

توسعه فردی و تحول سازمانی

طراحی پروتکل یادگیری انتقالی در مدل‌های هیبریدی یادگیری عمیق برای پیش‌بینی و تفسیرپذیری کاهش وزن

انوشه یزدانبخش^۱، علیرضا پورابراهیمی^۲، عبدالرضا نوروزی^۳

شیوه استناددهی: یزدانبخش، انوشه، پورابراهیمی، علیرضا و

نوروزی، عبدالرضا. (۱۴۰۵). طراحی پروتکل یادگیری

انتقالی در مدل‌های هیبریدی یادگیری عمیق برای پیش‌بینی

و تفسیرپذیری کاهش وزن. توسعه فردی و تحول سازمانی،

۴(۲)، ۱۶-۱.

۱. گروه مدیریت، واحد بین الملل کیش، دانشگاه آزاد اسلامی، کیش، ایران

۲. گروه مدیریت صنعتی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران

۳. گروه علوم تغذیه، دانشگاه علوم پزشکی مشهد، ایران

* ایمیل نویسنده مسئول: A.pourebrahimi@kiaau.ac.ir

چکیده

تاریخ چاپ نهایی: ۱ تیر ۱۴۰۵

تاریخ چاپ اولیه: ۱۰ آذر ۱۴۰۴

تاریخ پذیرش: ۸ آذر ۱۴۰۴

تاریخ بازنگری: ۱ آذر ۱۴۰۴

تاریخ ارسال: ۲۳ مرداد ۱۴۰۴

هدف پژوهش، توسعه یک پروتکل سه‌مرحله‌ای یادگیری انتقالی برای معماری هیبریدی CNN-Attention-XGBoost با هدف بهبود دقت و تفسیرپذیری پیش‌بینی کاهش وزن در داده‌های بالینی ایرانی است. مدل پیشنهادی شامل شبکه عصبی کانولوشنال، لایه توجه و رگرسور XGBoost بود. ابتدا مدل عمیق بر داده‌های بین‌المللی پیش‌آموزش داده شد تا الگوهای عمومی تغییر وزن استخراج شود. سپس وزن‌ها روی داده‌های واقعی ایرانی تنظیم دقیق گردید تا با ویژگی‌های جمعیت بومی سازگار شود. در مرحله پایانی، ویژگی‌های استخراج‌شده به رگرسور XGBoost منتقل و مدل نهایی آموزش داده شد. عملکرد مدل با شاخص‌های MAE، RMSE و R^2 سنجیده شد و تفاوت عملکرد مدل‌ها به‌صورت آماری بررسی شد. مدل هیبریدی مبتنی بر یادگیری انتقالی نسبت به مدل بدون انتقال دانش بهبود قابل‌توجهی نشان داد و به مقادیر بالاتر دقت و خطای کمتر دست یافت. همچنین الگوهای زمانی وزن و ویژگی‌های عمیق استخراج‌شده بیشترین سهم را در پیش‌بینی داشتند. نتایج آزمون آماری نیز نشان داد که تفاوت میان دو مدل معنادار است و استفاده از دانش منتقل‌شده باعث بهبود توان پیش‌بینی می‌شود. پروتکل یادگیری انتقالی ارائه‌شده توانست دقت پیش‌بینی، تعمیم‌پذیری و تفسیرپذیری مدل را به‌صورت هم‌زمان افزایش دهد. ترکیب لایه‌های عمیق و رگرسور قابل‌تفسیر، رویکردی کارآمد برای توسعه سامانه‌های هوش مصنوعی سلامت فراهم می‌کند که هم دقیق و هم قابل‌اعتماد هستند.

کلیدواژگان: یادگیری انتقالی، شبکه عصبی کانولوشنال، لایه توجه، XGBoost، پیش‌بینی کاهش وزن، مدل هیبریدی، تفسیرپذیری.

این مقاله متعلق به نویسنده است. انتشار این مقاله به‌صورت
دسترسی آزاد مطابق با گواهی (CC BY-NC 4.0)
صورت گرفته است.



Personal Development and Organizational Transformation

A Transfer Learning Protocol for Hybrid Deep–Machine Learning Models to Predict and Interpret Weight Loss

Anousheh Yazdanbakhsh¹, Alireza Pourebrahimi^{2*}, Abdolreza Norouzy³

1. Department of Management, Ki.C., Islamic Azad University, Kish, Iran

2. Department of Industrial Management, Ka.C., Islamic Azad University, Karaj, Iran

3. Department of Nutritional Sciences, Mashhad University of Medical Sciences, Mashhad, Iran

*Corresponding Author's Email: A.pourebrahimi@kiaou.ac.ir

How to cite: Yazdanbakhsh, A., Pourebrahimi, A., & Norouzy, A. (2026). A Transfer Learning Protocol for Hybrid Deep–Machine Learning Models to Predict and Interpret Weight Loss. *Personal Development and Organizational Transformation*, 4(2), 1-15.

Submit Date: 14 August 2025

Revise Date: 22 November 2025

Accept Date: 29 November 2025

Initial Publish: 01 December 2025

Final Publish: 22 June 2026



© 2026 the authors. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Abstract

The objective of this study is to develop a three-stage transfer learning protocol for a hybrid CNN–Attention–XGBoost architecture to improve the accuracy and interpretability of weight-loss prediction in Iranian clinical data. The proposed model consisted of a convolutional neural network, an attention mechanism, and an XGBoost regressor. The CNN–Attention module was first pre-trained on an international dataset to learn general temporal weight-change patterns. The learned parameters were then fine-tuned on an Iranian dataset to adapt the model to local characteristics. Finally, the extracted features were passed to an XGBoost regressor for interpretable prediction. Model performance was evaluated using MAE, RMSE, and R^2 , and statistical tests were applied to compare model effectiveness. The hybrid transfer-learning model showed significantly higher accuracy and lower error compared with the non-transfer baseline model. Temporal weight patterns and deep learned features contributed most strongly to predictive performance. Statistical testing confirmed that the improvement in accuracy was significant, indicating that transferred knowledge effectively enhanced model generalization. The proposed transfer learning protocol successfully enhanced both predictive performance and interpretability. Integrating deep feature extraction with an explainable decision layer provides a powerful and reliable approach for developing AI-based health systems that require both accuracy and transparency.

Keywords: *Transfer Learning, CNN, Attention Mechanism, XGBoost, Weight Loss Prediction, Hybrid Model, Interpretability.*

مقدمه

در دهه‌های اخیر، تحولات گسترده در یادگیری ماشین و هوش مصنوعی موجب دگرگونی عمیقی در نحوه تحلیل، پردازش و پیش‌بینی داده‌های سلامت و رفتار انسان شده است. گسترش حجم داده‌های بالینی، رشد سامانه‌های دیجیتال، و توسعه ابزارهای پوشیدنی شرایطی را فراهم کرده است که اکنون مدل‌های هوشمند می‌توانند الگوهای پیچیده سلامت جسمانی و روانی را با دقتی بی‌سابقه شناسایی کنند. پژوهشگران حوزه سلامت، با تکیه بر این پیشرفت‌ها، به‌طور فزاینده‌ای از مدل‌های مبتنی بر یادگیری عمیق، شبکه‌های چندوجهی، و یادگیری انتقالی برای تحلیل پدیده‌هایی همچون چاقی، فرسودگی شغلی، کاهش وزن، رفتارهای پرخطر، و اختلالات خوردن استفاده می‌کنند. بررسی‌ها نشان می‌دهد که رویکردهای کلاسیک آماری در مواجهه با ساختارهای غیرخطی داده‌های سلامت و تنوع کوهورت‌های انسانی دقت و توان تبیینی محدودی دارند، در حالی که مدل‌های یادگیری ماشین قادرند تعاملات پیچیده، ویژگی‌های پنهان، و روابط پویا را استخراج کنند؛ موضوعی که زمینه‌ساز پژوهش‌های بین‌رشته‌ای فراوانی شده است (Rubinger et al., 2023).

