Educational Measurement and Evaluation Studies

Print ISSN: 2476-2865 Online ISSN: 2783-0942

National Organization for Educational Testing

Comparing the Performance of Decision Tree and Artificial Neural Network in the Classification of University Entrance Exam Candidates

Habib Naderi[1], Keyvan Salehi[2] , Maryam Parsaeeian [3]

1. Assistant Professor, Faculty of Mathematical Sciences, Statistics and Computer, University of Sistan and Baluchestan,Sistan and Baluchestan, Zahedan, Iran. Email: h.h.naderi@gmail.com
2. Associate Professor, Faculty of Educational Sciences and Psychology, University of Tehran, Tehran, Iran ; (Corresponding author), Email: keyvansalehi@ut.ac.ir
3. Ph.D. Student of Educational Measurement and Evaluation, Faculty of Educational Sciences and Psychology, University of Tehran, Tehran, Iran. Email: maryam.parsaeian@ut.ac.ir

Article Info

Article Type:

Research Article

Received:

2022/10/04

Received in revised form:

2023/01/28

Accepted:

2023/02/25

Published online:

2023/03/22

ABSTRACT

Objective: Comparison of the capabilities of three different algorithms of decision tree and neural network model in order to choose the appropriate classification model to evaluate the performance of university entrance exam candidates was the purpose of this study.

Methods: It is based on a quantitative approach and a survey method, and the gender variables of academic records and balance scores of each course were used as effective variables in the classification.

Results: Considering all the variables without gender based on the neural network model, it was found that the specialized courses of mathematics, physics and chemistry, followed by general courses of Persian literature and Theology, respectively, are the most important ones in the classification. The distance between receiving the diploma and the university entrance exam had the least effect. It was also found that if only variables that related to the test courses are considered, the order of importance of chemistry and physics courses will be shifted.

Conclusion: Based on the overall accuracy index, the neural network model algorithm with an accuracy greater than 0.95 had a higher performance than the decision tree algorithms. On the other hand, the inclusion of educational background variables had a favorable effect on the accuracy of the neural network algorithm.

Keywords: *Evaluating the performance of volunteers, decision tree, C5, CHAID, CART, neural network*

Cite this article: Naderi, Habib; Salehi, Keyvan; Parsaeeian, Maryam (2023). Comparing the Performance of Decision Tree and Artificial Neural Network in the Classification of University Entrance Exam Candidates**.** *Educational Measurement and Evaluation studies*, 13 (41): 39-68 pages.DOI: 10.22034/EMES.2023.545992.2335

© The Author(s).

Publisher: National Organization of Educational Testing (NOET)

مطالعات اندازه‌گیری و ارزشیابی آموزشی

سازمان سنجش آموزش کشور

شاپا چاپی: ۲۸۶۵-۲۴۷۶ شاپا الکترونیکی: ۰۹۴۲-۲۷۸۳

مقایسهٔ عملکرد درخت تصمیم و شبکهٔ عصبی مصنوعی در طبقه‌بندی داوطلبان آزمون سراسری

حبیب نادری[۱]، کیوان صالحی[۲]، مریم پارساییان[۳]

۱. استادیار، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سیستان و بلوچستان، سیستان و بلوچستان، زاهدان، ایران. رایانامه: h.h.naderi@gmail.com
۲. دانشیار، دانشکدهٔ روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران؛ (نویسنده مسئول)، رایانامه: keyvansalehi@ut.ac.ir
۳. دانشجوی دکتری سنجش و اندازه‌گیری، دانشکدهٔ روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. رایانامه: maryam.parsaeian@ut.ac.ir

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

دریافت: ۱۴۰۱/۰۷/۱۲
اصلاح: ۱۴۰۱/۱۱/۰۸
پذیرش: ۱۴۰۱/۱۲/۰۶
انتشار: ۱۴۰۲/۰۱/۰۲

چکیده

هدف: مقایسهٔ قابلیت‌های سه الگوریتم مختلف درخت تصمیم و الگوی شبکهٔ عصبی به منظور انتخاب الگوی طبقه‌بندی مناسب برای ارزیابی عملکرد داوطلبان آزمون سراسری

روش پژوهش: مبتنی بر رویکرد کمّی و با روش پیمایشی است و از متغیرهای جنسیت، سوابق تحصیلی و نمره‌های تراز هر یک از درس‌ها به عنوان متغیرهای مؤثر در طبقه‌بندی استفاده شد.

یافته‌ها: با در نظر گرفتن همهٔ متغیرها و بدون جنسیت براساس الگوی شبکهٔ عصبی، مشخص شد که درس‌های تخصصی ریاضی، فیزیک و شیمی سپس درس‌های عمومی فارسی و دینی به ترتیب بیشترین اهمیت را در طبقه‌بندی دارند. فاصلهٔ دریافت دیپلم تا آزمون سراسری کم‌ترین تأثیر را داشت. همچنین مشخص شد در صورتی که تنها متغیرهای مربوط به درس‌های آزمون در نظر گرفته شود، ترتیب میزان اهمیت درس‌های شیمی و فیزیک جابه‌جا می‌شود.

نتیجه‌گیری: با استناد به شاخص دقت کلی، الگوریتم الگوی شبکهٔ عصبی با دقتی بیشتر از ۰/۹۵ عملکرد بالاتری نسبت به الگوریتم‌های درخت تصمیم داشت. از طرفی ورود متغیرهای سوابق تحصیلی در دقت الگوریتم شبکهٔ عصبی تأثیر مطلوبی داشت.

واژه‌های کلیدی: ارزیابی عملکرد داوطلبان، درخت تصمیم، CART ، CHAID ، C5، شبکهٔ عصبی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

استناد: نادری، حبیب؛ صالحی، کیوان؛ پارساییان، مریم (۱۴۰۲). مقایسه عملکرد درخت تصمیم و شبکه عصبی مصنوعی در طبقه‌بندی داوطلبان آزمون سراسری. مطالعات اندازه گیری و ارزشیابی آموزشی، ۱۳ (شماره ۴۱)، ۳۹-۶۸ صفحه.
DOI: 10.22034/EMES.2023.545992.2335

ناشر: سازمان سنجش آموزش کشور حق‌مؤلف © نویسندگان.

مقدمه

عملکرد تحصیلی جنبه‌ای حیاتی از تحقیقات آموزشی است که از الگوریتم‌های مختلف یادگیری ماشینی برای الگوسازی و پیش‌بینی آن استفاده شده است. یادگیری ماشینی زیرشاخه‌ای از هوش مصنوعی است که شامل توسعهٔ الگوریتم‌ها و الگوهای آماری است که می‌تواند از داده‌ها یاد بگیرد و پیش‌بینی یا تصمیم‌گیری کند. در سال‌های اخیر، یادگیری ماشینی به دلیل توانایی در پردازش سریع و دقیق حجم عظیمی از داده‌ها، به طور فزاینده‌ای در حوزه‌های مختلفی از جمله مالی، مراقبت‌های بهداشتی، بازاریابی و آموزش محبوب شده است. به عبارت دقیق‌تر، توانایی طبقه‌بندی سریع و دقیق مقادیر زیادی از داده‌ها، الگوریتم‌های یادگیری ماشین را به ابزاری جذاب برای پیش‌بینی موفقیت یا شکست دانش‌آموزان تبدیل می‌کند.

