Journal of Public Administration
Print ISSN: 2008-5877
Online ISSN: 2423-5342

Analyzing and Predicting Hiring Decisions Using Machine Learning and Deep Learning

Ahmad Jafarnjad*
*Corresponding Author, Prof., Department of Industrial Management, Faculty of Industrial Management and Technology, College of Management, University of Tehran, Tehran, Iran. E-mail: jafarnjd@ut.ac.ir

Arman Rezasoltani
Ph.D. Candidate, Department of Industrial Management, Faculty of Industrial Management and Technology, College of Management, University of Tehran, Tehran, Iran. E-mail: armanrezasoltani@ut.ac.ir

Amir Mohammad Khani
Ph.D. Candidate, Department of Industrial Management, Faculty of Industrial Management and Technology, College of Management, University of Tehran, Tehran, Iran. E-mail: amir.mo.khani@ut.ac.ir

Abstract

Objective
The aim of this article is to explore the use of machine learning and deep learning algorithms to predict the outcomes of hiring decisions. Selecting the right human resources is one of the most fundamental elements in any organization or institution, as it directly influences key performance indicators and overall productivity. Since the hiring process is inherently complex and success is difficult to predict, the application of modern techniques—such as machine learning and deep learning—is recommended to enhance the accuracy of selection decisions. This research aims to investigate the effectiveness of these techniques in helping organizations make better hiring choices and avoid costly mistakes.

Methods
In this study, several data points were collected, including age, education, work experience, technical abilities, and personality traits of job applicants. These data were analyzed using

Citation: Jafarnjad, Ahmad; Rezasoltani, Arman & Khani, Amir Mohammad (2025). Analyzing and Predicting Hiring Decisions Using Machine Learning and Deep Learning. *Journal of Public Administration*, 17(2), 295-327. (*in Persian*)

Journal of Public Administration, 2025, Vol. 17, No.2, pp. 295-327
Published by University of Tehran, Faculty of Management
https://doi.org/10.22059/JIPA.2025.390322.3649
Article Type: Research Paper

Received: December 14, 2024
Received in revised form: March 02, 2025
Accepted: April 14, 2025
Published online: Jun 10, 2025

machine learning and deep learning algorithms to predict each applicant's likelihood of success in various roles. Classification models were employed to simulate hiring behaviors and predict decisions at different levels of specificity as part of the machine learning analysis. Furthermore, deep learning models were used to explore complex and nonlinear relationships between applicant characteristics and hiring outcomes. All models were trained and tested on high-quality data obtained from trusted, peer-reviewed sources, which were rigorously processed to ensure accuracy and consistency.

Results

The findings indicate that the application of machine learning and deep learning models significantly improves the accuracy of predicting hiring outcomes. Among all the models evaluated, the CatBoost algorithm performed best, achieving an accuracy of 0.9533, a precision of 0.9540, a recall of 0.8925, and an F1 score of 0.9222, outperforming the other algorithms by a notable margin. The Random Forest and XGBoost models also delivered strong performances, with precision scores of 0.9213 and 0.9500, respectively. Feature analysis revealed that technical skills, recruitment strategy, and interview scores were the most influential factors in hiring decisions. Additionally, ensemble learning models-especially CatBoost-were able to identify and model the complex effects of applicants' personality traits, which traditional machine learning models often failed to capture.

Conclusion

This study demonstrates that machine learning and deep learning algorithms can significantly enhance decision-making in workforce selection. The CatBoost algorithm performed best due to its ability to model complex and nonlinear relationships between applicant characteristics and hiring outcomes. These technologies offer the potential to reduce hiring costs, improve the quality of new hires, and boost organizational productivity. However, to maximize the benefits of these methods, organizations must collect and process applicant data consistently and accurately. They must also regularly retrain and update machine learning models to ensure continued effectiveness and adaptability in the face of evolving labor market dynamics.

Keywords: Machine learning, Deep learning, Recruitment forecasting, Forecasting algorithms, Applicant characteristics, Recruitment data analysis.

مدیریت دولتی
شاپا چاپی: ۲۰۰۸-۵۸۷۷
شاپا الکترونیکی: ۲۴۲۳-۵۳۴۲

تحلیل و پیش‌بینی تصمیم‌های استخدامی با استفاده از یادگیری ماشین و یادگیری عمیق

احمد جعفرنژاد*

* نویسنده مسئول، استاد، گروه مدیریت صنعتی، دانشکدهٔ مدیریت صنعتی و فناوری، دانشکدگان مدیریت، دانشگاه تهران، تهران، ایران.
رایانامه: jafarnjd@ut.ac.ir

آرمان رضاسلطانی

دانشجوی دکتری، گروه مدیریت صنعتی، دانشکدهٔ مدیریت صنعتی و فناوری، دانشکدگان مدیریت، دانشگاه تهران، تهران، ایران. رایانامه: armanrezasoltani@ut.ac.ir

امیرمحمد خانی

دانشجوی دکتری، گروه مدیریت صنعتی، دانشکدهٔ مدیریت صنعتی و فناوری، دانشکدگان مدیریت، دانشگاه تهران، تهران، ایران. رایانامه: amir.mo.khani@ut.ac.ir

چکیده

هدف: هدف این مقاله، بررسی استفاده از یادگیری ماشین و الگوریتم‌های یادگیری عمیق برای پیش‌بینی نتایج تصمیم‌های استخدام است. انتخاب نیروی انسانی مناسب، یکی از مسائل مهم و اساسی در همهٔ سازمان‌ها و نهادهاست که بر ویژگی‌های عملکرد و بهره‌وری تأثیر مستقیمی دارد. از آنجایی که فرایند استخدام بسیار پیچیده و پیش‌بینی موفقیت در استخدام دشوار است، استفاده از تکنیک‌های مدرن مانند یادگیری ماشینی و یادگیری عمیق، برای بهبود دقت و صحت انتخاب پیشنهاد می‌شود. هدف این پژوهش بررسی توانایی این تکنیک‌ها برای کمک به سازمان‌ها در اتخاذ تصمیم‌های استخدامی بهتر و دور نگه‌داشتن آن‌ها از تصمیم‌گیری‌های پُرهزینه است.

روش: در این پژوهش از داده‌های مختلف شامل سن، تحصیلات، سابقهٔ کار، توانایی‌های فنی و مهارت‌ها و ویژگی‌های شخصیتی متقاضیان شغل استفاده شد. این داده‌ها به‌کمک الگوریتم‌های یادگیری ماشین و یادگیری عمیق، برای پیش‌بینی احتمال موفقیت متقاضیان در مشاغل مختلف تجزیه‌وتحلیل شدند. مدل‌های طبقه‌بندی برای شبیه‌سازی رفتارهای استخدام و پیش‌بینی تصمیم‌ها در سطوح مختلف ویژگی و به‌عنوان بخشی از مدل‌های یادگیری ماشین استفاده شد. علاوه‌براین، رابطهٔ بین ویژگی‌های متقاضی و نتایج استخدام از طریق مدل‌های یادگیری عمیق، برای شناسایی هرگونه روابط غیرخطی و پیچیده مورد بررسی قرار گرفت. مدل‌ها روی تمام داده‌های به‌دست‌آمده از منابع مورد اعتماد و بازبینی شده‌ای به‌دقت پردازش شدند، آموزش دیدند و ارزیابی شدند.

استناد: جعفرنژاد، احمد؛ رضاسلطانی، آرمان و خانی، امیرمحمد (۱۴۰۴). تحلیل و پیش‌بینی تصمیم‌های استخدامی با استفاده از یادگیری ماشین و یادگیری عمیق. *مدیریت دولتی*، ۱۷(۲)، ۲۹۵-۳۲۷.

مدیریت دولتی، ۱۴۰۴، دوره ۱۷، شماره ۲، صص. ۲۹۵-۳۲۷
ناشر: دانشکدهٔ مدیریت دانشگاه تهران
نوع مقاله: علمی پژوهشی

تاریخ دریافت: ۱۴۰۳/۰۹/۲۴
تاریخ ویرایش: ۱۴۰۳/۱۲/۱۲
تاریخ پذیرش: ۱۴۰۴/۰۱/۲۵
تاریخ انتشار: ۱۴۰۴/۰۳/۲۰
doi: https://doi.org/10.22059/JIPA.2025.390322.3649

یافته‌ها: نتایج این مطالعه نشان می‌دهد که استفاده از یادگیری ماشینی و مدل‌های یادگیری عمیق، در افزایش دقت پیش‌بینی تصمیم‌های استخدامی تأثیر بسزایی دارد. در بین مدل‌های ارزیابی شده، الگوریتم کتبوست (CatBoost) بهترین عملکرد را داشت و به صحت ۰/۹۵۳۳، دقت ۰/۹۵۴۰، بازیابی ۰/۸۹۲۵ و امتیاز F1 برابر با ۰/۹۲۲۲ دست یافت که به‌طور چشمگیری از سایر الگوریتم‌ها بهتر بود. به‌طور مشابه، مدل‌های جنگل تصادفی (Random Forest) و اکس‌جی‌بوست (XGBoost) نیز عملکرد خوبی داشتند و به‌ترتیب به دقت ۰/۹۲۱۳ و ۰/۹۵۰۰ دست یافتند. نتایج تحلیل ویژگی‌ها نشان داد که مهارت‌های فنی، استراتژی استخدام و امتیاز مصاحبه، مهم‌ترین عوامل مؤثر در تصمیم‌گیری استخدام بودند. علاوه‌براین، مدل‌های یادگیری جمعی، به‌ویژه کتبوست، قادر بودند اثرهای پیچیده‌تر ویژگی‌های شخصیتی متقاضیان را شناسایی کنند؛ در حالی‌که مدل‌های یادگیری ماشین سنتی قادر نبودند آن را درک کنند.

نتیجه‌گیری: در این مطالعه استفاده از یادگیری ماشین و یادگیری عمیق برای کمک به تصمیم‌گیری بهتر دربارهٔ انتخاب نیروی کار نشان داده شده است. کتبوست بهترین عملکرد را داشت؛ زیرا می‌توانست روابط پیچیده‌تر و غیرخطی‌تر بین ویژگی‌های متقاضی و نتایج استخدام را شبیه‌سازی کند. این الگوریتم‌ها باعث کاهش هزینه‌ها، بهبود کیفیت استخدام و بهره‌وری سازمان می‌شوند. با این حال، برای دستیابی به بهترین نتایج، سازمان‌ها باید با دقت و به‌طور مداوم، داده‌های متقاضی را جمع‌آوری و پردازش کنند و به‌طور منظم، مدل‌های یادگیری ماشین را برای حفظ و بهبود دقت پیش‌بینی‌ها با تغییر بازار کار به‌روز کنند.

کلیدواژه‌ها: یادگیری ماشین، یادگیری عمیق، پیش‌بینی استخدام، الگوریتم‌های پیش‌بینی، ویژگی‌های متقاضی، تحلیل داده‌های استخدامی.

مقدمه

تصمیم‌های استخدام یکی از فرایندهای بزرگ در سازمان‌های امروزی و بسیاری از مؤسسه‌هاست که بر عملکرد شرکت‌ها تأثیر می‌گذارد. این تصمیم‌ها می‌تواند از طریق بهبود رضایت و بهره‌وری کارکنان، به بهبود عملکرد سازمانی منجر شود. انتخاب نیروی کار مناسب می‌تواند به سازمان کمک کند تا با هزینه‌های ناشی از اشتباه در استخدام و احتمال اشتباه‌های مرتبط با استخدام افراد نامناسب مقابله کند (عباس‌پور، رحیمیان، غیاثی ندوشن و نرگسیان، ۱۳۹۷). با این حال، یکی از مشکلات بزرگ در استخدام، پیش‌بینی استخدام موفق است. عدم‌تطابق بین ویژگی‌های متقاضی و الزامات شغلی با انتخاب صحیح و ناکارآمدی که بسیاری از سازمان‌ها با آن مواجه می‌شوند، به انتخاب نادرست و کاهش کارایی منجر می‌شود (ایمانی، قلی‌پور، آذر و پورعزت، ۱۳۹۸؛ عارف نژاد و سپهوند، ۱۳۹۶)؛ بنابراین ضروری است که سازمان‌ها بتوانند به‌طور دقیق نتایج استخدام را پیش‌بینی کنند. به همین دلیل، روش‌های یادگیری ماشینی و یادگیری عمیق، بهتر می‌توانند نتیجهٔ تصمیم‌های استخدام را پیش‌بینی کنند (القحفه، مثنا، الجبری و سانگ[1]، ۲۰۲۴؛ العکاشه، فیصل مالک، هجران و زکی[1]، ۲۰۲۳). امروزه، در خصوص متقاضیان کار، داده‌های بسیاری در دسترس است؛ به‌طور مثال، سن، تحصیلات، تجربهٔ کاری، مهارت‌های فنی یا ویژگی‌های شخصیتی. این داده‌ها می‌توانند به‌عنوان ورودی برای مدل‌های یادگیری ماشینی و یادگیری عمیق، برای پیش‌بینی دقیق موفقیت در به‌دست‌آوردن شغل برای متقاضیان عمل کنند (گرونبرگ، پیترز، فرانسیس، بک و ماتز[1]، ۲۰۲۴). مزیت روش‌های یادگیری ماشین و الگوریتم‌های فراابتکاری، در توانایی آن‌ها در حل مسائل پیچیده و بهینه‌سازی فرایندها با دقت و کارایی بیشتر، به‌ویژه در شرایطی که داده‌ها غیرخطی و چندمتغیره هستند، نهفته است (قاسمیان صاحبی، توفیقی، عظاوی و زارع[1]، ۲۰۲۳؛ رضا، منیر، المطیری، یونس و فرید[1]، ۲۰۲۲). این مطالعه مدل‌های پیش‌بینی اشتغال را با یادگیری ماشین و الگوریتم‌های یادگیری عمیق بررسی و توسعه می‌دهد.

این پژوهش به تحلیل و پیش‌بینی تصمیمهای استخدامی با استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق می‌پردازد. مسئلهٔ اصلی پژوهش حاضر، پیش‌بینی نتیجهٔ استخدامی متقاضیان با استفاده از ویژگی‌های مختلف آن‌ها از قبیل سن، تحصیلات، تجربهٔ کاری و مهارت‌های فنی است. انتخاب درست نیروی انسانی، یکی از فرایندهای بسیار مهم در هر سازمان است که بر موفقیت یا شکست آن تأثیر زیادی دارد؛ بنابراین، درک اینکه کدام ویژگی‌ها در انتخاب موفق متقاضیان نقش دارند و چگونه می‌توان این ویژگی‌ها را با استفاده از داده‌های موجود پیش‌بینی کرد، امری ضروری است. پیش‌بینی نتیجهٔ استخدامی، به سازمان‌ها این امکان را می‌دهد که در فرایند انتخاب متقاضیان تصمیم‌های بهتری بگیرند و در نتیجه، میزان موفقیت و بهره‌وری خود را افزایش دهند. علاوه‌براین، این مقاله تلاش دارد تا به شکاف‌های موجود در تحقیقات قبلی پاسخ دهد. به‌طور خاص، برخی از پژوهش‌های پیشین، به استفاده از الگوریتم‌های

1. Al-Quhfa, Mothana, Aljbri & Song
1. Al Akasheh, Faisal Malik, Hujran & Zaki
1. Grunenberg, Peters, Francis, Back & Matz
1. Ghasemian Sahebi, Toufighi, Azzavi & Zare
1. Raza, Munir, Almutairi, Younas & Fareed

یادگیری ماشین پرداخته‌اند؛ اما هنوز بسیاری از روش‌های پیشرفته‌تر، مانند یادگیری عمیق و یادگیری جمعی در این زمینه به‌طورکامل مورد استفاده قرار نگرفته‌اند. این پژوهش قصد دارد به بهبود مدل‌های پیش‌بینی استخدامی از طریق استفاده از الگوریتم‌های یادگیری عمیق و یادگیری جمعی بپردازد.