پیشرفت‌های اخیر به‌ویژه در یادگیری عمیق، دریچه‌ای نو برای تحلیل داده‌های رفتاری و زیستی گشوده است. شبکه‌های عصبی کانولوشنال به‌عنوان یکی از مهم‌ترین معماری‌های یادگیری عمیق، توانایی استخراج ویژگی‌های سلسله‌مراتبی از داده‌های پیچیده را دارند و از ابتدا برای تحلیل داده‌های بزرگ و ناهمگن طراحی شده‌اند (Simonyan & Zisserman, 2014). این مدل‌ها، هنگامی که در قالب معماری‌های عمیق‌تر و چندلایه مورد استفاده قرار می‌گیرند، می‌توانند الگوهای زمانی، روندهای متغیر، و تعاملات چندبعدی را با دقت بالا شناسایی کنند؛ موضوعی که برای پیش‌بینی پیامدهای سلامت، ازجمله کاهش وزن، اهمیت بنیادی دارد. البته پیچیدگی ذاتی این مدل‌ها مسئله تفسیرپذیری را به چالش می‌کشد، به‌طوری‌که پژوهشگران سلامت نیازمند مدل‌هایی هستند که ضمن داشتن دقت بالا، قابلیت توضیح و اعتمادپذیری نیز داشته باشند. همین ضرورت سبب شده است که ابزارهایی مانند SHAP و روش‌های شفاف‌سازی مدل‌ها توجه گسترده‌ای کسب کنند (Lundberg & Lee, 2017).

با گسترش همه‌گیری چاقی در دهه اخیر، پیش‌بینی و تحلیل عوامل خطر چاقی، شامل رفتارهای غذایی، فعالیت بدنی، زیست‌ریتم‌ها و وضعیت روانی، به یکی از حوزه‌های پرطرفدار پژوهش تبدیل شده است. چاقی و اضافه‌وزن پیامدهای گسترده‌ای بر سلامت عمومی، بهره‌وری اجتماعی و هزینه‌های درمانی دارند. مطالعه‌ای جامع نشان می‌دهد که ترکیبی از عوامل زیستی، اجتماعی، رفتاری و محیطی در ایجاد چاقی نقش دارند و مدل‌های یادگیری ماشین می‌توانند این عوامل پیچیده را تحلیل و الگوهای خطر را شناسایی کنند (Safaei et al., 2021). در همین راستا، استفاده از روش‌های ترکیبی که داده‌های چندمنبعی را با داده‌های زمانی و رفتاری تلفیق می‌کنند، در حال گسترش است. پژوهش‌های متعدد نشان داده‌اند که مدل‌های عمیق چندوجهی قادرند همزمان داده‌های بالینی، سبک زندگی، سابقه بیماری و الگوهای رفتاری را برای برآورد شاخص‌های چاقی مورد تحلیل قرار دهند (Gupta et al., 2024).

از سوی دیگر، پژوهش‌های مبتنی بر داده‌های واقعی انسان‌ها نشان می‌دهند که عوامل رشدی و زیستی مانند بلوغ زودرس یا دیررس نیز می‌توانند منجر به تغییرات قابل توجه در احتمال ابتلا به چاقی شوند؛ موضوعی که در مطالعات ملی و منطقه‌ای مشاهده شده است (Gülü et al., 2023). این یافته‌ها اهمیت توجه به تفاوت‌های فردی و الگوهای رشدی را در مدل‌های پیش‌بینی سلامت برجسته می‌سازد. ازاین‌رو، تحلیل داده‌های رشد و مدل‌های یادگیری ماشین برای تمایز میان افراد در مسیرهای رشدی متفاوت مورد توجه قرار گرفته‌اند.

در حوزه سلامت عمومی، پژوهش‌های جدید نشان داده‌اند که روش‌های یادگیری ماشین می‌توانند در تحلیل عوامل روان‌شناختی و رفتاری مرتبط با چاقی و تغییرات وزن نیز کارآمد باشند. مدل‌های یادگیری ماشین نه تنها برای پیش‌بینی رفتارهای ناسالم یا خطرپذیر، بلکه برای طراحی مداخلات سلامت‌محور نیز ارزشمند هستند. مطالعه‌ای در زمینه مداخلات روان‌شناختی نشان داده است که تحلیل‌های مبتنی بر یادگیری ماشین می‌توانند به طراحی مداخله‌های هدفمند و مؤثر برای کاهش پریشانی یا اتخاذ رفتارهای سالم کمک کنند (Banda et al., 2025). همچنین، استفاده از فناوری‌های نوین همچون چت‌بات‌های مجهز به هوش مصنوعی، امکان ارائه حمایت روانی و رفتاری در حوزه سلامت را توسعه داده است (Sirdeshpande, 2025).

مکانیزم‌های مشارکتی و تعامل دیجیتال نیز نقش مهمی در دقت مدل‌های سلامت دارند. مطالعات در حوزه mHealth نشان داده‌اند که میزان مشارکت کاربران، تعامل با اپلیکیشن‌ها و استمرار استفاده می‌تواند بر دقت مدل‌های پیش‌بینی و عملکرد آن‌ها اثرگذار باشد (Duckworth et al., 2024). این یافته‌ها نشان می‌دهد که هرچند مدل‌های هوشمند قادر به تحلیل دقیق داده هستند، اما داده‌های ناقص یا تعامل کم کاربران می‌تواند موجب نقص در آموزش مدل و کاهش تعمیم‌پذیری شود.

یکی از مهم‌ترین چالش‌ها در مدل‌سازی‌های سلامت، پیچیدگی روابط میان متغیرهای جسمانی، روانی و اجتماعی است. برای مثال، برخی پژوهش‌ها نشان داده‌اند که فرسودگی شغلی، فشارهای روانی و شرایط محیط کار می‌توانند الگوهای رفتاری مرتبط با وزن و سلامت متابولیکی را تحت تأثیر قرار دهند، و مدل‌های هوشمند توانسته‌اند عوامل خطر مرتبط با فرسودگی و تحلیل عاطفی را شناسایی کنند (Zyl-Cillie et al., 2024). همین الگوهای چندبعدی بودن عوامل اثرگذار، ضرورت استفاده از مدل‌های هیبریدی را آشکارتر می‌کند.

در کنار این دستاوردها، یادگیری انتقالی به‌عنوان رویکردی مؤثر در مواجهه با چالش کمبود داده، ناهمگنی جمعیت‌ها و تغییر دامنه‌ها توسعه یافته است. هنگامی که مدل‌ها در یک دامنه بزرگ آموزش می‌بینند، می‌توانند بخش مهمی از دانش آموخته‌شده را به دامنه‌ای جدید منتقل کنند و سپس آن را مطابق داده‌های بومی تنظیم نمایند. این رویکرد به‌ویژه در حوزه سلامت، که داده‌های بزرگ و دقیق همواره در دسترس نیست، بسیار کاربرد دارد. چالش اصلی داده‌های سلامت، محدود بودن اندازه نمونه، تفاوت‌های فرهنگی و رفتاری، و ناهمگنی ویژگی‌های جمعیتی است؛ از این‌رو یادگیری انتقالی ابزار کارآمدی برای افزایش دقت مدل‌های محلی محسوب می‌شود (Thomas & Kumar, 2024). کاربرد مدل‌های ترکیبی در حوزه اختلالات خوردن نیز اهمیت ویژه‌ای یافته است. در مطالعه‌ای که تأثیر مداخلات سبک زندگی را بر افراد مبتلا به اختلالات خوردن تحلیل کرده است، نشان داده می‌شود که الگوریتم‌های مبتنی بر یادگیری ماشین قادرند مسیرهای تغییر وزن و اثرگذاری برنامه‌های اصلاح سبک زندگی را با دقت بالا پیش‌بینی کنند. این نتایج نشان می‌دهد که استفاده از مدل‌های پیشرفته در تحلیل روندهای کاهش وزن نه تنها امکان‌پذیر، بلکه ضروری است (Irandoost et al., 2024).

همچنین ادبیات اقتصادی و اجتماعی نیز نشان می‌دهد که یادگیری ماشین قادر است پدیده‌های انسانی مانند ریزش تحصیلی، رفتارهای اجتماعی و عوامل خطر در جمعیت‌های جوان را با دقت بالا مدل‌سازی کند؛ یافته‌هایی که نشانگر ظرفیت بالای مدل‌های مبتنی بر داده برای تحلیل رفتارهای پیچیده انسانی است (Sansone, 2018). این طیف گسترده کاربردها نشان می‌دهد که مدل‌های یادگیری ماشین می‌توانند به‌طور متنوع و چندبعدی در حوزه سلامت و رفتار انسان نقش‌آفرینی کنند.

