1052.
71) Shen, D., Wang, Y., Chen, X., Srivastava, V. and S.L, Toffolatti. 2022. Advances in multi-omics study of filamentous plant pathogens. *Frontiers in Microbiology*.
72) Silva, J.C.F., Teixeira, R.M., Silva, F.F., Brommonschenkel, S.H. and E.P, Fontes. 2019. Machine learning approaches and their current application in plant molecular biology: A systematic review. *Plant Science*, 284: 37-47.

73) Singh, A., Ganapathysubramanian, B., Singh, A.K. and S, Sarkar. 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science*, 21(2): 110-124.
74) Singh, D. 2022. AI in plant omics. *Computational and Structural Biotechnology Journal*, 20: 5913-5927.
75) Sperschneider, J., Dodds, P.N., Gardiner, D.M., Singh, K.B. and J.M, Taylor. 2018. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Molecular Plant Pathology*, 19(9): 2094-2110.
76) Srivastava, N. 2014. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1): 1929-1958.
77) Sweetlove, L.J. 2017. Engineering metabolic pathways in plants. *Annual Review of Plant Biology*, 68: 723-750.
78) Tattaris, M. 2016. Remote sensing for plant disease. *Remote Sensing*, 8(8): 614.
79) Van Wallendael, A. 2019. A molecular view of plant responses to abiotic stress. *Journal of Experimental Botany*, 70(16): 4135-4152.
80) Varshney, R.K. 2018. Can genomics deliver climate-change ready crops? *Current Opinion in Plant Biology*, 45: 205-211.
81) Wager, S. 2013. Dropout training as adaptive regularization. *Advances in Neural Information Processing Systems*, 26: 351-359.
82) Waldrop, M.M. 2019. Challenges in AI validation. *Nature*, 575(7781): 136-138.
83) Wallace, J.G. 2018. Genomic and metabolic prediction. *Plant Physiology*, 177(3): 1146-1155.
84) Wang, L. 2023. Deep learning for plant stress phenotyping. *Plant Communications*, 4(1): 100498.
85) Weckwerth, W. 2020. PANOMICS meets germplasm. *Plant Biotechnology Journal*, 18(7): 1507-1525.
86) Wold, E.S. 2018. Multi-omics data integration. *Briefings in Bioinformatics*, 19(4): 702-715.
87) Yan, J. and X, Wang. 2022. Unsupervised and semi-supervised learning: the next frontier in machine learning for plant systems biology.
88) Yang, X. 2021. Deep learning for plant phenomics. *Molecular Plant*, 14(5): 714-736.
89) Zandalinas, S.I. 2021. Plant responses to combined stress. *Plant Physiology*, 186(1): 66-78.
90) Zhang, H. 2020. Abiotic stress responses in plants. *Plant Cell Reports*, 39(5): 637-651.
91) Zhou, Z.H. 2021. Machine Learning. *Artificial Intelligence Review*, 54(2): 1513-1542.
92) Zhu, J.K. 2016. Abiotic stress signaling and responses in plants. *Cell*, 167(2): 313-324.
93) Zhu, X. and A.B, Goldberg. 2009. Introduction to semi-supervised learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 3(1): 1-130.

Iranian Journal of Plant & Biotechnology
Spring 2025, Vol 20, No 1, Pp 38-52

Iranian Journal of Plant & Biotechnology (IJPB)

Islamic Azad University
Garmsar Branch

Investigating Plant Defense Metabolomics Mechanisms Based on Artificial Intelligence

Mozhgan Zangeneh[1] and Mohamadreza Salehi Salmi (Corresponding author)[2]*
1- Ph.D, Department of Horticultural Sciences, Faculty of Agriculture, Agricultural Sciences and Natural Resource University of Khuzestan, Mulathani, Iran, m.zangeneh@asnrukh.ac.ir
2- Associate Professor, Department of Horticultural Sciences, Faculty of Agriculture, Agricultural Sciences and Natural Resource University of Khuzestan, Mulathani, Iran, Salehi@asnrukh.ac.ir
Received: May 2025 Accepted: July 2025

Abstract

Introduction: Plant defense against biotic stresses (living factors such as pathogens and pests) and abiotic stresses (environmental factors such as drought, salinity, and extreme temperatures) is of great importance in ensuring global food security. With the continuous growth of the world's population, the demand for agricultural products is increasing, highlighting the urgent need to develop efficient strategies for plant protection. The ability to protect crops from diseases, pests, and adverse environmental conditions is essential for maintaining agricultural productivity and securing global food supply. Artificial intelligence-powered metabolomics has emerged as a transformative approach, providing unprecedented insights into plant metabolic responses to biotic and abiotic stresses. This review examines current trends and future directions of AI-based metabolomics techniques, analyzing their role in deciphering metabolic networks, identifying stress-responsive biomarkers, and uncovering hidden patterns within omics datasets. Compared to traditional methods that face challenges in data integration and result interpretation, machine learning and deep learning algorithms enable rapid and accurate analysis of complex datasets. These advanced technologies not only facilitate the discovery of key metabolites involved in plant defense mechanisms but also enhance stress tolerance prediction and the development of resistant crop varieties. The integration of multi-omics data (genomics, transcriptomics, and proteomics) with metabolomics through AI-driven platforms offers a more comprehensive understanding of plant stress responses. However, challenges such as data standardization, model interpretability, and the need for high-performance computational resources remain unresolved. This paper also explores the potential of AI in optimizing agricultural practices, improving crop resilience and ensuring food security under climate change conditions.

Conclusion: By addressing current limitations and leveraging emerging technologies, this field holds revolutionary promise for plant sciences and global food security challenges.

Keywords: Crop resilience, Data analysis, Multi-omics, Stress, Sustainable agriculture