کارهای فزاینده‌ای در زمینهٔ پیش‌بینی تصمیم‌های استخدام، براساس الگوریتم‌های یادگیری ماشین و یادگیری عمیق انجام شده است. به‌طور مثال، الگوریتم‌هایی مانند درخت‌های تصمیم، جنگل‌های تصادفی و ماشین‌های بردار پشتیبان (SVM) برای پیش‌بینی نتیجهٔ استخدام استفاده شده‌اند (ایمیانوان و همکاران[1]، ۲۰۲۴). برای نشان دادن اثربخشی الگوریتم‌های مختلف مورداستفاده برای پیش‌بینی استخدام، رالاپالی و کومار[2] (۲۰۲۴) مطالعه‌ای انجام دادند و دریافتند که استفاده از الگوریتم‌های پیچیده‌تر، مانند اکس‌جی‌بوست[3] و کتبوست[4]، به دقت بیشتر پیش‌بینی‌ها منجر می‌شود. این نتایج نشان می‌دهد که فرایند استخدام را می‌توان با استفاده از الگوریتم‌های یادگیری ماشین شبیه‌سازی کرد. پژوهش‌های دیگر، استفاده از مدل‌های یادگیری عمیق، به‌ویژه شبکه‌های عصبی عمیق را بررسی می‌کنند؛ پژوهش‌هایی که مدل‌ها را با استفاده از روش‌هایی مانند آپتونا[5] بهینه می‌کنند تا تصمیم‌های استخدام را پیش‌بینی کنند و دقت آن‌ها را افزایش دهند (طاهردوست[6]، ۲۰۲۳). علاوه‌براین، برخی از مطالعات، مانند ما، ژانگ و ایلر[7] (۲۰۲۰)، از مدل‌های یادگیری عمیق، برای شبیه‌سازی رفتار کارفرما در بازارهای کار آنلاین استفاده می‌کنند و بررسی می‌کنند که مدل‌های یادگیری عمیق می‌توانند دقت پیش‌بینی را به‌طور چشمگیری افزایش دهند. در این مقاله سعی شده است که برای پیش‌بینی تصمیم استخدام با استفادهٔ هم‌زمان از روش‌های یادگیری ماشینی و یادگیری عمیق، مدل‌های دقیق‌تری ارائه شود.

این مطالعه در حال توسعه و بهبود مدل‌های پیش‌بینی اشتغال براساس الگوریتم‌های مختلف یادگیری ماشین و یادگیری عمیق است. این مدل‌ها می‌توانند پیش‌بینی‌های دقیق‌تری با نتیجهٔ استخدام متقاضیان ارائه دهند و در نتیجه به تصمیم‌گیری بهتر در فرایند استخدام کمک کنند. در این مقاله، الگوریتم‌ها برای افزایش دقت مدل‌ها و اعمال تکنیک‌های پیشرفته‌تر مانند ADASYN و RFECV و عملکرد بهتر نسبت به پژوهش‌های قبلی تنظیم شده‌اند. این پژوهش، به‌دلیل هزینه‌های بالقوهٔ سازمان‌ها و از دست‌دادن استعدادهای مهم، در صورت تصمیم‌گیری اشتباه در فرایند استخدام حائز اهمیت است؛ به این معنا که کاهش خطا در انتخاب کارکنان می‌تواند به پیش‌بینی دقیق‌تر استخدام منجر شود که برای سازمان‌ها از نظر عملکرد کلی سودمند است. علاوه‌براین، پژوهش حاضر می‌تواند وضعیت پیشرفت این حوزه را در چندین زمینهٔ یادگیری ماشینی و یادگیری عمیق ارتقا دهد و همچنین، راهی برای حل مشکلات پیچیده در زمینهٔ پیش‌بینی استخدام ارائه دهد. در ابتدا، روش‌های یادگیری ماشین و یادگیری عمیق برای پیش‌بینی تصمیم‌های

1. Imianvan et al.
2. Rallapalli & Kumar
3.XGBoost
4. CatBoost
5. Optuna
6. Taherdoost
7. Ma, Zhang & Ihler

استخدام در این مقاله ارائه شده است. در ادامه، به بهینه‌سازی مدل‌ها و انتخاب ویژگی‌ها با روش‌های مختلف پرداخته می‌شود؛ سپس نتایج آزمون‌ها ارائه و عملکرد مدل‌های مختلف با هم مقایسه می‌شود و در نهایت، پیشنهادهایی در رابطه با بهبود عملکرد مدل‌ها ارائه خواهد شد. این مقاله سعی دارد مدلی از انتخاب استخدام ایجاد کند که با استفاده از الگوریتم‌های جدید و بهبود فرایند پیش‌بینی، پیش‌بینی‌های دقیق‌تری را با خطای کمتر انجام دهد.

مبانی نظری

پیش‌بینی تصمیم‌های استخدامی

پیش‌بینی تصمیم‌های استخدام بسیار مهم است و این به تجزیه‌وتحلیل پیش‌بینی‌کننده مربوط می‌شود؛ به‌ویژه زمانی که از الگوریتم‌های یادگیری ماشینی، مانند رگرسیون و درخت‌های تصمیم استفاده می‌کنیم. به‌طور مثال، جایانتی و واسا[1] (۲۰۲۲) اظهار می‌کنند که این روش‌ها سازمان را در تجزیه‌وتحلیل داده‌های تاریخی برای اتخاذ استراتژی‌هایی برای استخدام در آینده تسهیل می‌کنند. کاربرد آن‌ها در انتخاب کارکنان، می‌تواند دقت طبقه بندی کارکنان را در موقعیت‌های شغلی خاص افزایش دهد (پامپوکتسی، آودیمیوتیس، ماراگوداکیس و آولونیتیس[2]، ۲۰۲۱).

پساچ و همکاران[3] (۲۰۲۰) استفاده از یک چارچوب تحلیلی را که شامل مدل‌های یادگیری ماشینی می‌شود، برای کمک به استخدام‌کنندگان منابع انسانی در تصمیم‌گیری برای موفقیت در کاریابی و بهینه‌سازی شیوه‌های استخدام، ترویج داده‌اند. ادغام هوش مصنوعی در فرایندهای استخدام نیز، به‌عنوان راهی برای خودکارسازی و تسریع مراحل مختلف استخدام با استقبال مواجه شده است. هانت و اوریلی[4] (۲۰۲۲) خاطرنشان می‌کنند که پیشرفت‌های هوش مصنوعی، تصمیم‌گیری سریع‌تر را تسهیل می‌کند و در عین حال، هرگونه تعصب انسانی را در طول جدول زمانی استخدام نامزدها کاهش می‌دهد. نتایج بیشتر از براون، جورج و مهافی کولتگن[5] (۲۰۱۸) مدل‌سازی شایستگی را نشان می‌دهد که با تجزیه‌وتحلیل پیش‌بینی‌کننده غنی شده است و به معیارهای انتخاب دقیق‌تر کمک می‌کند تا نقش‌های شغلی را با ویژگی‌های لازم کارمند هم‌سو کند. یافته‌ها نشان می‌دهد که چگونه کسب‌وکارها از هوش مصنوعی در طول کووید استفاده کردند که گویای گرایش به‌سمت شیوه‌های استخدام مبتنی بر داده است.

مطالعهٔ الشرف، فتوح و عید[6] (۲۰۲۲) مدلی خودکار را ارائه می‌کند که قادر است فرسایش کارکنان را پیش‌بینی کند و بر ارزش افزوده ترکیب تکنیک‌های یادگیری ماشینی کامل در تصمیم‌گیری منابع انسانی تأکید می‌کند. علاوه‌براین، اُزدمیر و نالبانت[7] (۲۰۲۰) روش‌شناسی‌هایی را پیشنهاد می‌کنند که به انتخاب کارکنان بر اساس تحلیل‌های سیستماتیک معیارهای انتخاب اولویت می‌دهد و بر ارزش رویکردهای داده‌محور تأکید می‌کند. استفاده از یادگیری ماشین برای

1. Jayanti and Wasesa
2. Pampouktsi, Avdimiotis, Maragoudakis & Avlonitis
3. Pessach et al.
4. Hunt & O'Reilly
5. Brown, George & Mehaffey-Kultgen
6. Alsheref, Fattoh & Ead
7. Özdemir & Nalbant

پیش‌بینی فرسایش کارکنان، همانند یک حلقهٔ بازخورد حیاتی عمل می‌کند که با یادگیری از خروج‌های گذشته و تنظیم استراتژی‌های آینده، شیوه‌های استخدام را افزایش می‌دهد. یاداو[1] (۲۰۲۱) تأکید می‌کند که تمرکز روی یک رویکرد داده‌محور، می‌تواند به سازمان‌ها کمک کند تا مدل‌های مربوطه را برای پیش‌بینی جابه‌جایی کارکنان بسازند که برای اصلاح استراتژی‌های استخدام بسیار مهم است؛ زیرا کسب‌وکارها برای حفظ استعدادها تلاش می‌کنند. در نتیجه، زیربنای نظری مدل‌های پیش‌بینی در استخدام و انتخاب کارکنان، به‌طور عمیقی در یادگیری ماشین و تجزیه‌وتحلیل پیش‌بینی‌کننده ریشه دارد. این فناوری‌ها، نه‌تنها دقت و کارایی فرایندهای استخدام را افزایش می‌دهند؛ بلکه سازوکارهایی را برای سازمان‌ها فراهم می‌کنند تا با نیازهای نیروی کار پویا سازگار شوند. همان طور که این روش‌ها به تکامل خود ادامه می‌دهند، ادغام آن‌ها در شیوه‌های استاندارد منابع انسانی، احتمالاً به‌طور فزاینده‌ای رایج می‌شود.

یادگیری ماشین و یادگیری عمیق

یادگیری ماشین، عمدتاً به‌دلیل تجزیه‌وتحلیل پیش‌بینی‌کنندهٔ آن تا حد زیادی با فرایندهای استخدام بهینه هم‌سو شده است. جایانتی و واسا (۲۰۲۲) خاطرنشان می‌کنند که روش‌های یادگیری ماشین برای پیش‌بینی فرایند استخدام، بر داده‌های گذشته تکیه می‌کنند تا شرکت‌ها بتوانند از اطلاعات برای تصمیم‌گیری در مورد استخدام بر اساس حقایق و نه بر اساس گمان استفاده کنند. ابزارهای اصلی که به‌طور الگوریتمی بینشی را در مورد اینکه نامزدهای استخدام ممکن است در نقش‌های خاص بهتر عمل کنند، عبارت‌اند از: رگرسیون، درختان تصمیم‌گیری و جنگل‌های تصادفی (جایانتی و واسا، ۲۰۲۲؛ کومار و گرگ[2]، ۲۰۱۸). توانایی ارزیابی سیستماتیک جنبه‌های مختلف یک متقاضی، شرکت‌ها را قادر می‌سازد تا نه‌تنها فرایند استخدام را سرعت دهند، بلکه امکان تصمیم‌گیری مغرضانه که یکی از زمینه‌های مهم نگرانی در فرایندهای استخدام امروزی است، به حداقل رسانند (هانت و اوریلی، ۲۰۲۰).

یادگیری عمیق یکی از زیرمجموعه‌های یادگیری ماشین است و با کمک به مدل در یادگیری از داده‌ها به‌روش‌های پیچیده‌تر و ظریف‌تر، به قدرت پیش‌بینی سیستم می‌افزاید (سارکر[3]، ۲۰۲۱). یادگیری عمیق می‌تواند از تکنیک‌هایی مانند شبکه‌های عصبی، برای پردازش داده‌های بدون ساختار، مانند رزومه‌ها یا مصاحبه‌های ویدئویی، برای استخراج بینش‌هایی در خصوص مناسب بودن نامزدها استفاده کند؛ در حالی‌که تحلیل سنتی نمی‌تواند به‌راحتی آن را آشکار کند (البارودی، منصوری و الامیر[4]، ۲۰۲۴). این سیستم‌ها قادر به تجزیه‌وتحلیل الگوها و هم‌بستگی‌ها هستند و به تصمیم‌گیری‌های آگاهانه‌تر در زمینهٔ انتخاب و حفظ نامزدها منجر می‌شوند و در نتیجه، کارایی کلی استخدام را به‌طور چشمگیری افزایش می‌دهند. پساچ و همکاران (۲۰۲۰) یک چارچوب کلی را برجسته کردند که ML و برنامه‌ریزی ریاضی را برای تقویت ابزارهای پشتیبانی تصمیم در دسترس استخدام‌کنندگان در منابع انسانی ترکیب می‌کند. این تکنیک، هم بر پیش‌بینی و هم بر عامل تفسیرپذیری در پس نحوهٔ عملکرد الگوریتم‌ها تأکید می‌کند که برای متخصصان

1. Yadav
2. Kumar & Garg
3. Sarker
4. Albaroudi, Mansouri & Alameer

منابع انسانی لازم است تا به مدل‌های مورد استفاده ایمان داشته باشند. این جنبه‌های توضیح‌پذیری، برای خرید و اطمینان از استفادهٔ اخلاقی از هوش مصنوعی در فرایندهای استخدام بسیار مهم هستند. در نتیجه، مبانی نظری زیربنای یادگیری ماشین و یادگیری عمیق در پیش‌بینی استخدام، بر یک تغییر پارادایم در شیوه‌های منابع انسانی تأکید می‌کند. این فناوری‌ها بینش‌هایی را دربارهٔ چگونگی اصلاح و انطباق قضاوت انسانی با تجزیه‌وتحلیل دقیق مبتنی بر داده‌ها اختیار سازمان‌ها قرار می‌دهند. با سرمایه‌گذاری مستمر در هوش مصنوعی و فرایندهای تحلیل پیش‌بینی‌کننده در فرایندهای استخدام، بهره‌وری و انصاف و هم‌سویی استراتژیک افزایش یافته است.

پیشینهٔ تجربی پژوهش

در جدول ۱، خلاصه‌ای از برخی مقاله‌های مرتبط با پیش‌بینی تصمیم‌های استخدامی و پیش‌بینی‌های مربوط به اشتغال با استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق آورده شده است. این جدول شامل اطلاعاتی از جمله نویسندگان، عنوان مقاله، هدف مقاله، نتایج و الگوریتم‌های مورد استفاده در هر پژوهش است.