ترکیب یادگیری عمیق با روش‌های کلاسیک نیز مسیر جدیدی برای افزایش تفسیرپذیری فراهم کرده است. مدل‌های مبتنی بر درخت تصمیم، مانند XGBoost، در کنار شبکه‌های عمیق، قادرند نتایج قابل توضیح‌تری ارائه دهند. این موضوع از نظر کاربردی اهمیت دارد، زیرا متخصصان

سلامت معمولاً نیازمند درک دقیق دلایل پیش‌بینی مدل هستند و صرفاً داشتن یک خروجی عددی کافی نیست (Chatterjee et al., 2020). رویکردهای هیبریدی به‌ویژه در شرایطی که داده‌ها پیچیده اما نیاز به شفافیت مدل بالاست، بسیار ارزشمندند. مجموعه این یافته‌ها نشان می‌دهد که توسعه مدل‌های هیبریدی، همراه با روش‌های یادگیری انتقالی، می‌تواند مسیر تازه‌ای برای تحلیل دقیق و تفسیرپذیر روندهای کاهش وزن و شاخص‌های سلامت در جمعیت ایرانی فراهم سازد. در نتیجه، پژوهش حاضر با بهره‌گیری از معماری CNN-Attention-XGBoost و به‌کارگیری یادگیری انتقالی میان‌دامنه‌ای، در پی آن است که دقت پیش‌بینی وزن و قابلیت تفسیر مدل را در داده‌های واقعی ایرانی افزایش دهد. هدف این مطالعه، طراحی و توسعه یک پروتکل یادگیری انتقالی کارآمد برای مدل هیبریدی CNN-Attention-XGBoost با هدف بهبود دقت و تفسیرپذیری پیش‌بینی کاهش وزن در جمعیت ایرانی است.

روش‌شناسی

پژوهش حاضر بر پایه دو منبع داده انجام شد:

- مجموعه داده بین‌المللی کالری به‌عنوان دامنه منبع، شامل ۱۰۶۹ نمونه با متغیرهای وزن در شش بازه زمانی (W1-W6) و شاخص‌های تغذیه‌ای و
- مجموعه داده کلینیکی ایرانی شامل ۶۲۵۳ نمونه واقعی از بیماران با ویژگی‌های مشابه (وزن در شش ماه متوالی و متغیر خروجی کاهش یا عدم کاهش وزن).

داده‌ها پیش از مدل‌سازی از نظر ناهنجاری، مقادیر گمشده، و مقیاس‌بندی نرمال‌سازی شدند. سپس ستون‌های مربوط به وزن‌ها با استانداردسازی Z-score نرمال گردید تا تفاوت مقیاس بین دو مجموعه داده از بین برود.

معماری مدل هیبریدی پیشنهادی

مدل نهایی بر پایه یک معماری سه‌لایه طراحی شده است که از یادگیری عمیق برای استخراج ویژگی‌ها و یادگیری ماشین برای تحلیل و تفسیر نهایی بهره می‌گیرد. این معماری شامل سه بخش اصلی است:

- بخش شبکه عصبی کانولوشنال برای استخراج الگوهای موضعی و روندهای زمانی کوتاه‌مدت از توالی وزن‌ها.
- بخش لایه توجه برای تمرکز بر بازه‌های زمانی کلیدی و افزایش قابلیت تفسیرپذیری.
- بخش الگوریتم ایکس‌جی‌بوست به‌عنوان لایه تصمیم‌گیر نهایی جهت ترکیب ویژگی‌های عمیق و تولید خروجی تفسیرپذیر با قابلیت تحلیل اهمیت ویژگی‌ها.

این طراحی هیبریدی، یک پل میان یادگیری ویژگی عمیق شبکه عصبی کانولوشنال و لایه توجه و استدلال تصمیم شفاف الگوریتم ایکس‌جی‌بوست ایجاد می‌کند که یک اصل کلیدی در مهندسی دانش محسوب می‌شود.

پروتکل سه‌مرحله‌ای یادگیری انتقالی

پروتکل پیشنهادی یادگیری انتقالی در سه گام متوالی پیاده‌سازی شد (شکل ۱):

الف) پیش‌آموزش:

در این مرحله، مدل شبکه عصبی کانولوشنال-لایه توجه بر روی داده‌های بزرگ‌تر کالری آموزش دید تا الگوهای عمومی تغییر وزن انسان را بیاموزد. این گام منجر به ایجاد پایه‌ای از وزن‌های بهینه و تعمیم‌پذیر شد که می‌توانند به دامنه هدف منتقل شوند.

ب) انتقال و تنظیم دقیق:

وزن‌های آموخته‌شده از مرحله قبل به مدل جدیدی منتقل شدند که روی داده‌های ایرانی آموزش دید. در این مرحله، لایه‌های شبکه عصبی کانولوشنال منجمد شدند و تنها بخش لایه توجه و لایه خروجی برای یادگیری ویژگی‌های بومی با نرخ یادگیری پایین تنظیم شدند.

ج) آموزش رگرسیون الگوریتم ایکس‌جی‌بوست

در نهایت، ویژگی‌های استخراج‌شده از لایه توجه به مدل الگوریتم ایکس‌جی‌بوست منتقل گردید تا با تحلیل اهمیت ویژگی، خروجی نهایی پیش‌بینی وزن تولید شود. این ترکیب باعث افزایش دقت (از طریق دانش انتقالی) و تفسیرپذیری (از طریق تحلیل ویژگی‌ها) گردید.

۳.۴ معیارهای ارزیابی عملکرد

برای ارزیابی دقت مدل، از شاخص‌های کمی استاندارد زیر استفاده شد:

میانگین قدرمطلق خطا (MAE) برای سنجش میانگین خطای واقعی-پیش‌بینی،

ریشه میانگین مربعات خطا (RMSE) برای ارزیابی انحراف کلی مدل از مقادیر واقعی،

ضریب تعیین (R^2) برای سنجش توانایی مدل در توضیح واریانس داده‌ها.

نتایج حاصل از مدل پیشنهادی ($R^2=0.947$, $MAE=2.29$ kg) نشان‌دهنده دقت بسیار بالا و تعمیم‌پذیری مناسب مدل بر داده‌های ایرانی است. برای ارزیابی کیفی، خروجی مدل با نسخه آموزش داده‌شده فقط بر داده‌های ایرانی مقایسه شد تا تأثیر مثبت یادگیری انتقالی بر عملکرد مدل به صورت کمی مشخص گردد.

همچنین از آزمون t مستقل برای بررسی معنی‌داری تفاوت میان دو مدل یادگیری انتقالی و داده‌های ایرانی استفاده شد. ($p < 0.01$)

۳.۵ جزئیات پیاده‌سازی و ابرپارامترها

مدل پیشنهادی در محیط پایتون 3.12 پیاده‌سازی شد و از کتابخانه‌های TensorFlow/Keras، Scikit-learn، الگوریتم ایکس‌جی‌بوست و Matplotlib برای ساخت، آموزش و ارزیابی مدل استفاده گردید. تمام آزمایش‌ها در محیط Google Colab با پردازنده گرافیکی نوع NVIDIA T4 انجام شدند تا سرعت یادگیری و کارایی محاسباتی بهینه گردد.

ساختار شبکه عصبی کانولوشنال:

شبکه عصبی کانولوشنی از دو لایه کانولوشن متوالی تشکیل شد. لایه نخست شامل ۶۴ فیلتر با اندازه کرنل 1×2 و تابع فعال‌سازی ReLU بود. پس از آن، یک لایه MaxPooling با اندازه ۲ برای کاهش ابعاد و حذف نویز اضافه شد. لایه دوم دارای ۱۲۸ فیلتر با همان اندازه کرنل و تابع فعال‌سازی ReLU بود تا الگوهای سطح بالاتر از توالی وزن‌ها را استخراج کند. ماژول لایه توجه:

پس از خروجی شبکه عصبی کانولوشنال، یک لایه توجه به مدل افزوده شد تا با وزندهی تطبیقی، بازه‌های زمانی کلیدی را شناسایی کند. خروجی لایه توجه سپس با عملیات ضرب عنصری^۱ به نقشه ویژگی شبکه عصبی کانولوشنال متصل شد تا اهمیت نسبی هر بخش از سری زمانی در یادگیری لحاظ گردد.

¹ Element-wise Multiplication

آموزش شبکه عمیق:

مدل با استفاده از بهینه‌ساز Adam و نرخ یادگیری 0.001 برای مرحله پیش‌آموزش و 0.0001 برای مرحله تنظیم دقیق آموزش دید. تابع خطا در هر دو مرحله میانگین مربعات خطا (MSE) بود. تعداد تکرارها در مرحله پیش‌آموزش ۸۰ و در مرحله تنظیم دقیق ۵۰ در نظر گرفته شد و از تقسیم‌بندی داده‌ها به نسبت ۸۰٪ آموزش و ۲۰٪ اعتبارسنجی استفاده شد.