در این مطالعه، دو الگوریتم محبوب یادگیری ماشینی یعنی درخت تصمیم و شبکهٔ عصبی مصنوعی در طبقه‌بندی داوطلبان آزمون سراسری با یکدیگر مقایسه شده‌اند. درخت تصمیم برای الگوسازی تصمیمات و پیامدهای احتمالی آنها در پیش‌بینی عملکرد تحصیلی استفاده می‌شود. به‌علاوه می‌توان از درخت تصمیم برای شناسایی مهم‌ترین عوامل مؤثر بر عملکرد دانش‌آموزان استفاده کرد. این عوامل ممکن است شامل جمعیت‌شناسی، وضعیت اجتماعی-اقتصادی، پیشرفت تحصیلی قبلی و عادات مطالعه باشد. از سوی دیگر، شبکه عصبی برای یادگیری الگوها و روابط پیچیده در داده‌ها هم برای کارهای طبقه‌بندی (مانند پیش‌بینی قبولی یا رد دانش‌آموز در یک درس) و هم برای کارهای رگرسیون (مانند پیش‌بینی نمرهٔ نهایی دانش‌آموز) استفاده کرد.

هدف این مقاله، مقایسهٔ ارزیابی عملکرد داوطلبان آزمون سراسری گرو علوم ریاضی و فنی با استفاده از روش‌های داده‌کاوی و با توجه به متغیرهای مختلف تأثیرگذار از جمله جنسیت، سوابق تحصیلی (معدل کتبی دیپلم و فاصلهٔ دریافت دیپلم و شرکت در آزمون سراسری) و نمره‌های تراز هر یک از درس‌ها است. علاوه بر این متغیرها با توجه به رویکرد عدالت‌محوری و توجه به مناطق مختلف کشور و با این استدلال که شرکت‌کنندگان در آزمون سراسری، امکانات آموزشی یکسانی نداشته‌اند تا بتوانند در آزمونی یکسان به رقابت بپردازند، تأثیر متغیر سهمیه نهایی (شامل مناطق یک، دو و سه و شاهد) نیز در نمرهٔ داوطلبان در نظر گرفته شد (صادقی جعفری، روشن و شکوری گنجوی، ۱۳۹۰). بدین‌منظور به مقایسه قابلیت‌های سه الگوریتم مختلف درخت تصمیم (CHAID، C5 و CART) و الگوی شبکهٔ عصبی پرداخته شده است. نتایج این مطالعه بینش ارزشمندی در مورد نقاط قوت و ضعف الگوریتم درخت تصمیم و شبکهٔ عصبی مصنوعی در پیش‌بینی عملکرد داوطلبان آزمون سراسری ارائه می‌کند. مؤسسات آموزشی می‌توانند این اطلاعات را برای بهبود الگوهای پیش‌بینی خود و درک بهتر عوامل مؤثر بر عملکرد دانش‌آموزان استفاده کنند.

مبانی نظری و پیشینه پژوهش

تعریف عملکرد تحصیلی از دیرباز مورد توجه دست اندرکاران تعلیم و تربیت بوده است. عملکرد تحصیلی دانشجویان به معنای میزان دستیابی دانشجو به اهداف آموزشی است (بنیک و کامر[1]، ۲۰۱۹). ماو[2] (۱۹۹۷)، نویدی (۱۳۸۲)، گلشنی فومنی (۱۳۷۵) و افروز (۱۳۷۵) پیشرفت تحصیلی را مترادف با عملکرد تحصیلی بالا و اصطلاح افت تحصیلی را مترادف با عملکرد تحصیلی پایین یا پس‌رفت تحصیلی مطرح کرده‌اند. در واقع اگر عملکرد تحصیلی به صورت پیوستار در نظر گرفته شود، در یک سر آن پیشرفت تحصیلی و سر دیگر آن افت تحصیلی قرار می‌گیرد. گاهی مشاهده می‌شود که محققان نیز در بسیاری از موارد این اصطلاحات را به طور مترادف به کار برده‌اند. فتی، آذری، برادران و اطلسی (۱۳۹۲) نیز کلید واژه‌های «افت تحصیلی»، «وضعیت تحصیلی»، «پیشرفت تحصیلی» و «موفقیت تحصیلی» را مترادف در نظر گرفته‌اند. بهترین شاخص برای نشان دادن افت یا پیشرفت تحصیلی از نظر رحیمی (۱۳۸۷)، مقایسهٔ عملکرد قبلی فراگیران با عملکرد فعلی آنها است. لاورنس و دیپا[3] (۲۰۱۳)، پیشرفت تحصیلی را به عنوان سطح موفقیت واقعی یا مهارت علمی توصیف کرده‌اند که در حوزهٔ دانشگاهی به دست آمده باشد. پیشرفت تحصیلی مقدار یادگیری آموزشگاهی فراگیران است که به جلوه‌هایی از وضعیت تحصیلی دانشجو اشاره می‌کند (بخشایش، ۱۳۹۳). از سوی دیگر افت تحصیلی به معنای عدم موفقیت در آموزش است (عبدالرزاق، کمال، موحسن، طارق، آل‌زبیدی و آل‌موسوی[4]، ۲۰۱۷). افت تحصیلی به این معناست که فراگیر از سطحی بالاتر به سطحی پایین‌تر در تحصیل و آموزش نزول کرده باشد. پس می‌توان آن را کاهش عملکرد فراگیر دانست (رحیمی، ۱۳۸۷). افت تحصیلی هنگامی اتفاق می‌افتد که فراگیر نتواند نمرهٔ ضعیف خود را به سطح مورد انتظار برساند و ملاک‌های آموزشی را برآورده نکند (ویدیاستوتی، کورنیاوان و چاندرا[5]، ۲۰۱۷).

مرور منظم مطالعات مرتبط نشان داد که معدل کل و ارزیابی داخلی[6]، دو ویژگی اصلی است که اغلب برای ارزیابی عملکرد تحصیلی استفاده می‌شود (آنجلین[7]، ۲۰۱۳؛ کریستین و ایوب[8]، ۲۰۱۴؛ لی، راسک و سانگ[9]، ۲۰۱۳؛ جیشان، راشو، هاک و رحمان[10]، ۲۰۱۵). از سی مقالهٔ مطالعه شده در این زمینه، ۱۳ مقاله از این دو ویژگی استفاده کرده‌اند. دلیل اصلی بیشتر محققان برای استفاده از معدل کل این است

1. Banik & Kumar
2. Mau
3. Lawrence & Deepa
4. Abdulrazzaq, Kamal, Muhsen, Tareq, Al Zubaidi & Al Mousawi
5. Widyastuti, Kurniawan & Chandra
6. internal assessment
7. Angeline
8. Christian & Ayub
9. Li, Rusk, & Song
10. Jishan, Rashu, Haque, & Rahman