جدول ۱. پیشینهٔ تجربی پژوهش

نویسندگان	عنوان مقاله	هدف	نتایج	الگوریتم مورد استفاده
علی شاه و همکاران[1] (۲۰۲۰)	یک شبکهٔ عصبی عمیق پیشرفته برای پیش‌بینی غیبت در محل کار	تجزیه‌وتحلیل الگوهای غیبت با استفاده از یادگیری عمیق برای اطلاع از تصمیم‌های استخدام	این مدل به صحت ۹۰/۶درصد دست یافت که نشان می‌دهد یادگیری عمیق می‌تواند به‌طور مؤثر رفتار کارکنان مرتبط با استخدام را پیش‌بینی کند	شبکه‌های عصبی عمیق
وینوتا و یوگیش[2] (۲۰۲۰)	پیش‌بینی اشتغال‌پذیری فارغ التحصیلان مهندسی با استفاده از الگوریتم‌های یادگیری ماشین	پیش‌بینی کسب مهارت و قابلیت استخدام فارغ‌التحصیلان مهندسی	این مطالعه نشان می‌دهد که طبقه‌بندی‌کننده‌های ANN در پیش‌بینی قابلیت استخدام بیشترین دقت را دارند که بر اهمیت هم‌سویی نتایج آموزشی با نیازهای صنعت تأکید می‌کند.	رگرسیون لجستیک، درخت تصمیم، K نزدیک‌ترین همسایه، ماشین بردار پشتیبان، بیز ساده، ANN
ما و همکاران (۲۰۲۰)	یک مدل انتخاب عمیق برای پیش‌بینی نتیجهٔ استخدام در بازارهای کار آنلاین	توسعهٔ مدلی برای پیش‌بینی نتایج استخدام در بازارهای کار آنلاین با استفاده از یادگیری ماشین	این مطالعه نتیجه می‌گیرد که مدل انتخاب عمیق پیشنهادی به‌طور چشمگیری از مدل لاجیت مشروط سنتی در پیش‌بینی اولویت‌های استخدام کارفرما در بازارهای کار آنلاین بهتر عمل می‌کند.	مدل لاجیت شرطی (CLM)، مدل انتخاب عمیق، نزول گرادیان تصادفی (SGD)

1. Shah
2. Vinutha & Yogish

نویسندگان	عنوان مقاله	هدف	نتایج	الگوریتم مورد استفاده
رضا و همکاران (۲۰۲۲)	پیش‌بینی ساییدگی کارکنان با استفاده از رویکردهای یادگیری ماشین	هدف این مطالعه پیش‌بینی فرسایش کارکنان با استفاده از تکنیک‌های یادگیری ماشین است که برای مدیریت منابع انسانی بسیار اهمیت دارد.	طبقه‌بندی‌کننده درخت اضافی بهینه‌سازی‌شده امتیاز صحت ۹۳درصد را برای پیش‌بینی ساییدگی کارکنان به‌دست آورد و از سایر مدل‌ها بهتر عمل کرد. عوامل کلیدی برای فرسایش عبارت بودند از: درآمد ماهانه، نرخ ساعتی، سطح شغل و سن.	طبقه‌بندی‌کنندهٔ درختان اضافی (ETC)، جنگل تصادفی، درختان تصمیم، ماشین‌های بردار پشتیبان (SVM)
الشرکاوی، حلمی و یحیی[1] (۲۰۲۲)	پیش‌بینی اشتغال‌پذیری دانش‌آموختگان فناوری اطلاعات با استفاده از الگوریتم‌های یادگیری ماشین	توسعهٔ یک مدل پیش‌بینی برای قابلیت استخدام فارغ‌التحصیلان فناوری اطلاعات	این تحقیق نشان می‌دهد که یادگیری ماشینی می‌تواند به‌طور مؤثری قابلیت استخدام را پیش‌بینی کند و مهارت‌های فارغ‌التحصیلان را با تقاضاهای بازار کار هم‌سو کند.	درخت تصمیم، بیز ساده گوسی، رگرسیون لجستیک، جنگل تصادفی، ماشین بردار پشتیبان
گائو، لیانگ و چن[2] (۲۰۲۲)	الگوریتم بهبودیافته لجن‌کپکی چندجمعیتی و کاربرد آن در پیش‌بینی پایداری اشتغال فارغ‌التحصیلان	ارائهٔ روشی برای پیش‌بینی ثبات اشتغال در بین فارغ‌التحصیلان	این مطالعه یک الگوریتم جدید را معرفی می‌کند که دقت پیش‌بینی را برای ثبات اشتغال بهبود می‌بخشد و بینش‌هایی را برای تدوین سیاست در بخش‌های اشتغال ارائه می‌دهد.	الگوریتم بهبودیافته لجن‌کپکی، ماشین بردار پشتیبان.
هانت و اوریلی (۲۰۲۲)	استخدام سریع در خرده‌فروشی: استفاده از هوش مصنوعی در استخدام همکاران خط مقدم با حقوق ساعتی در طول همه‌گیری کووید ۱۹	کشف اینکه چگونه هوش مصنوعی می‌تواند فرایندهای استخدام را در طول همه‌گیری بهبود بخشد.	هوش مصنوعی می‌تواند تصمیم‌های استخدام را تسریع کند و تعصبات انسانی را کاهش دهد و کارایی در فرایندهای استخدام را بهبود بخشد.	روش‌های مبتنی بر هوش مصنوعی
الباسام[3] (۲۰۲۳)	قدرت هوش مصنوعی در استخدام: بررسی تحلیلی استراتژی‌های استخدام مبتنی بر هوش مصنوعی	بررسی استراتژی‌های استخدام هوش مصنوعی و پیامدهای آن‌ها برای کیفیت و کارایی استخدام.	استراتژی‌های هوش مصنوعی می‌توانند کارایی استخدام را افزایش دهند؛ اما نگرانی‌های اخلاقی‌ای را در مورد تعصب و تبعیض ایجاد می‌کنند.	تکنیک‌های مختلف هوش مصنوعی
تیلاک، لالیتا دوی،	ایجاد انقلاب در قابلیت استخدام فارغ‌التحصیلان	استفاده از یادگیری ماشین برای پیش‌بینی استخدام	مدل‌های یادگیری ماشین برای پیش‌بینی استخدام در پردیس‌های	درخت تصمیم‌گیری، جنگل تصادفی،

1. ElSharkawy, Helmy & Yehia
2. Gao, Liang & Chen
3. Albassam

نویسندگان	عنوان مقاله	هدف	نتایج	الگوریتم مورد استفاده
کالایسلوی و تجا[1] (۲۰۲۳)	دانشگاه: استفاده از مدل‌های پیشرفته یادگیری ماشین برای بهینه‌سازی استراتژی‌های استخدام و استخدام در دانشگاه	در پردیس‌های دانشگاهی	دانشگاهی و تعیین شایستگی نامزدها بسیار مؤثرند.	رگرسیون لجستیک
منگ[2] (۲۰۲۳)	تحلیل و پیش‌بینی اشتغال دانشجویان براساس الگوریتم طبقه‌بندی درخت تصمیم	نتایج استخدام دانشجویان را با استفاده از طبقه‌بندی درخت تصمیم براساس سابقهٔ تحصیلی و تجربهٔ کاری آن‌ها پیش‌بینی می‌کند.	زمینهٔ علمی را برای هدایت اشتغال و برنامه‌های استعدادیابی در مؤسسات آموزش عالی فراهم می‌کند.	درخت تصمیم
ناگوویتسین[3] (۲۰۲۳)	پیش‌بینی اشتغال دانش‌آموزان در آموزش معلمان با استفاده از الگوریتم‌های یادگیری ماشین	یک برنامهٔ مبتنی بر یادگیری ماشینی برای پیش‌بینی اشتغال دانش‌آموزان در سیستم آموزشی ایجاد می‌کند.	الگوریتم درخت تصمیم بالاترین صحت (۸۹ درصد) را در شناسایی دانش‌آموزانی ارائه کرد که ممکن است در رشتهٔ خود به کار گرفته نشده باشند.	درخت تصمیم، رگرسیون لجستیک، کتبوست
حنیف، معروپ، کمارالدین و سامی[4] (۲۰۲۴)	رویکرد یادگیری ماشین در پیش‌بینی آگهی‌های شغلی جعلی	توسعهٔ یک مدل برای شناسایی آگهی‌های شغلی جعلی	این پژوهش با موفقیت آگهی‌های شغلی جعلی را شناسایی کرد و کاربرد یادگیری ماشین را در بهبود یکپارچگی بازار کار نشان داد.	رگرسیون لجستیک، ماشین بردار پشتیبان، درخت تصمیم، نایو بیز
معدنچیان[5] (۲۰۲۴)	از استخدام تا حفظ: ابزارهای هوش مصنوعی برای تصمیم‌گیری منابع انسانی	بررسی تأثیر ابزارهای هوش مصنوعی بر تصمیم‌گیری‌های منابع انسانی از استخدام تا حفظ کارکنان	هوش مصنوعی و تحلیل‌های پیش‌بینی می‌توانند فرایند استخدام و حفظ کارکنان را کارآمدتر کنند. چالش‌هایی چون سوگیری و حریم خصوصی مطرح است.	سیستم‌های ردیابی متقاضی مبتنی بر هوش مصنوعی، تجزیه‌وتحلیل پیشگویانه
جها، لیکیتا، چاندانا، ردی و بهارگاوی[6]	پیش‌بینی شغلی با استفاده از یادگیری ماشینی	از ML برای پیش‌بینی مسیر شغلی دانش‌آموزان براساس عملکرد تحصیلی،	مشاوره شغلی شخصی می‌تواند به دانش‌آموزان کمک کند تا تصمیم‌های شغلی آگاهانه‌ای را در	تکنیک های یادگیری ماشین

1. Thilak, Lalitha Devi, Kalaiselvi & Teja
2. Meng
3. Nagovitsyn
4. Hanif, Maarop, Kamaruddin & Samy
5. Madanchian
6. Jha, Likhitha, Chandana, Reddy & Bhargavi

نویسندگان	عنوان مقاله	هدف	نتایج	الگوریتم مورد استفاده
(۲۰۲۴)		برنامه‌های فوق برنامه و ویژگی‌های شخصیتی استفاده می‌کند.	یک نیروی کار در حال تکامل بگیرند.	
سومان، کائور، شارما و کومار[1] (۲۰۲۴)	سیستم مبتنی بر یادگیری ماشین برای پیش‌بینی پذیرش و اشتغال در بخش مهندسی و فناوری	تصمیم‌گیری مبتنی بر داده را با پیش‌بینی مسیر شغلی برای دانشجویان مهندسی پل می‌کند.	LSTM الگوهای پیچیده را ثبت می‌کند. GRU دقت پیش‌بینی شغلی را افزایش می‌دهد. به دانش‌آموزان کمک می‌کند تا تصمیم‌های شغلی مبتنی بر داده را اتخاذ کنند.	حافظهٔ بلندمدت ـ کوتاه‌مدت (LSTM)، واحدهای بازگشتی دروازه‌دار (GRU)

درحالی‌که تحقیقات زیادی در خصوص الگوریتم‌های یادگیری ماشین برای پیش‌بینی تصمیم‌های استخدامی، مانند درخت‌های تصمیم‌گیری و SVM انجام شده است، این مدل‌ها هنوز، هنگام برخورد با مجموعه داده‌های بزرگ و بسیار پیچیده آسیب می‌بینند. کاهش دقت نیز در شرایطی به‌وجود می‌آید که ویژگی‌های غیرخطی و تعاملات پیچیده‌تر وجود دارد. علاوه‌براین، بیشتر این مطالعات با هدف تولید مدل‌هایی برای برخی ویژگی‌های خاص در داده‌ها انجام شده‌اند که ممکن است در دنیای واقعی قابلیت تعمیم نداشته باشند. در مقابل، به نظر می‌رسد که توانایی یادگیری عمیق برای مدل‌سازی روابط غیرخطی پیچیده، پیش‌بینی‌ها را به‌طور چشمگیری بهبود می‌بخشد. به‌طور هم‌زمان، این روش‌ها پتانسیل ارائهٔ نتایج شبیه‌سازی بهتر تصمیم‌های استخدام را دارند (ما و همکاران، ۲۰۲۰). این دست تکنیک‌ها، چنین پژوهش‌هایی را پیش می‌برند و به توسعهٔ مدل‌هایی با دقت بیشتر و قابلیت تعمیم بهتر کمک می‌کنند. یکی دیگر از مواردی که در بیشتر این پژوهش‌ها استفاده نشده و به‌طور خاص نشان داده شده، این است که چگونه این کار باید انجام شود که پاسخ آن، افزودن داده‌های شخصیتی و رفتاری با ویژگی‌های معمولی مانند تحصیلات و تجربه است. استنتاج از این داده‌ها، می‌تواند به شبیه‌سازی دقیق‌تر فرایند انتخاب متقاضی کمک کند و در ویژگی‌های خاص روانی ـ اجتماعی می‌تواند نوآوری جدیدی را پیشنهاد دهد.

براساس بررسی مطالعات قبلی در خصوص پیش‌بینی تصمیم‌گیری استخدام، پژوهش‌های فعلی در این زمینه، به‌طور عمده بر الگوریتم‌های یادگیری ماشین سنتی (به‌طور مثال، درخت تصمیم، جنگل تصادفی و ماشین بردار پشتیبان) تکیه کرده‌اند که دقت چشمگیری را به‌دست آوردند؛ اما کمابیش کمبودهای ذاتی دارند. به‌طور مثال، بیشتر این مدل‌ها برای مدیریت داده‌های پیچیده‌تر یا ویژگی‌های غیرخطی کافی نیستند. علاوه‌براین، بیشتر این مطالعات از داده‌های محدود و ویژگی‌های خاصی استفاده می‌کنند که به‌طور گسترده در دنیای واقعی با داده‌های پیچیده‌تر و گسترده‌تر قابل استفاده نیستند. این تحقیق همچنین به دو دلیل برجسته است: گسترش آن در عملکرد این نوع مدل‌های پیش‌بینی استخدام از طریق یادگیری ماشین و الگوریتم‌های یادگیری عمیق. ما می‌توانیم دقت پیش‌بینی‌ها را با استفاده از روش‌های پیشرفته، از جمله یادگیری عمیق و یادگیری جمعی و بهینه‌سازی جدید مانند ADASYN و RFECV به‌طور

1. Suman, Kaur, Sharma & Kumar

چشمگیری بهبود دهیم. این امر پیش‌بینی دقیق‌تر تصمیم‌های استخدام را تسهیل می‌کند؛ زیرا این روش‌ها در شبیه‌سازی ویژگی‌های پیچیدهٔ غیرخطی بهتر عمل می‌کنند. این پژوهش همچنین مدل‌های قدیمی‌تر را با چندین مدل یادگیری جمعی و یادگیری عمیق مقایسه می‌کند و می‌تواند شکاف‌های تحقیقات قبلی را پُر کند.

روش‌شناسی پژوهش

هدف اصلی این پژوهش استفاده از مدل‌های یادگیری ماشین برای پیش‌بینی تصمیم نهایی در فرایند استخدام است. برای انجام این کار، از مدل‌های یادگیری ماشین با ویژگی‌های متقاضیان و پیش‌بینی نتیجه استخدام استفاده شد. مجموعه دادهٔ این پژوهش از وبسایت Kaggle استخراج شده است و شامل متغیرهای مربوط به عواملی است که بر فرایند استخدام تأثیر می‌گذارد؛ به‌طور مثال، سن، سطح تحصیلات، تجربه در سال، فاصله از شرکت، نمرهٔ مصاحبه، نمرهٔ مهارت‌های فنی، امتیاز ویژگی شخصیتی و استراتژی استخدام (ربیع الخاروآ[1]، ۲۰۲۴). ابتدا، داده‌ها برای وجود مقادیر از دست رفته، نقاط پُرت و عدم‌تعادل کلاس بررسی و توسط ADASYN برای اصلاح عدم‌تعادل داده‌ها استفاده شدند. برای کاهش پیچیدگی مدل و حذف ویژگی‌های بی‌اهمیت، از روش RFECV استفاده شد و در نهایت ۸ ویژگی مهم برای تحلیل‌های بیشتر انتخاب شد. این انتخاب، ضمن کمک به کاهش ابعاد داده‌ها، مدل‌ها را کارآمدتر کرد. برای حل این مشکل، از ۱۲ مدل یادگیری ماشین برای پیش‌بینی نتیجهٔ استخدام، پس از پیش‌پردازش داده‌ها استفاده شد. این مدل‌ها شامل جنگل تصادفی[2]، اکس‌جی‌بوست[3]، لایت‌جی‌بی‌ام[4]، کتبوست[5]، تقویت گرادیان[6]، درخت‌های فوق تصادفی[7]، ماشین بردار پشتیبان (SVM)، رگرسیون لجستیک[8]، بیز ساده[9]، نزدیک‌ترین همسایگان (KNN) [10]، درخت تصمیم[11] و آدابوست[12] بودند. بنابراین با استفاده از آپتونا، فراپارامترها برای دستیابی به دقت بیشتر با مدل‌ها بهینه شد. علاوه‌براین، یک مدل DNN توسعه داده شد و ساختار آن با استفاده از آپتونا برای بررسی قابلیت یادگیری عمیق تعیین شد. معیارهای صحت، دقت، بازیابی و امتیاز F1 در نهایت برای محاسبهٔ عملکرد مدل‌ها برای هر مدل استفاده شد. در این مطالعه از زبان برنامه‌نویسی پایتون استفاده شد و تمامی مدل‌ها روی رایانه‌ای مجهز به پردازندهٔ Intel Core i7-13700H، ۱۶ گیگابایت رم و تحت پایتون ۱۲.۳ اجرا شدند.

1. Rabie El Kharoua
2. Random Forest
3. XGBoost
4. LightGBM
5. CatBoost
6. Gradient Boosting
7. Extra Trees
8. Logistic Regression
9. Naïve Bayes
10. K-Nearest Neighbors - KNN
11. Decision Tree
12. AdaBoost

ویژگی‌های داده‌ها و پیش‌پردازش

مجموعه داده استفاده‌شدهٔ در این پژوهش ۱۵۰۰ نمونه است که ۱۰ متغیر مستقل و یک متغیر هدف را شامل می‌شود. متغیر هدف[1]، نشان‌دهندهٔ تصمیم نهایی در فرایند استخدام است. این متغیر دو کلاس دارد: متقاضیانی که استخدام نشده‌اند (۰) و متقاضیانی که استخدام شده‌اند (۱).