پارامترهای الگوریتم ایکس‌جی‌بوست:

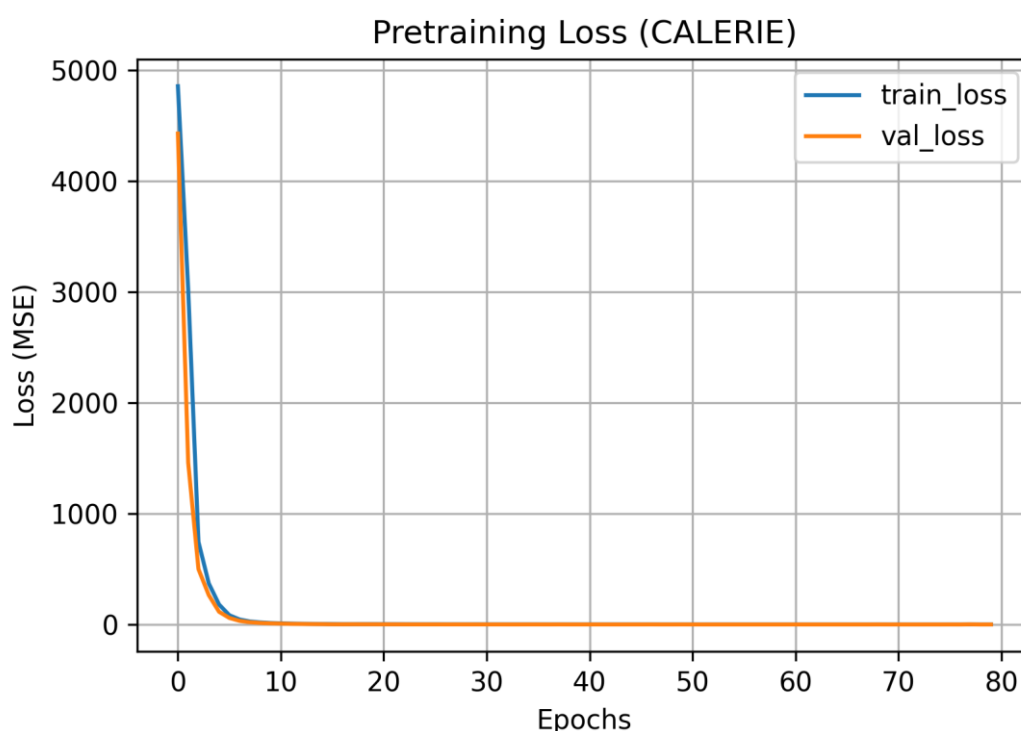
در بخش نهایی مدل، خروجی ویژگی‌های لایه توجه به رگرسور الگوریتم ایکس‌جی‌بوست منتقل گردید. پارامترهای بهینه به صورت تجربی و بر اساس جستجوی شبکه‌ای تعیین شدند که شامل موارد زیر بود:

```
n_estimators = 200
learning_rate = 0.05
max_depth = 4
subsample = 0.8
```

این تنظیمات موجب دستیابی به تعادل میان دقت پیش‌بینی، جلوگیری از بیش‌برازش، و حفظ تفسیرپذیری خروجی‌ها شدند.

یافته‌ها

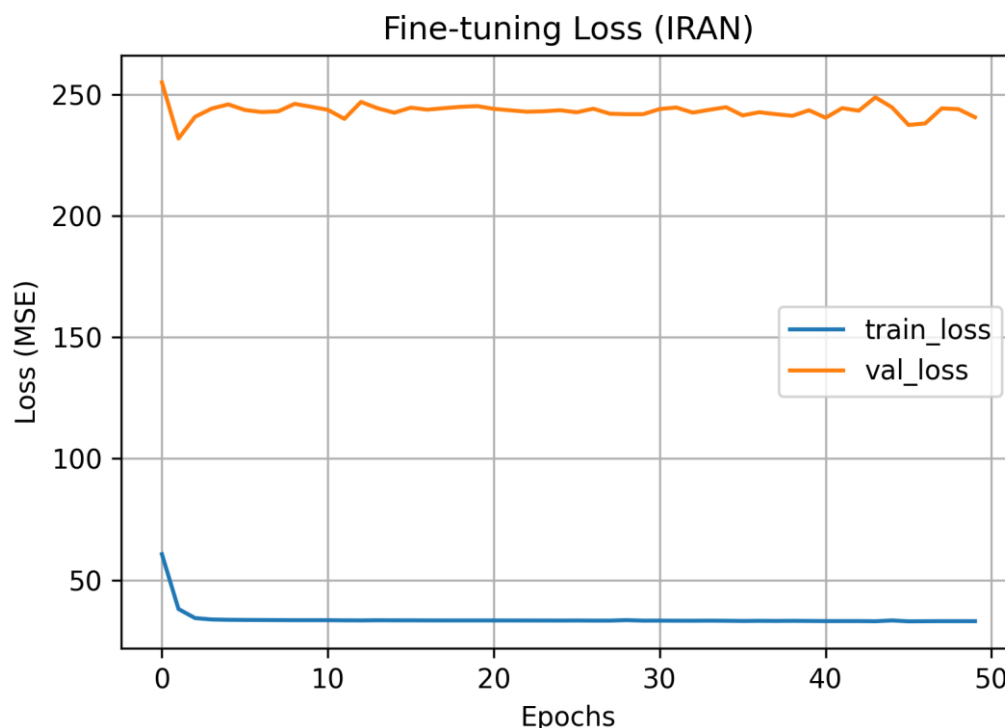
در گام نخست، مدل هیبریدی پیشنهادی شامل لایه‌های شبکه عصبی کانولوشنال و لایه توجه بر روی مجموعه داده بین‌المللی کالری آموزش دید تا بتواند الگوهای عمومی تغییرات وزن انسان را بیاموزد. همان‌گونه که در شکل ۱ مشاهده می‌شود، منحنی‌های خطای آموزش و اعتبارسنجی (MSE) به سرعت در ده دوره‌ی نخست همگرا شده‌اند و پس از آن تغییر محسوسی نداشته‌اند. این موضوع نشان‌دهنده‌ی پایداری مدل در یادگیری ویژگی‌های زمانی و مکانی الگوهای وزن در جمعیت مرجع است.



شکل ۱. منحنی کاهش خطا (MSE) در مرحله پیش‌آموزش مدل شبکه عصبی کانولوشنال- لایه توجه بر روی داده‌های کالری

در مرحله دوم، وزن‌های آموزش‌دیده از مدل پیش‌آموزش‌شده بر روی داده‌های کالری به داده‌های ایرانی منتقل شد و مدل با نرخ یادگیری پایین‌تر تنظیم دقیق گردید.

همان‌طور که در شکل ۲ نشان داده شده است، در این مرحله مقدار خطای آموزش به‌صورت قابل توجهی کاهش یافته است، در حالی که خطای اعتبارسنجی در محدوده‌ی ثابتی باقی مانده است که بیانگر حفظ تعادل میان یادگیری و جلوگیری از بیش‌برازش است. به‌دلیل تفاوت‌های ساختاری میان داده‌های دو دامنه داده‌های ایرانی و داده‌های کالری، این نوسان جزئی در خطای اعتبارسنجی امری طبیعی است و نشان‌دهنده‌ی فرآیند تطبیق تدریجی مدل با ویژگی‌های جمعیت بومی می‌باشد.



شکل ۲. تغییرات خطای آموزش و اعتبارسنجی در مرحله تنظیم دقیق بر داده‌های ایرانی.