کـه ارزش قابل ملاحظـه‌ای برای پویایی تحصیلی و شـغلی آینده دارد. همچنین می‌توان آن را نشـانه‌ای از ظرفیت علمی تحقق یافته دانسـت (بین‌مات، بونیامین، آرسـد و کاسـیم[1]، ۲۰۱۳). ایبراهیم وروسلی[2] (۲۰۰۷) با اسـتفاده از مقدار ضریب همبسـتگی (۰/۸۷)، نتیجه گرفتند که مهم‌ترین متغیر ورودی برای ارزیابـی عملکـرد تحصیلی، معدل کل اسـت. همچنین طبق مطالعه کریسـتین و ایـوب (۲۰۱۴) معدل کل بیشـترین تأثیـر را در تعیین بقای دانش‌آموز دارد، خـواه بتواند تحصیل خود را به پایان برسـاند یا نه. ویژگی مهم دیگر شامل ارزیابی‌های جمعیت‌شـناختی[3] و ارزیابی خارجی[4] دانش‌آموزان است. مشخصات جمعیت‌شناختی دانش‌آموزان شامل جنسیت، سن، سابقهٔ خانوادگی و معلولیت است (ایبراهیم و روسلی، ۲۰۰۷؛ بین‌مات، بونیامین، آرسـد و کاسیم، ۲۰۱۳؛ کریسـتین و ایوب، ۲۰۱۴). میت، بورگس، کوبیک و سـیبل[5] (۲۰۰۴) نشان دادند که بیشتر دانش‌آموزان دختر در مقایسه با دانش‌آموزان پسر دارای سبک‌ها و رفتارهای یادگیری مثبت مختلفی هسـتند (میت و همکاران، ۲۰۰۴). دانش‌آموزان دختر خودنظم‌دهی و دقت بیشـتری در مطالعات خود دارند. آنها همچنین راهبردهـای یادگیری موثری در مطالعات خود دارند (سیم‌سـک و بالابان[6]، ۲۰۱۰). دختران دانش‌آموز خود انگیزشـی، سـازمان‌دهی و مرور ذهنی دارند و به طـور مؤثر از آنها اسـتفاده می‌کنند. بنابر این نقش جنسـیت به عنوان یکی از ویژگی‌هـای مهم تأثیرگذار بر عملکرد دانش‌آموزان تأیید شـده اسـت (عثمان‌بگانویک و سـال‌جیک[7]، ۲۰۱۲). به عبارت دیگر دلیل اینکه بیشتر محققان از جنسـیت دانش‌آموزان به عنوان یک متغیر استفاده می‌کنند این است که دخترها و پسـرها در فرآیند یادگیری خود سـبک‌های متفاوتی دارند (بین‌مات، بونیامین، آرسد و کاسیم، ۲۰۱۳). همچنیـن سـابقهٔ تحصیلی هـم در پیش‌بینی عملکرد دانش‌آموز نقـش موثری دارد (آبوتیـر و ایل‌هالیز[8]، ۲۰۱۲؛ اولادوکان، آدبانجو و چارلز اوابا[9]، ۲۰۰۸؛ رامش، پرکاوی و رامر[10]، ۲۰۱۳).

بیـش از چهار دهه اسـت که ارزیابی عملکرد دانش‌آموزان دورهٔ متوسـطه به منظور ورود به دانشـگاه‌های ایران با برگزاری سـالانهٔ آزمون سراسری برای پنج گروه آزمایشـی علوم ریاضی و فنی، علوم تجربی، علوم انسـانی، زبان‌های خارجی و هنر به‌طور متمرکز انجام می‌شـود. با توجه به اهمیت نتایج این آزمون برای شـرکت‌کنندگان، از جمله آزمون‌های سرنوشت‌سـاز محسـوب می‌شـود و بنابر این نحـوهٔ گزینش دقیق داوطلبان در آن اهمیت ویژه‌ای دارد. با توجه به محتوای آزمون که از درس‌های دورهٔ متوسـطهٔ دوم اسـت،

1. Bin Mat, Buniyamin, Arsad, & Kassim
2. Ibrahim & Rusli
3. demographic
4. external assessment
5. Meit, Borges, Cubic & Seibel
6. Simsek & Balaban
7. Osmanbegovic & Suljic
8. Abu Tair & El-Halees
9. Oladokun, Adebanjo, & Charles-Owaba
10. Ramesh, Parkavi, & Ramar

می‌توان آن را به عنوان ابزاری برای سنجش پیشرفت تحصیلی دانش‌آموزان و نوع عملکرد آنها در دورهٔ متوسطهٔ دوم نیز در نظر گرفت.

سازمان سنجش به‌منظور سنجش دقیق داوطلبان، ملاک‌های مختلفی از جمله نمرهٔ علمی کسب شده در آزمون را در نظر می‌گیرد. داوطلبان با پاسخ به سؤال‌های مربوط به هر خرده‌آزمون، میزان توانایی خود را در حوزه‌های مختلف درسی نشان می‌دهند. به منظور مقایسه و ارزیابی داوطلبان لازم است برای هر داوطلب نمره‌ای واحد به نام شاخص نمرهٔ کل تعریف شود. نمرهٔ کل داوطلبان در سازمان سنجش آموزش کشور طی فرآیند سه مرحله‌ای محاسبهٔ نمرهٔ خام و سپس محاسبهٔ نمرهٔ تراز برای تمام درس‌ها و در نهایت محاسبهٔ نمرهٔ کل انجام می‌شود. برای محاسبهٔ نمرهٔ خام، تعداد پاسخ‌های غلط از سه برابر تعداد پاسخ‌های صحیح کم شده و حاصل در سه برابر تعداد سؤال‌های آن درس تقسیم می‌شود و در نهایت نتیجه در عدد صد ضرب می‌شود. برای تعیین نمرهٔ تراز کافی است توزیع تجربی نمره‌های خام هر درس را به دست آورده و با فرض نرمال بودن توزیع نمره‌های خام، چندک‌های متناظر این نمره‌ها (Z_{ij}) از توزیع نرمال استخراج شود. لازم به ذکر است که با توجه به ویژگی چندک‌ها، در همهٔ درس‌ها تعیین نمره‌های Z_{ij} حداکثر در بازه $(-5, +5)$ قرار می‌گیرد. یعنی نمرهٔ همهٔ داوطلبان در محدودهٔ بسیار نزدیک به هم و به صورت اعشاری پخش هستند. با هدف توزیع نمره‌ها در بازه وسیع‌تر و از بین بردن اعشار از تبدیل خطی زیر استفاده می‌شود:

$$NT_{ij} = 2250\, Z_{ij} + 5000$$

که در رابطهٔ بالا اندیس‌های i و j به ترتیب بیانگر شمارهٔ آزمودنی i در درس j است. در انتها نمرهٔ کل با گرفتن میانگین وزنی نمره‌های تراز شدهٔ هر فرد در تمام خرده‌آزمون‌ها حاصل می‌شود (گروه پژوهشی آمار کاربردی، ۱۳۹۴).

بررسی مبانی نظری نشان می‌دهد که یکی از روش‌های ارزیابی، طبقه‌بندی[1] است. مسائل طبقه‌بندی به شناسایی خصوصیاتی منجر می‌شود که مشخص می‌کند که هر مورد به کدام گروه تعلق دارد. این الگو، هم می‌تواند برای فهم داده‌های موجود و هم برای پیش‌بینی اینکه هر نمونهٔ جدید چگونه کار می‌کند استفاده شود. طبقه‌بندی در واقع ارزیابی ویژگی‌های مجموعه‌ای از داده‌ها و سپس اختصاص آنها به مجموعه‌ای از گروه‌های از پیش تعریف شده است. هر یک از الگوریتم‌های موجود طبقه‌بندی، برای بررسی داده‌ها و یافتن قوانین طبقه‌بندی، رویکردها و روش‌های متفاوتی دارند. این روش‌ها را می‌توان به سه نوع الگوریتم خطی، غیرخطی و مبتنی بر قاعده تقسیم کرد. الگوهای خطی و غیرخطی به طور عمده از رویکرد ریاضی پیروی می‌کنند و توابعی را برای امتیازدهی و جداسازی داده‌ها و کلاس‌ها ایجاد می‌کنند. با این تفاوت که

1. Classification

روش‌های خطی، کلاس‌های مختلف را با توابع خطی از هم مشـخص می‌کند ولی در روش‌های غیرخطی از توابع پیچیده‌تری اسـتفاده می‌شود. برخلاف رویکردهای ریاضی، خروجی الگوهای مبتنی بر قانون به صورت قوانین "اگر .... آنگاه" است. الگوریتم‌های مهم طبقه‌بندی را می‌توان به صورت جدول (۱) فهرست کرد (شاهیری و حوسین[1]، ۲۰۱۵).