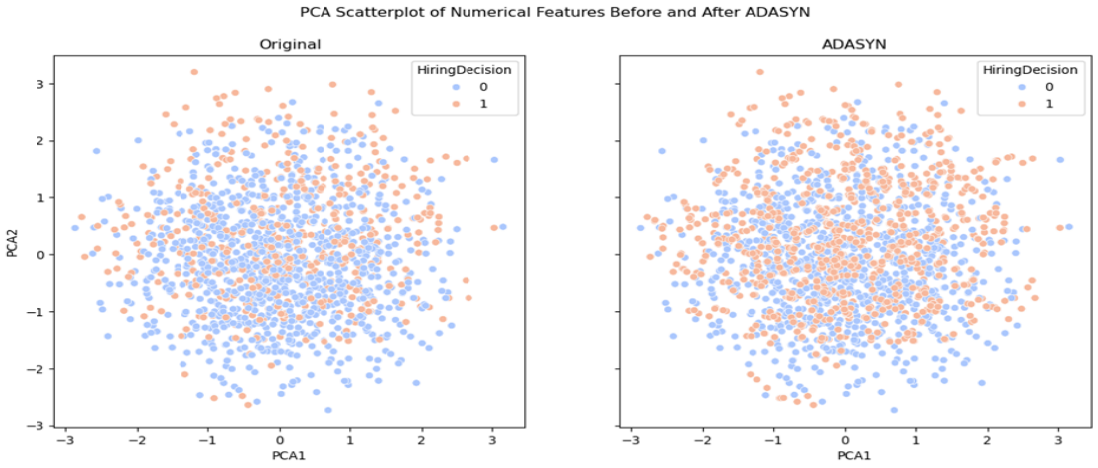
جدول ۲. توصیف ویژگی‌ها

ویژگی	توضیحات	نوع داده	محدوده / دسته‌بندی‌ها
سن	سن متقاضی	عدد صحیح	۲۰ تا ۵۰ سال
جنسیت	جنسیت متقاضی	دودویی	۰: مرد، ۱: زن
سطح تحصیلات	بالاترین مدرک تحصیلی متقاضی	دسته‌ای	۱: کارشناسی (نوع ۱)، ۲: کارشناسی (نوع ۲)، ۳: کارشناسی ارشد، ۴: دکتری
تجربهٔ کاری	تعداد سال‌های تجربهٔ کاری متقاضی	عدد صحیح	۰ تا ۱۵ سال
تعداد شرکت‌های قبلی	تعداد شرکت‌هایی که متقاضی در آن‌ها کار کرده است	عدد صحیح	۱ تا ۵ شرکت
فاصله از شرکت	فاصلهٔ محل سکونت متقاضی تا شرکت (کیلومتر)	اعشاری	۱ تا ۵۰ کیلومتر
امتیاز مصاحبه	امتیاز متقاضی در مصاحبه	عدد صحیح	۰ تا ۱۰۰
امتیاز مهارت	امتیاز مهارت‌های فنی متقاضی	عدد صحیح	۰ تا ۱۰۰
امتیاز شخصیتی	امتیاز ویژگی‌های شخصیتی متقاضی	عدد صحیح	۰ تا ۱۰۰
استراتژی جذب نیرو	استراتژی مورد استفاده در جذب نیرو	دسته‌ای	۱: تهاجمی، ۲: متعادل، ۳: محافظه‌کار
تصمیم استخدام	نتیجه فرایند استخدام (هدف مدل)	دودویی	۰: عدم‌استخدام، ۱: استخدام

این مطالعه از مجموعه داده‌ای استفاده کرده است که شامل اطلاعاتی در خصوص فرایند استخدام و نمونه‌هایی است که طی مراحل جذب، استخدام شده‌اند و استخدام نشده‌اند. از آنجایی که یافتن متقاضیان واجد شرایط اهمیت دارد، با بررسی مجموعه داده مشخص می‌شود که دارای یک کلاس نامتعادل است؛ به‌طوری که تعداد نمونه‌های مربوط به عدم‌استخدام (کلاس ۰) به‌طور چشمگیری بیشتر از نمونه‌های استخدام‌شده (کلاس ۱) است. مدل‌های یادگیری ماشینی به عدم‌تعادل در داده‌ها حساس هستند و کلاس غالب را پیش‌بینی می‌کنند (یعنی در مورد ما غیراستخدام شده). بنابراین لازم است از روشی برای متعادل کردن داده‌ها استفاده شود تا اطمینان حاصل شود که مدل می‌تواند الگوهای استخدام را به‌درستی بیابد. پس از تجزیه‌وتحلیل اولیه، ۱۰۳۵ نمونه در کلاس ۰ (غیر استخدامی) وجود داشت؛ اما تنها ۴۶۵ نمونه در کلاس ۱ (استخدام شده) بود. این نشان می‌دهد که مجموعه داده نامتعادل است و ممکن است نامزدهای مناسب کافی برای استخدام وجود نداشته باشد تا مدل‌های یادگیری ماشین بدون استفاده از روش‌های متعادل‌کننده به‌درستی شناسایی

1. Hiring Decision

کنند. برای انجام این کار از روش نمونه‌گیری مصنوعی تطبیقی[1] استفاده شد. روش نمونه‌گیری مصنوعی تطبیقی، نوعی تکنیک نمونه‌گیری افزایشی با هدف تولید نمونه‌های جدید برای طبقهٔ اقلیت است تا عدم‌تعادل در توزیع داده‌ها را اصلاح کند (یاداو[2]، ۲۰۲۱). برخلاف روش‌های کلاسیک مانند SMOTE، این روش به‌طور تطبیقی نمونه‌های بیشتری را در مناطقی تولید می‌کند که تراکم کلاس اقلیت کمتر است. این کار باعث بهبود عملکرد مدل در شناسایی متقاضیان مناسب برای استخدام می‌شود (دی و پراتاپ[3]، ۲۰۲۳). پس از اعمال روش نمونه‌گیری مصنوعی تطبیقی، تعداد نمونه‌های کلاس۱ (استخدام) به ۹۶۱ نمونه افزایش یافت؛ در حالی که تعداد نمونه‌های کلاس ۰ (عدم‌استخدام) بدون تغییر و برابر با ۱۰۳۵ نمونه باقی ماند.



شکل۱. توزیع داده‌ها قبل و بعد از اعمال روش نمونه‌گیری مصنوعی تطبیقی

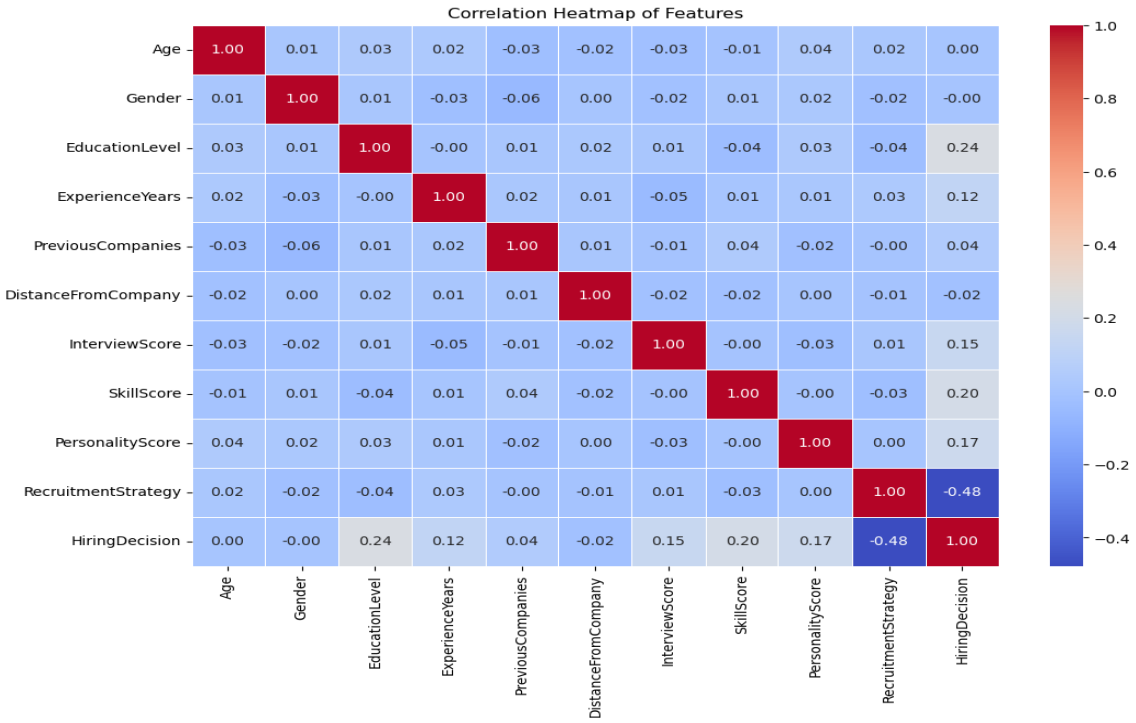
شکل بالا توزیع داده‌ها را قبل و بعد از اعمال روش نمونه‌گیری مصنوعی تطبیقی نشان می‌دهد. برای نمایش بهتر داده‌ها در فضای دو بُعدی، از تحلیل مؤلفهٔ اصلی (PCA)[4] استفاده شد. تحلیل مؤلفهٔ اصلی یک روش کاهش ابعاد است که داده‌های چندبُعدی را به دو جزء اصلی تبدیل می‌کند تا نمایش بصری آن‌ها را ساده‌تر و قابل تفسیرتر کند (میگندا، مولر و شنک[5]، ۲۰۲۱). توزیع اصلی داده‌ها در نمودار سمت چپ نشان داده شده است، جایی که کلاس یک (نمونه اشتغال) بسیار کمتر از کلاس صفر (نمونه غیراستخدامی) است. برخلاف نمودار سمت چپ، نمودار سمت راست توزیع داده را پس از اعمال نمونه‌گیری مصنوعی تطبیقی نشان می‌دهد، جایی که چگالی داده کلاس ۱ افزایش می‌یابد. این تغییر نشان می‌دهد که نمونه‌گیری مصنوعی تطبیقی، داده‌های کلاس اقلیت را افزایش داده و توزیع متعادلی برای خروجی ایجاد کرده است که می‌تواند عملکرد مدل‌های یادگیری ماشین را در پیش‌بینی اشتغال افزایش دهد. این مجموعه داده، از کیفیت بالایی برخوردار است و دادهٔ گم‌شده، نویزی یا ناقصی ندارد. در مراحل اولیهٔ تجزیه‌وتحلیل،

1. ADASYN
2. Yadav
3. Dey & Pratap
4. Principal Component Analysis
5. Migenda, Möller & Schenck

داده‌ها به‌دقت مورد بررسی شدند و مشخص شد که تمام مقادیر موجود در متغیرهای مختلف کامل و سازگارند. این فرایند پیش‌پردازش داده‌ها را ساده کرد و تمرکز بیشتر روی تحلیل منجر شد و مدل‌سازی دقیق‌تری را فراهم آورد. این ویژگی مجموعه داده، یعنی عدم‌وجود مقادیر از دست رفته و کیفیت داده بالا، اعتبار نتایج مدل‌های یادگیری ماشین را افزایش و پیچیدگی مراحل پیش‌پردازش را کاهش می‌دهد. با این حال، توزیع نامتعادل داده‌ها همچنان به‌عنوان یک چالش در این مطالعه شناسایی و مدیریت شد. به‌منظور مقیاس‌بندی داده‌ها از استاندارد اسکالر[1] استفاده شد. همهٔ متغیرها به بازه‌ای با میانگین صفر و انحراف معیار یک تبدیل شدند.

تحلیل هم‌بستگی

تحلیل هم‌بستگی به‌منظور شناسایی روابط خطی میان متغیرهای موجود در مجموعه داده انجام شده است (کزازی، خانی و بیرامی، ۱۴۰۱). هدف این تحلیل، شناسایی متغیرهایی است که بر متغیر هدف تأثیر بیشتری دارند و همچنین، متغیرهایی را که ارتباط تکراری یا ضعیفی با سایر متغیرها دارند، حذف می‌کند (وانگ، تانگ، ژانگ و گان[2]، ۲۰۱۸). در این پژوهش، از ماتریس هم‌بستگی به همراه نمودار نقشهٔ حرارتی برای نمایش بصری روابط بین متغیرها استفاده شده است.



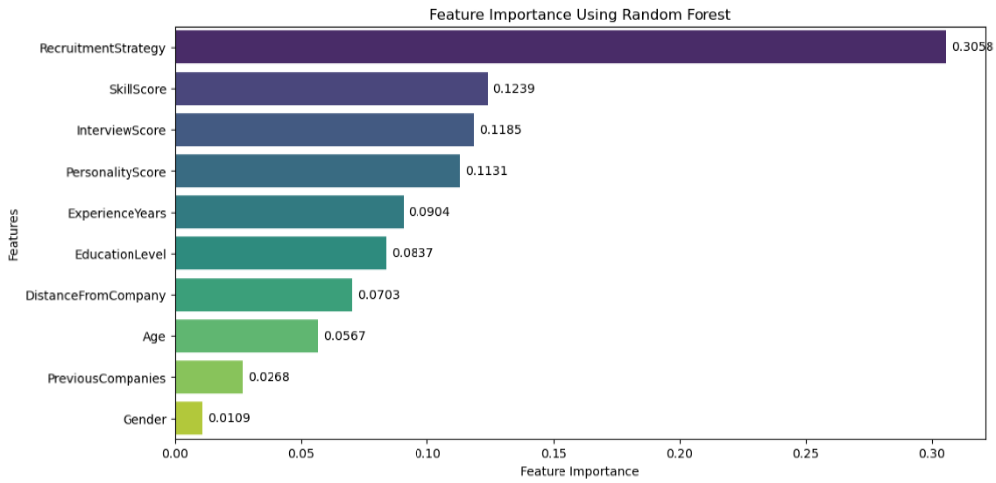
شکل۲. نقشهٔ حرارتی هم‌بستگی

1. Standard Scaler
2. Wang, Tang, Zhang & Guan

شکل بالا ماتریس هم‌بستگی ویژگی‌های مجموعه داده را نمایش می‌دهد. این نقشهٔ حرارتی میزان ارتباط بین متغیرهای مستقل و متغیر وابسته (تصمیم استخدام) را نشان می‌دهد و مشخص می‌کند که چگونه هر ویژگی با سایر ویژگی‌ها در ارتباط است (خلجستانی، پیری و ستوده، ۱۴۰۳). مقدار +۱ نشان‌دهندهٔ هم‌بستگی مثبت کامل، مقدار ۱- نشان‌دهندهٔ هم‌بستگی منفی کامل و مقدار نزدیک به صفر گویای عدم‌ارتباط میان دو متغیر است. بررسی این ماتریس نشان می‌دهد که ویژگی «استراتژی جذب نیروی انسانی» با متغیر تصمیم استخدام، بیشترین هم‌بستگی منفی را دارد (مقدار۰/۴۸-). این موضوع نشان می‌دهد که نوع استراتژی به‌کارگرفته‌شده در فرایند استخدام، بر تصمیم نهایی تأثیر مستقیمی دارد و تغییر در این استراتژی، می‌تواند شانس پذیرش متقاضیان را تحت تأثیر قرار دهد. در مقابل، ویژگی «سطح تحصیلات» با تصمیم استخدام بیشترین هم‌بستگی مثبت را دارد (مقدار ۰/۲۴)؛ یعنی افرادی که مدارک تحصیلی بالاتری دارند، برای پذیرفته شدن در فرایند استخدام احتمال بیشتری خواهند داشت. علاوه‌براین، ویژگی‌های «امتیاز مهارت‌های فنی»، «امتیاز ویژگی‌های شخصیتی» و «امتیاز مصاحبه» نیز با متغیر تصمیم استخدام هم‌بستگی مثبتی دارند. این امر نشان می‌دهد که عملکرد متقاضیان در آزمون‌های فنی، ویژگی‌های شخصیتی و مصاحبه‌های استخدامی، عوامل تأثیرگذاری در انتخاب نیروی انسانی محسوب می‌شوند. در مقابل، برخی از ویژگی‌ها مانند «سن»، «فاصلهٔ محل سکونت تا شرکت» و «تعداد شرکت‌های قبلی که فرد در آن‌ها کار کرده است» با تصمیم استخدام هم‌بستگی بسیار ضعیفی دارند. این موضوع نشان می‌دهد که این متغیرها بر فرایند تصمیم‌گیری در پذیرش یا رد یک متقاضی تأثیر چندانی ندارند و احتمالاً نقش کم‌رنگ‌تری در معیارهای ارزیابی استخدامی ایفا می‌کنند.