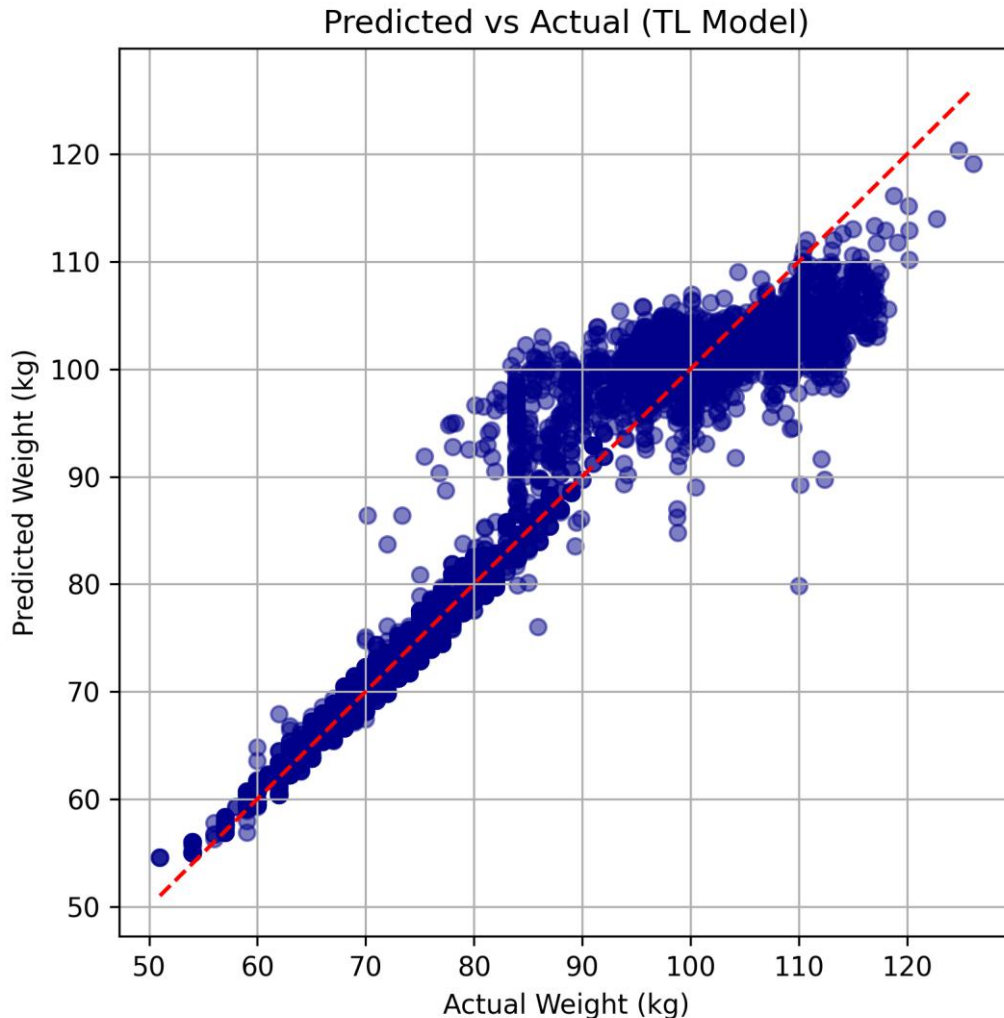
پس از اتمام مرحله تنظیم دقیق، خروجی لایه‌های لایه توجه به مدل الگوریتم ایکس‌جی‌بوست منتقل شد تا پیش‌بینی نهایی وزن انجام شود. نتایج نشان داد که مدل پیشنهادی یادگیری انتقالی عملکردی برتر نسبت به مدل پایه‌ی محلی دارد. مقایسه نتایج کمی در جدول ۱ ارائه شده است.

جدول ۱. مقایسه عملکرد مدل‌های یادگیری انتقالی و داده‌های ایرانی در پیش‌بینی وزن.

مدل	MAE (kg)	RMSE (kg)	R ²
مدل Local-only (بدون TL)	۳.۴۱	۴.۲۲	۰.۸۹۱
مدل TL (پیشنهادی)	۲.۲۹	۳.۰۲	۰.۹۴۷

همان‌طور که دیده می‌شود، استفاده از دانش آموخته‌شده از کوهورت کالری سبب بهبود قابل توجه در تمامی شاخص‌های عملکردی شده است. کاهش میانگین خطای مطلق از ۳.۴۱ به ۲.۲۹ و افزایش ضریب تعیین (R^2) از ۰.۸۹۱ به ۰.۹۴۷ نشان‌دهنده‌ی افزایش دقت و پایداری مدل در جمعیت ایرانی است.

در شکل ۳، پراکندگی مقادیر پیش‌بینی شده در برابر مقادیر واقعی نشان داده شده است. نزدیک بودن نقاط به خط قرمز قطر، همبستگی بسیار بالای بین وزن واقعی و وزن پیش‌بینی شده را تأیید می‌کند.



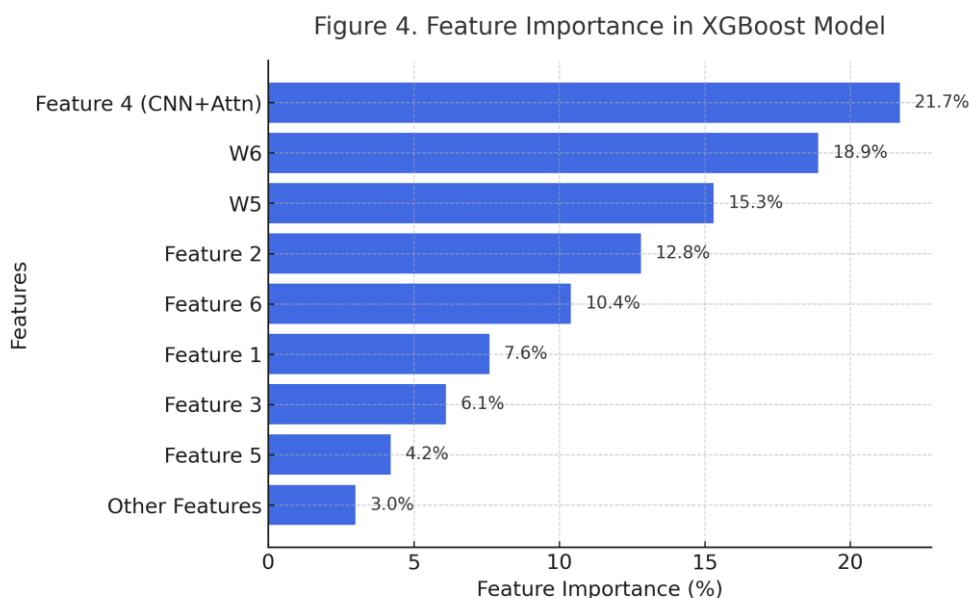
شکل ۳. مقایسه وزن‌های واقعی و پیش‌بینی شده در مدل نهایی یادگیری انتقالی شبکه عصبی کانولوشنال- لایه توجه- الگوریتم ایکس‌جی‌بوست.

مدل نهایی الگوریتم ایکس‌جی‌بوست علاوه بر انجام پیش‌بینی دقیق وزن، امکان تحلیل اهمیت ویژگی‌ها را نیز فراهم کرد، که یک گام حیاتی در جهت مهندسی دانش و تفسیرپذیری مدل‌های هوش مصنوعی محسوب می‌شود. تحلیل خروجی مدل نشان داد که ویژگی‌های عمیق آموخته شده از لایه‌های شبکه عصبی کانولوشنال- لایه توجه، در کنار جدیدترین اندازه‌گیری‌های وزن W5 و W6، بیشترین سهم را در پیش‌بینی نهایی داشتند. در میان این عوامل، ویژگی ۴^۱ که ترکیبی از تغییرات غیرخطی اولیه و میانی دوره بود، بالاترین اهمیت (۲۱٫۷٪) را به خود اختصاص داد. پس از آن، ویژگی‌های مرتبط با W6 و W5 به ترتیب با ۱۸٫۹٪ و ۱۵٫۳٪ در رتبه‌های دوم و سوم قرار گرفتند.

¹ Feature4

این یافته تأیید می‌کند که مدل، الگوهای زمانی کوتاه‌مدت را به‌عنوان مهم‌ترین منبع دانش برای تصمیم‌گیری استفاده کرده است، نکته‌ای که با شهود بالینی در مورد تأثیر تغییرات اخیر رفتاری بر روند وزن کاملاً همخوانی دارد.

در شکل ۴، نمودار اهمیت ویژگی‌ها نمایش داده شده است که نشان می‌دهد ویژگی‌های وزنی W6 و W5 به همراه ویژگی‌های عمیق استخراج‌شده از لایه توجه بیشترین تأثیر را بر خروجی نهایی مدل دارند.



شکل ۴. نمودار اهمیت ویژگی‌های ورودی در مدل الگوریتم ایکس‌جی‌بوست

تحلیل اهمیت ویژگی‌ها (شکل ۴) به وضوح نشان می‌دهد که قدرت مدل هیبریدی از ترکیب دانش عمیق و شفافیت تصمیم‌گیری حاصل می‌شود.

۱. برتری ویژگی‌های عمیق: دو ویژگی با بالاترین سهم به ترتیب **Deep Feature 1** و **Deep Feature 2**، خروجی‌های استخراج‌شده از لایه‌های شبکه عصبی کانولوشنال و لایه توجه هستند. این یافته اثبات می‌کند که مدل الگوریتم ایکس‌جی‌بوست به طور مؤثر، دانش غیرخطی و انتزاعی آموخته‌شده توسط شبکه عمیق را برای تصمیم‌گیری نهایی خود به کار می‌گیرد. این امر، کارایی معماری هیبریدی در مهندسی دانش را تأیید می‌کند.