جدول ۱. الگوریتم‌های رایج طبقه‌بندی

مبتنی بر قانون	غیرخطی	خطی
درخت تصمیم	ماشین بردار پشتیبان[2]	رگرسیون لوجستیک
CART	شبکه‌های عصبی	تحلیل تشخیص خطی
C4.5, C5	شبکه‌های بیزی	-
CHAID	نزدیک‌ترین همسایه[3] (KNN)	-
QUEST	-	-

از آنجا که تعداد متغیرهای مورد اسـتفاده در این پژوهش زیاد اسـت امـکان طبقه‌بندی صحیح داوطلبان با اسـتفاده روش‌های خطی مانند رگرسـیون لوجستیک ناممکن اسـت لذا روش‌های غیرخطی منجر به دقـت بالاتـری در طبقه‌بندی داوطلبان می‌شـود. مطالعات انجام شـده بیانگر ظرفیـت درخت تصمیم و الگوریتم‌های شـبکهٔ عصبی برای پیش‌بینی عملکرد تحصیلی اسـت. لذا در این مطالعه از این الگوریتم‌ها برای طبقه‌بندی داوطلبان کنکورسراسری استفاده شده است که سه الگوریتم مهم درخت تصمیم و الگوی شـبکه عصبی پرسـپترون برای این پژوهش در نظر گرفته شـد که در ادامه مبانی نظری هر یک از روش‌ها ارائه می‌شوند.

درخت تصمیم

اسـتفاده از درخت تصمیم به دلیل سادگی در اسـتفاده و فهم، قابلیت برای کار با داده‌های بزرگ، قابلیت ترکیب با روش‌های دیگر و همچنین اسـتفادهٔ مجدد آسـان از درخت تصمیم سـاخته شده مورد استقبال پژوهشـگران بسـیاری قرار گرفته اسـت. از فنّ درخت تصمیم هم می‌توان برای کشف و استخراج دانش از هر پایگاه داده و هم برای ایجاد الگوهای پیش‌بینی اسـتفاده کرد (کاظمی‌نژاد، اعتمادشهیدی و موسوی، ۲۰۰۵). بـه عبارت دیگر درخت تصمیم یکی از ابزارهای متداول برای دسـته‌بندی و پیش‌بینی اسـت که

1. Shahiri & Husain
2. Support Vector Machine
3. K-Nearest Neighbor

قابلیت تولید قانون را نیز دارد (شـهرابی و ذوالقدرشـجاعی، ۱۳۹۰). درخت تصمیم براساس اینکه متغیر پاسخ پیوسته یا گسسته باشد به دو دسته رگرسیونی و طبقه‌بندی تقسیم می‌شود (طالبی و اکبری، ۱۳۹۲). درخت تصمیم اغلب متشـکل از چندین گره در سـه بخش اصلی برگ[1]، ریشـه[2] و شـاخه[3] است که با نام گره‌هـای ورودی و خروجی شـناخته می‌شـود. به عبارت دقیق‌تر برگ‌ها گره‌هایی اسـت که تقسـیم‌های متوالی در آنجا پایان می‌یابد و بیانگر یک کلاس هستند. به گره اول هر درخت تصمیم ریشه گفته می‌شود. در هـر گره داخلی به تعداد جواب‌های ممکن تقسـیم صورت می‌گیرد و در نتیجهٔ این تصمیمات در بخش میانی درخت شـاخه ایجاد می‌شـود که هر شـاخه بیانگر ارزش ویژگی مورد نظر است (میچل[4]، ۱۹۹۷). قوانین ایجاد شـده در درخت تصمیم به صورت اگر و آنگاه بیان می‌شـود. همچنین در هر درخت تصمیم، مهم اسـت که کدام یک از ویژگی‌ها (یا ابعاد) در سـطوح بالاتری از درخت قرار گیرد تا طبقه‌بندی بهتری انجام شـود. لذا ویژگی که توزیع دسـته‌ها در گره‌های حاصل از آن همگن باشـد نسـبت به سایر ویژگی‌ها برتری دارد. منظور از همگن بودن گره این است که همهٔ رکوردهای موجود در آن متعلق به یک دستهٔ خاص باشد و چون در آن صورت آن گره به برگ تبدیل می‌شود بنابر این گره‌ای به عنوان گره همگن در نظر گرفته می‌شـود که کم‌ترین میزان ناخالصی را داشـته باشد. دو روش مهم برای محاسـبهٔ ناخالصی گره شاخص جینی و آنتروپی اسـت (صنیعی‌آباده، محمودی، طاهرپور، ۱۳۹۱). الگوریتم‌های مختلفی برای سـاخت درخت تصمیم توسـعه داده شـده که از آن جملـه می‌توان به ID3 (کوییـنلـن[5]، ۱۹۸۶)، C4.5(کوییـنلن، ۱۹۹۳)، CART[6] (بریمان[7] و همکاران، ۱۹۸۴)، CHAID[8] (کاس[9]، ۱۹۸۱) و QUEST[10] (لوح و شـیح[11]، ۱۹۹۷) اشاره کرد (شهرابی و ذوالقدرشجاعی، ۱۳۹۰).

درخت تصمیم در الگوریتم‌های C4.5 ،ID3 و C5 براسـاس مفهوم بی‌نظمی ساخته می‌شود. به این معنی که الگوریتم درصدد آن اسـت که میزان بی‌نظمی در گره‌های بالایی درخت حداقل باشـد. بنابر این برای تمام ویژگی‌های داده‌های اولیه، بی‌نظمی با اسـتفاده از آنتروپی شـانون محاسبه شده و سپس آن ویژگی که بیشـترین سودمندی را دارد به عنوان ریشـه انتخاب می‌شود (شهرابی و ذوالقدرشجاعی، ۱۳۹۰). تنها تفاوتـی که این الگوریتم‌ها با هم دارند این اسـت که در الگوریتم ID3 فقط قانون تقسـیم وجود دارد ولی الگوریتم C4.5، قانون هَرَس نیز دارد. بدین معنی که بعد از سـاخت درخت، تعدادی از شـاخه‌ها به دلیل

1. leaf
2. root
3. branch
4. Mitchell
5. Quinlan
6. Classification and Regression Tree
7. Breiman
8. Chi-squared Automatic Interaction Detector
9. Kass
10. Quick Unbased Efficient Statistical Tree
11. Loh and Shih

وجود نقـاط دور افتاده در داده‌های آموزشـی، ناهنجاری‌هایی ایجاد می‌کنـد. در روش هَرَس کردن با به کارگیـری معیارهای آماری، شـاخه‌هایی با اطمینان کمتر حذف می‌شـود که این امـر منجر به رده‌بندی سـریع‌تر و توانایی بیشـتر درخت بـرای رده‌بندی صحیح داده‌های آزمون می‌شـود (مشـکانی و ناظمی، ۱۳۸۸). بـا الگوریتم C5 می‌توان مجموعه داده‌ها را به بیش از دو قسـمت تقسـیم کرد که امکان هرس را پس از رشـد درخت فراهم می‌کند و قوانین تقسـیم گره‌ها با اندازه‌گیری ناخالصی انتخاب می‌شود (وندلر و گرات‌آپ[1]، ۲۰۲۱).

