Calibration, Evaluation, and Validation of the Most Appropriate Machine Learning Algorithms for Predicting Corn Seed Yield under Biofertilizer Applications: An Interdisciplinary Approach

M. Jahan[1*]

Department of Agrotechnology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

Introduction

The escalating degradation of water and soil resources, coupled with global warming, population growth, and climate change, necessitates a reevaluation of food production systems. Corn (*Zea mays* L.), as the third most strategic crop worldwide, faces increasing demand due to rising food needs. Optimizing seed yield under varying environmental conditions is critical, with arbuscular mycorrhizal fungi (AMF) and plant growth-promoting rhizobacteria (PGPR) playing pivotal roles by enhancing nutrient uptake (e.g., phosphorus and nitrogen) and improving plant resilience to stresses and pathogens. However, the complex eco-physiological interactions influencing seed yield remain poorly understood, underscoring the need for advanced modeling tools. Traditional regression models often fail to capture non-linear and interactive relationships among factors such as photosynthesis rates, nutrient levels, and plant anatomical traits. Consequently, machine learning models like Deep Neural Networks (DNN) and Transformers have emerged as powerful alternatives for precision agriculture. This study aims to identify key features affecting corn seed yield using stepwise regression and compare eight machine learning algorithms—Enhanced DNN, Transformer, XGBoost, SVM, ANN, ANFIS, LightGBM, and SVR—to determine the most accurate model for predicting yield and unveiling hidden eco-physiological relationships.

Materials and Methods

The experiments were conducted over two consecutive years at the research farm of Ferdowsi University of Mashhad, Iran (latitude 36°15'N, longitude 59°28'E, altitude 985 m), located in the Kashafroud River basin. The region features a semi-arid climate with an average annual rainfall of 252 mm and a mean temperature of 15°C, with loamy soil of moderate organic carbon content. The dataset comprised 96 samples and 73 features, collected over two growing seasons, including plant traits (e.g., leaf chlorophyll content via SPAD index, nitrogen concentration) and soil characteristics. Of these, 32 were primary features (e.g., photosynthesis rate, leaf area index, canopy temperature), and 41 were engineered interaction features (e.g., PmaxMean_LeafAreaIndex, CanopyTemp1_RootColonization), designed based on domain knowledge of agro-ecophysiology, soil ecology, AMF, and PGPR. Data were randomly split (random_state=42) into 70% training (67 samples), 15% validation (14 samples), and 15% test (15 samples) sets, with standardization applied using StandardScaler. Feature selection employed stepwise backward regression, reducing 73 features to 13 key variables (e.g., Canopy Temp_3, % P plant) based on adjusted R^2, multicollinearity, and variance inflation factor (VIF). Amongst 15 machine learning algorythms, the eight machine learning models were configured with specific architectures: ANFIS with TensorFlow-based layers, Transformer with 2 attention heads, ANN with 3 hidden layers, SVR with RBF kernel, etc., and evaluated using R^2, RMSE, MAE, and Willmott's d, with Shapiro-Wilk test for residual normality.

Results and Discussion

Stepwise regression yielded a model with an adjusted R^2 of 58.53% and a predictive R^2 of 51.06%, identifying 13 features (e.g., Canopy Temp_3 with coefficient 0.2451, p=0.002). Machine learning results showed ANFIS (R^2=0.555), Transformer (R^2=0.545), and ANN (R^2=0.518) as top performers, while SVM (R^2=0.325) lagged. The Nemeyi diagram ranked ANN first (mean rank 1.50, Willmott's d=0.848), followed by Transformer, with critical distance (1.11) indicating significant differences (α=0.05) between top and bottom models. Taylor diagrams highlighted Transformer's superiority (RMSE=2.834, R^2=0.545), with ANFIS and ANN close behind. SHAP plots revealed that interactions like Leaf Area Index_Dry Matter Yield were critical, with SVR sharing 9 features with the regression model. Correlation analysis clustered features into physiological (e.g., SPAD_Mean) and yield-related (e.g., Dry Matter Yield) groups, with strong positive correlations (e.g., SPAD_2 and cob diameter, r=0.82) and negative ones (e.g., Canopy Temp_Mean and Root Colonization, r=-0.82). Skewed distributions in KDE plots underscored the need for non-linear models. Force plots (e.g., LightGBM sample) showed features like Leaf Area

Index_Dry Matter Yield (+280.047) driving predictions, while loss curves for Enhanced DNN indicated effective convergence. The study suggests that neural network-based models excel in capturing complex eco-physiological interactions, with canopy temperature and root-nitrogen interactions as key predictors, offering insights for sustainable corn production despite data size limitations.

Keywords: Artificial Neural Network, Arbuscular mycorrhizal fungi, Plant growth-promoting rhizobacteria, Eco-physiological interactions, Feature engineering, Input efficiency

واسنجی، ارزیابی و تعیین اعتبار مناسب‌ترین الگوریتم‌های یادگیری ماشین برای پیش‌بینی عملکرد دانه ذرت تحت شرایط کاربرد کودهای زیستی: رهیافتی بین‌رشته‌ای

محسن جهان[1*]

۱- گروه اگروتکنولوژی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران

(*- نویسنده مسئول: jahan@ferdowsi.um.ac.ir)

چکیده

پیش‌بینی عملکرد دانه بخشی کلیدی در کشاورزی است که به کشاورزان و برنامه ریزان کمک می‌کند تا منابع را به طور کاراتر مدیریت کرده و بهره‌وری را بالا ببرند. این مطالعه به بررسی و مقایسه هشت الگوریتم یادگیری ماشین برای پیش‌بینی عملکرد دانه ذرت (*Zea mays* L.) با استفاده از ۷۳ ویژگی گیاهی و خاکی، شامل ۳۲ ویژگی اصلی و ۴۱ ویژگی تعاملی مهندسی‌شده، پرداخته است. آزمایش‌ها در مزرعه تحقیقاتی دانشگاه فردوسی مشهد طی دو سال انجام شد و داده‌ها با تقسیم تصادفی (۷۰٪ آموزش، ۱۵٪ اعتبارسنجی، ۱۵٪ تست) و استانداردسازی (StandardScaler) پردازش شدند. ابتدا، مدل رگرسیون گام‌به‌گام پس‌رونده ۱۳ ویژگی کلیدی (مثل Canopy Temp_3 ، P plant %) را شناسایی کرد (R^2(adj)=58.53%). سپس، الگوریتم‌های ANFIS، Transformer، ANN، SVR، LightGBM، XGBoost، Enhanced DNN، و SVM از بین پانزده الگوریتم رایج، غربال و ارزیابی شدند. معیارهای R^2 ، RMSE، MAE، و Willmott's d نشان داد که ANFIS (R^2=0.555) ، Transformer (R^2=0.545)، و ANN (R^2=0.518)بهترین عملکرد را داشتند، در حالی که SVM (R^2=0.325) ضعیف‌ترین بود. نمودارهای SHAP و Decision Plot نشان داد که تعاملات مثل Leaf Area Index_Dry Matter Yield و Canopy Temp_4 در پیش‌بینی نقش کلیدی دارند. تحلیل همبستگی، دو خوشه اصلی (فیزیولوژیکی و عملکردی) را آشکار کرد، و توزیع‌های اریب داده‌ها لزوم مدل‌های غیرخطی را تأیید کرد. نتایج نشان داد مدل‌های شبکه عصبی (به‌ویژه با مکانیزم‌های Attention وTensorFlow) روابط اکوفیزیولوژیکی پیچیده را بهتر مدل‌سازی می‌کنند. ویژگی‌هایی مثل دمای کانوپی و برهمکنش محتوای نیتروژن گیاه–درصد کلونیزاسیون ریشه به‌عنوان متغیرهای اصلی شناسایی شدند که نشان‌دهنده اهمیت تعاملات بین جذب عناصر غذایی و پاسخ‌های فیزیولوژیکی گیاه بودند. این مطالعه نشان داد که ترکیب ویژگی‌های تعاملی با الگوریتم های پیشرفته یادگیری ماشین می‌تواند از یک سو، دقت پیش‌بینی را بالا ببرد و به افزایش کارآیی سیستم های تولید گیاهان زراعی در راستای اهداف کشاورزی پایدارکمک کند و از سوی دیگر، زمینه های لازم برای مدیریت زراعی دقیق، شناسایی مسیرهای اکوفیزیولوژیکی مؤثر در شکل گیری عملکرد نهایی و سازگاری با تغییرات اقلیمی را فراهم آورد.

کلمات کلیدی: برهمکنش اکوفیزیولوژیک، ریزوباکتری های تحریک کننده رشد گیاه، شبکه عصبی مصنوعی، قارچ های میکوریزا آربوسکولار، کارایی نهاده، مهندسی ویژگی.

مقدمه

تخریب منابع آب و خاک، افزایش جمعیت، گرمایش جهانی و تغییرات اقلیمی، تجدید نظر در روش ها و سیستم های تولید غذا را ناگزیر ساخته است (Godfray et al., 2024) . ذرت (*Zea mays* L.) سومین گیاه زراعی استراتژیک در سطح جهانی است و تقاضا برای آن به دلیل افزایش جمعیت و نیاز فزاینده به غذا رو به رشد می باشد. برای پاسخ به این نیاز، باید عملکرد دانه تحت شرایط محیطی متغیر بهینه شود. با این حال، تولید ذرت به شدت در معرض تنش‌های محیطی مانند خشکسالی، شوری و کمبود عناصر غذایی قرار دارد که این شرایط با تغییرات اقلیمی تشدید شده‌اند (Bracho-Mujica et al., 2024). در این راستا، روش‌های کشاورزی پایدار، از جمله استفاده از میکروارگانیسم‌های مفید خاک مانند قارچ‌های میکوریز آربوسکولار[1] و ریزوباکتری‌های محرک رشد گیاه[2]، به دلیل پتانسیل آن‌ها در بهبود رشد گیاه، عملکرد و تحمل تنش مورد توجه قابل‌توجهی قرار گرفته‌اند (Brar et al., 2024) . قارچ‌های میکوریز آربوسکولار و ریزوباکتری‌های محرک رشد گیاه با افزایش جذب عناصر غذایی مثل فسفر و نیتروژن، همچنین بهبود مقاومت گیاه به تنش‌ها و عوامل بیماری زا، نقش مهمی در بهبود عملکرد ذرت دارند. با وجود این، روابط پیچیده اکوفیزیولوژیکی بین این عوامل و عملکرد دانه، هنوز به‌خوبی شناخته نشده و نیاز به ابزارهای پیشرفته برای مدل‌سازی دارد.

پیش‌بینی دقیق عملکرد گیاهان زراعی از موضوعات کلیدی برای بهبود تولید غذا و پایداری اقتصادی می باشد. ذرت سومین گیاه زراعی مهم و استراتژیک در سطح جهانی بوده و پیش‌بینی دقیق عملکرد دانه آن برای بهینه‌سازی روش‌های کشاورزی، مدیریت منابع، و درک بهتر تعاملات اکوفیزیولوژیکی در سیستم‌های زراعی از اهمیت زیادی برخوردار است. با توجه به پیچیدگی داده‌های کشاورزی که شامل متغیرهای متعدد، روابط غیرخطی، و نویز زیاد است، اخیراً استفاده از مدل‌های یادگیری ماشین به‌عنوان ابزاری قدرتمند برای افزایش دقت پیش‌بینی‌ها مطرح شده است (Fashoto et al., 2021; Nayak et al., 2022, Pentos et al., 2024; Wang et al., 2024). مدل‌های سنتی (رگرسیونی) معمولاً نمی‌توانند روابط غیرخطی و تعاملی بین عوامل اکوفیزیولوژیکی مثل نرخ فتوسنتز، سطح عناصر غذایی، و ویژگی‌های آناتومیکی (ساختاری) گیاه رو به‌خوبی شبیه سازی کنند (Khaki et al., 2020; Pentos et al., 2022). به همین دلیل، الگوریتم‌های یادگیری ماشین مثل شبکه‌های عصبی عمیق[3] و الگوریتم‌های جدیدتر مثل ترنسفورمر[4] به‌عنوان ابزارهای قدرتمند برای این کار شناخته شده ند(Nayak et al., 2022). علاوه بر این، تکنیک‌های تفسیری مانند تحلیل SHAP (SHapley Additive exPlanations) ، بینش‌های عمیق‌تری درباره اهمیت هر ویژگی و اثرات علی تغییرات خاص ارائه می‌دهند (Abekoon et al., 2025). این روش‌ها به پژوهشگران اجازه می‌دهند نه‌تنها عملکرد ذرت را پیش‌بینی کنند، بلکه درک کنند چگونه کودهای زیستی و سایر نهاده های طبیعی بر ویژگی‌های فیزیولوژیکی کلیدی مانند نرخ فتوسنتز، شاخص سطح برگ و کارایی مصرف آب تحت شرایط مختلف محیطی تأثیر می‌گذارند.

قارچ‌های میکوریزا آربوسکولار با بیش از ۹۰ درصد گیاهان خشکی، از جمله ذرت، رابطه هم‌زیستی متقابل برقرار می‌کنند و جذب عناصر غذایی– به‌ویژه فسفر(P) – را افزایش داده و مقاومت گیاه را در برابر تنش‌های محیطی بهبود می‌بخشند .(Bhupenchandra et al., 2024) فراهمی فسفر، به‌عنوان دومین عنصر غذایی ضروری برای رشد گیاه، اغلب به دلیل تثبیت سریع در خاک‌های کشاورزی، محدود است و قارچ‌های میکوریزا آربوسکولار نقش مهمی در آزادسازی فسفر خاک و افزایش دسترسی آن برای گیاهان ایفا می‌کند (Mitra et al., 2023) . به‌طور مشابه، ریزوباکتری های تحریک کننده رشد گیاه با حل کردن عناصر غذایی، تولید مواد محرک رشد و بهبود توسعه سیستم ریشه، به حاصلخیزی خاک کمک کرده و رشد و عملکرد گیاه را در شرایط نامطلوب بهبود می‌بخشد (Khoso et al., 2024). استفاده ترکیبی از قارچ‌های میکوریزا آربوسکولار و ریزوباکتری های تحریک کننده رشد گیاه نشان داده است که محتوای کلروفیل برگ را افزایش می‌دهد، بهره‌وری آب را بهبود می‌بخشد و دمای کانوپی را کاهش می‌دهد، که در نهایت منجر به عملکرد بالاتر گیاه زراعی و تحمل بهتر تنش‌ها می‌شود (Salvador et al., 2022). با این حال، مکانیسم‌های اکوفیزیولوژیکی که این تعاملات هم‌افزا را پشتیبانی می‌کنند، و همچنین پویایی‌های اکولوژیکی آن‌ها

[1] Arbuscular Mycorehizal Fungi
[2] Plant Growth Promoting Rhizobacteria (PGPR)
[3] Deep Neural Networks (DNN)
[4] Transformer

در سیستم‌های زراعی ذرت، هنوز به‌خوبی شناخته نشده‌اند و نیازمند ابزارهای پیشرفته برای کمی‌سازی مشارکت و بهینه‌سازی کاربرد آنها است (Bao et al., 2022).

در این مطالعه، ابتدا بین ویژگی های رشد و نموی ذرت و ویژگی های خاک، ضرایب همبستگی محاسبه شدند. سپس، با استفاده از روش رگرسیون چند متغیره گام به گام پس رونده[1]، مهمترین و تأثیرگذارترین متغیرها بر عملکرد دانه ذرت شناسایی شدند. در ادامه، و به منظور بهره گیری از مزایای الگوریتم های یادگیری ماشین همچون شناسایی و تبیین روابط پیچیده غیر خطی، اعتبارسنجی متقابل، انتخاب دقیق ترین و پایدارترین مدل، و تعمیم پذیری بیشتر، هایپرپارامترهای پانزده الگوریتم یادگیری ماشین تنظیم و اجرا شدند. پس از ارزیابی های اولیه، هشت الگوریتم برتر شامل سه الگوریتم درختی و چهار الگوریتم مبتنی بر شبکه عصبی مصنوعی و یک الگوریتم بینابینی مورد واسنجی، ارزیابی، و مقایسه قرار گرفتند. هدف نهایی این بود که تعیین شود کدام مدل با استفاده از ویژگی‌های تعاملی (اینتراکشن) می‌تواند بهترین دقت و کارایی را داشته باشد و روابط اکوفیزیولوژیکی پنهان در شکل گیری عملکرد دانه را آشکار سازد.