اهمیت ویژگی‌ها

شکل ۳ اهمیت ویژگی‌های مؤثر بر فرایند استخدام را براساس مدل جنگل تصادفی نشان می‌دهد. در این تحلیل، اهمیت هر ویژگی در پیش‌بینی متغیر هدف محاسبه شده و ویژگی‌های مهم‌تر در بالای نمودار قرار گرفته‌اند.



شکل۳. اهمیت ویژگی‌ها با جنگل تصادفی

نتایج نشان می‌دهد که قوی‌ترین ویژگی فرایند استخدام، استراتژی استخدام است و دارای بیشترین وزن ۰/۳۰۵۸ است؛ از این رو الگوی جذب و انتخاب متقاضیان در انتخاب نهایی استخدام بسیار مهم است. در مرحلهٔ بعد، ویژگی‌های امتیاز مهارت‌های فنی، نمرهٔ مصاحبه و امتیاز ویژگی‌های شخصیتی، بسته به تأثیر آن‌ها، به‌ترتیب ۰/۱۲۳۹، ۰/۱۱۸۵ و ۰/۱۱۳۱ رتبه‌بندی می‌شوند. پیامدهای حاصل از نتایج این است که قابلیت‌های فنی، عملکرد مصاحبه و ویژگی‌های شخصیتی متقاضیان، عواملی هستند که برای جذب متقاضی باید در نظر گرفته شوند. ویژگی‌های سطح تحصیلات، سابقهٔ کار و فاصله از شرکت، در تصمیم‌گیری واقعی اهمیت متوسطی دارد. بنابراین به نظر می‌رسد که عملکرد مصاحبه و مهارت‌های فنی، نسبت به تجربه یا سطح تحصیلات متقاضیان روی انتخاب متقاضیان تأثیر بیشتری دارد؛ هرچند هر یک اهمیت خاص خود را دارند. از سوی دیگر، تعداد شرکت‌های قبلی و جنسیت، در تعیین احتمال کمترین تأثیر را دارد؛ زیرا تأثیر متغیر جنسیت با مقدار ۰/۰۱۰۹ ناچیز است و به این معناست که مدل یادگیری ماشین به‌طور کلی جنسیت را عامل مهمی برای انتخاب نامزد در نظر نمی‌گیرد.

انتخاب ویژگی‌ها

برای این مطالعه، مدل یادگیری ماشین و کاهش پیچیدگی داده‌ها با روش حذف ویژگی بازگشتی با اعتبارسنجی متقاطع (RFECV) بهینه‌سازی شد (دویدر، عمرانی، بالنگین، بنونا و آبیک[1]، ۲۰۲۲). این روش سعی می‌کند که اهمیت ویژگی‌ها را دریافت کند، آن‌هایی را که تأثیر زیادی روی متغیر هدف ندارند، حذف کند و فقط آن‌هایی را حفظ کند که در مدل‌سازی بیشترین اهمیت را دارند (آواستی، گوتام و شارما[2]، ۲۰۲۴). RFECV بازگشتی یک یا چند ویژگی را با کمترین تأثیر در هر مرحله حذف می‌کند، مدل را دوباره آموزش می‌دهد و عملکرد آن را ارزیابی می‌کند (آواد و فریهات[3]، ۲۰۲۳). با استفاده از اعتبارسنجی متقاطع، این فرایند انجام شد و در نهایت مجموعه‌ای از ویژگی‌هایی که بهترین عملکرد را برای مدل‌سازی به ارمغان می‌آورد، انتخاب شد (دویدر و همکاران، ۲۰۲۲).

پس از اجرای RFECV هشت ویژگی کلیدی که در تصمیم نهایی استخدام بیشترین تأثیر را دارند، انتخاب شدند. ویژگی‌هایی مورد استفاده عبارت‌اند از: سن، سطح تحصیلات، سابقهٔ کار، فاصله از محل کار، نمرهٔ مصاحبه، نمرهٔ مهارت‌های فنی، امتیاز ویژگی شخصیتی و استراتژی استخدام. نتایج ارائه شده در این انتخاب نشان می‌دهد که عوامل ویژه به قابلیت‌های فردی، مهارت‌ها، ارزیابی و استراتژی استخدام شرکت، روی تشکیل فرایند تصمیم‌گیری استخدام پیامدهای مستقیمی دارند.

RFECV مزایای کلیدی بسیاری دارد. اول، پیچیدگی مدل را با حذف ویژگی‌های غیرضروری کاهش می‌دهد و از این رو پردازش داده‌ها را سریع‌تر می‌کند. دوم، مدل‌ها بهتر عمل می‌کنند، وقتی ویژگی‌های بی‌اهمیت حذف شوند، داده‌ها نویز کمتری دارند (آواد و فریهات، ۲۰۲۳). این روش از برازش بیش از حد جلوگیری می‌کند و به مدل اجازه می‌دهد تا تعمیم‌پذیری بهتری نسبت به داده‌های جدید داشته باشد. نتایج یک بررسی نشان می‌دهد که ویژگی‌هایی مانند جنسیت و

1. Douider, Amrani, Balenghien, Bennouna & Abik
2. Awasthi, Gautam & Sharma
3. Awad & Fraihat

تعداد شرکت‌های قبلی، بر تصمیم‌گیری‌های استخدام تأثیر چندانی نمی‌گذارد و حذف آن‌ها دقت مدل را افزایش می‌دهد (دویدر و همکاران، ۲۰۲۲).

مدل‌های یادگیری ماشین و یادگیری عمیق

در این پژوهش، برای پیش‌بینی تصمیم نهایی در فرایند استخدام، ۱۲ مدل یادگیری ماشین مورد استفاده قرار گرفت. این مدل‌ها عبارت‌اند از: جنگل تصادفی، تقویت گرادیان، درخت‌های فوق تصادفی، ماشین بردار پشتیبان (SVM)، رگرسیون لجستیک، بیز ساده، نزدیک‌ترین همسایگان، درخت تصمیم، اکس‌جی‌بوست، لایت‌جی‌بی‌ام، کتبوست و آدابوست. برای بهینه‌سازی هایپرپارامترها، از آپتونا استفاده شد. آپتونا یک چارچوب بهینه‌سازی فرامدل است که از روش‌های جست‌وجوی هوشمند مانند جست‌وجوی بیزی[1] برای یافتن بهترین ترکیب پارامترهای مدل استفاده می‌کند (اکیبا، سانو، یاناسه، اوهتا و کویاما[2]، ۲۰۱۹). برخلاف روش‌های سنتی مانند جست‌وجوی شبکه‌ای[3] و جست‌وجوی تصادفی[4] که فضای پارامترها را به‌صورت جامع، اما غیرکارآمد بررسی می‌کنند، آپتونا با استفاده از بهینه‌سازی تطبیقی سریع‌تر به بهترین تنظیمات مدل دست پیدا می‌کند. جدول ۳ مدل‌های یادگیری ماشین و هایپرپارامترهای بهینه آن‌ها را نشان می‌دهد.

جدول۳. مدل‌های یادگیری ماشین و هایپرپارامترهای بهینه

مدل	توضیح مدل	هایپرپارامترهای بهینه
جنگل تصادفی	مدلی مبتنی بر ترکیب چندین درخت تصمیم که با نمونه‌گیری تصادفی و ترکیب نتایج باعث بهبود دقت پیش‌بینی می‌شود (حسینی، ۱۴۰۳).	n_estimators': 228, 'max_depth': 20, '} 'min_samples_split': 5, {'min_samples_leaf': 2
تقویت گرادیان	مدلی که به‌صورت مرحله‌ای درخت‌های تصمیم را اضافه می‌کند و خطاها را در هر مرحله کاهش می‌دهد (فهیمی، رجب‌زاده قطری، شعار و خادمی، ۱۴۰۲).	n_estimators': 263, 'max_depth': 8, '} 'min_samples_split': 9, {'min_samples_leaf': 1
درخت‌های فوق‌تصادفی	مشابه جنگل تصادفی است؛ اما با تصادفی‌سازی بیشتر برای جلوگیری از هم‌بستگی بین درخت‌ها (نادکارنی، ویجی و کامات[5]، ۲۰۲۳).	n_estimators': 288, 'max_depth': 19, '} 'min_samples_split': 6, {'min_samples_leaf': 1
ماشین بردار پشتیبان	مدلی که داده‌ها را در فضای برداری نگاشت کرده و از ابرصفحه‌ها برای تفکیک کلاس‌ها استفاده می‌کند (سان و زی[6]، ۲۰۲۳).	{'C': 0.3448, 'kernel': 'poly'}
رگرسیون لجستیک	مدلی آماری که احتمال وابستگی کلاس‌ها به ویژگی‌ها را بررسی و دسته‌بندی می‌کند (مسلمان‌زاده، کوشا و صاعدی، ۱۴۰۳).	{'C': 9.8795, 'solver': 'lbfgs'}

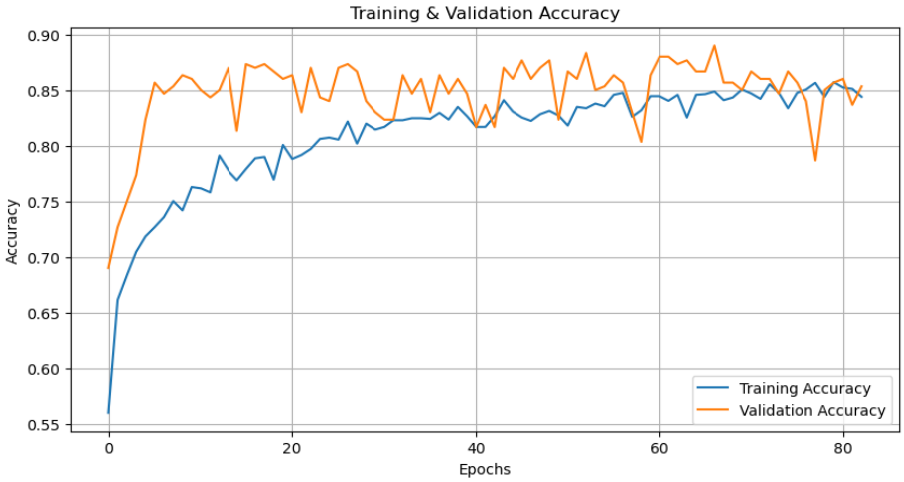
1. Bayesian Optimization
2. Akiba, Sano, Yanase, Ohta & Koyama
3. Grid Search
4. Random Search
5. Nadkarni, Vijay & Kamath
6. Sun & Xie

مدل	توضیح مدل	هایپرپارامترهای بهینه
بیز ساده	مدلی مبتنی بر نظریهٔ بیز که احتمال تعلق داده‌ها به هر کلاس را محاسبه می‌کند (ال هندی[1] و همکاران، ۲۰۲۰).	{var_smoothing': 4.2273e-08'}
نزدیک‌ترین همسایگان	مدلی که داده‌های جدید را براساس نزدیکی به نمونه‌های قبلی دسته‌بندی می‌کند (ژانگ[2]، ۲۰۲۱).	{n_neighbors': 24'}
درخت تصمیم	مدلی ساده، اما قدرتمند که داده‌ها را به‌صورت سلسله‌مراتبی براساس بیشترین اطلاعات دسته‌بندی می‌کند (منگ، بای، ژنگ، لیو و ژنگ[3]، ۲۰۲۳).	max_depth': 6, 'min_samples_split': '} {2, 'min_samples_leaf': 8
اکس‌جی‌بوست	مدلی که از تقویت گرادیان استفاده می‌کند و درخت‌های تصمیم را بهینه می‌کند (جعفرنژاد چقوشی، رضاسلطانی و خانی، ۱۴۰۳).	n_estimators': 106, 'learning_rate': '} {0.0165
لایت‌جی‌بی‌ام	مدلی سبک و سریع که بهینه‌سازی تقویت گرادیان را برای داده‌های حجیم انجام می‌دهد (جعفرنژاد و همکاران، ۱۴۰۳).	n_estimators': 132, 'learning_rate': '} {0.1679
کتبوست	مدلی مبتنی‌بر تقویت گرادیان که از روش‌های تطبیقی برای بهبود یادگیری استفاده می‌کند (جعفرنژاد و همکاران، ۱۴۰۳).	iterations': 242, 'learning_rate': '} {0.0167
آدابوست	مدلی که درخت‌های تصمیم ضعیف را با استفاده از وزن‌دهی تطبیقی بهبود می‌بخشد (دینگ، ژو، چن و لی[4]، ۲۰۲۲).	n_estimators': 222, 'learning_rate': '} {0.0301

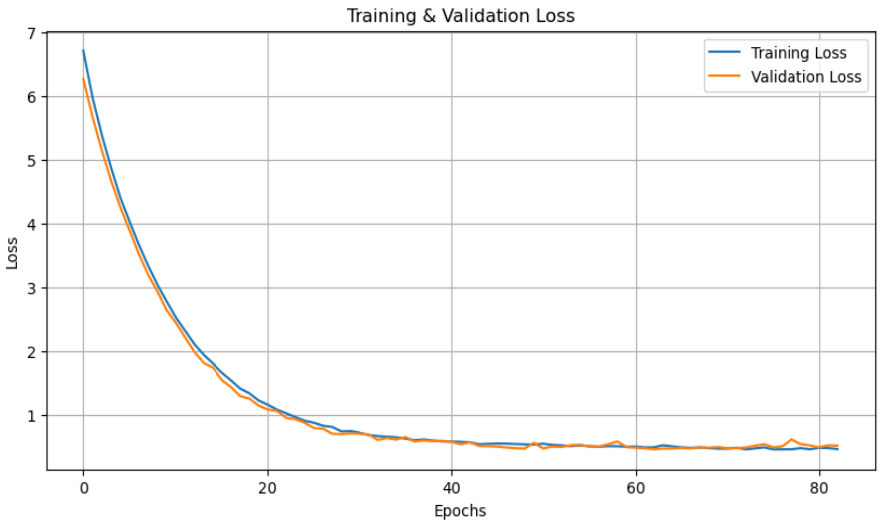
علاوه‌بر مدل‌های یادگیری ماشین، یک مدل شبکه عصبی عمیق (DNN) نیز توسعه داد شد. در اولین مدل‌سازی، مدل شبکهٔ عصبی عمیق برای ساختار مدل ۳ تا ۷ لایهٔ پنهان برای یادگیری رابطه بین ویژگی‌های ورودی تنظیم شد. برای هرلایه، تعداد نورون‌ها بین ۳۲ تا ۵۱۲ نورون تنظیم و بهترین مقدار برای تعداد نورون‌ها در فاصله ۳۲ واحد انتخاب شد. برای جلوگیری از بیش‌برازش[5]، از نرخ افت[6] در محدوده ۰/۱ تا ۰/۵ در هر لایه استفاده شد تا بهترین مقدار آن مشخص شود. علاوه‌براین، نرخ یادگیری با ۳ مقدار ۰/۰۰۱، ۰/۰۰۰۵ و ۰/۰۰۰۱ مورد ارزیابی قرار گرفت تا مشخص شود که کدام مقدار مناسب‌ترین برای به‌روزرسانی وزن‌های مدل است. در لایهٔ پنهان از یک تابع فعال‌سازی با ReLU و در لایهٔ خروجی به‌عنوان یک تابع فعال‌سازی با تابع Sigmoid استفاده شد. تابع هزینه دودویی متقاطع آنتروپی برای کاهش میزان خطا انتخاب شد و بهینه ساز Adam برای تنظیم وزن استفاده شد. علاوه‌براین، توقف اولیه نیز برای پایدارتر کردن آموزش مدل مورد استفاده قرار گرفت؛ به‌طوری که آموزش در ۲۰ دوره پس از اینکه خطای اعتبارسنجی از ۲۰ دوره متوالی[7] فراتر رفت، متوقف شود و بهترین حالت مدل ذخیره شود. اندازهٔ دستهٔ ورودی ۶۴ نمونه برای تمام مراحل آموزش مدل اولیه در ۲۰۰ دورهٔ مثبت تعیین شد. معماری مدل و ابرپارامترها با استفاده از آپتونا بهینه شدند.