در هر سـه الگوریتم ID3،C4.5 وC5 برای انتخاب مهم‌ترین خصیصه از شاخص آنتروپی استفاده می‌شود کـه معیـاری عددی از میزان اطلاعـات یا میزان تصادفی بودن هر متغیر تصادفی اسـت کـه به هر یک از ملاک‌ها وزن منحصر به فردی می‌دهد و به صورت زیر محاسبه می‌شود:

$$I(n) = -\sum_{i=1}^{c} p(i|n) log(i|n)$$

که در آن $p(i|n)$ کسـری از نمونه‌های متعلق به کلاس i اسـت که در گره n واقع اسـت. بنابر این طبق رابطهٔ آنتروپی می‌توان نتیجه گرفت که آن ویژگی در بالای درخت قرار می‌گیرد که بیشـترین بار اطلاعاتی را داشته باشد و باعث کاهش آنتروپی یا بی‌نظمی شود که بدین منظور بهرهٔ اطلاعات[2] محاسبه می‌شود.

منظور از بهرهٔ اطلاعات مربوط به هر ویژگی، مقدار کاهش آنتروپی اسـت که به واسـطهٔ جداسازی داده‌ها براسـاس این ویژگی، حاصل می‌شـود. به عبارت دیگـر بهرهٔ اطلاعات برای یک ویژگی نظیر A نسـبت به مجموعه داده‌هایS طبق معادلهٔ زیر محاسبه می‌شود:

$$Gain(A) = I(n) - I_{res}$$

که در آن؛

$$I_{res} = -\sum_{v} P(v) \sum_{c} P(c|v) log P(c|v)$$

بنابـر ایـن برای یافتن مهم‌تریـن ویژگی در مجموعهٔ داده‌ها، ابتدا لازم اسـت آنتروپی کل براسـاس هدف طبقه‌بندی حسـاب شود و سـپس برای هر یک از متغیرها، بهرهٔ اطلاعات محاسبه شود و متغیری که بهرهٔ اطلاعات آن از بقیهٔ متغیرها بیشتر است در ریشه قرار داده شود (شهرابی و ذوالقدرشجاعی، ۱۳۹۰).

CART یک درخت باینری اسـت. این بدان معناسـت که هر گره بدون برگ دقیقاً دو شـاخه دارد. در این

1. Wendler, T., & Gröttrup
2. Information gain

الگوریتم برای ساخت درخت از ضریب جینی[1] استفاده می‌شود که معیار برای ناخالصی است و پراکندگی تقسیم را توصیف می‌کند. به عبارت دیگر ضریب جینی برابر با واریانس متغیر تصادفی برنولی حاصل از استخراج نمونه‌ای تصادفی (با جای‌گذاری) از این گره و مشاهدهٔ کلاس آن می‌باشد. بنابر این می‌توان کاهش ناخالصی را به عنوان کاهش وایانس در نظر گرفت و هدف به دست‌آوردن گره‌هایی با واریانس صفر در برچسب کلاس می‌باشد (مشکانی و ناظمی، ۱۳۸۸). به زبان ریاضی ضریب جینی در گره $\sigma\sigma$ به صورت زیر محاسبه می‌شود:

$$Gini(\sigma) = 1 - \sum_{j} \left(\frac{N(\sigma, j)}{N(\sigma)} \right)^2$$

که در آن، j طبقه‌ای از متغیر هدف است، $N(\sigma, j)$ تعداد داده‌هایی است که در طبقهٔ j گره σ قرار دارد و $N(\sigma)$ بیانگر تعداد کل داده‌ها در گره σ است. بهره ضریب جینی در این تقسیم به صورت زیر محاسبه می‌شود:

$$GiniGain(\sigma, s) = Gini(\sigma) - \frac{N(\sigma_L)}{N(\sigma)} Gini(\sigma_L) - \frac{N(\sigma_R)}{N(\sigma)} Gini(\sigma_R)$$

که در آن σ_R و σ_L دو گره فرزند σ هستند و s نیز معیار تقسیم است و تقسیم دوتایی که بهرهٔ ضریب جینی آن ماکزیمم باشد انتخاب می‌شود.

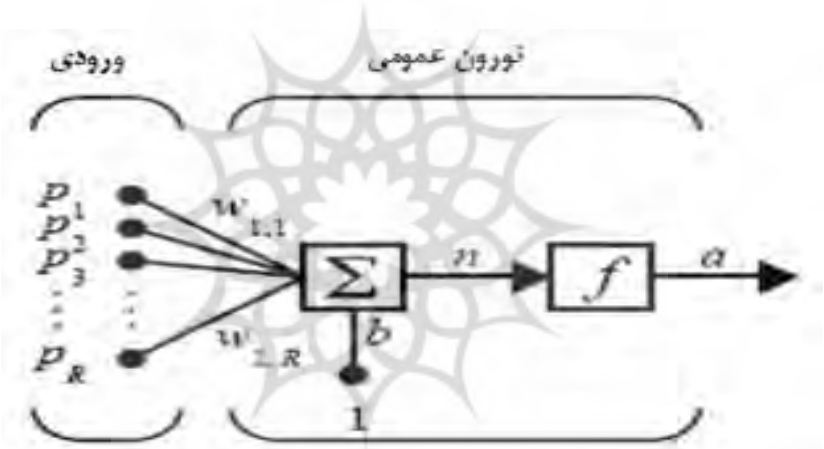
در الگوریتم CHAID مجموعه داده‌ها را می‌توان به بیش از دو گروه تقسیم کرد. در این الگوریتم، هَرَس انجام نمی‌شود. در عوض برای تصمیم‌گیری در مورد قانون تقسیم برای هر گره، از آزمون استقلال خی‌دو استفاده می‌شود. برای هر متغیر ورودی، کلاس‌ها براساس شباهت آماری خود در یک ابرکلاس[2] ادغام می‌شوند و در صورت عدم تفاوت آماری حفظ می‌شوند. سپس این متغیرهای ابرکلاس، از نظر میزان وابستگی (شباهت) با متغیر هدف با استفاده از آزمون خی‌دو مقایسه می‌شوند. سپس گره‌ای که بیشترین معنی‌داری را دارد به عنوان معیار تقسیم برای گره انتخاب می‌شود (وندلر و گرات‌آپ، ۲۰۲۱).

مزیت‌های درخت تصمیم عبارت‌اند از: ۱) نسبت به نقاط پرت نیرومند[3] است. ۲) نمایش راه حل در قالب رَوندنما (فلوچارت) است و بنابر این درک الگو و قوانین تصمیم‌گیری با استفاده از درخت تصمیم آسان است. ۳) می‌تواند با انواع مختلف داده، یعنی متغیرهای عددی و دسته‌ای، کار کند. ۴) در پش‌بینی سریع است. ۵) هیچ پیش‌فرضی در مورد توزیع متغیرها ندارد، بنابر این آماده‌سازی داده‌ها ساده‌تر است. ۶) می‌تواند مقادیر گمشده را مدیریت کند (وندلر و گرات‌آپ، ۲۰۲۱؛ وانگ و لی[4]، ۲۰۰۶).

1. Gini
2. super-class
3. Robust
4. Wang & Lee

شبکهٔ عصبی

شـبکهٔ عصبی مصنوعی شاخه‌ای از هوش مصنوعی و روشی مناسب برای تشخیص الگوهای ناشناخته در داده‌ها است. منشأ الهام این ابزارها، سامانه‌های زیستی و طبیعی بدن انسان هستند. خروجی مطلوب در شـبکهٔ عصبی مصنوعی نیز مانند شبکهٔ عصبی انسـان طی دو مرحله آموزش و یادگیری حاصل می‌شود. شـبکه‌های عصبی از مجموعه‌ای از لایه‌ها شامل نورون‌ها تشکیل شده‌اند که به صورت موازی با هم عمل می‌کنند و اسـاس عملکرد شبکه‌های عصبی را تشـکیل می‌دهند. به عبارت دقیق‌تر شبکهٔ عصبی شامل سـه لایهٔ ورودی، پنهان یا میانی و خروجی اسـت. لایهٔ ورودی برای معرفی پارامترها به شـبکه اسـت و لایهٔ انتقال‌دهنده محسوب می‌شود. لایهٔ پنهان بین لایهٔ ورودی و خروجی است و نقش پردازشگری اطلاعات را برعهده دارد. لایهٔ خروجی محل قرارگیری مقادیر خروجی شبکه است. شکل (۱) نورونی ساده با R ورودی را نشان می‌دهد.