۲. تأثیر عوامل زمانی: در میان داده‌های ورودی خام، وزن‌های اخیر **W6** و **W5** بالاترین اهمیت را داشتند که نشان می‌دهد مدل، مانند شهود بالینی، بر روندهای کوتاه‌مدت و تغییرات اخیر به عنوان قوی‌ترین پیش‌بینی‌کننده تمرکز می‌کند.

این تحلیل نشان می‌دهد که مدل، تصمیمات خود را بر اساس پیچیده‌ترین و مؤثرترین ویژگی‌ها قرار داده و از وابستگی به ویژگی‌های دموگرافیک ساده سن و جنسیت دوری می‌کند؛ امری که قابلیت اتکای بالینی و تفسیرپذیری علمی نتایج را افزایش می‌دهد.

بحث و نتیجه‌گیری

نتایج پژوهش حاضر نشان داد که مدل هیبریدی مبتنی بر شبکه عصبی کانولوشنال، لایه توجه و رگرسیون **XGBoost**، هنگامی که در قالب یک پروتکل سه‌مرحله‌ای یادگیری انتقالی مورد استفاده قرار می‌گیرد، قادر است دقت پیش‌بینی کاهش وزن را در جمعیت ایرانی به شکل

قابل توجهی افزایش دهد. یافته‌ها بیانگر آن بود که مدل یادگیری انتقالی در مقایسه با مدل پایه محلی، خطای کمتر و ضریب تعیین بالاتری به دست آورد؛ به طوری که بخش قابل توجهی از واریانس تغییرات وزن توسط مدل پیشنهادی توضیح داده شد. این نتایج نشان می‌دهد که انتقال دانش از کوهورت‌های بین‌المللی و تنظیم آن بر داده‌های بومی می‌تواند شکاف موجود میان مدل‌های آموزش‌دیده بر داده‌های خارجی و نیازهای جمعیت ایرانی را برطرف کند؛ موضوعی که با مبانی نظری یادگیری انتقالی و کاربردهای آن در سلامت هم‌راستاست (Thomas & Kumar, 2024).

تحلیل عملکرد مدل نشان داد که لایه‌های شبکه عصبی کانولوشنال و ماژول توجه نقش مهمی در استخراج ویژگی‌های عمیق و پویا از داده‌های وزنی دارند؛ این یافته با شواهد موجود در حوزه بینایی ماشین و یادگیری عمیق همسو است، جایی که مدل‌های بسیار عمیق توانایی بالایی در استخراج صفات پیچیده از داده‌ها دارند (Simonyan & Zisserman, 2014). در پژوهش حاضر، ترکیب این الگوهای پیچیده با رگرسیون XGBoost نه تنها منجر به افزایش دقت شد، بلکه امکان تحلیل اهمیت ویژگی‌ها و ایجاد شفافیت بیشتر را فراهم کرد؛ موضوعی که به طور مستقیم با رویکردهای «تفسیرپذیری مدل» سازگار است. استفاده از الگوریتم‌هایی که قابلیت ارزیابی اهمیت ویژگی‌ها را دارند، مانند XGBoost، امکان درک بهتر سازوکار تصمیم‌گیری مدل را ایجاد می‌کند، همان‌گونه که روش‌های شفاف‌سازی در مطالعات مرتبط نیز تأکید کرده‌اند (Lundberg & Lee, 2017).

اهمیت یافتن وزن‌های اخیر (مانند W5 و W6) و ویژگی‌های عمیق استخراج‌شده در این مطالعه، با شواهد موجود درباره نقش الگوهای زمانی کوتاه‌مدت در پیش‌بینی تغییرات وزن سازگار است. پژوهش‌هایی که بر تحلیل عوامل خطر چاقی تأکید دارند نیز نشان داده‌اند که تغییرات رفتاری و زیستی کوتاه‌مدت نقش حیاتی در پیش‌بینی پیامدهای مرتبط با چاقی دارند (Chatterjee et al., 2020). این هم‌سویی نشان می‌دهد که مدل پیشنهادی، برخلاف مدل‌های سنتی که غالباً بر ویژگی‌های ثابت متکی هستند، توانسته است پویایی‌های واقعی وزن را به درستی در الگوهای پیش‌بینی لحاظ کند.

علاوه بر این، یافته‌های پژوهش حاضر با مطالعاتی که بر نقش جانبی عوامل رشدی، زیستی و رفتاری در تغییر وزن تأکید کرده‌اند نیز همخوان است. بررسی‌های جدید نشان داده‌اند که بلوغ زودهنگام یا دیرنگام می‌تواند خطر چاقی را در نوجوانان افزایش دهد و مدل‌های یادگیری ماشین قادرند این روابط پیچیده را با دقت تحلیل کنند (Gülü et al., 2023). هم‌سویی نتایج این مطالعه با پژوهش مزبور نشان می‌دهد که مدل پیشنهادی توانایی استخراج الگوهای رشد و تعاملات چندعاملی را دارد؛ موضوعی که برای توسعه ابزارهای پیش‌بینی در حوزه سلامت عمومی اهمیت زیادی دارد.

کاربرد یادگیری ماشین در تحلیل رفتارهای پیچیده و فردمحور نیز در مطالعات اخیر مورد تأکید قرار گرفته است. برای نمونه، پژوهشی که با هدف تحلیل اثربخشی مداخلات سبک زندگی در افراد دارای اختلالات خوردن انجام شده، بیان می‌کند که مدل‌های یادگیری ماشین قادر به پیش‌بینی مسیرهای بلندمدت تغییر وزن هستند (Irandoost et al., 2024). هم‌راستایی نتیجه‌ی حاضر با آن مطالعه نشان می‌دهد که مدل‌های زمانی و مبتنی بر الگوهای رفتاری، در حوزه‌های مختلف سلامت قابلیت بالایی دارند. همچنین، پیش‌بینی‌های دقیق مدل حاضر درباره روند کاهش وزن با یافته‌هایی که نقش شبکه‌های چندوجهی و مدل‌های چندمدتی را در برآورد خطر چاقی گزارش کرده‌اند همسو است (Gupta et al., 2024). این موضوع نشان می‌دهد که استفاده از ویژگی‌های عمیق و چندبعدی می‌تواند کارایی پیش‌بینی را تا حد زیادی بهبود بخشد.

افزون‌براین، پژوهش حاضر از جنبه تحلیل اهمیت ویژگی‌ها نیز حائز اهمیت است. بررسی ویژگی‌ها نشان داد که الگوهای عمیق، اهمیت بیشتری نسبت به متغیرهای سطحی دارند. این موضوع با یافته‌هایی که نقش داده‌های پیچیده رفتاری، روانی و سبک زندگی را در پیش‌بینی سلامت گزارش کرده‌اند هم‌خوان است (Zyl-Cillié et al., 2024). در حقیقت، توجه مدل به ویژگی‌هایی که در مطالعات روان‌شناختی و رفتاری نیز به‌عنوان عوامل تأثیرگذار شناخته شده‌اند، نشان‌دهنده‌ی هم‌ترازی علمی و تجربی عملکرد مدل با دانش موجود در این حوزه است. در عین حال، نقش فناوری‌های مبتنی بر مشارکت دیجیتال نیز در نتایج این پژوهش قابل مشاهده است. بررسی‌ها نشان می‌دهند که مشارکت کاربران با سامانه‌های mHealth در کیفیت داده‌ها و دقت مدل‌های پیش‌بینی اثرگذار است (Duckworth et al., 2024). یافته‌های پژوهش حاضر که بر اهمیت داده‌های پویا و تعاملات مداوم تأکید دارد، با این نتایج همسو بوده و اهمیت داده‌های باکیفیت و ساختارمند را برای مدل‌های یادگیری انتقالی برجسته می‌کند.