مواد و روش‌ها

مکان اجرای آزمایش

آزمایش‌ها طی دو سال متوالی در مزرعه تحقیقاتی دانشکده کشاورزی دانشگاه فردوسی مشهد (موقعیت: عرض جغرافیایی ۳۶ درجه و ۱۵ دقیقه شمالی، طول جغرافیایی ۵۹ درجه و ۲۸ دقیقه شرقی، ارتفاع ۹۸۵ متر از سطح دریا) انجام شد. این منطقه در حوضه رودخانه کشفرود در شمال شرق ایران قرار دارد. آب‌وهوای منطقه، نیمه‌خشک با میانگین بارندگی سالانه ۲۵۲ میلی‌متر و دمای متوسط ۱۵ درجه سانتی‌گراد می باشد. خاک محل آزمایش، لومی با محتوای کربن آلی متوسط بود.

منشاء مجموعه داده

مجموعه داده شامل ۹۶ نمونه و ۷۳ ویژگی[2] بود که طی دو فصل رشد اندازه گیری و ثبت شده بودند. این داده‌ها شامل ویژگی‌های گیاهی مثل محتوای کلروفیل برگ (شاخص SPAD)، غلظت نیتروژن گیاه، و ویژگی‌های ریشه، به علاوه ویژگی‌های خاک بود (شرح کامل ویژگی‌ها به دلیل محدودیت صفحات در این متن آورده نشده است). از این ۷۳ ویژگی، ۳۲ ویژگی اصلی بودند (مثل نرخ فتوسنتز، شاخص سطح برگ، و دمای کانوپی) و ۴۱ ویژگی، ویژگی‌های تعاملی (اینتراکشن) مهندسی‌شده بودند (مثل PmaxMean_LeafAreaIndex و CanopyTemp1_RootColonization). داده‌ها با استفاده از تقسیم تصادفی (random_state=42) به ۷۰ درصد آموزش (۶۷ نمونه)، ۱۵ درصد اعتبارسنجی (۱۴ نمونه)، و ۱۵ درصد تست (۱۵ نمونه) تقسیم شدند. برای استانداردسازی دادها، از ماژول StandardScaler استفاده شد.

مهندسی ویژگی‌های تعاملی (اینتراکشن)

ویژگی‌های اینتراکشن به منظور مدل‌سازی روابط پیچیده بین متغیرها ساخته شدند. در داده‌های زراعی، تعاملات بین متغیرها معمولاً غیرخطی و به هم وابسته است. مثلاً، تأثیر دمای کانوپی روی عملکرد دانه ممکن است به میزان کولونیزاسیون ریشه بستگی داشته باشد. معیار انتخاب ویژگی‌های اینتراکشن، دانش حوزه‌ای[3] و در نظر گرفتن روابط و دانش اگرواکوفیزیولوژیکی، اکولوژی خاک، اکوفیزیولوژی قارچ های میکوریز آربوسکولار و ریزوباکتری های محرک رشد گیاه (که جزو تیمارهای آزمایش بودند) بود (برای آگاهی از جزئیات آزمایش های مزرعه ای به منبع https://doi.org/10.1007/s44396-025-00001-0 مراجعه شود). برای مثال RootColonization_SPAD1 تعامل بین کولونیزاسیون ریشه و محتوای کلروفیل را نشان می‌دهد، که رابطه بین جذب عناصر غذایی توسط ریشه و کارایی سیستم فتوسنتزی گیاه را منعکس می‌کند. CanopyTemp1_RootColonization این ویژگی تعامل بین دمای کانوپی و کولونیزاسیون ریشه را نشان می‌دهد و می‌تواند تأثیر تنش دمایی بر کارکرد ریشه را مشخص کند. ۴۱ ویژگی اینتراکشن به ۳۲ ویژگی اصلی اضافه شد تا الگوریتم های یادگیری ماشین بتوانند الگوهای پنهان بین داده ها را دقیق تر پیدا کنند.

---

[1] Bckward Elimination Stepwise Regression Method
[2] Feature
[3] Domain Knowledge

پیش‌پردازش

در مرحله آماده سازی داده ها، همه ویژگی های عددی با استفاده از ماژول StandardScaler در پایتون استاندارد شدند. این کار مقیاس ویژگی ها را یکسان کرده و باعث می‌شود مدل‌هایی مثل SVR و Enhanced DNN که به مقیاس حساس هستند، عملکرد بهتری داشته باشند.

انتخاب ویژگی ها (فیچرها): از روش رگرسیون گام به گام پس‌رونده برای انتخاب ویژگی های مؤثر استفاده شد، به این ترتیب که ابتدا مدل با همه ۷۳ ویژگی اجرا شد و سپس طی چندین مرحله و به‌صورت تکراری[1] ویژگی های کم‌اهمیت حذف شدند. در این مرحله، معیارهایی مثل ضریب تبیین، هم خطی چندگانه[2] و فاکتور تورم واریانس[3] در نظر گرفته شدند. در نهایت، از بین ۷۳ ویژگی، ۱۳ ویژگی مهم‌تر انتخاب شدند که فهرست آنها در جدول ۱ آورده شده است.

جدول ۱. سیزده ویژگی گیاهی–خاکی انتخاب شده توسط مدل رگرسیون گام به گام پس رونده

Table 1. Thirteen plant-soil features selected by the stepwise backward regression model

Feature	ویژگی
Canopy Temp_3	دمای کانوپی در مرحله کاکل دهی
% P plant	درصد فسفر اندام های گیاهی
% N soil	درصد نیتروژن کل خاک
PlantHeight_CobNumber	اثر متقابل ارتفاع بوته و تعداد بلال
Nplant_SpecificRootLength	اثر متقابل درصد نیتروژن اندام های گیاهی و طول ویژه ریشه
Leaf Area Index_SPAD_Mean	اثر متقابل شاخص سطح برگ و میانگین عدد اسپد
Leaf Area Index_Dry Matter Yield	اثر متقابل شاخص سطح برگ و عملکرد ماده خشک
Root Colonization_Dry Matter Yield	اثر متقابل درصد کلونیزاسیون ریشه و عملکرد ماده خشک
Stem Diameter_Dry Matter Yield	اثر متقابل قطر ساقه و عملکرد ماده خشک
(Specific Root Length) ^2	مجذور طول ویژه ریشه
(Canopy Temp_Mean) ^2 * Leaf Area Index	اثر متقابل مجذور میانگین دمای کانوپی و شاخص سطح برگ
(Root Colonization) ^2 * Specific Root Length	اثر متقابل مجذور درصد کلونیزاسیون ریشه و طول ویژه ریشه
(Specific Root Length) ^2 * Dry Matter Yield	اثر متقابل مجذور درصد کلونیزاسیون ریشه و عملکرد ماده خشک

دلیل این که با وجود مدل رگرسیونی گام به گام پس رونده، چندین الگوریتم یادگیری ماشین مورد استفاده و ارزیابی قرار گرفتند این است که، مدل های رگرسیونی بصورت جعبه سیاه عمل کرده و توانایی کشف روابط پیچیده و غیرخطی را ندارند. علاوه براین، مدل های رگرسیونی همواره با خطر بیش برازش[4] و کم برازش[5] مواجه هستند.

الگوریتم ها، تنظیمات و پارامترها

به منظور شناسایی الگوریتم های متناسب با مجموعه داده، ابتدا پانزده الگوریتم (چهارده الگوریتم مستقل و یک الگوریتم ترکیبی) مورد بررسی و ارزیابی های اولیه قرار گرفتند. به این ترتیب که، ابتدا هایپرپارامترهای این الگوریتم ها تنظیم شدند و سپس بر روی داده ها آموزش داده شدند. در ادامه، هایپرپارامترها[6] از طریق روش هایی شامل تعیین اعتبار متقابل، گرید سرچ و آزمون و خطا بهینه شدند. در نهایت، خروجی این الگوریتم ها مورد وارسی و ارزیابی قرار گرفت. در ادامه، به سبب عدم تناسب روش شناختی برخی الگوریتم ها با ماهیت مجموعه داده، شش الگوریتم

[1] iterative
[2] Multicolinearity
[3] Variance Inflation Factor (VIF)
[4] overfitting
[5] underfitting
[6] تنظیماتی هستن که قبل از آموزش مدل باید توسط کاربر تعیین شوند (و مدل بر خلاف پارامترها، آنها را از داده ها یاد نمی‌گیرد)، مثلاً تعداد لایه‌ها یا نورون‌ها در شبکه عصبی، نرخ یادگیری(Learning Rate) ، یا تعداد درخت‌ها در رندوم فورست.

شامل CNN (Convolutional Neural Network, LSTM (Long Short-Term Memory), Random Forest, CatBoost, TabNet, K-NN (K-Nearest Neighbors) حذف شدند. الگوریتم ترکیبی مورد بررسی که حاصل تلفیق دو الگوریتم ترنسفورمر و اس وی آر بود، به دلیل مشابهت معیارهای ارزیابی آن با تک تک الگوریتم های سازنده آن، از مجموعه مورد بررسی حذف شد و تجزیه و تحلیل داده ها با هشت الگوریتم که شرح آنها در ادامه می آید، انجام گرفت.

۱. سیستم استنتاج فازی عصبی تطبیقی Adaptive Neuro-Fuzzy Inference System (ANFIS)

معماری: یک مدل ترکیبی با TensorFlow شامل:

- لایه ورودی با تعداد نورون‌های برابر با تعداد ویژگی‌ها (تعداد ستون‌های X).
- یک لایه مخفی با ۳۲ نورون و فعال‌ساز sigmoid (شبیه‌سازی توابع عضویت).
- یک لایه مخفی با ۱۶ نورون و فعال‌ساز ReLU[1] (قوانین فازی).
- لایه خروجی با ۱ نورون (بدون فعال‌ساز).
- نرخ یادگیری: ۰/۰۰۱ با بهینه‌ساز Adam
- تنظیمات: تعداد epoch[2] ها ۱۰۰ با batch size[3] برابر ۸.

۲. ترنسفورمر Transformer

معماری: ۲ لایه attention با ۲ سر (head) و feed-forward dimension[4] برابر با ۳۲

- نرخ Dropout[5] برابر با ۰/۱
- نرخ یادگیری: ۰/۰۰۱ با بهینه‌ساز AdamW
- تنظیمات: تعداد epoch ها برابر با ۱۵۰ با early stopping patience برابر با ۱۵[6] epoch
- نحوه تنظیم: تنظیم دستی برای تعداد لایه‌ها، نورون‌ها، نرخ یادگیری، تعداد سرها و feed-forward dimension (بدون استفاده از GridSearchCV[7] یا جستجوی تصادفی).

۳. شبکه عصبی مصنوعی Artificial Neural Network (ANN)

معماری: ۳ لایه مخفی با ۶۴، ۳۲ و ۱۶ نورون،

- نرخ یادگیری ۰/۰۰۱، فعال‌ساز ReLU برای لایه‌های مخفی، بدون فعال‌ساز برای لایه خروجی.
- نرخ Dropout برابر با ۰/۲ به منظور کاهش بیش‌برازش،
- آموزش با ۱۰۰ epoch و batch size برابر ۸.
- بهینه‌سازی برای نرخ یادگیری و تعداد لایه‌ها، با استفاده از GridSearchCV

---

[1] یک تابع فعال‌سازی (Rectified Linear Unit) که خروجی نورون را اگر مثبت باشد، همان عدد را برمی‌گرداند، و اگر منفی باشد صفر می‌کند مثلاً: $f(x)=\max(0,x)$ یادگیری را بطور ساده و سریع بهتر می‌کند.

[2] در یادگیری ماشین، اپوک یعنی یک دور کامل که الگوریتم کل داده‌های آموزشی را یک‌بار می‌بیند و پردازش می‌کند. مثلاً اگر ۱۰۰۰ نمونه داده آموزشی داشته باشیم و Batch Sizeبرابر با ۱۰۰ باشد، در هر اپوک مدل ۱۰ بار به‌روزرسانی می‌شود تا کل داده‌ها رو ببیند. بنابراین، تعداد اپوک ها نشان می‌دهد که مدل چند بار کل داده‌ها را مرور کرده است.

[3] تعداد نمونه‌هایی که در هر مرحله از آموزش به مدل داده می‌شود تا پارامترها به‌روزرسانی شوند. مثلاً اگر بچ ساز ۳۲ باشد، یعنی مدل هر بار روی ۳۲ نمونه داده آموزش می‌بیند. اگر بچ سایز کوچیک‌تر باشد حافظه کمتری نیاز است، ولی آموزش ممکن است کندتر انجام گیرد.

[4] اندازه لایه مخفی در یه شبکه عصبی پیش‌خور (Feed-Forward) یعنی در آن لایه ۳۲ نورون وجود دارد که داده‌ها را پردازش می‌کنند.

[5] درصد نورون‌هایی که در هر لایه شبکه عصبی به‌صورت تصادفی در حین آموزش غیرفعال (خاموش) می‌شوند تا از بیش‌برازش (Overfitting) جلوگیری شود، مثلاً نرخ ۰/۰۵ یعنی ۵۰ درصد نورون‌ها بطور تصادفی خاموش می‌شوند.

[6] یعنی اگر مدل به مدت ۱۵ دور آموزش (Epoch) روی داده‌های اعتبارسنجی (Validation) بهبود پیدا نکند (مثلاً خطا کم نشود)، آموزش متوقف می‌شود تا از بیش‌برازش (Overfitting) جلوگیری گردد.

[7] جستجوی شبکه‌ای با اعتبارسنجی متقابلGridSearch Cross Validation : تمام ترکیب‌های ممکن از پارامترهای الگوریتم (مثلاً عمق درخت، تعداد درخت، نرخ یادگیری و غیره) را تولید کرده و آنها را تست می‌کند و در ادامه بهترین آنها را بر اساس دقت (مثلاً R^2) برای متغیر هدف (مثلاً عملکرد دانه) پیدا می‌کند.

٤. **بردار پشتیبان رگرسیون Support Vector Regression (SVR)**

معماری: کرنل RBF (Radial Basis Function)

- پارامترها C=1.0 ، epsilon=0.1
- تنظیمات: مقیاس‌بندی داده‌ها قبل از آموزش
- نحوه تنظیمGridSearchCV : برای C و epsilon

٥. **لایت جی بی ام LightGBM**

- پارامترها: تعداد estimators برابر با ۲۰۰، نرخ یادگیری ۰/۱، حداکثر عمق ۵
- تنظیمات: استفاده از تکنیک early stopping با ۵۰ دور صبر
- نحوه تنظیم: جستجوی تصادفی (Random Search[1]) برای تعداد estimators[2] و عمق.

٦. **ایکس جی بوست Extrem Gradient Boosting (XGBoost)**

- پارامترها: تعداد estimators برابر با ۱۰۰، نرخ یادگیری ۰/۰۵، حداکثر عمق ۴
- تنظیمات: استفاده از regularization[3] (L2) به منظور کاهش بیش‌برازش
- نحوه تنظیم: GridSearchCV برای نرخ یادگیری و تعداد estimators.

۷. **شبکه عصبی عمیق بهبودیافته Enhanced Deep Neoral Network (EDNN)**

معماری: ۵ لایه مخفی با ۱۲۸، ۶۴، ۳۲، ۱۶، و ۸ نورون، تعداد epoch ها ۲۰۰

- استفاده از مکانیزم ReduceLROnPlateau برای تنظیم نرخ یادگیری لایه‌های مخفی و خطی برای خروجی
- نرخ Dropout برابر با ۰/۲ برای جلوگیری از بیش‌برازش
- تنظیمات: استفاده از Batch Normalization[4] به منظور پایداری یادگیری، استفاده از بهینه‌ساز Adam[5] (نرخ یادگیری ۰/۰۰۰۱)، توقف زودهنگام Early stopping با صبر ۵۰ دوره
- تنظیم هایپرپارامترها: با GridSearchCV برای نرخ یادگیری و تعداد نورون‌ها

۸. **رگرسیون بردار پشتیبان Support Vector Machine (SVM)**

معماری: کرنل خطی (linear) با پارامتر C=1.0

- تنظیمات: مقیاس‌بندی داده‌ها (StandardScaler) قبل از آموزش. مدل با داده‌های استانداردشده آموزش دید.
- نحوه تنظیم: تنظیم دستی برای کرنل و مقدار C، بدون استفاده از GridSearchCV یا جستجوی تصادفی.

ارزیابی و تعیین اعتبار

---

[1] با انتخاب تصادفی اما هدفمند، امکان بهینه‌سازی سریع‌تر را در مسائل با فضای پارامتر گسترده فراهم می‌سازد.

[2] مدل‌ها یا الگوریتم‌هایی که در یادگیری ماشین برای پیش‌بینی یا برآورد کردن استفاده می‌شوند. مثلاً در یک جنگل تصادفی، هر درخت تصمیم یک برآوردگر است، در Scikit-learn، برآوردگرها اشیایی هستند که داده را می‌گیرند و با متدهایی مثل fit و predict کار می‌کنند.