1. El Hindi
2. Zhang
3. Meng, Bai, Zhang, Liu & Zhang
4. Ding, Zhu, Chen & Li
5. Overfitting
6. Dropout Rate
7. Epochs

به‌عنوان ساختار مدل بهینه، تعداد زیادی از ۵ لایه پنهان انتخاب شد که هر کدام دارای ترکیب متفاوتی از نورون‌ها و نرخ‌های فروپاشی در هر لایه بودند که پس از فرایند بهینه‌سازی اجرا شدند. به‌طور مشخص، لایهٔ اول شامل ۳۸۴ نورون با نرخ افت ۰/۵، لایهٔ دوم ۳۲ نورون با نرخ افت ۰/۵، لایهٔ سوم ۴۱۶ نورون با نرخ افت ۰/۲، لایهٔ چهارم ۱۲۸ نورون با نرخ افت ۰/۱ و لایهٔ پنجم ۳۲۰ نورون با نرخ افت ۰/۴ تنظیم شد. علاوه‌براین، مقدار ۰/۰۰۱ به‌عنوان نرخ یادگیری بهینه انتخاب شد.



شکل٤. روند تغییرات دقت در طول آموزش و اعتبارسنجی مدل شبکه عصبی عمیق



شکل۵. روند تغییرات میزان خطا در طول آموزش و اعتبارسنجی مدل شبکه عصبی عمیق

نمودارهای بالا نشان می‌دهد که چگونه دقت و خطا در طول آموزش و اعتبارسنجی برای یک مدل شبکه عصبی عمیق تغییر می‌کند. نمودار اول دقت داده‌هایی را که در جلسات آموزشی استفاده می‌شود، نشان می‌دهد و نمودار دوم خطای مدل را برای هر دو داده نشان می‌دهد. از نمودار دقت می‌توان دریافت که دقت مدل در مجموعه داده‌های آموزشی و اعتبار سنجی، به‌تدریج افزایش یافت و پس از حدود ۴۰ جلسه نسبتاً پایدار شد؛ اما دقت آموزش و اعتبارسنجی در شروع آموزش با تفاوت‌های بسیار زیادی همراه است و اولی با داده‌های موجود در فرایند تطبیق می‌یابد. این تفاوت با گذراندن جلسه‌های بیشتر و همگرا شدن دو منحنی کاهش می‌یابد که نشان می‌دهد مدل تعمیم‌پذیری بهتری به‌دست آورده است. با وجود این، دقت اعتبارسنجی پس از ۴۰ جلسه نوسان دارد؛ به این معنا که در برخی موارد مدل تغییرات ناگهانی را در پیش‌بینی برخی پارامترها تجربه می‌کند.

از نمودار نرخ خطا می‌توانیم ببینیم که خطای مدل هر دو مجموعه داده با تعداد چرخه‌های آموزشی کاهش می‌یابد. با این حال، خطا در مدل در ابتدای آموزش بسیار زیاد است، جایی که با پیشرفت فرایند یادگیری کاهش می‌یابد و در نهایت پس از حدود ۴۰ چرخه به مقدار تقریبی می‌رسد. منحنی‌های خطای آموزش و اعتبارسنجی در آخرین مراحل بسیار نزدیک به هم هستند؛ یعنی مدل بیش از حد برازش نمی‌کند و روی داده‌های جدید عملکرد خوبی دارد. این موضوع نشان می‌دهد که مدل شبکهٔ عصبی عمیق توانایی یادگیری الگوهای پیچیده از داده‌های آموزشی و عملکرد خوب در داده‌های اعتبار سنجی را دارد. مفاهیم مثبت واقعی (TP)[1] و منفی واقعی (TN)[2] برای ارزیابی پیش‌بینی مدل‌های یادگیری ماشین و یادگیری عمیق برای پیش‌بینی تصمیم‌های استخدام استفاده می‌شوند (ویوویچ[3]، ۲۰۲۱). TP تعداد متقاضیانی است که مدل به‌درستی پیش‌بینی کرده و استخدام می‌شوند، TN تعداد متقاضیانی است که مدل به‌درستی پیش‌بینی کرده و استخدام نمی‌شوند. در واقع، مثبت کاذب (FP)[4] به تعداد مواردی اشاره دارد که مدل نتوانسته است پیش‌بینی کند که یک فرد قرار است استخدام شود، در حالی‌که واقعاً این طور نبود. علاوه‌براین، منفی کاذب (FN)[5] تعداد متقاضیانی را توصیف می‌کند که مدل پیش‌بینی کرده بود استخدام نمی‌شوند؛ اما به‌طور کلی آن‌ها استخدام شده‌اند (نایدو، زووا و سیباندا[6]، ۲۰۲۳). شاخص‌های صحت، دقت، بازخوانی و امتیاز F1 براساس این معیارها محاسبه شدند؛ سپس از اعتبارسنجی متقاطع پنج‌تایی، برای ارزیابی دقیق‌تر عملکرد مدل‌های یادگیری ماشین در پیش‌بینی تصمیم‌های استخدام استفاده شد. در اینجا داده‌ها به ۵ قسمت مساوی تقسیم شدند و هربار یکی از این قسمت‌ها به‌عنوان مجموعه آزمایش و بقیهٔ قسمت‌ها یک مجموعه آموزشی درنظر گرفته شد. فرایند فوق ۵ بار برای هریک از داده‌ها تکرار شد تا اطمینان حاصل شود که همهٔ داده‌ها به‌عنوان مجموعه آزمایشی ارائه می‌شوند. سپس از نتایج هر تکرار میانگین گرفته شد تا عملکرد مدل را بتوان روی داده‌های جدید، بدون سوگیری از هیچ بخشی از آن داده، ارزیابی کرد. این روش از برازش بیش از حد جلوگیری می‌کند و اعتبار نتایج مدل را افزایش می‌دهد.

1. True Positive
2. True Negative
3. Vujovic
4. False Positive
5. False Negative
6. Naidu, Zuva & Sibanda

جدول٤. شاخص‌های ارزیابی مدل‌های یادگیری ماشین

شاخص	تعریف	فرمول
صحت[1]	نسبت پیش‌بینی‌های صحیح به کل نمونه‌ها	$\frac{TP+TN}{TP+FP+FN+TN}$
دقت[2]	نسبت پیش‌بینی‌های درست برای یک کلاس به کل نمونه‌هایی که به‌عنوان آن کلاس پیش‌بینی شده‌اند	$\frac{TP}{TP+FP}$
بازیابی[3]	نسبت پیش‌بینی‌های درست برای یک کلاس به کل نمونه‌های واقعی آن کلاس	$\frac{TP}{TP+FN}$
امتیاز F1	میانگین‌های رمونیک بین دقت و بازیابی برای ایجاد تعادل بین آن‌ها.	$\frac{2\times Precision\times Recall}{Precision+Recall}$

یافته‌های پژوهش

در این پژوهش، عملکرد ۱۲ مدل یادگیری ماشین و یک مدل شبکه عصبی عمیق برای پیش‌بینی تصمیم استخدام مورد بررسی قرار گرفت و معیارهای صحت، دقت، بازیابی و امتیاز F1 برای هر مدل محاسبه شد. جدول ۵ نتایج مربوط به ۴ معیار برای هر مدل را نشان می‌دهد.

جدول ۵. مقدارهای بهینه‌ای برای هر یک از ویژگی‌های تأثیرگذار بر خرابی

مدل	صحت	دقت	بازیابی	امتیاز F1
جنگل تصادفی	۰/۹۴	۰/۹۲۱۳۴۸	۰/۸۸۱۷۲۰	۰/۹۰۱۰۹۹
تقویت گرادیان	۰/۹۴	۰/۹۳۱۰۳۴	۰/۸۷۰۹۶۸	۰/۹
درخت‌های فوق‌تصادفی	۰/۹۱۶۶۶۷	۰/۸۹۵۳۴۹	۰/۸۲۷۹۵۷	۰/۸۶۰۳۳۵
ماشین بردار پشتیبان	۰/۸۷۶۶۶۷	۰/۸۱۱۱۱۱	۰/۷۸۴۹۴۶	۰/۷۹۷۸۱۴
رگرسیون لجستیک	۰/۷۸	۰/۶۰۹۷۵۶	۰/۸۰۶۴۵۲	۰/۶۹۴۴۴۴
بیز ساده	۰/۸۰	۰/۶۲۷۹۰۷	۰/۸۷۰۹۶۸	۰/۷۲۹۷۳۰
نزدیک‌ترین همسایگان	۰/۸۴۶۶۶۷	۰/۷۰۰۸۵۵	۰/۸۸۱۷۲۰	۰/۷۸۰۹۵۲
درخت تصمیم	۰/۸۸۶۶۶۷	۰/۸۳۹۰۸۰	۰/۷۸۴۹۴۶	۰/۸۱۱۱۱۱
اکس‌جی‌بوست	۰/۹۳	۰/۹۵۰۰۰۰	۰/۸۱۷۲۰۴	۰/۸۷۸۶۱۳
لایت‌جی‌بی‌ام	۰/۹۳۳۳۳۳	۰/۹۱۹۵۴۰	۰/۸۶۰۲۱۵	۰/۸۸۸۸۸۹
کتبوست	۰/۹۵۳۳۳۳	۰/۹۵۴۰۲۳	۰/۸۹۲۴۷۳	۰/۹۲۲۲۲۲
آدابوست	۰/۸۸	۰/۷۶۶۳۵۵	۰/۸۸۱۷۲۰	۰/۸۲۰۰۰۰
شبکه عصبی عمیق	۰/۸۹	۰/۸۸۴۶۱۵	۰/۷۴۱۹۳۵	۰/۸۰۷۰۱۸

1. Accuracy
2. Precision
3. Recall

نتایج آزمون‌ها نشان داد که مدل‌های مبتنی بر درخت تصمیم مانند کتبوست، جنگل تصادفی و تقویت گرادیان، بهتر از مدل‌های دیگر بودند. مدل کتبوست با صحت ۰/۹۵۳۳۳۳، دقت ۰/۹۵۴۰۲۳، بازیابی ۰/۸۹۲۴۷۳ و امتیاز F1 برابر با ۰/۹۲۲۲۲۲ بهترین مدل این پژوهش است. دلیل عملکرد برتر این مدل، نسبت به سایر روش‌ها این است که می‌تواند با داده‌های طبقه‌بندی شده و عدم‌تعادل داده‌ها مقابله کند. با استفاده از ترتیب پردازش بهینه در داده‌ها و تکنیک‌های کاهش بیش از حد برازش، کتبوست به دستیابی بهتر به دقت و بازیابی و یک مدل نهایی پایدار بهتر منجر شد.

جنگل تصادفی عملکرد بسیار خوبی با صحت ۰/۹۴ و امتیاز F1 برابر با ۰/۹۰۱۰۹۹ نشان داده است. دلیل موفقیت برای این مدل، مجموعه‌ای از درخت‌های تصمیم‌گیری تصادفی و ترکیبی از آن پیش‌بینی‌ها به‌منظور افزایش تعمیم‌پذیری مدل است. عملکرد این مدل خوب است؛ زیرا به‌ویژه در مشکلات با چندین ویژگی و پیچیده، به‌خوبی عمل می‌کند و از بیش برازش پیش از حد جلوگیری می‌کند. در این مطالعه، تقویت گرادیان و لایت‌جی‌بی‌ام نیز مدل‌های بسیار قوی‌ای بودند. با صحت ۰/۹۴ و امتیاز F1 برابر با ۰/۹، تقویت گرادیان می‌تواند الگوهای پیچیده داده را به‌خوبی یاد بگیرد.

تقویت گرادیان و یادگیری گام‌به‌گام در این مدل برای کاهش خطای مدل، در هر مرحله اعمال می‌شود. از سوی دیگر، مدل لایت‌جی‌بی‌ام نیز با صحت ۰/۹۳۳۳۳۳ و امتیاز F1 برابر با ۰/۸۸۸۸۹ عملکرد مناسبی نشان داده است.

لایت‌جی‌بی‌ام یکی از محبوب‌ترین مدل‌های یادگیری ماشینی است؛ زیرا زمان مناسبی را صرف پردازش مجموعه داده‌های بزرگ با سرعت کافی می‌کند.

بهترین مدل از بین مدل‌ها، مدل اکس‌جی‌بوست با صحت ۰/۹۳ و امتیاز F1 برابر با ۰/۸۷۸۶۱۳ بود. نمونه‌برداری هوشمند و تنظیم دقیق وزن‌ها، این مدل را ترکیبی از درخت‌های تقویت‌کننده گرادیان و تصمیم‌گیری می‌کند که از الگوهای تشکیل‌شده از داده‌ها استفاده می‌کند.

مدل‌های ساده‌تر مانند رگرسیون لجستیک و بیز ساده، به‌وضوح عملکرد بدی داشتند. این مدل‌ها در شناسایی تصمیم‌های پیچیده استخدام چندان موفق نبودند. با استفاده از رگرسیون لجستیک صحت ۰/۷۸ و امتیاز F1 برابر با ۰/۶۹۴۴۴۴ بود. این مدل با توجه به خطی بودن نمی‌تواند رابطهٔ غیرخطی و پیچیده بین متغیرها را بیاموزد. با این حال، بیز ساده نیز در مقایسه با مدل‌های دیگر ضعیف عمل کرد (صحت ۰/۸۰ و امتیاز F1 برابر با ۰/۷۲۹۷۳۰)؛ زیرا براساس آن فرض استقلال بین متغیرها، در بسیاری از موارد درست نیست (صحت ۰/۸۹، دقت ۰/۸۸۴۶۱۵، بازیابی ۰/۷۴۱۹۳۵ و امتیاز F1 برابر با ۰/۸۰۷۰۱۸).

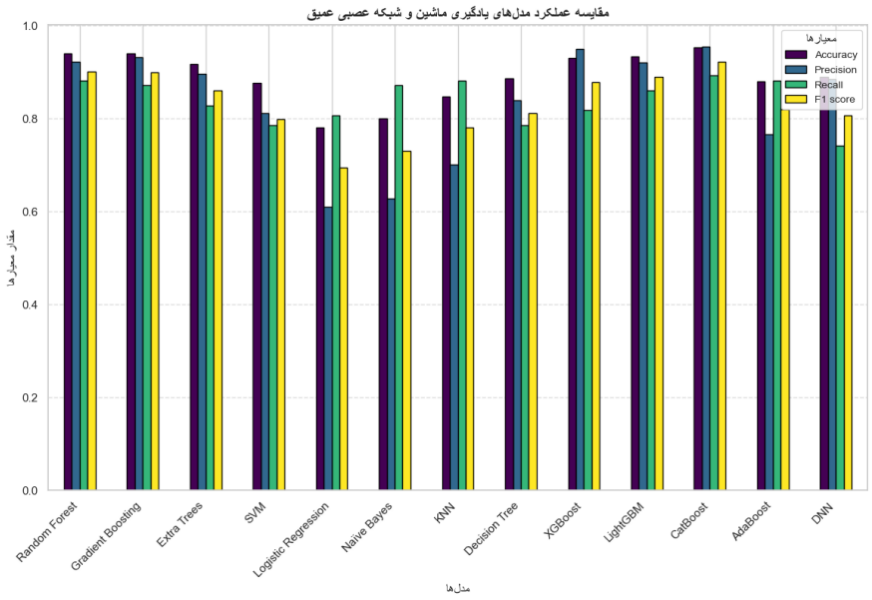
مدل شبکه عصبی عمیق به‌طور معقولی کار کرد. با این حال، از مدل‌های یادگیری جمعی بهتر عمل نکرد؛ شاید دلیل آن، حساسیت به وزن‌های اولیه، پیچیدگی در تنظیم فراپارامتر یا نیاز شبکهٔ عصبی به داده‌های بیشتر برای یادگیری بهتر باشد. در حالی که الگوهای داده‌های پیچیده را به‌خوبی یاد گرفت، مدل‌های یادگیری جمعی در پیش‌بینی متقاضیان واجد شرایط بهتر عمل کردند. با این حال، اگر داده‌های بیشتری در دسترس بود و تنظیمات معماری شبکه بهینه‌تر می‌شد، عملکرد آن بهبود پیدا می‌کرد.