شکل ۱. نمونه‌ای از شبکهٔ عصبی تک‌لایه (دموت و بیل[1]، ۲۰۰۰)

خروجی نورون با معادلهٔ زیر تعیین می‌شود:

$$a = f\left(\sum_i w_i p_i + b\right)$$

کـه در ایـن رابطه b مقدار ثابت اسـت و به عنوان مؤلفه مرز تشـخیص اسـتفاده می‌شـود و می‌تواند نوع طبقه‌بنـدی داده‌هـا را تغییر دهد. بنابر این وجود مقدار b در الگو ضروری اسـت. به f تابع محرک یا تابع فعال‌سـازی می‌گوینـد که معمولاً تابعی غیرخطی اسـت. لازم به ذکر اسـت که پارامترهـای b و w قابل تنظیم هسـتند و تابع fهم توسـط پژوهشگر انتخاب می‌شـود. به طور کلی به این سه مرحله در شکل بالا

1. Demuth , & Beale

یعنی ورودی، وزن‌ها و تابع محرک به اصطلاح یک لایه گفته می‌شود که باید آموزش[1] داده شود. آموزش دادن به لایه، به معنی پیدا کردن وزن‌هایی است که ورودی را به خروجی مورد نظر تبدیل کند (هاسون[2]، ۱۹۹۵). بنابر این، شبکهٔ عصبی نوعی روش محاسباتی است که به کمک فرآیند یادگیری و با استفاده از پردازشگرهایی به نام نورون (که هر نورون بیانگر یک ویژگی است) تلاش می‌کند روابط درونی بین داده‌ها را شناسایی کند و نگاشتی بین فضای ورودی (لایه ورودی) و فضای مطلوب (لایه خروجی) داده‌ها ایجاد کند. بدین منظور فرآیند یادگیری را لایه به لایه انجام می‌دهد تا در نهایت اختلاف بین مقادیر مطلوب و مقادیر پیش‌بینی شدهٔ شبکه به حداقل ممکن برسد (هان و همکاران، ۲۰۱۱).

از طرفی نتیجهٔ (خروجی) هر شبکهٔ عصبی به صورت باینری ارائه می‌شود و بنابر این اساس کار هر نورون در شبکه‌های عصبی، ترکیب گزاره‌های منطقی و ایجاد گزاره‌ای مرکب است که ترکیب عطفی و فصلی آن در شبکه‌های عصبی با یک نورون و ترکیب فصلی ضمنی برای شبکه‌های عصبی با بیش از یک نورون می‌تواند طبقه‌بندی را انجام دهد. در جدول (۲) ترکیب‌های عبارت‌ها به صورت ریاضی نشان داده شده است. برای توضیح بیشتر در مورد گزاره‌های منطقی، دو متغیر ورودی X و Y را فرض کنید که هر یک شامل ویژگی‌هایی هستند، در صورتی نتیجهٔ حاصل از ترکیب عطفی درست خواهد بود اگر و فقط اگر هر دو عبارت سادهٔ تشکیل‌دهنده آن درست باشد و نتیجهٔ حاصل از ترکیب فصلی، درست خواهد بود اگر و فقط اگر دست‌کم یکی از عبارت‌های تشکیل‌دهندهٔ آن درست باشد و در آخر نتیجهٔ حاصل از ترکیب فصلی ضمنی درست خواهد بود اگر و تنها اگر دو عبارت سادهٔ تشکیل‌دهندهٔ آن مانند هم نباشند.

جدول ۲. گزاره‌های منطقی در ریاضیات

عبارت ساده اول	عبارت ساده دوم	ترکیب عطفی	ترکیب فصلی	ترکیب فصلی ضمنی
X	Y	$X \wedge Y$	$X \vee Y$	XOR
۰	۰	۰	۰	۰
۰	۱	۰	۱	۱
۱	۰	۰	۱	۱
۱	۱	۱	۱	۰

از آنجا که همیشه با یک نورون نمی‌توان به جواب رسید، شبکه‌های عصبی پیشرفته‌تری مطرح شدند. به طور کلی دو نوع شبکهٔ عصبی پیشخور[3] (FFN) و بازگشتی[4] (RNN) وجود دارد که تفاوت آنها در این است

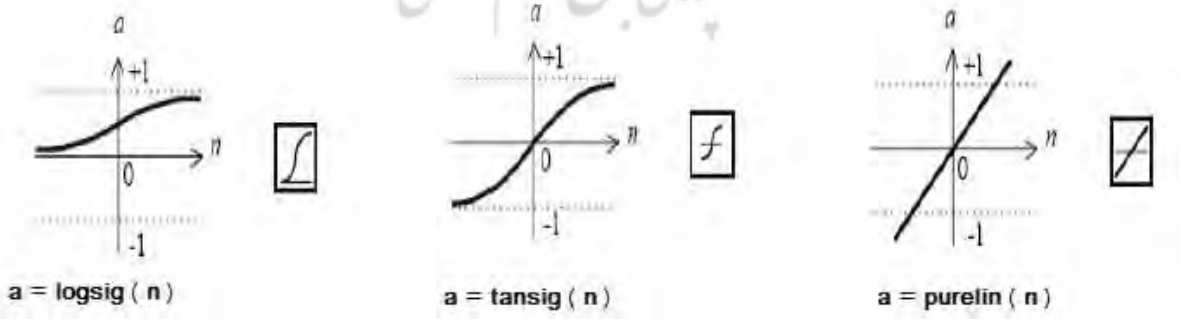
1. train
2. Hassoun
3. Feed Forward Network
4. Recurrent Neural Network

که در شبکه‌های عصبی بازگشتی، دست‌کم یک سیگنال برگشتی از یک نورون به همان نورون یا نورون‌های همان لایه و یا لایهٔ قبل وجود دارد (منهاج، ۱۳۹۷).

یکی از شبکه‌های عصبی پیشخور، شبکه عصبی پرسپترون[1] است که قادر است با انتخاب مناسب تعداد لایه‌ها و سلول‌های عصبی، نگاشتی غیرخطی با دقت دلخواه ایجاد کند. شبکه‌های پرسپترون به صورت تک‌لایه و چندلایه ارائه می‌شود. شبکه‌های پرسپترون چندلایه، شبکه‌هایی کاملاً مرتبط هستند، بدین معنی که هر نورون در هر لایه، به تمام نورون‌های لایهٔ قبل متصل است (منهاج، ۱۳۹۷). شکل (۲) مثالی از یک شبکه پرسپترون با سه لایهٔ پنهان است که در هر یک از این لایه‌ها چهار نورون وجود دارد. در لایهٔ ورودی آن، سه نورون و یک نورون در لایهٔ خروجی وجود دارد.

شکل ۲. نمونه‌ای از شبکه عصبی پرسپترون با سه لایه میانی

تابع f که برای افزایش قابلیت شبکه اضافه می‌شود، توابع مختلفی دارد که از متداول‌ترین آنها در شبکهٔ عصبی پرسپترون چندلایه می‌توان به توابع لگاریتم سیگموئیدی، تانژانت سیگموئیدی و تابع محرک خطی اشاره کرد (دموت و بیل[2]، ۲۰۰۰).