[3] تکنیکی در یادگیری ماشین برای جلوگیری از بیش‌برازش است که با اضافه کردن یک جریمه به تابع خطا (Loss Function)، مدل را ساده‌تر نگه می‌دارد. مثلاً در رگرسیون، L1 (Lasso) یا L2 (Ridge) وزن‌های مدل را محدود می‌کنند تا خیلی پیچیده نشود.

[4] به اختصار BatchNorm یک تکنیک در یادگیری ماشین است که برای بهبود سرعت و پایداری آموزش شبکه‌های عصبی استفاده می‌شود. به طور خلاصه: در هر لایه شبکه عصبی، خروجی‌ها را در هر دسته (Batch) نرمال‌سازی می‌کند، یعنی میانگین آنها را نزدیک صفر و واریانس را نزدیک یک می‌کند. سپس، با دو پارامتر (مقیاس و جابجایی) که یاد گرفته می‌شوند، خروجی را تنظیم می‌کند. این کار باعث می‌شود که گرادیان‌ها پایدارتر بشوند، آموزش سریع‌تر انجام گیرد و مدل کمتر به مقادیر اولیه حساس باشد. معمولاً در شبکه‌های عمیق مثلCNN ها قبل یا بعد از لایه‌های فعال‌ساز مثل ReLU استفاده می‌شود. استفاده از بچ نورمالیزاسیون کاهش مشکلاتی مثل ناپدید شدن گرادیان، کاهش بیش‌برازش، و امکان استفاده از نرخ یادگیری بالاتر را به دنبال دارد.

[5] یک الگوریتم بهینه‌سازی مبتنی بر گرادیان که ترکیبی از مومنتوم و تطبیق نرخ یادگیری است. این الگوریتم، به‌جای یک نرخ یادگیری ثابت، برای هر پارامتر نرخ یادگیری را به‌صورت تطبیقی تنظیم می‌کند. سریع و کارآمد است و برای شبکه‌های عصبی عمیق مورد استفاده قرار می گیرد.

عملکرد مدل‌ها با R^2 (ضریب تعیین)، RMSE (جذر میانگین مربعات خطا)، MAE (میانگین قدرمطلق خطا) و Willmott's d (شاخص توافق ویلموت) سنجیده شد. به منظور تحلیل خطاها، از آزمون شاپیرو–ویلک[1] برای بررسی نرمال بودن باقیمانده‌ها استفاده شد.

نرم افزارهای مورد استفاده

تحلیل ها با استفاده از پایتون نسخه ۳.۹ در محیط Spyder (Scientific PYthon Development EnviRonment) 6.0.4 انجام شد و از کتابخانه‌هایی مانند TensorFlow برای ساخت و آموزش مدل ها، scikit-learn برای پیش‌پردازش داده‌ها و تقسیم‌بندی آن‌ها، وMatplotlib و Pandas برای تولید نمودارها بهره گرفته شد. استانداردسازی داده‌ها با استفاده از StandardScaler از کتابخانه scikit-learn انجام شد تا ویژگی‌ها میانگین صفر و انحراف معیار یک داشته باشند، که این موضوع برای آموزش مؤثر شبکه های عصبی ضروری است. تفسیر خروجی های الگوریتم ها و رسم نمودارهای وابستگی جزئی و نمودارهای تصمیم با استفاده از کتابخانه SHAP (Shapley Additive exPlanations) صورت گرفت.

نتایج و بحث

به منظور توسعه یک مدل رگرسیون خطی برای پیش‌بینی عملکرد دانه در گیاه ذرت، از روش رگرسیون گام‌به‌گام حذفی پس‌رونده با سطح اهمیت حذف[2] ۰/۰۵ استفاده شد. ابتدا مجموعه داده‌ای با ۷۳ ویژگی شامل متغیرهای فیزیولوژیکی، مورفولوژیکی، و شیمیایی خاک و گیاه جمع‌آوری شد. برای کاهش هم‌خطی‌چندگانه و بهبود تعمیم‌پذیری مدل، مدل های مختلف ارزیابی و مقایسه شدند تا این که مدل نهایی با ۱۳ ویژگی انتخاب شد (جدول ۲). این مدل ضریب تعیین تعدیل‌شده[3] برابر ۶۴/۲۰ درصد و ضریب تعیین پیش‌بینی[4] برابر ۵۱/۰۶ درصد را به دست آورد، که نشان‌دهنده دقت قابل قبول و تعمیم‌پذیری مناسب مدل است. معادلهٔ رگرسیونی و ضرایب ویژگی ها در معادلهٔ ۱ نشان داده شده است. ضرایب ویژگی ها به همراه اشتباه استاندارد ضرایب، پی ولیو و مقدار عامل تورم واریانس هر ویژگی و ضریب تبیین مدل به ترتیب در جدول های ۲ و ۳ نشان داده شده است. از آنجا که قبل از تحلیل، داده ها استاندارد شدند بنابراین، ضرایب رگرسیونی قابل مقایسه با یکدیگر بوده و میزان اهمیت هر متغیر در پیش بینی عملکرد دانه را نشان می دهند.

معادله ۱: Equation 1:

$$\begin{aligned}\text{Seed Yield} = {} & 0.0000 + 0.2451\ \text{Canopy Temp\_3} - 0.488\ \%\text{ P plant} + 0.395\ \%\text{N soil} \\ & + 0.3387\ \text{PlantHeight\_CobNumber} + 0.298\ \text{Nplant\_SpecificRootLength} \\ & + 0.479\ \text{Leaf Area Index\_SPAD\_Mean} - 0.911\ \text{Leaf Area Index\_Dry Matter Yield} \\ & + 0.669\ \text{Root Colonization\_Dry Matter Yield} + 0.642\ \text{Stem Diameter\_Dry Matter Yield} \\ & + 0.742\ (\text{Specific Root Length})\ \hat{}\ 2 + 0.453\ (\text{Canopy Temp\_Mean})\hat{}2*\text{Leaf Area Index} \\ & - 0.643\ (\text{Root Colonization})\hat{}2*\text{SpecificRootLength} - 0.531\ (\text{Specific Root Length})\hat{}2*\text{Dry Matter Yield}\end{aligned}$$

جدول ۲. ضرایب ویژگی ها به همراه اشتباه استاندارد ضرایب، پی ولیو و مقدار عامل تورم واریانس[5] هر ویژگی و ضریب تبیین مدل

Table 2. Feature coefficients with standard errors, p-values, and variance inflation factors for each feature and the model's coefficient of determination

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.0000	0.0657	0.00	1.000	
Canopy Temp_3	0.2451	0.0757	3.24	0.002	1.31
% P plant	-0.488	0.159	-3.07	0.003	5.77
%N soil	0.395	0.167	2.36	0.021	6.40

[1] Shapiro-Wilk
[2] Alpha to remove
[3] R-squared adjusted
[4] R-squared predicted
[5] Variance Inflation Factor (VIF)

PlantHeight_CobNumber	0.3387	0.0824	4.11	0.000	1.56
Nplant_SpecificRootLength	0.298	0.132	2.27	0.026	3.97
Leaf Area Index_SPAD_Mean	0.479	0.130	3.69	0.000	3.86
Leaf Area Index_Dry Matter Yield	-0.911	0.349	-2.61	0.011	27.88
Root Colonization_Dry Matter Yield	0.669	0.263	2.54	0.013	15.86
Stem Diameter_Dry Matter Yield	0.642	0.252	2.55	0.013	14.53
(Specific Root Length) ^2	0.742	0.320	2.32	0.023	23.40
(Canopy Temp_Mean)^2*Leaf AreaIndex	0.453	0.193	2.35	0.021	8.50
(Root Colonization)^2*SpecificRootLength	-0.643	0.248	-2.59	0.011	14.09
(Specific Root Length)^2*Dry Matter Yield	-0.531	0.250	-2.12	0.037	14.35

جدول ۳. ضریب تبیین تعدیل شده و پیش بینی مدل رگرسیونی گام به گام در پیش بینی عملکرد دانه ذرت

Table 3. Adjusted and predicted coefficients of determination for the stepwise regression model in predicting corn grain yield

S	R-sq	R-sq(adj)	R-sq(pred)
0.643993	64.20%	58.53%	51.06%

ضرایب همبستگی بین ویژگی های تشخیص داده شده توسط مدل رگرسیونی

ضرایب همبستگی بین ۱۳ ویژگی برتر تشخیص داده شده توسط مدل گام به گام پس رونده، به همراه دندروگرام حاصل از خوشه بندی سلسله مراتبی متغیرها و نمودار پراکندگی و توزیع‌های تخمین چگالی هسته‌ای[1] آنها در شکل ۱ نشان داده شده است. این شکل، الگوها، گروه‌بندی‌ها و وابستگی‌های بین متغیرها را به خوبی نشان می‌دهد.

کلاسترینگ سلسله‌مراتبی (دندروگرام)

ترتیب کلاسترینگ (از چپ به راست در دندروگرام) شباهت در الگوهای همبستگی را منعکس می‌کند، به‌طوری که ویژگی‌های نزدیک به هم در یک خوشه گروه‌بندی شده اند. دندروگرام حاصل، ویژگی‌ها را در دو خوشه، مطابق دانش اکوفیزیولوژی گروه‌بندی کرد:

- خوشه ۱ : شامل SPAD_Mean، SPAD_2 و Nplant_SpecificRootLength است. این متغیرها که به خصوصیات فیزیولوژیکی گیاه (مثل محتوای کلروفیل و ویژگی‌های ریشه) مربوط هستند، ارتباط قوی‌ای با یکدیگر دارند.
- خوشه ۲ : شامل Dry Matter Yield، Leaf Area Index_SPAD_Mean و Seed Yield است. این خوشه متغیرهای مرتبط با عملکرد ماده خشک، محتوای کلروفیل برگ ها و شاخص سطح برگ را گروه‌بندی می‌کند و وابستگی متقابل آنها رو مورد تأکید قرار می دهد.

تحلیل ضرایب همبستگی (نیمه بالای ماتریس)

ضرایب همبستگی و مقادیر p-value روابط معنی‌دار را نشان می دهند. بر این اساس، همبستگی ها در سه گروه به شرح زیر قابل تفسیر هستند:

- همبستگی‌های مثبت قوی: بین محتوای کلروفیل در مرحله خروج برگ پرچم (SPAD_2) و قطر بلال و درصد فسفر گیاه در انتهای فصل رشد، همبستگی های قوی وجود داشت (به ترتیب r=0.82 و r=0.80). SPAD_Dry Matter Yield و Leaf Area Index_SPAD_Mean: r = 0.37، $p < 0.001$. این همبستگی مثبت متوسط، تولید زیست‌توده را با سطح برگ و محتوای کلروفیل مرتبط می‌کند.
- همبستگی‌های منفی قوی: Canopy Temp_Mean و % Root Colonization: r = -0.82، $p < 0.001$. حاکی از آن است که دماهای بالاتر سایه انداز گیاهی به شدت با کاهش کلونیزاسیون ریشه مرتبط است، که احتمالاً به دلیل تنش گرمایی بر کارکرد کلی گیاه از جمله کاهش میزان فتوسنتز و تولید آسمیلات ها و به دنبال آن کاهش تراوشات ریشه ای مؤثر بر فعالیت و جمعیت های

[1] kernel density estimation (KDE)

ریزجانداران ساکن ریزوسفر، تأثیر می‌گذارد. در منابع متعدد به نقش قارچ های میکوریزا آربوسکولار در بهبود روابط آب–خاک–گیاه و به دنبال آن میزان تبخیر و تعرق اشاره شده است (Bao et al., 2022, Bhupenchandra et al., 2024; Brar et al., 2024).

- همبستگی‌های متوسط: وجود همبستگی بین Nplant_SpecificRootLength و SPAD_2 ($r = 0.60$ ، $p < 0.001$.) که نشان‌دهنده رابطه منطقی بین طول مخصوص ریشه (مورفولوژی ریشه) و محتوای کلروفیل می باشد. افزایش طول مخصوص ریشه و نیز محتوای کلروفیل برگ به عنوان نتیجه کاربرد ریزوباکتری های تحریک کننده رشد گیاه، از سوی محققان متعدد مورد تأکید قرار گرفته است (Bhupenchandra et al., 2024).

نمودارهای KDE در قطر(توزیع چگالی احتمال)

نمودارهای KDE با دو خوشه رنگی نشان داده شده اند، که کلاسترینگ تجمعی با n_clusters=2 را منعکس می‌کنند و داده‌ها را به زیرگروه‌های معنی‌دار تقسیم کرده اند.

- توزیع‌های چندوجهی: ویژگی‌هایی مثل Canopy Temp_Mean و Leaf Temp چندین قله دارند، که نشان‌دهنده وجود زیرگروه‌های متمایز (شامل تیمارها و شرایط آزمایشی مختلف) در داده‌ها است.
- توزیع‌های اریب: ویژگی هایی مثل Seed Yield و Dry Matter Yield توزیع‌های اریب دارند، که نشان‌دهنده غیرخطی بودن روابط این داده‌ها است و حاکی از آن است که برای مدل‌سازی نیاز به مدل های قوی تر (مانند مدل های یادگیری ماشین) می باشد.

نمودارهای پراکندگی در نیمه پایینی ماتریس (روابط زوجی)

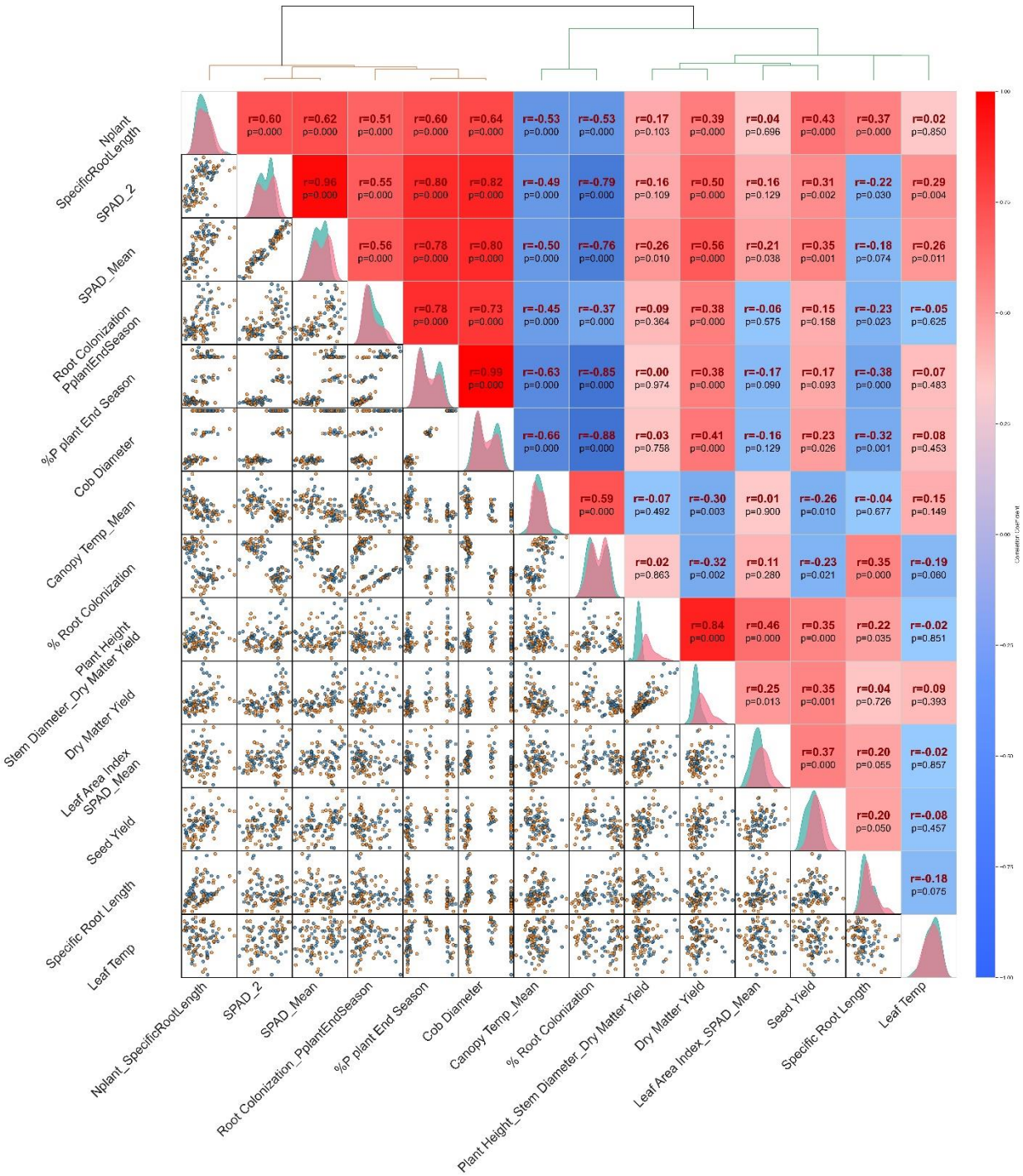
- روندهای خطی قوی: نمودار پراکندگی بین SPAD_Mean و SPAD_2 یک رابطه خطی قوی را نشان می‌دهد، که با همبستگی بالای این دو ویژگی سازگار است.
- الگوهای پراکنده: نمودار پراکندگی بین Canopy Temp_Mean و Seed Yield پراکندگی بیشتری دارد، که نشان‌دهنده یک رابطه ضعیف‌تر یا غیرخطی هست (با وجود همبستگی منفی).