مدل ماشین بردار پشتیبان با صحت ۰/۸۷۶۶۶۷ و امتیاز F1 برابر با ۰/۷۹۷۸۱۴ عملکرد متوسطی نشان داده است. با این حال، در برخی موارد، این مدل ممکن است از انتخاب یک هسته مناسب و از وابستگی به مقیاس ویژگی‌ها رنج ببرد.

سایر میانگین عملکردهای مورد انتظار از مدل‌های درخت‌های فوق تصادفی و آدابوست نیز با صحت‌های ۰/۹۱۶۶۶۷ برای درخت‌های فوق تصادفی و ۰/۸۸ برای آدابوست مشاهده شده است. مانند جنگل تصادفی، مدل درخت‌های فوق تصادفی نیز تا حد زیادی همین کار را انجام می‌دهد؛ اما نقاط برش ویژگی‌ها به‌طور تصادفی انتخاب می‌شوند. با این حال، این مدل پس از این رویکرد در مقایسه با جنگل تصادفی عملکرد خوبی دارد؛ اما قابلیت تعمیم کمتری دارد.

مدل تقویت تطبیقی (آدابوست) با امتیاز F1 برابر با ۰/۸۲۰۰۰۰ تا حدی مؤثر است؛ اما با دقت کمی نسبت به سایرین کاهش یافته است؛ زیرا به نویز داده‌ها بسیار حساس است. مدل نزدیک‌ترین همسایگان با صحت ۰/۸۴۶۶۶۷ و امتیاز F1 برابر با ۰/۷۸۰۹۵۲ عملکرد متوسطی داشته است. زمانی که داده‌ها به‌شدت پراکنده هستند، این به کاهش عملکرد در این مدل منجر می‌شود؛ زیرا به نزدیک‌ترین نمونه‌های آموزشی متکی است. علاوه‌براین، این مدل در مقایسه با سایر مدل‌ها زمان نسبتاً طولانی‌تری برای پردازش داده‌ها دارد.

در این مطالعه، مدل درخت تصمیم نیز مورد بررسی قرار گرفت و صحت ۰/۸۸۶۶۶۷ و امتیاز F1 برابر با ۰/۸۱۱۱۱۱ را به‌دست آورد و از مدل‌های ساده بهتر بود؛ اما نسبت به مدل‌های یادگیری گروهی عملکرد ضعیفی داشت؛ زیرا این مدل معمولاً روی داده‌های آموزشی بیش از حد برازش می‌کند و باعث می‌شود آن را کمتر به داده‌های دیگر تعمیم دهد.



شکل ٦. مقایسه عملکرد مدل‌ها

بحث و نتیجه‌گیری

در این مطالعه از الگوریتم‌های یادگیری ماشین و یادگیری عمیق برای پیش‌بینی تصمیم‌های استخدام استفاده شد. مشخص شد که عمدتاً مدل‌های یادگیری جمعی مبتنی‌بر تقویت گرادیان، جنگل تصادفی و یادگیری عمیق، از مدل‌های سنتی برای پیش‌بینی نتایج استخدام بهتر عمل کردند. به‌طور کلی، این مدل‌ها نتایجی را ارائه می‌دهند که از آن‌ها می‌توان دریافت که یادگیری جمعی و یادگیری عمیق، به‌ویژه در پردازش داده‌های پیچیده و غیرخطی، بر دقت پیش‌بینی می‌افزایند و خطا را در مقایسه با مدل‌های قدیمی‌تر، مانند درخت‌های تصمیم‌گیری و ماشین‌های بردار پشتیبان کاهش می‌دهند. نتایج حاصل از تحلیل داده‌های متقاضیان شغل، مانند سن، تحصیلات، تجربهٔ کاری، مهارت‌های فنی و ویژگی‌های شخصیتی، نشان داد که این ویژگی‌ها می‌توانند به‌عنوان ورودی‌های مدل‌های پیش‌بینی برای شبیه‌سازی رفتارها و انتخاب‌های استخدامی به‌کار روند. به‌ویژه، الگوریتم‌های یادگیری جمعی توانسته‌اند رابطه‌های پیچیده و غیرخطی میان ویژگی‌های متقاضیان و نتایج تصمیم‌های استخدامی را مدل‌سازی کنند. این موضوع نشان می‌دهد که هنگام انتخاب متقاضیان با استفاده از این الگوریتم‌ها، می‌توان به دقت بیشتری در مقایسه با روش‌های سنتی دست یافت.

مطالعات متعددی در زمینهٔ کاربرد الگوریتم‌های یادگیری ماشین و یادگیری عمیق در فرایند استخدام انجام شده است. القحفه و همکاران (۲۰۲۴) در پژوهش خود به استفاده از مدل‌های یادگیری ماشین برای پیش‌بینی موفقیت متقاضیان شغلی پرداختند. نتایج آن‌ها نشان داد که استفاده از داده‌های متقاضیان برای پیش‌بینی نتایج استخدام، می‌تواند دقت تصمیم‌گیری را به‌طور چشمگیری افزایش دهد. به‌طور خاص، این پژوهش نشان داد که ویژگی‌های فردی متقاضیان، مانند سطح تحصیلات، تجربهٔ کاری، مهارت‌های فنی، و ویژگی‌های شخصیتی در پیش‌بینی موفقیت در یک شغل نقش مهمی دارند. مطالعهٔ مشابهی را العکاشه و همکاران (۲۰۲۳) انجام دادند و به این نتیجه رسیدند که ویژگی‌های فردی و تجربه‌های قبلی متقاضیان، بر موفقیت آن‌ها در فرایند استخدام تأثیر زیادی دارد. این پژوهش تأکید کرد که داده‌های مربوط به شخصیت متقاضیان و تجربه‌های شغلی آن‌ها می‌تواند به‌عنوان ورودی‌های مؤثر در مدل‌های پیش‌بینی‌کننده به‌کار روند. این نتایج در پژوهش حاضر نیز تأیید شد و نشان داده شد که مدل‌های یادگیری ماشین و یادگیری عمیق، قادرند اطلاعات پیچیده و مختلف را تحلیل کنند و پیش‌بینی‌هایی دقیق از عملکرد آیندهٔ متقاضیان ارائه دهند. گرونبرگ و همکاران (۲۰۲۴) در مطالعه‌ای دیگر نشان دادند که الگوریتم‌های یادگیری عمیق، به‌دلیل قابلیتی که در تحلیل داده‌های پیچیده و شبیه‌سازی روابط غیرخطی میان متغیرهای مختلف دارند، می‌توانند به‌طور خاص در تحلیل و پیش‌بینی نتایج استخدامی مفید واقع شوند. این پژوهش به‌ویژه بر تأثیر ویژگی‌های شخصیتی و روان‌شناختی متقاضیان در تصمیم‌های استخدامی تأکید داشت. نتایج مطالعهٔ آن‌ها نشان داد که مدل‌های یادگیری عمیق می‌توانند الگوهای رفتاری پیچیده متقاضیان را شبیه‌سازی کنند و پیش‌بینی دقیقی از موفقیت یا عدم‌موفقیت آن‌ها در شغل مدنظر ارائه دهند. در مجموع، نتایج این پژوهش‌ها و پژوهش حاضر نشان می‌دهد که الگوریتم‌های یادگیری ماشین و یادگیری عمیق، نه تنها قادرند داده‌های پیچیدهٔ متقاضیان را تحلیل کنند؛ بلکه می‌توانند در کاهش اشتباه‌های انسانی و

انتخاب‌های نادرست در فرایند استخدام، به‌طور مؤثری عمل کنند. در واقع، این الگوریتم‌ها به‌ویژه در شرایطی که داده‌های زیاد و پیچیده‌ای برای تصمیم‌گیری وجود دارد، می‌توانند تصمیم‌های دقیق‌تری اتخاذ کنند.

استفاده از مدل‌های پیشرفتهٔ یادگیری ماشین و یادگیری عمیق در این پژوهش مزایای زیادی داشت. یکی از این مزایا، توانایی این مدل‌ها در شبیه‌سازی الگوهای غیرخطی پیچیده و تعاملات متغیرهای مختلف بود که برای مدل‌های سنتی قابل پردازش نبود. این توانایی باعث شد تا نتایج این پژوهش دقت بیشتری داشته باشد و به شبیه‌سازی دقیق‌تری از فرایندهای استخدامی منجر شود. علاوه‌براین، روش‌های پیشرفتهٔ متعادل‌سازی داده‌ها و انتخاب ویژگی‌هایی مانند ADASYN و RFECV در این تحقیق، به کاهش ابعاد داده‌ها و همچنین انتخاب ویژگی‌های مؤثرتر کمک کردند. این اقدام‌ها موجب بهبود عملکرد کلی مدل‌ها شدند و نتایج پیش‌بینی دقیق‌تری به همراه داشتند. این پژوهش نشان داد که استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق می‌تواند به‌طور چشمگیری دقت پیش‌بینی در فرایند تصمیم‌گیری استخدامی را افزایش دهد. مدل‌های یادگیری جمعی و یادگیری عمیق، قادرند رفتارهای پیچیدهٔ کارکنان و متقاضیان را دقیق‌تر شبیه‌سازی کنند؛ از این رو، در مقایسه با مدل‌های سنتی نتایج بهتری ارائه می‌دهند. این نوآوری می‌تواند در بهبود فرایندهای استخدامی و کاهش هزینه‌های ناشی از اشتباه‌ها در انتخاب نیروی انسانی مفید واقع شود.

نتایج و پیشینهٔ مطالعهٔ حاضر، به‌وضوح مزیت‌های یادگیری ماشین و الگوریتم‌های یادگیری عمیق را در فرایند استخدام نشان می‌دهد که برخی نیز باید در کانون توجه قرار گیرد؛ زیرا این مطالعه محدودیت‌هایی دارد. کیفیت و یکپارچگی داده‌ها، یکی از چالش‌های بسیار مهم است. همهٔ مدل‌های یادگیری ماشینی و یادگیری عمیق، برای دستیابی به پیش‌بینی‌های قابل اعتماد به داده‌های تمیز و دقیق نیاز دارند؛ یعنی اگر داده‌ها نادرست یا ناقص باشند، پیش‌بینی‌های مدل‌ها، به نتایج نادرستی منتهی می‌شود که می‌تواند به تصمیم‌گیری‌های نادرستی در فرایند استخدام منجر شود. علاوه‌براین، مسئلهٔ دیگر، مشکل زمان و کار فشردهٔ آموزش مدل‌های یادگیری عمیق است. استخراج ویژگی‌های پیچیده و روابط غیرخطی از داده‌های متقاضی در این مدل‌ها، زمان و منابع محاسباتی زیادی را می‌طلبد. این مسئله می‌تواند در برخی سازمان‌ها که به منابع محاسباتی دسترسی محدودی دارند، چالش‌برانگیز باشد و امکان استفاده از این الگوریتم‌ها را محدود کند.

نتایج این مطالعه نشان داد که مدل‌های یادگیری ماشینی و یادگیری عمیق، برای بهبود دقت پیش‌بینی‌های تصمیم‌گیری استخدام پتانسیل زیادی دارند و همچنین، به سازمان‌ها کمک می‌کنند تا منابع انسانی آگاهانه‌ای را انتخاب کنند. با این حال، چنین مدل‌هایی به بررسی بیشتری در برخی جنبه‌ها نیاز دارند؛ از این رو تحقیقات آینده ممکن است ویژگی‌های ورودی را به‌گونه‌ای گسترش دهند که متغیرهایی مانند تعداد فعالیت‌های تعاملی در مصاحبه‌های آنلاین، تحلیل احساسات پاسخ‌های نامزد و داده‌های رفتاری در رابطه با تعاملات اجتماعی را شامل شود. علاوه‌براین، یک راه امیدوارکننده برای تحقیقات آینده، در تأثیرهای فرهنگی و سازمانی بر پویایی فرایند استخدام نهفته است. از این رو، چندین الگوریتم پردازش زبان طبیعی، مزیت‌های منحصربه‌فرد و مؤثری دارند تا هم‌سویی ارزش‌های سازمانی را با ویژگی‌های متقاضی برای انتخاب بهتر سرمایهٔ انسانی تحلیل کنند. از سوی دیگر، رویکردهای مشترک با استفاده از

مدل‌های یادگیری ماشین با روش‌های تصمیم‌گیری چند معیاره، مانند پردازش سلسله‌مراتبی تحلیلی یا روش‌های مبتنی بر تئوری فازی، ممکن است رویکردهای ترکیبی را برای محاسبه بهبودیافته برای پیش‌بینی فرایند استخدام ارائه دهند. علاوه‌براین، یکی از چالش‌های اصلی در استفاده از هوش مصنوعی در فرایندهای استخدام، تعصب‌های ناعادلانهٔ احتمالی است که ممکن است در مدل‌های یادگیری ماشین ایجاد شود. تحقیقات بیشتر می‌تواند الگوریتم‌ها و روش‌های منصفانه را برای کاهش تعصب در این مدل‌ها، برای تصمیم‌گیری شفاف‌تر و منصفانه‌تر بررسی کند.

در سطح عملی، نتیجهٔ این تحقیق چندگانه است: کمک به سازمان‌ها، ارائه‌دهندگان راه‌حل‌های استخدام، فرایندهای استخدام و مدیران منابع انسانی که می‌توانند انتخاب خود را بهبود بخشند. در میان این برنامه‌ها، بهینه‌سازی فرایند انتخاب از طریق توسعه و استفاده از مدل‌های پیش‌بینی، شاید مهم‌ترین آن‌ها باشد. چنین مدل‌های پیش‌بینی‌کننده‌ای، به بررسی سریع‌تر متقاضیان و کاهش زمان و هزینه‌های مربوط به کل فرایند استخدام کمک می‌کند. نتایج همچنین به کاهش نرخ جابه‌جایی شغلی کمک خواهند کرد؛ زیرا سازمان با تجزیه‌وتحلیل و درک عوامل موفقیت برای متقاضیان، می‌تواند انتخاب مناسب‌تری از منابع انسانی داشته باشد. ارائهٔ بینش به مدیران منابع انسانی بر اساس حقایق به تسهیل قضاوت بهتر در خصوص استخدام و تخصیص منابع کمک می‌کند و در نتیجه، کارایی و بهره‌وری سازمانی را افزایش می‌دهد.

منابع

ایمانی، حسین؛ قلی‌پور، آرین؛ آذر، عادل و پورعزت، علی اصغر (۱۳۹۸). شناسایی مؤلفه‌های سیستم تأمین منابع انسانی در راستای ارتقای سلامت نظام اداری. *مدیریت دولتی*، ۱۱(۲)، ۲۵۱-۲۸۴.

خلجستانی، سعید؛ پیری، حبیب و ستوده، رضا (۱۴۰۳). ارائه الگوی پیش‌بینی حساسیت جبران خدمات مدیرعامل با استفاده از الگوریتم‌های فراابتکاری (ژنتیک و ازدحام ذرات). *مدیریت دولتی*، ۱۶(۳)، ۵۶۲- ۶۰۰.

جعفرنژاد چقوشی، احمد؛ رضاسلطانی، آرمان و خانی، امیرمحمد (۱۴۰۳). مقایسه مدل‌های یادگیری جمعی برای پیش‌بینی رتبهٔ کشوری دانش‌آموزان در کنکور سراسری. *مدیریت صنعتی*، ۱۶(۳)، ۴۵۷- ۴۸۱.

حسینی، سیدعابد (۱۴۰۳). تجزیه‌وتحلیل سیگنال‌های مغزی به کمک آنتروپی پراکندگی سلسله‌مراتبی و جنگل تصادفی در کاربرد بازاریابی عصبی. *هوش محاسباتی در مهندسی برق*، ۱۵(۱)، ۴۱- ۵۶.