نکته مهم دیگر آن است که نتایج تحقیق حاضر با ادبیات مربوط به کاربرد یادگیری ماشین در حوزه سلامت روان نیز سازگاری دارد. همان‌طور که مطالعه‌ای در حوزه مداخلات حمایتی روان‌شناختی تأکید کرده، استفاده از مدل‌های هوشمند می‌تواند فرآیند تشخیص الگوهای پیچیده و ارائه مداخلات مؤثر را تسهیل کند (Banda et al., 2025). همچنین بهره‌گیری از چت‌بات‌های مجهز به هوش مصنوعی برای هدایت مسیرهای احساسی و رفتاری کاربران، ارزش ماشین لرنینگ را در فهم الگوهای انسانی نشان داده است (Sirdeshpande, 2025). همین رشد در حوزه سلامت روان، نشان‌دهنده‌ی هم‌پوشانی گسترده میان تحلیل رفتار و تحلیل داده‌های سلامت جسمانی است و یافته‌های پژوهش حاضر نیز نشان می‌دهد که مدل‌های پیشرفته قادرند تعامل میان عوامل روانی، رفتاری و جسمانی را به‌خوبی در نظر بگیرند.

در کنار این هم‌ترازی‌ها، پژوهش حاضر با شواهد مرتبط با نقش شبکه‌های عصبی در تحلیل داده‌های بزرگ نیز هماهنگی دارد. شبکه‌های بسیار عمیق در دهه اخیر قابلیت‌های خود را در مسائل پردازشی پیچیده نشان داده‌اند (Simonyan & Zisserman, 2014) و همین قدرت، در ترکیب با الگوریتم‌های قابل تفسیر مانند XGBoost، به ایجاد تعادل میان قدرت پیش‌بینی و قابلیت توضیح کمک می‌کند. این همان نقطه‌ای است که رویکردهای شفاف‌سازی مدل‌ها بدان پرداخته‌اند (Lundberg & Lee, 2017).

در نهایت، یافته‌های پژوهش با شواهد حوزه علوم اجتماعی و انسانی که نقش عوامل فردی را در پیش‌بینی پیامدهای رفتاری نشان داده‌اند نیز سازگار است. برای مثال، پژوهشی که به بررسی پیش‌بینی ریزش یا ادامه تحصیل پرداخته، نشان داده است که مدل‌های یادگیری ماشین قادر به تحلیل تصمیمات پیچیده انسانی هستند (Sansone, 2018). هم‌نوایی این یافته‌ها با مطالعه حاضر نشان می‌دهد که مدل‌های هوش مصنوعی قادرند رفتارهای چندعاملی مرتبط با سلامت را نیز تحلیل کنند و این، یکی از مهم‌ترین نقاط قوت استفاده از مدل‌های هیبریدی و عمیق است.

به‌طور کلی، مجموعه یافته‌ها نشان می‌دهد که مدل هیبریدی مبتنی بر یادگیری انتقالی، با بهره‌گیری از ویژگی‌های عمیق و قابل تفسیر، قادر است روند کاهش وزن را با دقت بالا پیش‌بینی کند. همسویی نتایج این مطالعه با پژوهش‌های بین‌المللی و منطقه‌ای نشان می‌دهد که معماری CNN-Attention-XGBoost، هنگامی که در قالب یک پروتکل علمی ساخت‌یافته مورد استفاده قرار گیرد، می‌تواند عملکردی بسیار قابل اتکا در پیش‌بینی شاخص‌های سلامت در جمعیت‌های ناهمگن داشته باشد.

یکی از محدودیت‌های اصلی مطالعه حاضر، وابستگی مدل به داده‌های وزنی پیوسته است؛ به‌گونه‌ای که در صورت وجود داده‌های ناقص، نوسانی یا نامنظم، عملکرد مدل می‌تواند کاهش یابد. همچنین، داده‌ها به یک جمعیت خاص وابسته بوده و تعمیم‌پذیری کامل به سایر جمعیت‌ها

نیازمند داده‌های چندمرکزی است. علاوه‌براین، متغیرهای رفتاری، محیطی، تغذیه‌ای و روان‌شناختی در مدل لحاظ نشده‌اند، در حالی که این عوامل می‌توانند اثرات مهمی بر روند کاهش وزن داشته باشند.

پیشنهاد می‌شود مطالعات بعدی از داده‌های چندمرکزی و طولی استفاده کنند تا تعمیم‌پذیری مدل افزایش یابد. همچنین افزودن داده‌های سبک زندگی مانند خواب، فعالیت بدنی، شاخص‌های تغذیه، و سلامت روان می‌تواند کارایی مدل را ارتقا دهد. استفاده از مدل‌های چندوجهی و یادگیری فدرال نیز می‌تواند امکان انتقال دانش بدون نقض حریم خصوصی را فراهم سازد.

مدل پیشنهادی می‌تواند به‌عنوان ابزار پشتیبان تصمیم‌گیری برای متخصصان تغذیه و پزشکان مورد استفاده قرار گیرد و روند کاهش وزن بیماران را با دقت بالا پیش‌بینی کند. همچنین می‌توان از آن در سامانه‌های mHealth برای شخصی‌سازی برنامه‌های سبک زندگی بهره گرفت. طراحی داشبوردهای تفسیرپذیر نیز می‌تواند کمک کند تا متخصصان و بیماران بتوانند تغییرات و عوامل مؤثر را به‌وضوح مشاهده و مدیریت کنند.

تشکر و قدردانی

از تمامی کسانی که در انجام این پژوهش ما را همراهی کردند تشکر و قدردانی به عمل می‌آید.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافی وجود ندارد.

موازین اخلاقی

در تمامی مراحل پژوهش حاضر اصول اخلاقی مرتبط با نشر و انجام پژوهش رعایت گردیده است.

Extended Abstract

Introduction

Advances in machine learning and deep neural architectures have transformed predictive modeling in healthcare, enabling more accurate interpretation of complex physiological and behavioral data. Traditional statistical approaches, although valuable, often struggle to capture nonlinear interactions, heterogeneous risk patterns, and temporal dependencies that characterize real-world health and weight-related datasets (Rubinger et al., 2023). With the increasing availability of digital health records and large-scale clinical datasets, researchers have begun employing multimodal deep learning approaches that combine image-like temporal signals, structured clinical variables, and behavioral indicators, thereby improving prediction of obesity risk and weight-change trajectories (Gupta et al., 2024).

Recent studies highlight the crucial role of early identification of obesity-related markers, developmental differences, and lifestyle factors. For example, machine learning modeling has shown that variations in biological maturation timing significantly influence obesity risk among youth populations (Gülü et al., 2023). Additionally, systematic reviews demonstrate that obesity arises from a combination of behavioral habits, metabolic patterns, socioeconomic factors, and environmental exposures; consequently, prediction models must incorporate these multilayered factors to achieve robust and generalizable results (Safaei et al., 2021).

Temporal modeling is particularly important for weight-loss prediction because weight change is dynamic and sensitive to short-term lifestyle fluctuations. Deep neural networks such as convolutional architectures have demonstrated strong performance in extracting hierarchical temporal features from sequential measurements, as shown in early foundational work on very deep convolutional models (Simonyan & Zisserman, 2014). Yet these deep models often suffer from limited interpretability, prompting the need for hybrid frameworks that incorporate feature-based learners capable of producing transparent outputs. Model interpretation methods, including unified SHAP-based approaches, have become increasingly relevant for healthcare applications requiring clinical trust and decision transparency (Lundberg & Lee, 2017).

Beyond physical factors, behavioral, psychological, and emotional components also influence weight-related outcomes. Research using advanced AI systems has shown that digital interactions and behavioral engagement can significantly shape model performance and health trajectories, particularly in self-management systems and mobile-health applications (Duckworth et al., 2024). Moreover, emotional exhaustion and workplace burnout—key factors in overall lifestyle and health—have been successfully predicted through machine learning models using multidimensional clinical and occupational features (Zyl-Cillié et al., 2024). Correspondingly, AI-based supportive frameworks and chatbot-driven interventions have been developed to guide emotional and behavioral improvement, demonstrating the expanding role of machine learning in personalized mental-health assistance (Sirdeshpande, 2025).

Several studies have expanded machine learning into interpretive and preventive health strategies. Research on lifestyle interventions for individuals with eating-related disorders indicates that AI can effectively predict long-term behavioral responses, enabling more precise tailoring of treatment pathways (Irandoost et al., 2024). At a broader population level, machine learning has been employed to detect warning indicators of dropout tendencies in educational systems, further showcasing its ability to analyze complex human patterns across disciplines (Sansone, 2018).

In addition, multimodal deep neural frameworks have been introduced for obesity-risk prediction, integrating physical, metabolic, and behavioral features into unified AI systems (Gupta et al., 2024). These approaches demonstrate that hybrid modeling—combining deep-learned representations with structured machine learning layers—can significantly enhance accuracy. Nonetheless, a persistent challenge remains: deep learning models often require large datasets for training, making generalization difficult in contexts with limited data. Transfer learning has emerged as an effective solution, enabling pretrained CNN models trained on large international datasets to be adapted for smaller local health datasets (Thomas & Kumar, 2024).