شکل ۳. توابع محرک رایج در شبکه عصبی پرسپترون چندلایه

1. Demuth & Beale
2. Perceptron

در همهٔ شـبکه‌های عصبی دو مسـئلهٔ انتخاب، یکی در مورد معماری شـبکه و دیگـری در مورد الگوریتم آموزشـی مناسب مطرح است. منظور از معماری شـبکه، انتخاب بهینهٔ تعداد لایه‌ها، تعداد نورون‌های هر لایه و نوع تابع فعال‌سازی هر نورون است و الگوریتم آموزشی مناسب شبکه‌های عصبی مبتنی بر مجموعه داده‌ها و ویژگی‌های آنان است.

مهم‌تریـن انتخاب ویژگی هر شـبکهٔ عصبـی، توانایی یادگیری اسـت. یعنی براسـاس الگوریتم یادگیری مشخصی، وزن‌ها تغییر کنند تا حدی که میزان اتلاف شبکه کمینه (مینیمم) شود. برای یادگیری در شبکهٔ پرسـپترون از قاعدهٔ پس انتشـار خطا[1] استفاده می‌شود. در واقع پس انتشـار خطا با مقایسهٔ بین خروجی مشـاهده شـده (واقعی) و خروجی مطلوب، ضرایب وزنی شـبکه را تغییر می‌دهد به طوری که در دفعات بعدی، خروجی صحیح‌تری حاصل شـود. در این روش برای هر الگو، مجموع مربعات خطا طبق رابطهٔ زیر محاسبه می‌شود:

$$E_p = \frac{1}{2}\sum_k (t_{pk} - o_{pk})^2$$

که در آن t_{pk} و o_{pk} به ترتیب بیانگر مقادیر مطلوب و مقادیر مشـاهده شـده هستند (مشکانی و ناظمی، ۱۳۸۸). پس از محاسـبهٔ خطا، مقدار آن به صورت معکوس (از لایهٔ خروجی به سـمت لایهٔ اول) در شـبکه انتشـار می‌یابد. سـپس مقدار هر وزن در جهت کاهش خطا تغییر می‌کند. مهم‌ترین بخش الگوریتم پس انتشار در آموزش شبکه عصبی پرسپترون چندلایه، به روز رسانی وزن‌های شبکه است (منهاج، ۱۳۹۷).

با توجه به اهمیت ارزیابی و تصمیم‌گیری دقیق دربارهٔ عملکرد داوطلبان آزمون سراسـری، در این مطالعه، قابلیت‌های سـه الگوریتم مختلف درخت تصمیم (CHAID، C5 و CART) و الگوی شبکهٔ عصبی مقایسه شـده اسـت. بدین منظور براساس مبانی نظری مطرح شـده در پیشـینهٔ پژوهش از متغیرهای جنسیت، سـوابق تحصیلی (معدل کتبی دیپلم و فاصله دریافت دیپلم تا شـرکت در آزمون سراسری) و نمره‌های تراز هر یک از درس‌ها به عنوان متغیرهای مؤثر در طبقه‌بندی و پیش‌بینی سـطح عملکرد داوطلبان اسـتفاده شد و بر این اساس به سؤال‌های زیر پاسخ داده می‌شود:

۱. مناسـب‌ترین الگوی طبقه‌بندی داوطلبان آزمون سراسری براساس همهٔ درس‌ها با احتساب متغیرهای سوابق تحصیلی کدام است؟

۲. مناسـب‌ترین الگوی طبقه‌بندی داوطلبان آزمون سراسری براساس همهٔ درس‌ها با کنترل سایر متغیرها کدام است؟

لازم به ذکر اسـت ایدهٔ نویسـندگان برای افزودن متغیـر فاصله دریافت دیپلم تا کنکور سراسـری به فرآیند

1. backpropagation

طبقه‌بندی این بود که هر چه داوطلب جوان‌تر باشد تا زمان کنکور خود فاصله کمتری دارد لذا آمادگی بیشتری برای کنکور دارد.

روش پژوهش

با توجه به اینکه در این مطالعه از داده‌های آزمون سراسری از پایگاه داده‌های سازمان سنجش آموزش کشور استفاده شده است، پژوهش از جنبهٔ گردآوری داده‌ها جزو مطالعات توصیفی و از جنبهٔ تحلیلی از نوع تحلیل ثانویه به شمار می‌آید. داده‌های این پژوهش متعلق به همهٔ فارغ‌التحصیلان نظام جدید متوسطهٔ ساکن استان البرز به تعداد ۸۳۰۵ نفر است که در گروه آزمایشی علوم ریاضی و فنی در آزمون سراسری شرکت کرده‌اند.

با توجه به تعداد زیاد متغیرهای این پژوهش، طبقه‌بندی داوطلبان به صورت غیرخطی دقیق‌تر می‌شود لذا در این پژوهش از دو روش درخت تصمیم و شبکهٔ عصبی برای طبقه‌بندی داوطلبان گروه ریاضی و فنی استفاده شد. به منظور استفاده و تحلیل بهتر از درخت تصمیم لازم است که همهٔ متغیرهای کمّی به متغیرهای طبقه‌بندی شده با طبقه‌های چهارتایی براساس چارك اول و دوم و سوم تبدیل شوند. همچنین برای به کارگیری درست شبکهٔ عصبی لازم است که مقیاس داده‌ها یکسان شود لذا پیش از استفاده، همهٔ متغیرهای کمّی استاندارد شدند.

براساس مبانی نظری مطرح شده در پیشینهٔ پژوهش از متغیرهای جنسیت، سوابق تحصیلی (معدل کتبی دیپلم و فاصلهٔ دریافت دیپلم تا شرکت در آزمون سراسری)، سهمیهٔ نهایی (شامل مناطق یک، دو و سه و شاهد) و نمره‌های تراز هر یک از درس‌ها به عنوان متغیرهای مؤثر در طبقه‌بندی و پیش‌بینی سطح عملکرد داوطلبان استفاده شد. قبل از استفاده از متغیرهای جنسیت و سهمیه که کیفی بودند تأثیر آنها در نمرهٔ کل داوطلبان با استفاده از آزمون تی(t) مستقل و تحلیل واریانس بررسی شد. نتایج نشان داد که دخترها و پسرها در نمرهٔ کل تفاوت معناداری با یکدیگر دارند که با مبانی نظری مطرح شده مطابقت دارد ولی بین داوطلبان سهمیه‌های مختلف. تفاوت معناداری وجود ندارد لذا تنها متغیر جنسیت وارد الگو شد. هدف تحقیق حاضر طبقه‌بندی نامزدها براساس سطح عملکرد آنهاست. برای دستیابی به این هدف، نمره کل به عنوان معیاری برای ارزیابی سطح عملکرد عمل می‌کند و بنابر این متغیر هدف تلقی می‌شود. علاوه بر این، برای اهداف طبقه‌بندی، ضروری است که متغیر پاسخ به دو یا چند کلاس مجزا تقسیم شود که هر یک به طور مناسب برچسب‌گذاری شده باشند. لذا چارک‌های نمره‌های کل نهایی داوطلبان استان البرز محاسبه و مطابق جدول (۳) به داده‌های مربوط به نمره کل چهار برچسب ضعیف، متوسط، خوب و خیلی خوب داده شد. از طرفی با توجه به ماهیت روش طبقه‌بندی که در دو مرحلهٔ آموزش و آزمون انجام می‌شود

در این پژوهش هشتاد درصد مجموعهٔ داده‌ها به عنوان داده‌های آموزشی و بیست درصد داده‌ها به عنوان دادهٔ آزمون در نظر گرفته شـد تا قابلیت اعتبار الگو سـنجیده شود. در واقع با استفاده از داده‌های آزمون، دقت الگویی که آموزش داده شده است سنجیده می‌شود.