نتیجه‌گیری و توصیه‌ها

- **تأثیرات مثبت**: Dry Matter Yield و Leaf Area Index_SPAD_Mean همبستگی‌های مثبتی با عملکرد دانه دارند، که نشان می‌دهد زیست‌توده بالاتر و سطح برگ/ محتوای کلروفیل بیشتر، به افزایش تولید دانه کمک می‌کند.
- **تأثیرات منفی**: Canopy Temp_Mean همبستگی منفی با عملکرد دانه دارد، که حاکی از آن است دماهای بالای سایه انداز (احتمالاً به دلیل تنش گرمایی) ممکن است عملکرد را کاهش دهد.
- **روابط ضعیف**: متغیرهایی مثل %P plant End Season و RootColonization_PplantEndSeason همبستگی‌های ضعیف‌تری با عملکرد دانه دارند، که نشان‌دهنده تأثیر مستقیم و محدود روی عملکرد می باشد.

به طور کلی، نمودار همبستگی زوجی با کلاسترینگ، روابط پیچیده‌ای بین ۱۳ ویژگی را آشکار می‌کند:

- انتخاب ویژگی: متغیرهایی با همبستگی بالا)مثل SPAD_Mean، Dry Matter Yield) گزینه‌های مناسبی برای مدل‌سازی پیش‌بینی عملکرد دانه هستند.
- بینش‌های خوشه‌ای: دو خوشه اصلی، خصوصیات فیزیولوژیکی (خوشه ۱) را از معیارهای مرتبط با عملکرد (خوشه ۲) جدا می‌کنند، که این موضوع می تواند به نوبه خود زمینه لازم برای تحلیل‌ها یا مداخلات هدفمند (مدیریت اگرواکولوژیک از جمله نوع و میزان نهاده) را فراهم آورد.
- ملاحظات مدل‌سازی: توزیع‌های اریب و الگوهای چندوجهی نشان می‌دهند که برای توضیح دقیق این روابط، نیاز به تبدیل داده‌ها یا مدل‌های غیرخطی هست.

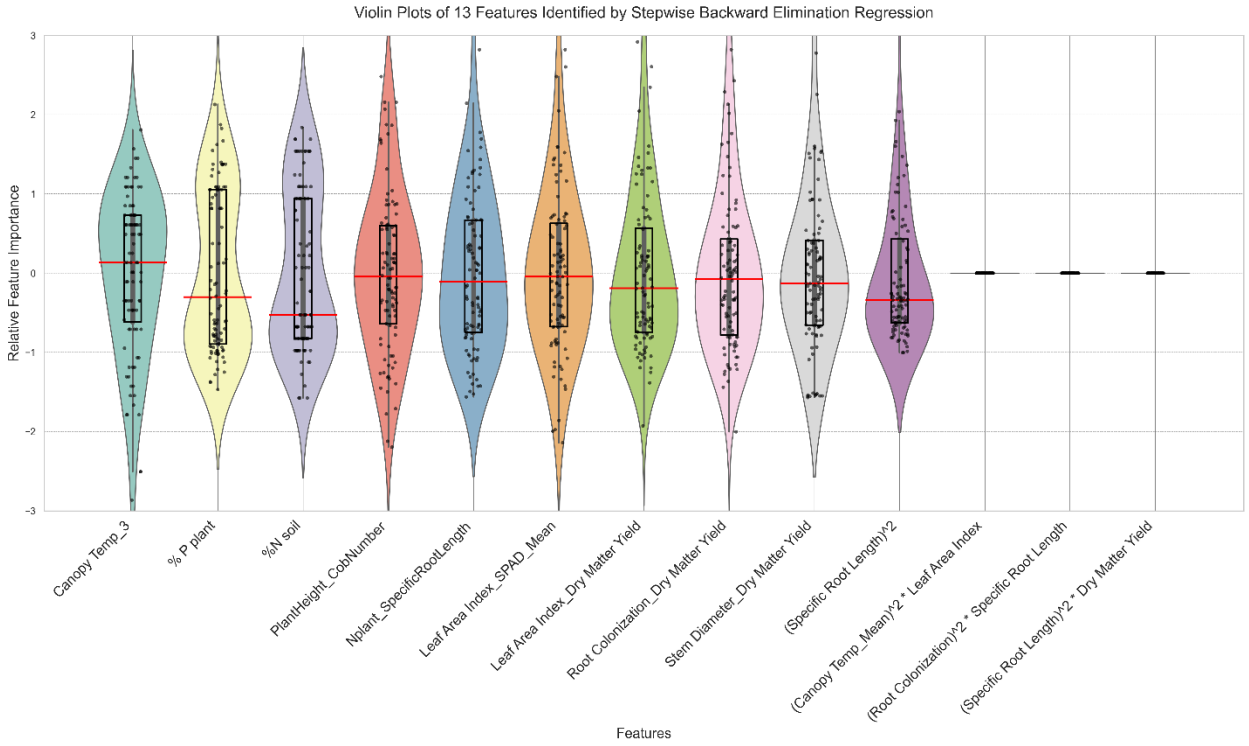


شکل ۱. ماتریس ضرایب همبستگی، نمودار توزیع چگالی و دندروگرام بین سیزده ویژگی مؤثر بر عملکرد دانه ذرت

Figure 1. Matrix of correlation coefficients, kernel density estimation plots, and dendrogram among thirteen features influencing corn grain yield

نمودار شیپ ویولونی اهمیت نسبی سیزده متغیر تشخیص داده شده توسط مدل رگرسیونی

شکل ۲، نمودار ویولن‌پلات توزیع و اهمیت نسبی سیزده ویژگی کلیدی شناسایی‌شده توسط مدل رگرسیونی گام به گام پس رونده، را برای پیش‌بینی عملکرد دانه ذرت نشان می‌دهد. این ویژگی‌ها شامل متغیرهای فیزیولوژیکی و مورفولوژیکی هستند که پراکندگی داده‌ها و میانگین اهمیت آن‌ها نشان داده شده است. محور عمودی نشان‌دهنده اهمیت نسبی ویژگی‌هاست و محور افقی نام ویژگی‌ها را نشان می‌دهد. ویژگی‌هایی مانند"Canopy Temp_3" ، "PlantHeight_CobNumber"و "Root Colonization_Dry Matter Yield" دارای پراکندگی گسترده‌تری هستند که بیانگر تنوع بیشتر تأثیر آن‌ها در مدل‌های مختلف است. همچنین، میانه داده های هر ویژگی با خط قرمز مشخص شده که نشان می‌دهد برخی ویژگی‌ها مانند "%N soil" و "Root Colonization_Dry Matter Yield" دارای پراکنش بیشتر داده های با مقادیر منفی هستند.



شکل ۲. نمودار شیپ ویولونی اهمیت نسبی سیزده متغیر تشخیص داده شده توسط مدل رگرسیونی گام به گام

Figure 2. SHAP violin plot shows relative importance of 13 features identified via Stepwise Elimination Regression

برای سه ویژگی ترکیبی (انتهای سمت راست نمودار)، شامل (Canopy Temp_Mean)^2 * Leaf Area Index, (Root Colonization)^2 * Specific Root Length و(Specific Root Length)^2 * Dry Matter Yield ، مشاهده می‌شود که ویولن‌پلات‌ها به صورت خطوط افقی صاف رسم شده‌اند. این پدیده ناشی از پراکندگی بسیار کم یا عدم وجود داده‌های معنی‌دار در این ویژگی‌ها پس از استانداردسازی است. در فرآیند رگرسیون گام به گام حذفی، این فیچرها ممکن است به دلیل همبستگی ضعیف با متغیر هدف (عملکرد دانه ذرت) یا داده‌های ناقص حذف شده باشند. با این حال، برای حفظ یکپارچگی تحلیل و نمایش کامل سیزده فیچر انتخاب‌شده، این موارد با مقادیر صفر پر شده‌اند، که منجر به شکل‌گیری ویولن‌های تک‌بعدی شده است.

ارزیابی دقت و توان الگوریتم های یادگیری ماشین در پیش بینی عملکرد دانه ذرت

در حوزه علوم داده و یادگیری ماشین، مفاهیم ارزیابی مدل، اعتبارسنجی مدل، دقت مدل و توان مدل از جمله اصطلاحات کلیدی هستند که علیرغم ارتباط مفهومی، دارای تمایزات ساختاری و کارکردی مشخصی می‌باشند. ارزیابی مدل[1]، فرآیند جامع سنجش عملکرد مدل با به‌کارگیری معیارهای کمّی مختلف (نظیر R^2 ، RMSE، MAE) بر روی مجموعه داده‌های آزمایشی است و هدف های آن تعیین قابلیت تعمیم‌پذیری مدل در شرایط عملیاتی، و سنجش اثربخشی مدل در حل مسئله هدف می باشد. ارزیابی، با تمامی مفاهیم دیگر مرتبط است، چرا که چارچوبی جامع برای سنجش کیفیت مدل ارائه می‌دهد و دربرگیرنده سنجه‌هایی مانند دقت مدل می‌باشد. ارزیابی، رویکردی کل‌نگر دارد و به یک جنبه خاص محدود نمی‌شود و عموماً در مرحله نهایی توسعه مدل اجرا می‌شود. اعتبارسنجی[2]، فرآیند نظام‌مند تنظیم هایپرپارامترها و ارزیابی مقدماتی مدل بر روی مجموعه داده‌های اعتبارسنجی به منظور جلوگیری از بروز پدیده بیش‌برازش می باشد. اهداف آن، بهینه‌سازی ساختار مدل پیش از ارزیابی نهایی و اطمینان از تعادل میان پیچیدگی و تعمیم‌پذیری مدل است. همچون ارزیابی مدل، از داده‌های مستقل استفاده می‌کند (ممکن است معیارهای دقت را در فرآیند خود به کار گیرد). بر خلاف ارزیابی، اعتبارسنجی ماهیتاً فرآیندی تکرارشونده و اصلاحی است و تمرکز اصلی آن بر تنظیم مدل و نه قضاوت نهایی، است. دقت مدل[3]، معیار کمّی خاصی که میزان انطباق پیش‌بینی‌های مدل با مقادیر واقعی را اندازه‌گیری می‌کند. در مسائل رگرسیون عموماً از R^2 ، RMSE و در مسائل طبقه‌بندی از Accuracy استفاده می‌شود. هدف، سنجش کارایی پیش‌بینی مدل و ارائه معیاری عینی برای مقایسه مدل‌های مختلف است. زیرمجموعه‌ای از فرآیند ارزیابی مدل محسوب می‌شود و ممکن است در فرآیند اعتبارسنجی نیز مورد استفاده قرار گیرد. دقت، نگاهی تک‌بعدی به عملکرد مدل دارد و فاقد جنبه‌های فرآیندی دیگر مفاهیم است. توان مدل[4]، عبارت است از ظرفیت ذاتی مدل در یادگیری الگوهای پیچیده و انطباق با داده‌ها که تابعی از پارامترهای ساختاری نظیر تعداد لایه‌ها، نورون‌ها و عمق شبکه می‌باشد. اهداف آن شامل تعیین پتانسیل مدل در حل مسائل پیچیده و ایجاد تعادل میان انعطاف‌پذیری و خطر بیش‌برازش است. با معماری مدل در ارتباط است و بر معیارهای عملکردی مانند دقت تأثیر مستقیم دارد. بطور کلی، توان مدل ماهیتاً توصیف‌کننده ویژگی‌های ذاتی مدل است و برخلاف سایر مفاهیم، بیشتر متمرکز بر پتانسیل مدل است تا عملکرد آن.

ضریب تبیین، جذرمیانگین مربعات خطا، قدر مطالق خطا و شاخص توافق هشت مدل مورد بررسی در جدول ۴ نشان داده شده اند. سه الگوریتم آنفیس، ترنسفورمر و ای ان ان، با ضریب تبیین بیشتر از ۰/۵۰ به عنوان الگوریتم های برتر در پیش بینی عملکرد دانه ذرت شناخته شدند. مور و همکاران (Moore et al., 2013) دامنه زیر را برای تفسیر ضریب تبیین (شاخصی از نکویی برازش[5]) پیشنهاد کرده اند: $R^2 < 0.25$: بسیار ضعیف؛ $0.25 \le R^2 < 0.50$: ضعیف؛ $0.50 \le R^2 < 0.75$: متوسط؛ $R^2 \ge 0.75$: قابل توجه. بر اساس راهنمای کوهن و جانسون، ضرایب تبیین مساوی یا بزرگتر از ۰/۲۶ در محدودهٔ قابل توجه[6] قرار می گیرند. اگرچه ضریب تعیین به‌طور گسترده‌ای به‌عنوان شاخصی از برازش مدل در رگرسیون خطی و چندگانه به‌کار می‌رود، اما تفسیر آن در زمینه یادگیری ماشین – به‌ویژه با الگوریتم‌های غیرخطی باید با احتیاط صورت گیرد. دلیل این امر آن است که R^2 لزوماً ظرفیت پیش‌بینی یا توان تعمیم مدل‌های پیچیده را به‌طور کامل نشان نمی‌دهد و در برخی موارد ممکن است خوش‌بینانه یا گمراه‌کننده باشد، به‌ویژه در شرایطی با ابعاد بالا یا بیش‌برازش. بنابراین توصیه می‌شود که علاوه بر R^2 ، از شاخص‌هایی مانندRMSE ، MAE یا اعتبارسنجی متقابل برای ارزیابی جامع‌تر عملکرد مدل استفاده گردد. (Kuhn and Johnson, 2013).

جدول ٤. ضریب تبیین، جذرمیانگین مربعات خطا، قدر مطالق خطا و شاخص توافق هشت مدل یادگیری ماشین در پیش بینی عملکرد دانه ذرت

Table 4. Coefficient of determination, root mean square error, mean absolute error, and Willmott's d index for eight machine learning models to predict corn grain yield

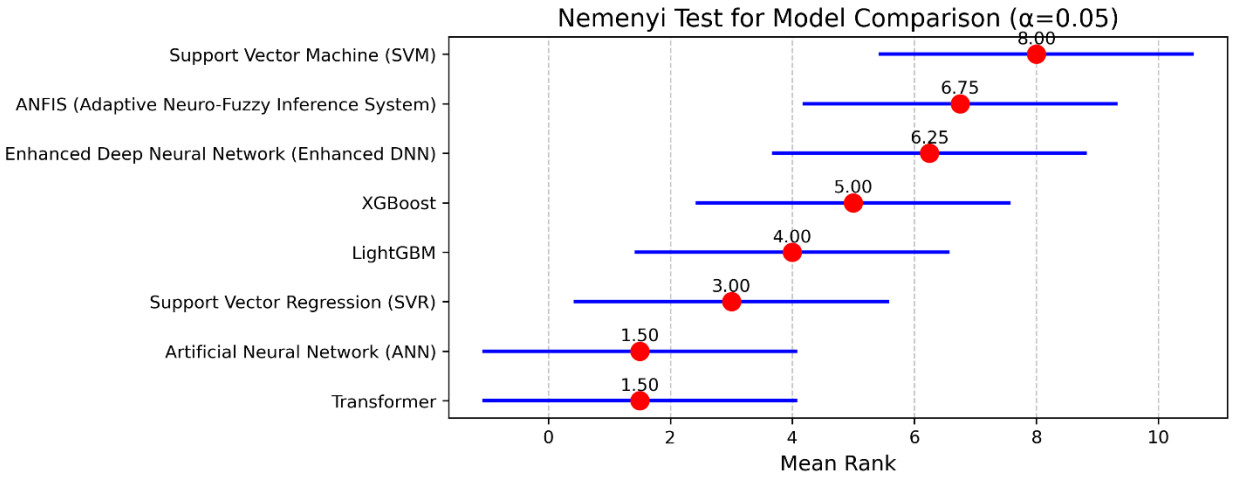
Algorythm	R^2	RMSE	MAE	Willmotte'd
ANFIS (Adaptive Neuro-Fuzzy Inference System)	0.555	2.804	2.368	0.826

[1] Model Evaluation
[2] Model Validation
[3] Model Accuracy
[4] Model Capacity
[5] goodness-of-fit
[6] Substantial

Transformer	0.545	2.834	2.331	0.837
Artificial Neural Network (ANN)	0.518	2.917	2.254	0.848
Support Vector Regression (SVR)	0.452	3.111	2.381	0.758
LightGBM	0.385	3.296	2.617	0.727
XGBoost	0.339	3.417	2.780	0.689
Support Vector Machine (SVM)	0.325	2.96	2.37	0.796
Enhanced Deep Neural Network (Enhanced DNN)	0.321	3.462	3.313	0.674

نمودار نمنیی[1]

شکل ۳ نمودار نمنیی برای مقایسه هشت مدل یادگیری ماشین را بر اساس معیارهای R^2 ، Willmott's d، MAE، RMSE نشان می دهد. محور افقی میانگین رتبه[2] هر مدل را نشـــان می‌دهد. رتبه کمتر یعنی عملکرد بهتر (زیرا معیارهای مثل R^2 و Willmott's d که هر چه بالاتر باشـند بهتر هسـتند، رتبه پایین‌تر می‌گیرند، و معیارهایی مثل MAE و RMSE که هر چه پایین‌تر باشـند بهتر هسـتند، نیز در رتبه پایین‌تر قرار می‌گیرند). در محور عمودی اسم مدل‌ها نشان داده شده است، از بهترین (بالا) به ضعیف‌ترین (پایین). خطوط افقی آبی فاصله بحرانی[3] را نشان می‌دهند. اگر خط افقی مربوط به دو مدل همپوشانی نداشته باشند، تفاوت عملکرد آنها از نظر آماری معنی‌دار است (باα=0.05).