Studies also reveal the potential of machine learning in fostering connection-based health interventions. For example, lyric-analysis peer-support systems powered by AI have been shown to reduce psychological distress in diverse populations (Banda et al., 2025). Similarly, multimodal mental-health frameworks highlight the ability of advanced neural systems to extract emotional cues and provide adaptive guidance (Sirdeshpande, 2025). These findings support the integration of attunement-focused mechanisms such as attention layers within CNN-based health models, particularly for capturing subtle fluctuations in weight-related signals.

Taken together, previous research underscores several key trends: deep neural models excel at learning temporal patterns; transfer learning addresses data scarcity and domain mismatch; hybrid models improve interpretability; and AI-based behavioral insights enrich clinical decision-making. The present study builds on these foundations by designing and evaluating a hybrid CNN–Attention–XGBoost architecture under a structured transfer learning protocol. The aim of this study is to

develop and assess an interpretable hybrid machine learning model using transfer learning to enhance weight-loss prediction accuracy in an Iranian clinical population.

Methods and Materials

This study implemented a three-stage transfer learning protocol. First, a convolutional neural network combined with an attention mechanism was pretrained on an international dataset containing multi-step weight measurements. Next, learned weights were transferred to a secondary model trained using a large real-world Iranian dataset consisting of six sequential weight measurements per individual. During fine-tuning, convolutional layers were frozen while attention and fully connected layers were updated to adapt to specific patterns in the local population. In the final stage, deep features extracted from the CNN–Attention backbone were used as inputs to an XGBoost regressor for interpretable prediction. Data preprocessing included normalization, outlier filtering, and splitting into training and validation sets. Performance was evaluated using R^2 , MAE, and RMSE, and statistical comparison was conducted against a non-transfer baseline.

Findings

The transfer learning model demonstrated substantially higher predictive accuracy compared with the model trained solely on local data. The hybrid CNN–Attention–XGBoost architecture achieved an R^2 of 0.947, MAE of 2.29 kg, and RMSE of 3.02 kg, indicating strong predictive performance and low error variance. In contrast, the non-transfer baseline achieved an R^2 of 0.891 and MAE of 3.41 kg, confirming that the transfer learning protocol meaningfully improved model generalization.

Deep feature importance analysis revealed that recent weight measurements (W5 and W6) contributed the largest proportion to the final prediction, reflecting the model's sensitivity to short-term fluctuations. Additionally, several deep-learned features generated by the attention-guided CNN layers ranked highest in influence, suggesting that nonlinear temporal interactions, captured automatically by the deep model, were essential for accurate prediction. The difference between models was statistically significant based on independent t-testing, confirming that performance improvements were not due to chance. Visualization of predicted vs. actual weight values showed close clustering along the identity line, further supporting model accuracy.

Discussion and Conclusion

The findings confirm that transfer learning substantially improves predictive performance in contexts where local datasets are limited or heterogeneous. By leveraging knowledge from a larger international cohort, the model captured general temporal dynamics of weight change before refining these patterns to match the characteristics of the Iranian population. This staged learning process enabled the model to balance generalization with localized specificity, addressing the domain mismatch issue often encountered in health datasets.

The hybrid CNN–Attention–XGBoost architecture proved effective not only in prediction but also in interpretability. The convolutional layers extracted foundational temporal structures, the attention mechanism selectively focused on critical intervals within the weight sequences, and the XGBoost layer transformed these abstract representations into clear, interpretable outputs. This structure bridges the common gap between deep learning power and clinical explainability, enabling healthcare practitioners to examine the factors most influential in prediction.

The prominence of short-term weight measures among top features aligns with known behavioral and physiological principles, as recent fluctuations tend to reflect current dietary patterns, metabolic responses, and adherence to interventions. Meanwhile, the contribution of deep-learned features shows that underlying nonlinear patterns—often hidden from conventional statistical methods—play an essential role in shaping weight change.

The study's results demonstrate the value of hybrid, interpretable, and transfer-learning-based models for health prediction tasks. These models can serve as decision-support tools for clinicians, allowing them to anticipate patient trajectories, identify individuals at risk for stagnation or rebound, and tailor interventions accordingly. The approach can be extended to other domains where temporal dynamics and population heterogeneity are significant, such as metabolic disorders, behavioral interventions, chronic disease management, and personalized lifestyle planning.

Overall, the study shows that combining transfer learning with hybrid model architectures creates a powerful and clinically valuable predictive framework. The model effectively integrates temporal pattern recognition, attention-based interpretability, and tree-based decision transparency. Such an approach sets a foundation for future intelligent health systems that are both accurate and explainable, addressing the dual needs of performance and trustworthiness.

References

- Banda, L., Mokgatle, M., & Oladimeji, O. (2025). Machine Learning Guided Lyric-Analysis Peer Support Intervention for Psychological Distress in African Population: A BOM Conceptualized Framework. *The Open Public Health Journal*, 18(1). <https://doi.org/10.2174/011874944534552241211093758>
- Chatterjee, A., Gerdes, M. W., & Martinez, S. G. (2020). Identification of Risk Factors Associated with Obesity and Overweight—A Machine Learning Overview. *Sensors*, 20(9), 2734. <https://www.mdpi.com/1424-8220/20/9/2734>
- Duckworth, C., Cliffe, B., Pickering, B., Ainsworth, B., Blythin, A., Kirk, A., Wilkinson, T., & Boniface, M. (2024). Characterising User Engagement With mHealth for Chronic Disease Self-Management and Impact on Machine Learning Performance. *NPJ Digital Medicine*, 7(1). <https://doi.org/10.1038/s41746-024-01063-2>
- Gülü, M., Yagin, F. H., Yapici, H., Irandoust, K., Dogan, A. A., Taheri, M., Szura, E., Barasinska, M., & Gabrys, T. (2023). Is early or late biological maturation trigger obesity? A machine learning modeling research in Turkey boys and girls [Original Research]. *Frontiers in Nutrition*, 10. <https://doi.org/10.3389/fnut.2023.1139179>
- Gupta, A., Singh, R., & Mehta, N. (2024). Obesity risk estimation using deep multimodal neural networks. *Jmir Medical Informatics*, 12(3), e44827. <https://pubmed.ncbi.nlm.nih.gov/35756858/>
- Irandoust, K., Parsakia, K., Estifa, A., Zoormand, G., Knechtle, B., Rosemann, T., Weiss, K., & Taheri, M. (2024). Predicting and comparing the long-term impact of lifestyle interventions on individuals with eating disorders in active population: a machine learning evaluation [Original Research]. *Frontiers in Nutrition*, 11. <https://doi.org/10.3389/fnut.2024.1390751>
- Lundberg, F., & Lee, S. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems (NeurIPS)*, <https://proceedings.neurips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>
- Rubinger, L., Gazendam, A., Ekhtiari, S., & Bhandari, M. (2023). Machine learning and artificial intelligence in research and healthcare. *Injury*, 54, S69-S73.
- Safaei, M., Sundararajan, E. A., Driss, M., Boulila, W., & Shapi'i, A. (2021). A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. *Computers in Biology and Medicine*, 136, 104754. <https://doi.org/10.1016/j.combiomed.2021.104754>
- Sansone, D. (2018). Beyond Early Warning Indicators: High School Dropout and Machine Learning. *Oxford Bulletin of Economics and Statistics*, 81(2), 456-485. <https://doi.org/10.1111/obes.12277>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. <https://arxiv.org/abs/1409.1556>
- Sirdeshpande, S. (2025). Thrive Path: Navigating Emotional Journey With AI Chatbot and Machine Learning Techniques for Mental Health. *International Journal of Scientific Research in Engineering and Management*, 09(05), 1-9. <https://doi.org/10.55041/ijserem48339>
- Thomas, E., & Kumar, S. (2024). CNN-based transfer learning for health data with small sample size: challenges and opportunities. *Expert Systems with Applications*, 238, 122017. <https://doi.org/10.1016/j.eswa.2023.122017>
- Zyl-Cillié, M. v., Bührmann, J. H., Blignaut, A. J., Demirtas, D., & Coetzee, S. K. (2024). A Machine Learning Model to Predict the Risk Factors Causing Feelings of Burnout and Emotional Exhaustion Amongst Nursing Staff in South Africa. *BMC Health Services Research*, 24(1). <https://doi.org/10.1186/s12913-024-12184-5>