عارف نژاد، محسن و سپهوند، رضا (۱۳۹۶). اثر جهت‌گیری مسیر شغلی متنوع بر قابلیت استخدامی کارکنان با نقش میانجی سرمایهٔ مسیر شغلی (مطالعهٔ موردی: ترخیص‌کاران گمرک شهید رجایی هرمزگان). *مدیریت دولتی*، ۹(۴)، ۶۸۷-۷۰۸.

عباس‌پور، عباس؛ رحیمیان، حمید؛ غیاثی ندوشن، سعید و نرگسیان، جواد (۱۳۹۷). ارائه مدل انتخاب کارکنان مستعد در سازمان‌های دولتی. *مدیریت دولتی*، ۱۰(۴)، ۶۰۵-۶۲۸.

فهیمی، محمدرضا؛ رجب‌زاده قطری، علی؛ شعار مریم، خادمی مریم (۱۴۰۲). مدل پیش‌بینی تقاضای زنجیره تأمین با تنوع محصولی بالا با استفاده از روش‌های یادگیری ماشین مبتنی بر تقویت گرادیان. *فصلنامه علمی ـ پژوهشی اقتصاد و مدیریت شهری*، ۱۲(۴۵)، ۴۷- ۶۶.

کزازی، ابوالفضل؛ خانی، امیر محمد و بیرامی، ثریا (۱۴۰۰). تأثیر مدیریت کیفیت زنجیره‌تأمین و عملکرد نوآوری بر عملکردعملیاتی کسب وکارهای فعال در صنایع غذایی استان گلستان. مطالعات مدیریت صنعتی، ۱۹(۶۲)، ۶۷-۹۸.

مسلمان‌زاده، فاطمه؛ کوشا، حمیدرضا و صاعدی،کاظم (۱۴۰۳). پیش‌بینی ماهیت حریق مبتنی بر یادگیری ماشین: رگرسیون لجستیک یک الگوریتم تفسیر پذیر. *پژوهش‌های نظری و کاربردی هوش ماشینی*، ۲(۱)،۱۰۴-۱۱۹.

References

Abbas Pour, A., Rahimian, H., Ghiasi Nodooshan, S. & Nargesian, J. (2018). Presenting a Model to Select Talented Employee in State Organizations. *Journal of Public Administration*, 10(4), 605-628. doi: 10.22059/jipa.2019.271575.2443. (*in Persian*)

Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.1907.10902

Al Akasheh, M., Faisal Malik, E., Hujran, O. & Zaki, N. (2023). A Decade of Research on Data Mining Techniques for Predicting Employee Turnover: A Systematic Literature Review. *Expert Systems with Applications*, 121794. https://doi.org/10.1016/j.eswa.2023.121794

Albaroudi, E., Mansouri, T. & Alameer, A. (2024). A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring. *AI,* 5(1), 383–404. MDPI. https://www.mdpi.com/2673-2688/5/1/19

Albassam, W. A. (2023). The Power of Artificial Intelligence in Recruitment: An Analytical Review of Current AI-Based Recruitment Strategies. *International Journal of Professional Business Review*, *8*(6). https://doi.org/10.26668/businessreview/2023.v8i6.2089

Ali Shah, S. A., Uddin, I., Aziz, F., Ahmad, S., Al-Khasawneh, M. A. & Sharaf, M. (2020). An Enhanced Deep Neural Network for Predicting Workplace Absenteeism. *Complexity*, *2020*, 1–12. https://doi.org/10.1155/2020/5843932

Al-Quhfa, H., Mothana, A., Aljbri, A. & Song, J. (2024). Enhancing Talent Recruitment in Business Intelligence Systems: A Comparative Analysis of Machine Learning Models. *Analytics*, *3*(3), 297–317. https://doi.org/10.3390/analytics3030017

Alsheref, F. K., Fattoh, I. E. & Ead, M.W. (2022). Automated Prediction of Employee Attrition Using Ensemble Model Based on Machine Learning Algorithms. *Computational Intelligence and Neuroscience,* 2022, 1–9. https://doi.org/10.1155/2022/7728668

Arefnejad, M. & Sepahvnd, R. (2018). The Effect of the Diverse Job Orientation on the Employability of Employees Considering the Mediating Role of the Career Path Capital (Case Study: Shahid Rajaee Customs' Clearance Officers in Hormozgan). *Journal of Public Administration*, *9*(4), 687-708. doi: 10.22059/jipa.2018.251033.2183 (*in Persian*)

Awad, M. & Fraihat, S. (2023). Recursive Feature Elimination with Cross-Validation with Decision Tree: Feature Selection Method for Machine Learning-Based Intrusion

Detection Systems. *Journal of Sensor and Actuator Networks*, *12*(5), 67. https://doi.org/10.3390/jsan12050067

Awasthi, N., Gautam, P. R. & Sharma, A. K. (2024). RFECV-DT: Recursive Feature Selection with Cross Validation using Decision Tree based Android Malware Detection. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–6. https://doi.org/10.1109/icccnt61001.2024.10725127

Brown, L., George, B. & Mehaffey-Kultgen, C. (2018). The development of a competency model and its implementation in a power utility cooperative: an action research study. *Industrial and Commercial Training*, 50(3), 123–135. https://doi.org/10.1108/ict-11-2017-0087

Dey, I. & Pratap, V. (2023). *A Comparative Study of SMOTE, Borderline-SMOTE, and ADASYN Oversampling Techniques using Different Classifiers*. https://doi.org/10.1109/icsmdi57622.2023.00060

Ding, Y., Zhu, H., Chen, R. & Li, R. (2022). An Efficient AdaBoost Algorithm with the Multiple Thresholds Classification. *Applied Sciences*, *12*(12), 5872. https://doi.org/10.3390/app12125872

Douider, M., Amrani, I., Balenghien, T., Bennouna, A., & Abik, M. (2022). Impact of Recursive Feature Elimination with Cross-validation in Modeling the Spatial Distribution of Three Mosquito Species in Morocco. *Revue D Intelligence Artificielle*, *36*(6), 855–862. https://doi.org/10.18280/ria.360605

ElSharkawy, G., Helmy, Y. & Yehia, E. (2022). Employability Prediction of Information Technology Graduates using Machine Learning Algorithms. *International Journal of Advanced Computer Science and Applications*, *13*(10). https://doi.org/10.14569/ijacsa.2022.0131043

Gao, H., Liang, G. & Chen, H. (2022). Multi-Population Enhanced Slime Mould Algorithm and with Application to Postgraduate Employment Stability Prediction. *Electronics*, *11*(2), 209. https://doi.org/10.3390/electronics11020209

Ghasemian Sahebi, I., Toufighi, S.P. Azzavi, M. & Zare, F. (2023). Presenting an optimization model for multi cross-docking rescheduling location problem with metaheuristic algorithms. *OPSEARCH*, 61(1), 137–162. https://doi.org/10.1007/s12597-023-00694-5

Grunenberg, E., Peters, H., Francis, M. J., Back, M. D. & Matz, S. C. (2024). Machine learning in recruiting: predicting personality from CVs and short text responses. *Frontiers in Social Psychology*, *1*. https://doi.org/10.3389/frsps.2023.1290295

Hosseini, S. A. (2024). Analysis of EEG Signals using Hierarchical Dispersion Entropy and Random Forest in the Neuromarketing Application. *Computational Intelligence in Electrical Engineering,* 15(1), 41-56. doi: 10.22108/isee.2023.133401.1561(*in Persian*)

Hunt, W. & O'Reilly, J. (2022). *Rapid Recruitment in Retail: Leveraging AI in the hiring of hourly paid frontline associates during the Covid-19 Pandemic*. https://doi.org/10.20919/alnb9606

Imani, H., Gholipour, A., Azar, A. & Pourezzat, A. A. (2019). Identifying Components of Staffing System to Develop Administrative Integrity. *Journal of Public Administration*, *11*(2), 251-284. doi: 10.22059/jipa.2019.277466.2504 (*in Persian*)

Imianvan, A. A., Robinson, S. A., Asuquo, D. E., George, U. D., Dan, E. A., Ejodamen, P. U. & Udoh, A. E. (2024). Enhancing Job Recruitment Prediction through Supervised Learning and Structured Intelligent System: A Data Analytics Approach. *Journal of Advances in Mathematics and Computer Science*, *39*(2), 72–88. https://doi.org/10.9734/jamcs/2024/v39i21869

Jafarnejad Chaghoshi, A., Rezasoltani, A. & Khani, A. M. (2024). Unleashing the Power of Ensemble Learning: Predicting National Ranks in Iran's University Entrance Examination. *Industrial Management Journal*, 16(3), 457-481. doi: 10.22059/imj.2024.381521.1008178 (*in Persian*)

Jayanti, L. P. S. D. & Wasesa, M. (2022). Application of Predictive Analytics To Improve The Hiring Process In A Telecommunications Company. Jurnal CoreIT: *Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi,* 8(1). https://doi.org/10.24014/coreit.v8i1.16915

Jha, K., Likhitha, D., Chandana, M. S., Reddy, M. R. P., & Bhargavi, M. (2024, July). Career Prediction Using Machine Learning. In *2024 8th International Conference on Inventive Systems and Control (ICISC)* (pp. 118-122). IEEE. https://doi.org/10.1109/icisc62624.2024.00027

Kazai, A., Khani, A. M. and birami, S. (2021). The effect of supply chain quality management and innovation performance on the operational performance of businesses operating in the food industry of Golestan province. *Industrial Management Studies,* 19(62), 67-98. doi: 10.22054/jims.2021.58750.2612 (*in Persian*)

Khaljastani, S., Piri, H. & Sotoudeh, R. (2024). Presenting a Prediction Model for CEO Compensation Sensitivity using Meta-heuristic Algorithms (Genetics and Particle Swarm). *Journal of Public Administration*, 16(3), 562-600. doi: 10.22059/jipa.2024.373930.3482 (*in Persian*)

Kumar, V. & Garg, M. L. (2018). Predictive Analytics: A Review of Trends and Techniques. *International Journal of Computer Applications*, 182(1), 31–37. https://doi.org/10.5120/ijca2018917434

Ma, Y., Zhang, Z. & Ihler, A. (2020). A Deep Choice Model for Hiring Outcome Prediction in Online Labor Markets. *International Journal of Computers Communications & Control*, *15*(2). https://doi.org/10.15837/ijccc.2020.2.3760

Madanchian, M. (2024). From Recruitment to Retention: AI Tools for Human Resource Decision-Making. *Applied Sciences*, *14*(24), 11750. https://doi.org/10.3390/app142411750

Meng, L., Bai, B., Zhang, W., Liu, L. & Zhang, C. (2023). Research on a Decision Tree Classification Algorithm Based on Granular Matrices. *Electronics*, *12*(21), 4470–4470. https://doi.org/10.3390/electronics12214470

Meng, Z. (2023). *Analysis and Prediction of College Students' Employment Based on Decision Tree Classification Algorithm*. 1–6. https://doi.org/10.1109/wconf58270.2023.10234985

Migenda, N., Möller, R. & Schenck, W. (2021). Adaptive dimensionality reduction for neural network-based online principal component analysis. *PLOS ONE*, *16*(3), e0248896. https://doi.org/10.1371/journal.pone.0248896

Hanif, A. M., Maarop, N., Kamaruddin, N. & Samy, G. N. (2024). Machine Learning Approach in Predicting Fraudulent Job Advertisement. *International Journal of Academic Research in Business and Social Sciences*, *14*(1), 1182–1193. http://dx.doi.org/10.6007/IJARBSS/v14-i1/20532

Mosalmanzadeh, F., Koosha, H. & Saedi, K. (2025). Predicting the nature of fire based on machine learning: Logistic regression is an interpretable algorithm. *Applied and basic Machine intelligence research,* 2(1), 104-119. doi: 10.22034/abmir.2025.22313.1068 (*in Persian*)

Nadkarni, S. B., Vijay, G. S. & Kamath, R. C. (2023). *Comparative Study of Random Forest and Gradient Boosting Algorithms to Predict Airfoil Self-Noise*. https://doi.org/10.3390/engproc2023059024

Nagovitsyn, R. (2023). Predicting Student Employment in Teacher Education Using Machine Learning Algorithms. *Education & Self Development*, *18*(2), 133–148. https://doi.org/10.26907/esd.18.2.10

Naidu, G., Zuva, T. & Sibanda, E.M. (2023). A Review of Evaluation Metrics in Machine Learning Algorithms. *Lecture Notes in Networks and Systems*, *724*, 15–25. https://doi.org/10.1007/978-3-031-35314-7_2

Ozdemir, Y. & Nalbant, K. G. (2020). Personnel Selection for Promotion using an Integrated Consistent Fuzzy Preference Relations - Fuzzy Analytic Hierarchy Process Methodology: A Real Case Study. *Asian Journal of Interdisciplinary Research*, 219–236. https://doi.org/10.34256/ajir20117

Pampouktsi, P., Avdimiotis, S., Maragoudakis, M. & Avlonitis, M. (2021). Applied Machine Learning Techniques on Selection and Positioning of Human Resources in the Public Sector. *Open Journal of Business and Management*, 09(02), 536–556. https://doi.org/10.4236/ojbm.2021.92030

Pessach, D., Singer, G., Avrahami, D., Chalutz Ben-Gal, H., Shmueli, E. & Ben-Gal, I. (2020). Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. *Decision Support Systems,* 134(1), 113290. https://doi.org/10.1016/j.dss.2020.113290

Rabie El Kharoua. (2024). *Predicting Hiring Decisions in Recruitment Data*. Doi.org. https://doi.org/10.34740/KAGGLE/DSV/8715385

Rallapalli, S. & Kumar, Y.M. (2024). *Optimizing Employee Promotion Decisions: A Novel Machine Learning Framework for Predictive Analysis by using GBM CatBoost*. 1–7. https://doi.org/10.1109/ssitcon62437.2024.10796145

Raza, A., Munir, K., Almutairi, M., Younas, F. & Fareed, M. M. S. (2022). Predicting Employee Attrition Using Machine Learning Approaches. *Applied Sciences*, *12*(13), 6424. https://doi.org/10.3390/app12136424

Sarker, I. H. (2021). Deep Learning: a Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(6). Springer. https://doi.org/10.1007/s42979-021-00815-1

Suman, S., Kaur, S. J., Sharma, A. & Kumar, S. (2024). *Machine Learning-Based System for Admission and Jobs Prediction in Engineering and Technology Sector*. https://doi.org/10.1109/ic2pct60090.2024.10486533

Sun, F. & Xie, X. (2023). Deep Non-Parallel Hyperplane Support Vector Machine for Classification. *IEEE Access*, *11*, 7759–7767. https://doi.org/10.1109/access.2023.3237641

Taherdoost, H. (2023). Deep Learning and Neural Networks: Decision-Making Implications. *Symmetry*, *15*(9), 1723. https://www.mdpi.com/2073-8994/15/9/1723

Thilak, K. D., Lalitha Devi, K., Kalaiselvi, K. & Teja, J. (2023). Revolutionizing University Graduate Employability: Leveraging Advanced Machine Learning Models to Optimize Campus Recruitment and Placement Strategies. *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, 1–6. https://doi.org/10.1109/rmkmate59243.2023.10369300

Vujovic, Ž. Đ. (2021). Classification Model Evaluation Metrics. *International Journal of Advanced Computer Science and Applications*, *12*(6). https://doi.org/10.14569/ijacsa.2021.0120670

Wang, L., Tang, X., Zhang, J. & Guan, D. (2018). Correlation Analysis for Exploring Multivariate Data Sets. *IEEE Access*, *6*, 44235–44243. https://doi.org/10.1109/access.2018.2864685

Yadav, S. (2021). *A Comparative Analysis of Sampling Techniques for Imbalanced Datasets in Machine Learning - IJIRCT*. Ijirct.org. https://www.ijirct.org/viewPaper.php?paperId=2411071

Zhang, S. (2021). Challenges in KNN Classification. *IEEE Transactions on Knowledge and Data Engineering*, *34*(10), 1–1. https://doi.org/10.1109/tkde.2021.3049250

Vinutha, K. & Yogisha, H. K. (2021, March). Prediction of employability of engineering graduates using machine learning techniques. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 742-745). IEEE.