شکل ۳. نمودار نمنیی مقایسه هشت الگوریتم یادگیری ماشین در پیش بینی عملکرد دانه ذرت

Figure 3. Nemeni plot comparing eight machine learning algorithms in predicting corn grain yield

تفسیر نمودار نمنیی

۱. **رتبه‌بندی کلی:** مدل ANN با میانگین رتبه ۱/۵۰ و شاخص دی برابر ۰/۸۴۸ بهترین عملکرد را نشان می‌دهد، که نشان‌دهنده توافق بالای پیش‌بینی‌ها با داده‌های واقعی است.

- مدل Transformer با ضریب تبیین ۰/۵۴۵ و میانگین رتبه ۱/۵۰ نیز عملکرد بسیار خوبی دارد.
- مدل ANN هم‌رتبه با Transformer اســت، که نشــان می‌دهد عملکرد آن خیلی به Transformer نزدیک بوده و تفاوت آماری معنی‌داری با آن ندارد (چون خطوط آنها همپوشانی دارند).
- مدل SVR (با میانگین رتبه ۳) رتبه سوم را به خود اختصاص داد، اگر چه که از ANN و Transformer کمی ضعیف‌تر است، با اینحال خوب عمل کرده است.

[1] Nemenyi diagram
[2] Mean Rank
[3] Critical Distance

- LightGBM (4.00) کاهشی قابل‌توجه نشان داد و عملکرد آن از SVR کمتر شد.
- XGBoost (5.00) عملکردی متوسط را از خود نشان داد.
- EnhancedDNN (6.25) نسبت به مدل‌های مبتنی بر شبکه عصبی ساده (مثلANN) ضعیف‌تر عمل کرد.
- ANFIS (6.75) نزدیک به Enhanced DNN است.
- SVM (8.00) با بالاترین میانگین رتبه، ضعیف‌ترین عملکرد را از خود نشان داد، که باضریب تبیین ۰/۳۲۵ و شاخص دی ۰/۷۹۶ هم‌خوانی دارد.

۲. **فاصله بحرانی**

فاصله بحرانی: فاصله بحرانی برابر با ۵/۱۱ نشان می‌دهد که مدل‌هایی که تفاوت میانگین رتبه آن‌ها بیشتر از این مقدار است (مثل ANN و SVM)، تفاوت آماری معنی‌داری با یکدیگر دارند. مدل‌هایی که خط افقی آنها بیشتر از این فاصله از هم جدا شده باشد (مثل Transformerو SVM)، با یکدیگر تفاوت آماری معنی‌دار دارند. Transformer و ANN با یکدیگر همپوشانی دارند، ولیSVM با هیچ‌کدام از مدل‌های بالا همپوشانی ندارد، یعنی ضعیف‌تر بودن آن معنی‌دار است.

۳. **الگوهای مشاهده‌شده**

الگوهای مشاهده‌شده: مدل‌های مبتنی بر شبکه عصبی (ANN، Transformer، Enhanced DNN) به طور کلی عملکرد بهتری نسبت به مدل‌های درختی (LightGBM، XGBoost) و مدل‌های فازی (ANFIS) دارند. این ممکن است به دلیل توانایی آن‌ها در مدل‌سازی روابط پیچیده و غیرخطی در داده‌های ذرت باشد. مدل‌های درختی مثل XGBoost و LightGBM در میانه جدول هستند، که نشان می‌دهد در این دیتاست عملکرد متوسطی دارند.

اس وی آر، جزو هیچ یک ازالگوریتم های های درختی و شبکه عصبی محسوب نمی شود، بلکه یک مدل مبتنی بر Support Vector Machine هست که با کرنل‌ها (مثل RBF) عمل می‌کند.

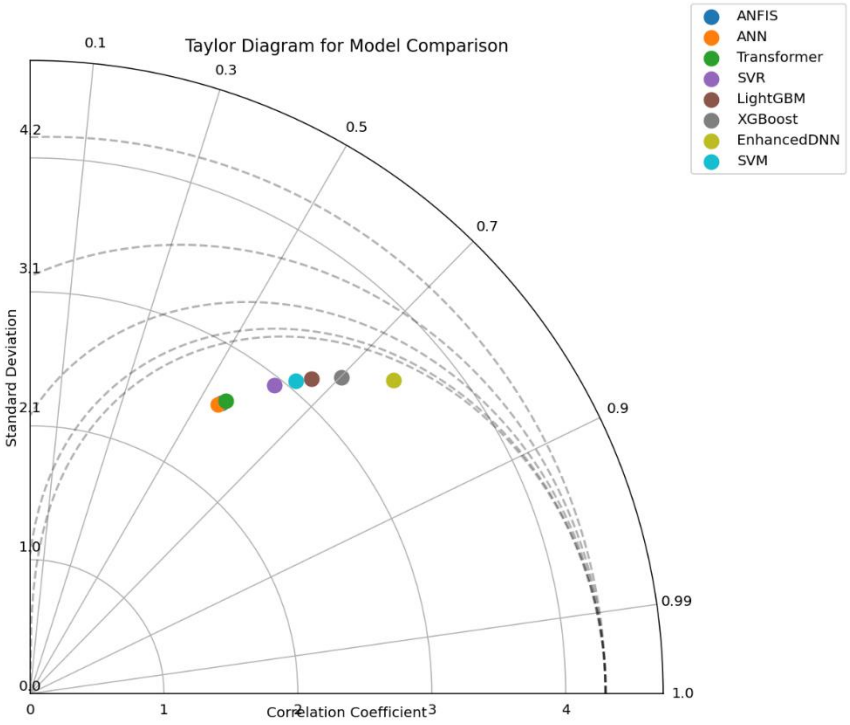
نمودار تیلور[1]

نمودار تیلور (شکل ۴)، هشت مدل یادگیری ماشین مورد بررسی را بر اساس انحراف معیار، ضریب همبستگی، جذرمیانگین مربعات خطا، ضریب تبیین، و شاخص توافق ویلموت مقایسه می‌کند. این نمودار تعادل بین انحراف معیار و ضریب همبستگی را نشان می‌دهد، در حالی که معیارهای RMSE و R^2 دقت پیش‌بینی‌ها را ارزیابی می‌کنند. مدل‌هایی که به نقطه مرجع نزدیک‌تر هستند، عملکرد بهتری دارند (نقطه مرجع در نمودار تیلور، نقطه‌ای با انحراف معیار بالا و ضریب همبستگی نزدیک به یک است که به‌عنوان معیار ایده‌آل عملکرد مدل‌ها در نظر گرفته می‌شود). نمودار تیلور نشان می‌دهد که مدل Transformer با RMSE=2.834 و R^2=0.545 بهترین عملکرد را در میان هشت مدل دارد. مدل‌های ANN (RMSE=2.917 , R^2=0.518)و ANFIS (RMSE=2.884 , R^2=0.529)، در رتبه‌های بعدی قرار دارند. در مقابل، مدل‌های LightGBM (RMSE=3.296 , R^2=0.385)و XGBoost (RMSE=3.417 , R^2=0.339), EnhancedDNN (RMSE=3.462 , ،) (R^2=0.321) و SVM (RMSE=3.442 , R^2=0.329) ضعیف‌ترین عملکرد را نشان می‌دهند.

در ادامه، تحلیل مفصلی از نمودار تیلور ارائه می شود:

۱. **انحراف معیار** :انحراف معیار مدل‌ها در محدوده تقریبی ۲/۵۷ تا ۳/۵۸ قرار دارد. مدل اِی ان ان با انحراف معیار ۲/۵۷۳ کمترین تغییرپذیری را در پیش‌بینی‌های خود نشان می‌دهد، در حالی که مدل‌های آنفیس (۲/۶۰۰) و ترنسفورمر (۲/۶۲۷) نیز انحراف معیار پایینی دارند. در مقابل، مدل انهنسد دی ان ان (۳/۵۸۲) و ایکس جی بوست (۳/۳۱۷) بالاترین انحراف معیار را دارند، که نشان‌دهنده پراکندگی بیشتر خطاها در پیش‌بینی‌ها می باشد.

[1] Taylor Diagram



شکل ٤. نمودار تیلور مقایسه هشت الگوریتم یادگیری ماشین در پیش بینی عملکرد دانه ذرت

Figure 4. Taylor diagram comparing eight machine learning algorithms in predicting corn grain yield

۲. **ضریب همبستگی:** در نمودار تیلور، ضریب همبستگی[1] که روی محور افقی (زاویه) نمایش داده می‌شود، نسبت انحراف معیار مدل به انحراف معیار مرجع[2] است، نه همبستگی بین پیش‌بینی‌های مدل و داده‌های واقعی (مثلاً ۰/۷۵۸ برای انهنسد دی ان ان به معنای عملکرد بهتر نیست، بلکه نشان‌دهنده انحراف معیار بالاتر است). این ویژگی خاص نمودار تیلور است. ضرایب همبستگی مدل‌ها به شرح زیر هستند:

ANFIS: 0.550 , ANN: 0.544 , Transformer: 0.555 , SVR: 0.621 , LightGBM: 0.667 , XGBoost: 0.701 , EnhancedDNN: 0.758 , SVM: 0.648

مدل انهنسد دی ان ان با ضریب همبستگی ۰/۷۵۸ بالاترین مقدار را دارد، اما این به معنای انحراف معیار بالاتر (پراکندگی بیشتر خطاها) است. در مقابل، مدل‌های اِی ان ان (۰/۵۴۴) و ترنسفورمر (۰/۵۵۵) به دلیل ضریب همبستگی پایین‌تر، به نقطه مرجع نزدیک‌ترند و عملکرد بهتری دارند.

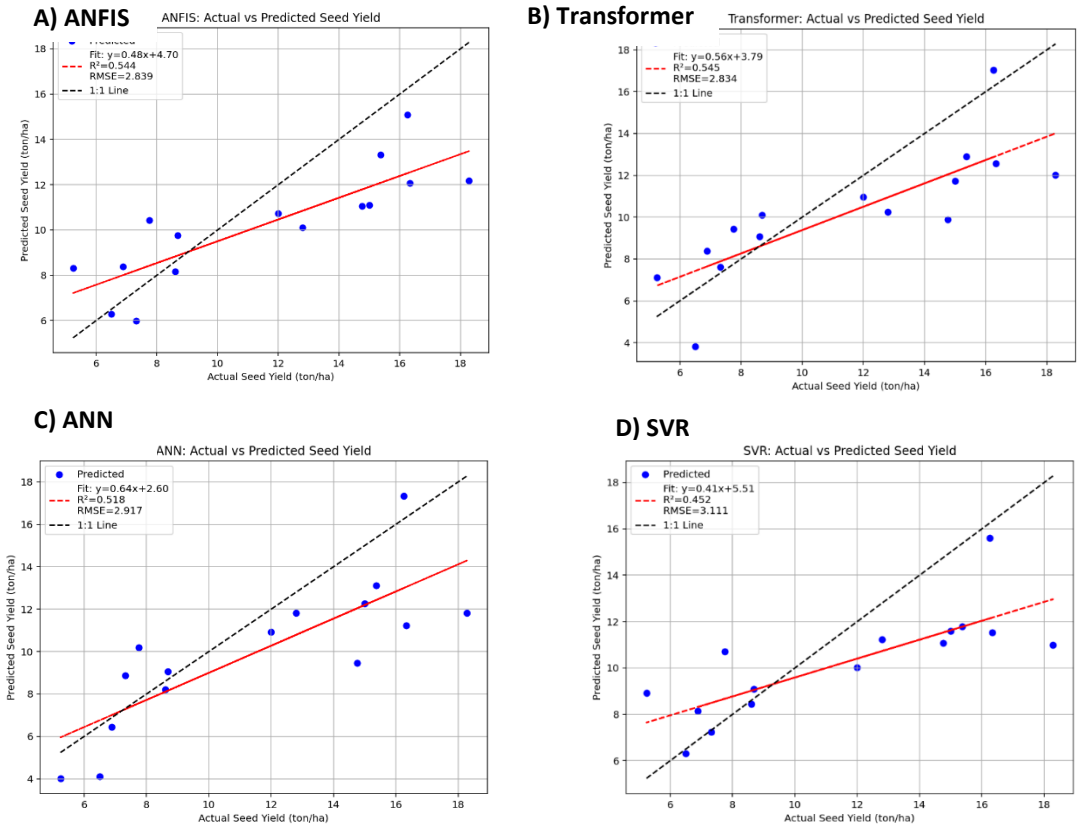
۳. **مقایسه مدل‌ها:** مدل‌های اِی ان ان، آنفیس و ترنسفورمر عملکرد بسیار نزدیکی دارند و بهترین مدل‌ها در این مقایسه‌اند. این مدل‌ها انحراف معیار پایینی دارند (به ترتیب ۲/۵۷۳، ۲/۶۰۰، و ۲/۶۲۷) و جذرمیانگین مربعات خطا پایینی (به ترتیب ۲/۹۱۷، ۲/۸۸۴، و ۲/۸۳۴) دارند. نقطه مربوط به آنفیس در نمودار تیلور به دلیل انحراف معیار نزدیک، زیر نقاط اِی ان ان و ترنسفورمر قرار گرفته است. این مدل‌ها همچنین ضریب تبیین بالایی دارند (ترنسفورمر ۰/۵۴۵، آنفیس ۰/۵۲۹ و ای ان ان ۰/۵۱۸)، که نشان‌دهنده دقت بالای پیش‌بینی‌هاست.

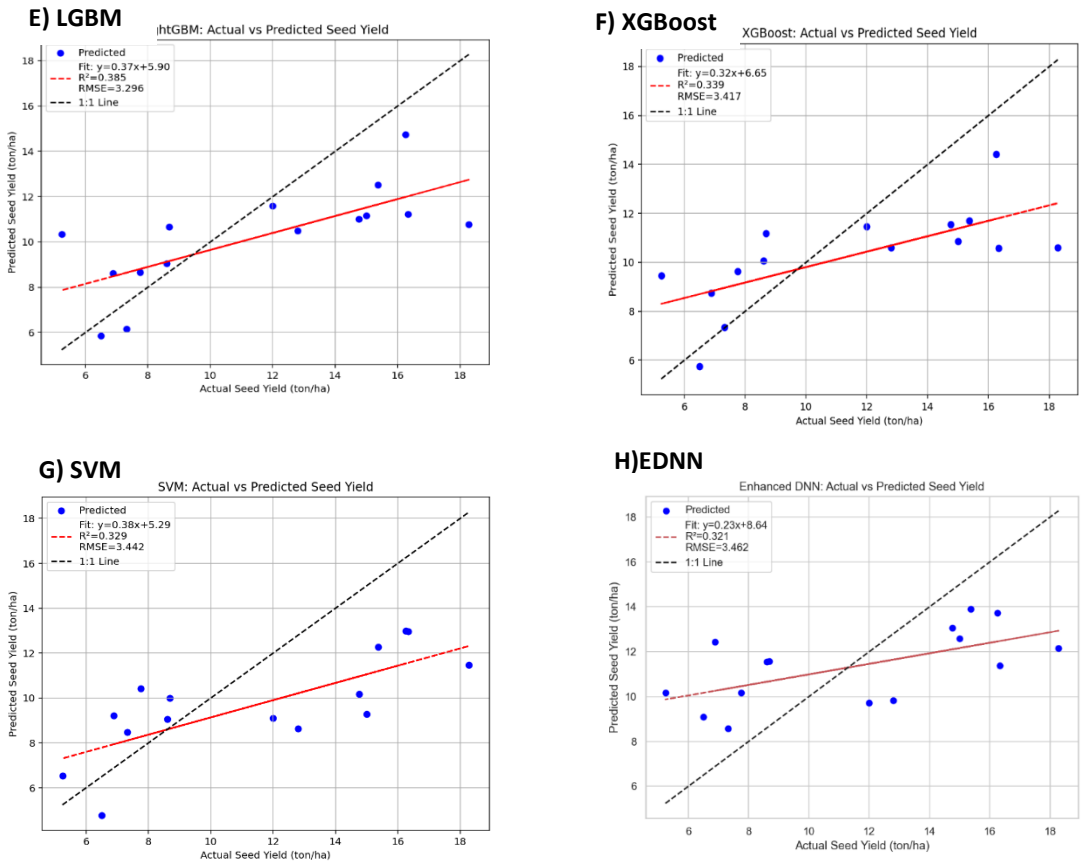
[1] Correlation Coefficient
[2] Reference

٤. **مشـــاهده مدل‌های افراطی:** مدل‌های لایت جی بی ام با انحراف معیار ۳/۱۵۳، و جذرمیانگین مربعات خطا ۲/۶۱۷، ایکس جی بوست با انحراف معیار ۳/۳۱۷، و جذرمیانگین مربعات خطا ۲/۷۸۰، انهنسد دی ان ان با انحراف معیار ۳/۵۸۲، و جذر میانگین مربعات خطا ۳/۱۳۶، اس وی آر با انحراف معیار ۳/۰۶۳، و جذر میانگین مربعات خطا ۲/۳۸۱، انحراف معیار و جذرمیانگین مربعات خطا بالایی دارند، که نشان‌دهنده پراکندگی بیشتر خطاها و دقت پایین‌تر پیش‌بینی‌هاست. این مدل‌ها همچنین R^2 پایینی دارند (به ترتیب ۰/۳۸۵، ۰/۳۳۹، ۰/۳۲۱، و ۰/۳۲۹)، که تأییدکننده عملکرد ضعیف‌تر آن‌ها است.

تحلیل و بحث نتایج الگوریتم های یادگیری ماشین

۱- **نمودار پراکنش نقاط پیش بینی شده در برابر نقاط مشاهده شده:** نمودار پراکنش نمونه های پیش بینی شده در مقابل نمونه های مشاهده شده برای هشت مدل مورد ارزیابی در شکل ۵ نشان داده شده است.



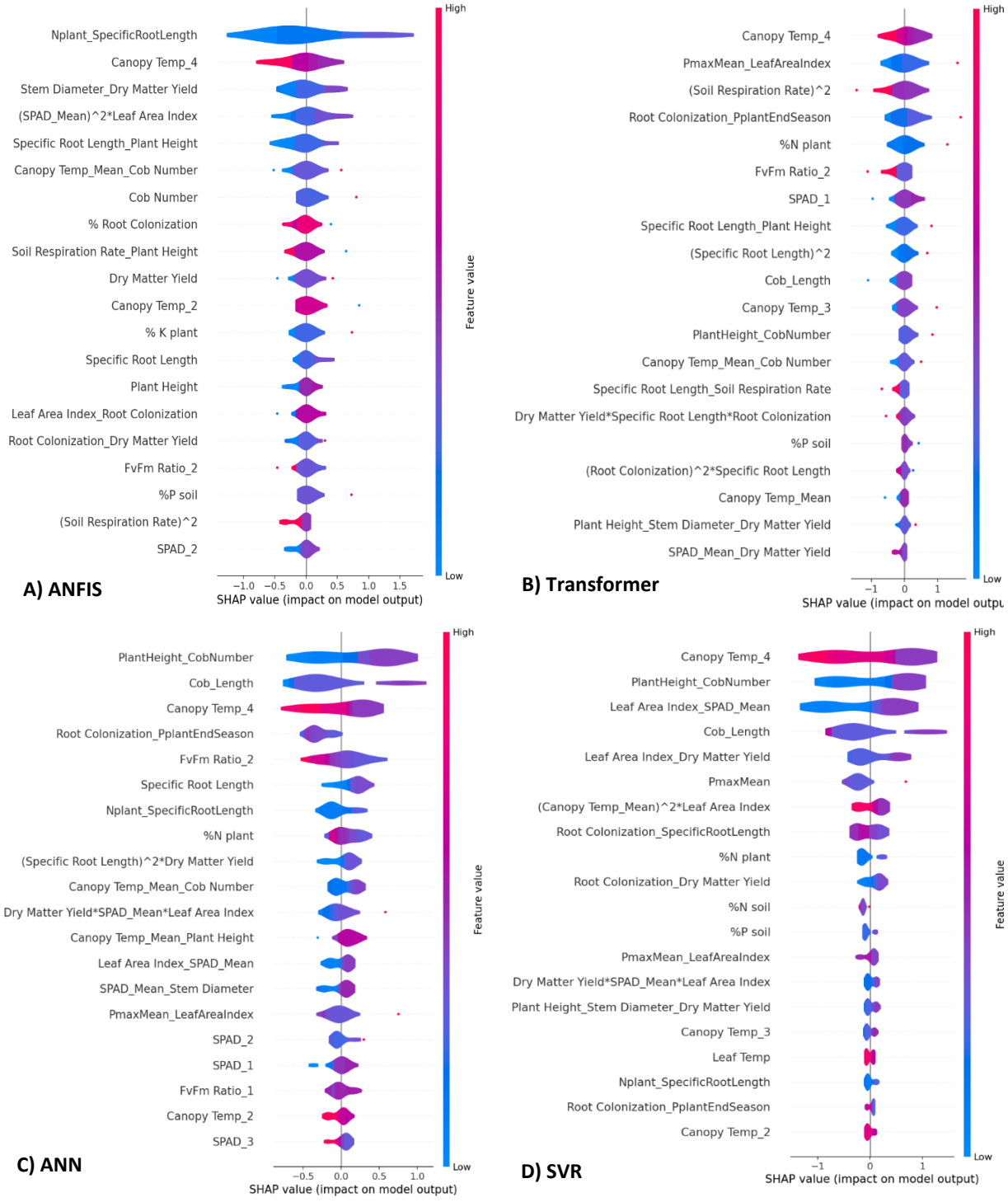


شکل ۵. نمودار پراکنش و خط رگرسیونی مقادیر پیش بینی شده عملکرد دانه ذرت در برابر مقادیر مشاهده شده (Aآنفیس، B ترنسفورمر، C اِی ان ان، D اس وی آر، E لایت جی بی ام، F ایکس جی بوست، G اس وی ام، H انهنسد دی ان ان).

Figure 5. Scatter plot and regression line of predicted versus observed corn grain yield values (A ANFIS, B Transformer, C ANN, D SVR, E LightGBM, F XGBoost, G SVM, H DNN)

Fashoto et al. (2021) اظهار داشتند که مدل‌های یادگیری ماشین در برآورد عملکرد محصول ذرت نسبت به روش‌های سنتی (رگرسیون خطی چندگانه و حذف معکوس) برتری دارند. گزارش شده است که مدل جنگل تصادفی در تخمین فعالیت‌های آنزیمی خارج‌سلولی خاک در زمین‌های شور ساحلی بازسازی‌شده کشاورزی عملکرد بهتری نسبت به رگرسیون‌های خطی چندگانه داشته است (Xie et al., 2021) .

۲- **نمودارهای شیپ ویولونی** اهمیت نسبی ۲۰ ویژگی برتر تشخیص داده شده توسط چهار مدل برتر (که ضریب تبیین بالای ۰/۵۰ داشتند) در شکل ۶ نشان داده شده است.



شکل ٦. نمودارهای شیپ ویولونی اهمیت نسبی ۲۰ ویژگی برتر تشخیص داده شده توسط چهار الگوریتم برتر (A آنفیس، B ترنسفورمر، C ای ان ان، D اس وی آر)

Figure 6. SHAP violin plots of the relative importance of the top 20 features identified by the four best algorithms (A ANFIS, B Transformer, C ANN, D SVR)

مدل آنفیس در پنج ویژگی با سیزده ویژگی تشخیص داده شده توسط مدل گام به گام مشترک بود (دمای کانوپی ۳ و ۴، نیتروژن گیاه در طول مخصوص ریشه، طول مخصوص ریشه در عملکرد ماده خشک، شاخص سطح برگ در عدد اسپد میانگین).
مدل ترنسفورمر در پنج ویژگی با سیزده ویژگی تشخیص داده شده توسط مدل گام به گام مشترک بود (دمای کانوپی ۳ و ۴، ارتفاع گیاه در تعداد بلال، مجذور کلونیزاسیون ریشه در طول مخصوص ریشه، دمای کانوپی میانگین).
شش ویژگی مشترک مدل ای ان ان با مدل گام به گام عبارت بودند از: ارتفاع گیاه در تعداد بلال، دمای کانوپی ۴ و ۲، محتوای نیتروژن گیاه در طول مخصوص ریشه، مجذور طول مخصوص ریشه در عملکرد ماده خشک، شاخص سطح برگ در عدد اسپد میانگین.
مدل اس وی آر بیشـــترین تعداد ویژگی مشـــترک با مدل گام به گام را دارا بود (۹ ویژگی) که عبارتند از: دمای کانوپی ۴، ۳ و ۲، ارتفاع گیاه در تعداد بلال، شاخص سطح برگ در عدد اسپد میانگین، شاخص سطح برگ در عملکرد ماده خشک، مجذور دمای کانوپی میانگین در شاخص سطح برگ، درصد کلونیزاسیون ریشه در عملکرد ماده خشک، درصد نیتروژن گیاه در طول مخصوص ریشه.
در جدول ۵ ویژگی های مشترک بین چهار الگوریتم ماشین برتر، آورده شده است.

جدول ۵. ویژگی های مشترک در چهار الگوریتم ماشین برتر با مدل رگرسیون گام به گام (به تفکیک هر الگوریتم)

Table 5. Common features among the four top machine learning algorithms and the stepwise regression model (by algorithm)

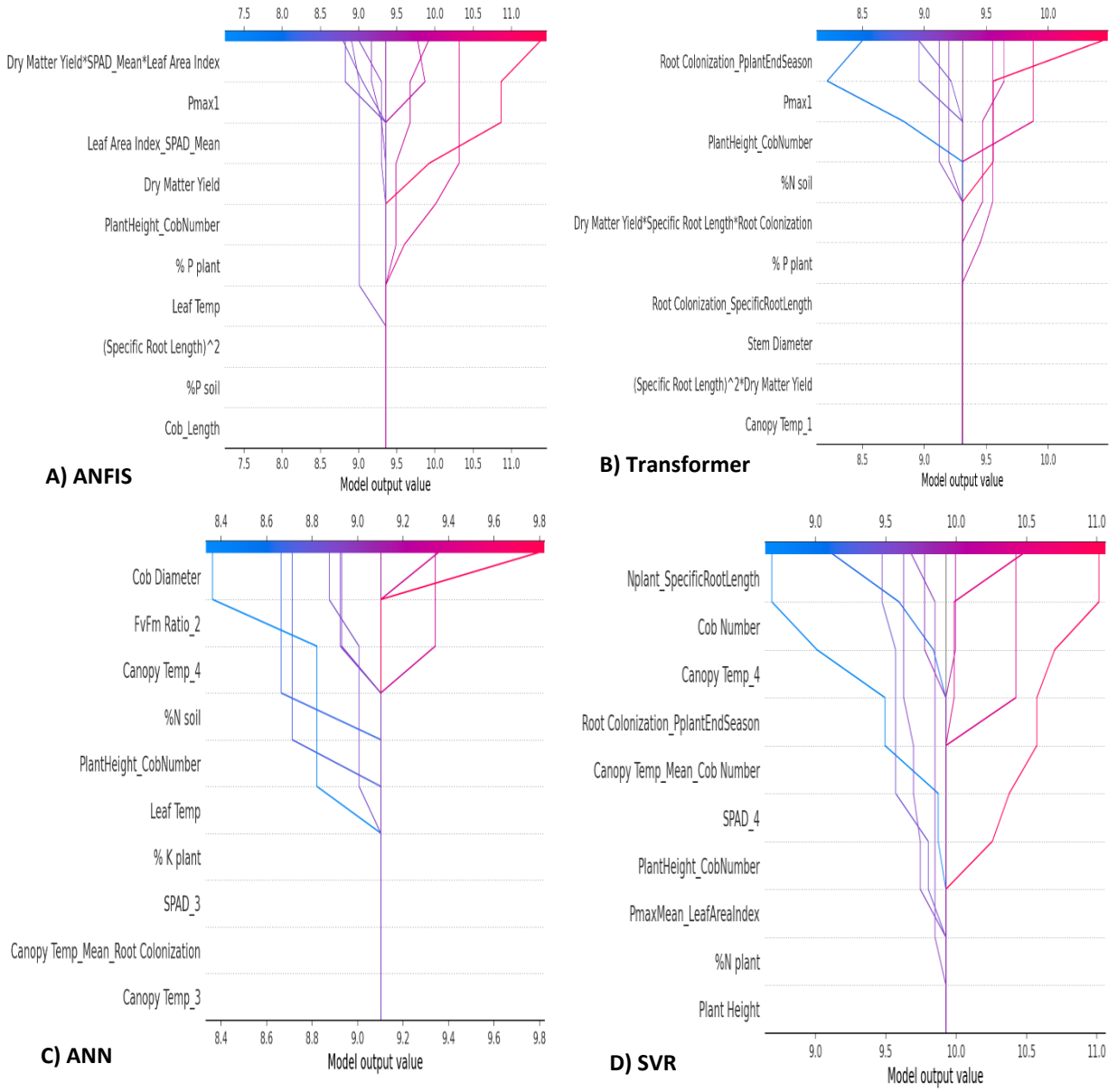
	ANFIS	Transformer	ANN	SVR
1	Canopy_Temp4	Canopy_Temp4	PlantHeight_CobNumber	Canopy_Temp4
2	Nplant_SpecificRootLength	Canopy_Temp3	Canopy_Temp4	PlantHeight_CobNumber
3	(Specific Root Length) ^2*Dry Matter Yield	PlantHeight_CobNumber	Nplant_SpecificRootLength (Specific Root Length) ^2*Dry Matter Yield	Leaf Area Index_SPAD_Mean
4	Leaf Area Index_SPAD_Mean	(Root Colonization) ^2*SpecificRootLenght	Leaf Area Index_SPAD_Mean	Leaf Area Index_Dry Matter Yield
5	Canopy_Temp2	Root Colonization_PplantEndSeason	Canopy_Temp2	(Canopy Temp_Mean) ^2*Leaf Area
6				Root Colonization_Dry Matter Yield
7				Canopy_Temp3
8				Nplant_SpecificRootLength
9				Canopy_Temp2

از بین سـیزده ویژگی تشـخیص داده شـده توسـط مدل رگرسـیونی گام به گام پس رونده، ۵، ۲، ۶ و ۵ ویژگی به ترتیب در مدل های آنفیس، ترنسفورمر، ای ان ان و اس وی آر نیز تشخیص داده شدند. الگوریتم اس وی آر با سادگی و بکارگیری کرنلRBF ، عملکرد خوبی (R^2 =۰/۴۵۲) از خود نشـان داد. ویژگی دمای کانوپی در مرحله پر شـدن دانه، به عنوان ویژگی مؤثر و با اهمیت بالا در پیش بینی عملکرد دانه، توسـط چهار الگوریتم برتر تشخیص داده شد. سه برهمکنش شامل محتوای نیتروژن گیاه و طول ویژه ریشه، شاخص سطح برگ و عدد اسپد، ارتفاع بوته و تعداد بلال و اثر سـاده دمای کانوپی در مرحله کاکل دهی در سـه الگوریتم برتر حضـور داشـتند که با مبانی اکوفیزیولوژیکی رشـد و نمو گیاهان زراعی همخوانی دارد (Taiz et al., 2018). بنابراین، می توان این ویژگی ها را مهمترین متغیرهای شکل دهنده عملکرد دانه ذرت دانست. عمدهٔ ویژگی‌های تشخیص داده شده نشان‌دهنده اهمیت تعامل بین کولونیزاسیون ریشه و سیستم فتوسنتزی بود که با مطالعات قبلی همخوانی دارد. تشخیص دو برهمکنش مجذور درصد کلونیزاسیون ریشه و طول ویژه ریشه، و درصد کلونیزاسیون ریشه و عملکرد ماده خشک توسط الگوریتم های ترنسفورمر و اس وی آر، می تواند نشان دهنده توانایی این دو الگوریتم در کشف اثرات پنهان کاربرد کودهای بیولوژیک بر عملکرد دانه ذرت باشد. در مطالعه ای مشخص شد که TempMax به‌عنوان پنجمین ویژگی مهم در پیش‌بینی عملکرد گندم زمستانه با استفاده از مقدار SHAP شناسایی شده است (Sirvastava et al., 2022). گزارش شده است که تحلیل SHAP نشان داد که مقادیر بالای ویژگی‌های تعداد روزهای رشد

و میانگین سطح برگ گیاه تأثیر منفی بر پیش‌بینی عناصر غذایی داشته‌اند، در حالی که مقادیر بالای تعداد برگ‌ها و ارتفاع گیاه اثر مثبتی بر پیش‌بینی عناصر غذایی داشته‌اند (Abekoon et al., 2025).

۳- **نمودارهای شیپ دسیژن پلات** برای نمونه های تست توسط چهار مدل برتر در شکل ۶ نشان داده شده است. محور عمودی، ترتیب ویژگی‌ها به ترتیب تأثیرگذاری روی تصمیم مدل را نشان می دهد (یعنی برای هر نمونه، مدل چطور برای خروجی نهایی تصمیم گرفته است). محور افقی، خروجی مدل (پیش‌بینی عملکرد دانه)، را نشان می دهد. خطوط رنگی (هر خط نشان دهنده یک نمونه است)، نشان دهنده تغییر مقدار خروجی مدل برای نمونه‌های مختلف بر اثر ورود تدریجی هر ویژگی است. به عبارت دیگر، خط از بالا (ویژگی اول) شروع می‌شود و در هر مرحله، ویژگی بعدی باعث بالا رفتن یا پایین رفتن مقدار خروجی مدل می‌شود. در پایان (پایین‌ترین نقطه)، خروجی نهایی مدل برای آن نمونه مشخص شده است. رنگ خطوط، آبی نشان دهنده خروجی پایین‌تر، و قرمز نشان دهنده خروجی بالاتر است. تحلیل ویژگی ها از بالا به پایین صورت می گیرد و تحلیل خروجی، از چپ به راست تا برسد به خط پایه که پیش بینی نهایی است. بطور کلی، تحلیل تصمیم مدل، بر اساس ترتیب ویژگی ها (از بالا به پایین) و اثر نهایی هر ویژگی روی پیش بینی (از چپ به راست) انجام می شود.

- در مدل آنفیس (شکل A ۷) اثر متقابل عملکرد ماده خشک و میانگین قرائت اسپد و شاخص سطح برگ، مهمترین اثر تعاملی بوده و باعث افزایش شدید عملکرد دانه شده است (خط رنگی از سمت چپ (حدود ۹) به بالای محور (حدود ۱۱) می‌رسد). دومین ویژگی، فتوسنتز ماکسیمم یک، است که تأثیر مثبت ملایم روی مقدار خروجی دارد، اما نسبت به اولی، اثرش کمتر بوده و نمودار آن در حوالی خروجی ۹/۵ متمرکز شده است. ویژگی سوم، اثر متقابل شاخص سطح برگ و قرائت اسپد میانگین است، که اثر آن مثبت و متوسط است. ویژگی چهارم، عملکرد ماده خشک است که همانند ویژگی سوم، اثر مثبت ملایمی دارد. ویژگی پنجم، اثر متقابل ارتفاع بوته و تعداد بلال است که اثر مثبت بر عملکرد دانه داشته و سبب افزایش آن شده است. ویژگی های ششم (درصد فسفر گیاه) و هفتم (دمای برگ) تأثیر متوسط (متعادل) بر عملکرد دانه دارند. ویژگی هشتم، مجذور طول مخصوص ریشه است که اثر آن تقریباً خنثی بود. ویژگی نهم (درصد فسفر خاک) و دهم (طول بلال) در انتهای لیست یعنی با کمترین اهمیت نسبی یا اثر ضعیف بر عملکرد دانه، واقع شده اند. به طور کلی، می توان گفت که در مدل آنفیس، ترکیب شاخص سطح برگ، قرائت اسپد و عملکرد ماده خشک، مهم ترین پیش بینی کننده عملکرد دانه تشخیص داده شده است. ویژگی های منفرد مثل فتوسنتز ماکزیمم یک و عملکرد ماده خشک، اثر متقابل شاخص سطح برگ و قرائت اسپد و همچنین اثر متقابل ارتفاع بوته و تعداد بلال، اثر افزایشی بر عملکرد دانه دارند. با در نظر گرفتن مسیر این ویژگی ها، می توان گفت که مدل آنفیس بر اساس ویژگی های فیزیولوژیکی (سرعت فتوسنتز ماکزیمم و قرائت اسپد) و مورفولوژیک (شاخص سطح برگ و ارتفاع بوته)، پیش بینی می کند.



شکل ۷. نمودارهای شیپ دسیژن پلات برای نمونه های تست توسط چهار الگوریتم برتر(A آنفیس، B ترنسفورمر، C ای ان ان، D اس وی آر)

Figure 7. SHAP Decision Plots for test samples by the four best algorithms (A ANFIS, B Transformer, C ANN, D SVR)

- در مدل ترنسفورمر (شکل ۷ B)، اثر متقابل درصد کلونیزاسیون ریشه و محتوای فسفر گیاه در انتهای فصل رشد، اولین و مهم ترین ویژگی اســـت، و برای بیشـــتر نمونه‌ها، اثر منفی دارد (خطوط آبی به چپ حرکت کرده‌اند ◄کاهش خروجی)، اگرچه برای برخی نمونه‌ها، اثر خنثی یا مثبت هم دیده می‌شود (رنگ به بنفش یا قرمز نزدیک شده). نرخ فتوسنتز ماکزیمم یک، اثر دوگانه دارد، به این صـورت که برای برخی نمونه‌ها باعث افزایش خروجی (رنگ قرمز)، و برای برخی سـبب کاهش خروجی شـده اسـت. این بدان معنی اســـت که بســـته به مقدار آن، عملکرد این ویژگی متغیر خواهد بود. اثر متقابل ارتفاع بوته و تعداد بلال، در اکثر موارد باعث افزایش خروجی شده که نشان‌دهنده اثر مثبت و مستقیم آن بر رشد گیاه است. درصد نیتروژن خاک و درصد فسفر گیاه، اثر متعادل دارند زیرا نمودار خطوط در این سطوح زیاد جابجا نشدند (خط ها صاف تر شدند). اثرات متقابل عملکرد ماده خشک و طول مخصوص ریشه و

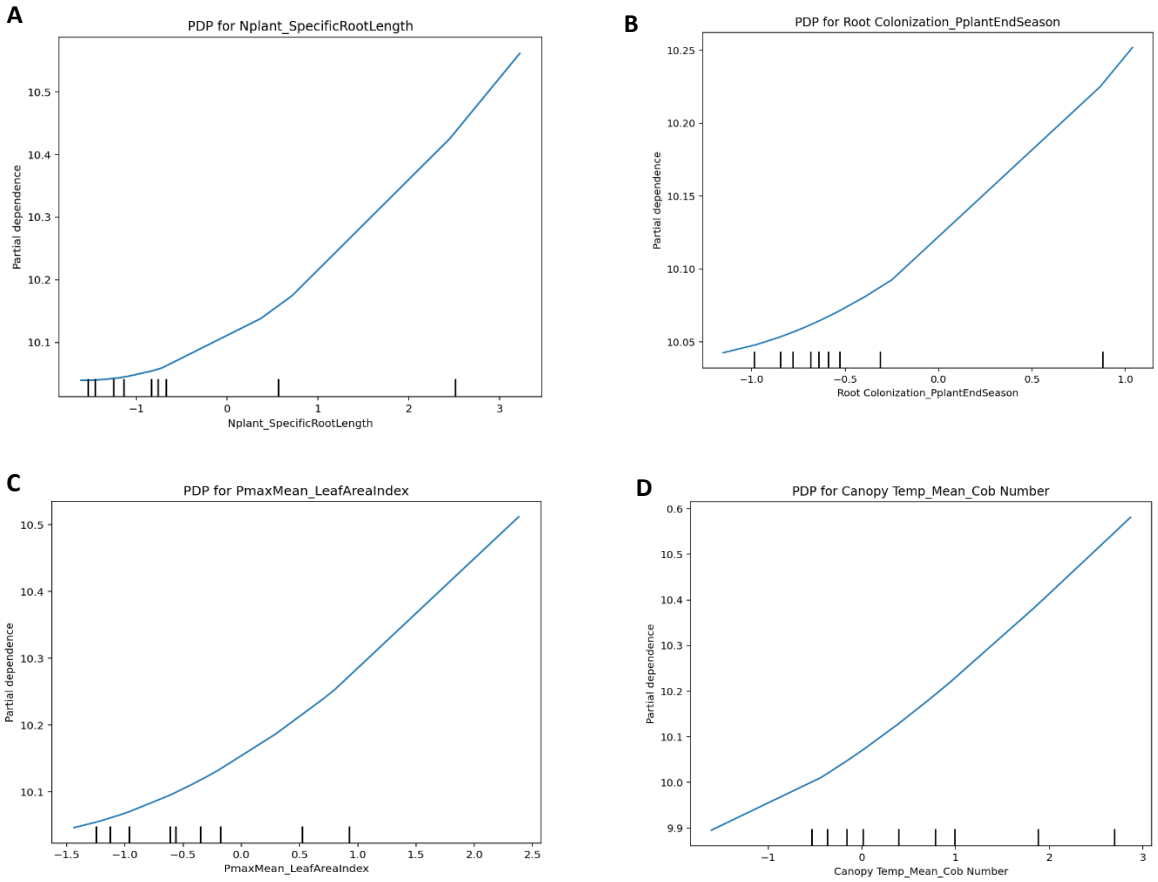
درصد کلونیزاسیون ریشه، به میزان قابل توجهی خروجی را تحت تأثیر قرار دادند و سبب افزایش آن شده اند. همین مطلب در مورد اثر متقابل، مجذور طول مخصوص ریشه و عملکرد ماده خشک نیز صادق است. دمای کانوپی یک (در مرحله خروج برگ پرچم)، آخرین ویژگی است و اثر خنثی دارد زیرا خط ها در این قسمت تغییر چندانی ندارند. در مجموع می توان گفت که مدل ترنسفورمر مقدار عملکرد دانه ذرت را با ویژگی های ریشه، سیستم فتوسنتزی و عملکرد ماده خشک انجام می دهد.

- از شرح مفصل نتایج نمودارهای تصمصم مدل های ای ان ان و اس وی آر صرف نظر شده و تنها به سازوکار نهایی آنها در پیش بینی اشاره می شود (به سبب محدودیت برای تعداد صفحات مقاله). الگوریتم ای ان ان (شکل C ۷) به وضوح یاد گرفت که با ترکیب ویژگی هایی شامل دمای کانوپی و خصوصیات فیزیولوژیکی برگ (نسبت فلورسانس کلروفیل متغیر به بیشینه دو؛ در مرحله کاکل دهی) و نیز ویژگی های ساختاری شامل قطر بلال می تواند پیش بینی نهایی را بالا یا پایین ببرد. در مجموع می توان نتیجه گرفت که الگوریتم ای ان ان، عملکرد دانه ذرت را با ویژگی های مرتبط با تنش (فیزیولوژیک) پیش بینی می کند.
- نمودار مسیر الگوریتم اس وی آر (شکل D ۷) نشان می دهد که تعداد بلال، دمای کانوپی چهار (در مرحله پر شدن دانه)، و اثر متقابل کلونیزاسیون ریشه و محتوای فسفر گیاه در انتهای فصل رشد، سبب افزایش پیش بینی شدند. این الگوریتم نیز همانند ترنسفورمر بر تأثیر برهمکنش میزان کلونیزاسیون ریشه و محتوای فسفر گیاه در انتهای فصل رشد، بر عملکرد دانه تأکید نمود. همچنین، هر دو الگوریتم، طول ویژهٔ ریشه (همانند درصد کلونیزاسیون ریشه) که نتیجهٔ مستقیم کاربرد کودهای زیستی است (Bao et al., 2022; Brar et al., 2024; Khoso et al., 2024)، را به عنوان ویژگی مهم و تأثیرگذار بر پیش بینی عملکرد دانه تشخیص دادند.

۳- **نمودارهای وابستگی جزئی**[1] چهار ویژگی منتخب از ده ویژگی برتر تشخیص داده شده توسط مدل اس وی آر در شکل ۸ نشان داده شده است (ده ویژگی برتر در جدول ۵ خلاصه شده‌اند). این نمودارها کمک می کنند به فهم این که یک ویژگی خاص چگونه به‌تنهایی (به صورت میانگین گرفته‌شده روی تمام نمونه‌ها) بر خروجی یک مدل یادگیری ماشین تاثیر می‌گذارد. محور افقی مقدار نرمال‌شده[2] ویژگی مورد نظر را نشان می دهد و محور عمودی وابستگی جزئی که نشان‌دهنده اثر مستقیم این ویژگی بر خروجی مدل است، را بدون اثر سایر متغیرها نشان می دهد. خطوط عمودی کوتاه[3] توزیع داده‌های واقعی در آن ویژگی را نشان می دهند.

همانگونه که در شکل ۸A مشاهده می شود با افزایش ویژگی برهمکنشی درصد نیتروژن گیاه و طول ویژه ریشه، مقدار عملکرد دانه نیز به طور صعودی و غیر خطی افزایش می یابد که این موضوع نشان دهنده اهمیت تغذیه نیتروژنی گیاه و تأثیر متقابل آن بر توسعه سیستم ریشه می باشد. نمودار ۸B نشان می دهد که برهمکنش درصد کلونیزاسیون ریشه و محتوای فسفر گیاه در انتهای فصل رشد، مقدار عملکرد دانه را بصورت تصاعدی و غیرخطی افزایش می دهد که این موضوع تأییدی بر نقش مثبت همزیستی قارچ های میکوریز آربوسکولار در بهبود کلی رشد و نمو و عملکرد دانه ذرت می باشد(Bao et al., 2022; Brar et al., 2024; Khoso et al., 2024) . همانطور که در نمودار مشخص است، میزان اثر در مقادیر بالاتر محسوس تر است. نمودار ۸C رابطه صعودی و غیرخطی برهمکنش میانگین نرخ فتوسنتز بیشینه و شاخص سطح برگ و عملکرد دانه را نشان می دهد. شیب خط در ابتدا کم و به‌تدریج تندتر می‌شود؛ به‌ویژه از مقدار نرمال‌شده‌ی حدود ۰/۵ به بعد شتاب افزایش عملکرد دانه زیاد می‌شود. بیشتر مشاهدات واقعی در نواحی بین ۱- تا ۱ قرار دارند، این به این معنی است که اکثر گیاهان در مجموعه داده دارای این ویژگیِ در محدوده نسبتاً متوسط هستند. تأثیر این ویژگی بر عملکرد دانه در مقادیر بالاتر، شدیدتر می‌شود. به عبارت دیگر، حتی افزایش‌های کوچک در این ویژگی می‌تواند در دامنه بالا تأثیر زیادی بر عملکرد دانه داشته باشد، که این موضوع از نظر فیزیولوژیکی کاملاً منطقی به نظر می رسد (Taiz et al., 2018). نمودار ۸D رابطه برهمکنش میانگین دمای کانوپی و تعداد بلال بر عملکرد دانه ذرت را نشان می دهد که رابطه ای است صعودی و غیرخطی. تأثیر مستقیم افزایش دمای کانوپی بر عملکرد دانه، ممکن است در شرایط خاص مثلاً در مراحل پر شدن دانه که مصادف با روزهای گرم تابستان است، از نظر فیزیولوژیکی قابل توجیه باشد. با اینحال، این رابطه ممکن است ناشی از وابستگی غیرمستقیم متغیر با صفات دیگر باشد و باید با احتیاط تفسیر شود.

[1] Partial Dependence
[2] z-score
[3] tick marks



شکل ۸. نمودارهای وابستگی جزئی A) اثر متقابل درصد نیتروژن گیاه و طول مخصوص ریشه، B) اثر متقابل درصد کلونیزاسیون ریشه و درصد فسفر گیاه در انتهای فصل رشد، C) اثر متقابل درصد نیتروژن گیاه و طول ویژه ریشه، D) اثر متقابل میانگین دمای کانوپی و تعداد بلال ذرت حاصل الگوریتم اس وی آر.

Figure 8. Partial dependence plots: A) Interaction of plant nitrogen percentage and specific root length, B) Interaction of root colonization percentage and plant phosphorus content at the end of the growing season, C) Interaction of mean maximum photosynthesis rate and leaf area index, D) Interaction of mean canopy temperature and cob number from the SVR algorithm

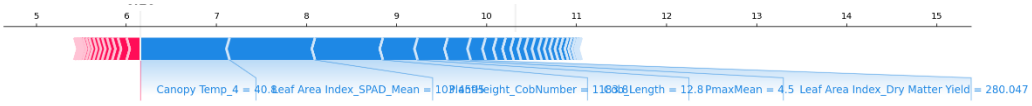
گزارش شده است که الگوریتم CNN-RNN برای پیش‌بینی عملکرد محصول سه ویژگی برجسته دارد که آن را به روشی بالقوه مفید برای مطالعات دیگر پیش‌بینی عملکرد محصول تبدیل می‌کند. (۱) الگوریتم CNN-RNN برای ضبط وابستگی‌های زمانی عوامل محیطی و بهبود ژنتیکی بذرها در طول زمان بدون نیاز به اطلاعات ژنوتیپ طراحی شده بود. (۲) مدل توانایی تعمیم پیش‌بینی عملکرد به محیط‌های آزمایش‌نشده را بدون کاهش قابل‌توجه دقت پیش‌بینی نشان داد. (۳) همراه با روش بک‌پروپگیشن، مدل توانست نشان دهد که شرایط آب‌وهوایی، دقت پیش‌بینی آب‌وهوا، شرایط خاک، و روش‌های مدیریت تا چه حد قادر به توضیح تغییرات در عملکرد محصولات هستند (Khaki et al., 2020) .

۴- **نمودار فورس پلات** برای مدل لایت جی بی ام به‌عنوان نمونه، با هدف نمایش ساختار این نوع نمودارها آورده شده است (این مدل عملکرد نسبتاً ضعیفی با R^2=0.385 دارد) (شکل ۹). این نمودار برای یک نمونه منفرد نشان می‌دهد که مدل از چه مقداری (مقدار پایه[1]) شروع کرده و هر ویژگی چطور روی خروجی نهایی اثر گذاشته و باعث افزایش یا کاهش خروجی شده است (شکل ۹). محور افقی

[1] base value

نشـان‌دهنده مقدار پیش‌بینی نهایی مدل اسـت (در اینجا حدوداً ۱۴/۸ در انتهای نمودار). مقدار پایه حدود ۵/۳ اسـت. رنگ آبی برای ویژگی هایی اسـت که باعث افزایش مقدار پیش‌بینی شـده‌اند و هرچه پیکان آبی کشـیده‌تر باشـد، سـهم آن ویژگی در افزایش خروجی بیشـتر اسـت. رنگ قرمز برای ویژگی هایی است که باعث کاهش مقدار پیش‌بینی شده‌اند. در شکل ۸ سمت چپ نمودار (قرمزها) اثر منفی دارند، و خیلی کوچک هسـتند. برهمکنش شـاخص سـطح برگ و عملکرد ماده خشـک، بزرگ‌ترین اثر مثبت را روی خروجی مدل دارد (Leaf Area Index_Dry Matter Yield = 280.047)، به عبارت دیگر، مقدار بالای این ویژگی باعث افزایش چشـمگیر عملکرد دانه (خروجی مدل) شده است. میانگین نرخ فتوسنتز بیشینه اثر افزایشی قوی دارد (PmaxMean = 4.5) که نشان دهنده اهمیت آن برای مدل می باشد. طول بلال زیاد باعث تقویت خروجی مدل شـده اسـت (Cob Length = 12.8). عدد بالای مربوط به برهمکنش ارتفاع گیاه و تعداد بلال اثر افزایشـی بر عملکرد دانه داشـت (PlantHeight_CobNumber = 11603.3). برهمکنش شـاخص سـطح برگ × میانگین SPAD نیز اثر مثبتی بر خروجی مدل گذاشته است (Leaf Area Index_SPAD_Mean = 108.45). دمای بالای پوشش گیاهی در مرحله پرشدن دانه اثر مثبت دارد (این کمی نامعمول است و ممکن است مدل رفتار خاصی را از این متغیر آموخته باشد، مثل همبستگی غیرخطی) (Canopy Temp_4 = 40.8). بطور کلی، مدل از مقدار پایه ۵/۳ تن در هکتار عملکرد دلنه شـروع کرده، و با تأثیر مثبت متغیرهای کلیدی (مثل عملکرد ماده خشک، شاخص سطح برگ، طول بلال و شدت فتوسنتز بیشینه)، خروجی را به مقدار نهایی حدود ۱۴/۸ تن در هکتار عملکرد دانه رسانده است و ویژگی‌های با اثر منفی، اندک و کم‌اثر بوده‌اند.

Sirvastava et al., (2022) پیشـنهاد کردند که فراتر از پیش‌بینی، از نمودارهای SHAP و فورس پلات برای تفسـیر خروجی‌های مدل پیشنهادی CNN استفاده شده که بینش‌های کلیدی در توضیح نتایج پیش‌بینی عملکرد (اهمیت متغیرها بر حسب زمان) ارائه داد. آن‌هاDUL ، سـرعت باد در هفته دهم، و مقدار تابش در هفته هفتم را به‌عنوان مهم‌ترین ویژگی‌ها در پیش‌بینی عملکرد گندم زمسـتانه یافتند. گزارش شـده اسـت که رطوبت، فسـفر(P) ، پتاسـیم(K) ، دما، بارندگی، نیتروژن(N) ، و pH به‌عنوان مهم‌ترین ویژگی‌هایی که بر خروجی مدل در تکنیک توصیه محصول در کشاورزی هوشمند تأثیر دارند، شناسایی شده‌اند (Abekoon et al., 2025).

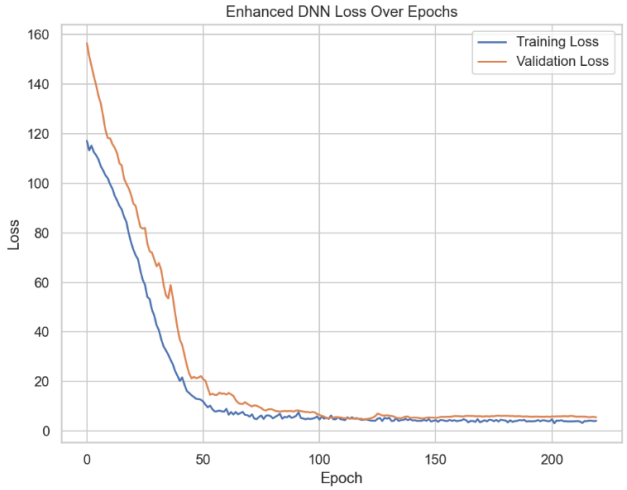


شکل ۹. نمودار فورس پلات تولی شده توسط مدل لایت جی بی ام
Figure 9. Force plot generated by the LightGBM model

۵- نمودار زیان در هر اپوک

اپوک[1] یک دوره کامل از پردازش تمام داده‌های آموزشـی توسـط مدل در فرآیند یادگیری ماشـین اسـت. نمودار زیان در هر اپوک[2] نشـان‌دهنده میزان خطای مدل در طول فرآیند یادگیری در هر دوره یا اپوک است. با افزایش تعداد اپوک‌ها، معمولاً مقدار زیان کاهش می‌یابد که نشان‌دهنده بهبود عملکرد مدل است، مگر اینکه مدل بیش‌ازحد به داده‌های آموزشی وابسته شود (اصطلاحاً اُورفیتینگ یا بیش برازش رخ دهد). این نمودار برای ارزیابی و تنظیم فرآیند یادگیری ماشین استفاده می‌شود. نمودار زیان در هر اپوک مدل شبکه عصبی عمیق ارتقاء یافته در شکل ۱۰ نشان داده شده است. انطباق بسیار نزدیک دو نمودار بویژه از اپوک ۱۰۰ به بعد حاکی از همگرایی زیان آموزش و اعتبارسنجی می باشد که تأکیدی بر یادگیری مؤثر توسط الگوریتم است.

[1] Epoch
[2] Loss per Epoch



شکل ۱۰. نمودار زیان در هر اپوک مدل شبکه عصبی عمیق ارتقاء یافته

Figure 10. Loss per epoch plot for the enhanced deep neural network model

نتیجه‌گیری کلی

یافته ها: آنفیس و ترنسفورمر به ترتیب با ضریب تبیین برابر با ۰/۵۴۴ و ۰/۵۴۵ و جذر میانگین مربعات خطا ۲/۸۳۹ و ۲/۸۳۴ به دلیل بکارگیری مکانیزم ترکیبی TensorFlow (در آنفیس) و Attention (در ترنسفورمر) و همچنین استفاده از بهینه ساز Adam در هر دو الگوریتم، تعاملات پیچیده را به خوبی مدل‌سازی کردند و به‌عنوان برترین الگوریتم های یادگیری ماشین در پیش بینی عملکرد دانه ذرت شناخته شدند. در رتبه بعدی، الگوریتم اِ ی ان ان قرار گرفت که با معماری سه لایه، فعال ساز ReLU و استفاده از بهینه ساز GriedSearch با تعیین اعتبار متقابل، روابط غیرخطی را به‌خوبی مدل‌سازی کرد و عملکردی بسیار نزدیک به آنفیس و ترنسفورمر از خود نشان داد (R^2 =۰/۵۱۸). ویژگی های دمای کانوپی در مرحله پر شدن دانه، محتوای نیتروژن گیاه و طول ویژه ریشه، شاخص سطح برگ و عدد اسپد، ارتفاع بوته و تعداد بلال و اثر ساده دمای کانوپی در مرحله کاکل دهی به عنوان مهمترین متغیرهای شکل دهنده عملکرد دانه ذرت، توسط چهار الگوریتم برتر تشخیص داده شد. عمدهٔ ویژگی‌های تشخیص داده شده نشان‌دهنده اهمیت تعامل بین کولونیزاسیون ریشه و سیستم فتوسنتزی بود که با مطالعات قبلی همخوانی دارد. این الگوریتم ها نشان دادند که با یادگیری الگوهای پنهان در داده ها، می توانند روابط اکوفیزیولوژیکی دخیل در شکل گیری عملکرد دانه ذرت را مشخص کرده و پیش بینی دقیق تر آن را ممکن سازند.

محدودیت ها: محدودیت‌هایی مثل اندازه نسبتاً کوچک مجموعه داده‌ها و تمرکز روی یک منطقه خاص وجود دارد که در تحقیقات آینده می تواند برطرف شود، با این حال، درتفسیر نتایج برای آزمایش هایی با داده‌های کمتر، جانب احتیاط باید رعایت شود. نتایج این مطالعه می تواند زمینه های لازم برای مدیریت زراعی مؤثر در جهت افزایش کارایی نهاده ها و سازگاری با شرایط تنش و تغییرات اقلیمی در تولید ذرت دانه ای را فراهم آورد.

توصیه ها: استفاده از RF برای سیستم‌های پشتیبانی تصمیم کشاورزی و ادغام داده‌های ماهواره‌ای برای پوشش مکانی بهتر، می تواند دقت و کارآیی مدل ها را افزایش دهد.

References

1. Abekoon, T., Sajindra, H., Rathnayake, N., Ekanayake, I., Jayakody, A. (2025). A novel application with explainable machine learning (SHAP and LIME) to predict soil N, P, and K nutrient content in cabbage cultivation. Smart Agricultural Technology, 11: 100879. https://doi.org/10.1016/j.atech.2025.100879
2. Bao, X., Zou, J., Zhang, B., Wu, L., Yang, T., Huang, Q. (2022). Arbuscular Mycorrhizal Fungi and Microbes Interaction in Rice Mycorrhizosphere. *Agronomy*, *12*(6): 1277. https://doi.org/10.3390/agronomy12061277
3. Bhupenchandra, I., Chongtham, S.K., Devi, A.G. et al. (2024). Unlocking the Potential of Arbuscular Mycorrhizal Fungi: Exploring Role in Plant Growth Promotion, Nutrient Uptake Mechanisms, Biotic Stress Alleviation, and Sustaining Agricultural Production Systems. J Plant Growth Regul. https://doi.org/10.1007/s00344-024-11467-9
4. Bracho-Mujica, G., Rötter, R. P., Haakana, M., Palosuo, T., Fronzek, S., et al. (2024). Effects of changes in climatic means, variability, and agro-technologies on future wheat and maize yields at 10 sites across the globe. Agricultural and Forest Meteorology, 346: 109887. https://doi.org/10.1016/j.agrformet.2024.109887
5. Brar, B., Bala, K., Saharan, B.S. et al. (2024). Bio-boosting agriculture: Harnessing the potential of fungi-bacteria-plant synergies for crop improvement. Discov. Plants, 1: 21 https://doi.org/10.1007/s44372-024-00023-0
6. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd Ed. New York: Routledge.
7. Fashoto, S., Mbunge, E., Opeyemi, O.G., Van Den Burg, J., (2021). Implementation of machine learning for predicting maize crop yields using multiple linear regression and backward elimination. Malays. J. Comp. 6: 679–697.
8. Godfray, H.C.J., Poore, J., Ritchie, H. (2024). Opportunities to produce food from substantially less land. BMC Biology, 22:138. https://doi.org/10.1186/s12915-024-01936-8
9. Khaki, S., Wang, A., Archontoulis, A.V. (2020). A CNN-RNN Framework for Crop Yield Prediction. Front. Plant Sci. Sec. Computational Genomics, https://doi.org/10.3389/fpls.2019.01750
10. Khoso, M.A., Wagan, S., Alam, I., Hussain, A., Ali, Q., Saha, S., Poudel, T.R., Manghwar, H., Liu, F. (2024). Impact of plant growth-promoting rhizobacteria (PGPR) on plant nutrition and root characteristics: Current perspective. Plant Stress, 11: 100341. https://doi.org/10.1016/j.stress.2023.100341.
11. Kuhn, M., Johnson, K. (2013). Applied Predictive Modeling. pringer Science & Business Media, 600 pages. ISBN: 978-1-4614-6848-6. https://doi.org/10.1007/978-1-4614-6849-3
12. Mitra, D., Panneerselvam, P., Senapati, A., Chidambaranathan, P., Nayak, A.K., Mohapatra, P.K.D. (2023). Arbuscular Mycorrhizal Fungi Response on Soil Phosphorus Utilization and Enzymes Activities in Aerobic Rice under Phosphorus-Deficient Conditions. Life (Basel). 30,3(5):1118. https://doi.org/10.3390/life13051118. PMID: 37240763; PMCID: PMC10221761.
13. Moore, D. S., Notz, W. I, & Flinger, M. A. (2013). The basic practice of statistics (6th ed.). New York, NY: W. H. Freeman and Company. Page (138).
14. Nayak, H.S., Silva, J.V., Parihar, C.M., Krupnik, T.J., Sena, D.R., Kakraliya, S.K., Jat, H.S., Sidhu, H.S., Sharma, P.C., Jat, M.L., et al. (2022). Interpretable machine learning methods to explain on-farm yield variability of high productivity wheat in Northwest India. Field Crop. Res. 287: 108640. https://doi.org/10.1016/j.fcr.2022.108640
15. Pentoś, K., Mbah, J. T., Pieczarka, K., Niedbała, G., & Wojciechowski, T. (2022). Evaluation of Multiple Linear Regression and Machine Learning Approaches to Predict Soil Compaction and Shear Stress Based on Electrical Parameters. Applied Sciences, 12(17): 8791. https://doi.org/10.3390/app12178791
16. Roell, Y. E., Beucher, A., Møller, P. G., Greve, M. B., Greve, M. H. (2020). Comparing a Random Forest based prediction of winter wheat yield to historical yield potential. Agronomy, 10(3), 3. https://doi.org/10.3390/agronomy10030395
17. Salvador, G.L.O., Araju, F.F., Pereira, A.P.A., Bonofacio, A., Araujo, A.S.F. (2022). Rhizobacteria and arbuscular mycorrhizal fungus presented distinct and specific effects on soybean growth when inoculated with organic compost. Rhizosphere, 22: 100513. https://doi.org/10.1016/j.rhisph.2022.100513
18. Srivastava, A. K., Safaei, N., Khaki, S., Lopez, G., Zeng, W., Ewert, F. et al. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-06249-w
19. Taiz, L., E. Zeiger, I. M. Moller, et al. (2018). Fundamentals of plant physiology. New York, USA: Oxford University. Press. ISBN 978160535790
20. Wang, Y., Zhang, Q., Yu, F., Zhang, N., Zhang, X., Li, Y., Wang, M., Zhang, J. (2024). Progress in Research on Deep Learning-Based Crop Yield Prediction. *Agronomy*, *14*(10): 2264. https://doi.org/10.3390/agronomy14102264
21. Xie, X., Wu, T., Zhu, M., Jiang, G., Xu, Y., Wang, X., Pu, L. (2021). Comparison of random forest and multiple linear regression models for estimation of soil extracellular enzyme activities in agricultural reclaimed coastal saline land. Ecological Indicators, 120: 106925. https://doi.org/10.1016/j.ecolind.2020.106925