جدول ۳. وضعیت نمره‌های کل براساس چارک‌های استان البرز

کمتر از ۳۸۶۲	بین ۳۸۶۲ و ۴۷۹۵	بین ۴۷۹۵ و ۶۱۳۸	بیشتر از ۶۱۳۸
ضعیف (D)	متوسط (C)	خوب (B)	خیلی خوب (A)

لازم بـه ذکر اسـت کـه تعداد لایه‌ها و تعـداد نورون‌ها در هر لایهٔ پنهان در شـبکهٔ عصبی بـا آزمون و خطا مشخص می‌شود (پاتل و میستری[1]، ۲۰۱۵).

یافته‌ها

با اسـتفاده از نرم‌افزار "*SPSS Modeler*" نسـخهٔ ۱۸، به سؤال‌های پژوهش در مورد انتخاب الگوی مناسب طبقه‌بندی پاسـخ داده شـد. با اینکه جنسـیت در نتیجهٔ آزمون تأثیری ندارد، اما آزمون تی «t" نشان داد که بین عملکرد دخترها و پسـرها تفاوت معنی‌داری وجود داشت، نتایج الگوریتم‌های استفاده شده در دو وضعیت با جنسـیت و بدون جنسیت مورد بحث و بررسی قرار می‌گیرد. نتیجهٔ بررسی یافته‌ها در سؤال اول پژوهـش مبنی بر تعیین دقت کلی الگوهای شـبکهٔ عصبی از نوع پرسـپترون چندلایـه و درخت تصمیم با الگوریتم‌های CHAID، C5 و CART در طبقه‌بندی داوطلبان آزمون سراسـری، در جدول (۴) آورده شده اسـت. با توجـه به نتایج موجود در جدول (۴) می‌تـوان نتیجه گرفت که در بین همـهٔ انواع الگوهای مورد مطالعه، الگوی شبکهٔ عصبی از نوع پرسپترون چندلایه، بیشترین دقت (۹۵/۱۵۳ و ۹۵/۶۳۱ و ۹۲/۸۱۹) را در طبقه‌بندی صحیح داوطلبان در هر دو حالت مقایسـهٔ تمام درس‌های تخصصی و عمومی با احتساب متغیرهای دیگر با جنسـیت و بدون جنسـیت و مقایسهٔ تمام درس‌های تخصصی و عمومی دارد. همچنین الگوریتـم "C5" نیز بیشـترین دقـت (۸۰/۰۱۲ و ۸۰/۴۳۱ و ۷۸/۵۱۶) را در بیـن الگوریتم‌های درخت تصمیم دارد. از طرفی دقت شـبکهٔ عصبی، بدون احتسـاب جنسیت، افزایش می‌یابد. به بیان دیگر شبکه عصبی با ورود متغیرهای سـوابق تحصیلی و بدون ورود متغیر جنسـیت، با دقت بیشـتری سطح عملکرد داوطلبان را تعیین می‌کند.

1. Patel & Mistry

جدول ٤. برآورد دقت کلی هر یک از الگوریتم‌ها براساس سؤال‌های پژوهش

ردیف	متغیرهای الگو	شبکه عصبی	C5	CART	CHAID
سؤال ۱	تمام درس‌های تخصصی و عمومی با احتساب متغیرهای دیگر با جنسیت	۹۵/۱۵۳	۸۰/۰۱۲	۶۹/۵۳۹	۶۹/۲۴۰
سؤال ۱	تمام درس‌های تخصصی و عمومی با احتساب متغیرهای دیگر بدون جنسیت	۹۵/۶۳۱	۸۰/۴۳۱	۶۹/۵۳۹	۶۹/۲۴۰
سؤال ۲	تمام درس‌های تخصصی و عمومی	۹۲/۸۱۹	۷۸/۵۱۶	۶۹/۲۴۰	۶۹/۲۴۰

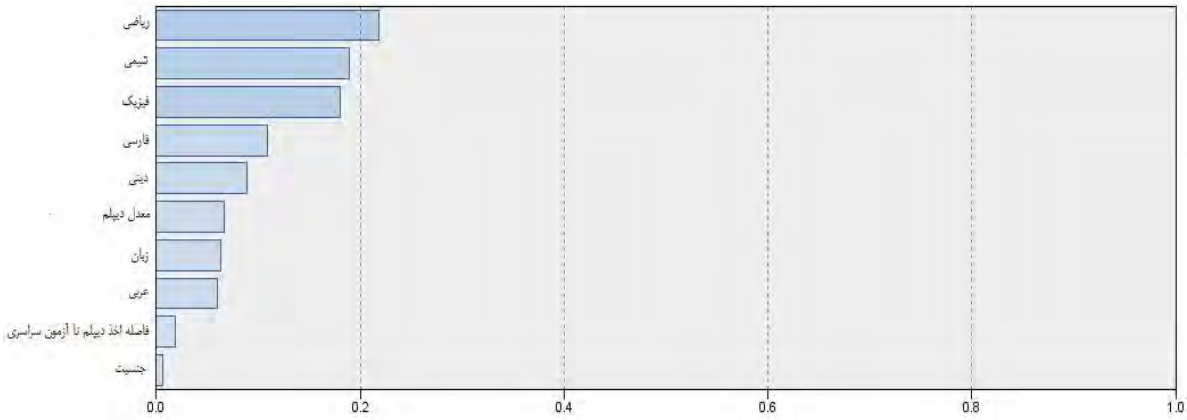
در ادامه با توجه به دقت بالای شبکهٔ عصبی در مقایسه با درخت تصمیم، تنها به ارائهٔ نتایج مربوط به شبکهٔ عصبی اکتفا می‌شود. دقت الگوی شبکهٔ عصبی پرسپترون بر حسب فراوانی و بدون در نظر گرفتن جنسیت در هر یک از مجموعه داده‌های آموزشی و آزمون در جدول (۵) آورده شده است.

جدول ۵. نتایج حاصل از ماتریس اغتشاش شبکهٔ عصبی برای تمام درس‌ها به همراه متغیرهای دیگر بدون جنسیت

	داده‌های آموزش	داده‌های آزمون
درست	۶۳۸۴ (٪۹۶/۲۳)	۱۶۰۷ (٪۹۶/۱۶)
غلط	۲۵۰ (٪۳/۷۷)	۶۴ (٪۳/۸۴)

براساس از جدول (۵) نتیجه گرفته می‌شود که یادگیری شبکه به خوبی انجام گرفته است، به طوری که وقتی داده‌های آزمون به الگو داده شد تنها کمتر از چهار درصد از داده‌های آزمون به غلط پیش‌بینی شد و ۹۶ درصد به درستی در طبقه مناسب قرار داده شد.

در مورد سؤال (۱) پژوهش که تمام درس‌های تخصصی و عمومی به همراه متغیرهای جنسیت، سوابق تحصیلی و فاصلهٔ دریافت دیپلم تا زمان آزمون سراسری در نظر گرفته شد، الگوی شبکهٔ عصبی از نوع پرسپترون با یازده نورون در لایهٔ پنهان حاصل شد. میزان اهمیت متغیرهای استفاده شده در شبکهٔ عصبی با احتساب جنسیت در شکل (۴) مشخص شده است.



شکل ٤. رتبه‌بندی همهٔ متغیرها اعم از درسی و غیردرسی (سؤال اول پژوهش) در شبکهٔ عصبی با احتساب جنسیت

همان‌گونه که مشـاهده می‌شود با استفاده از الگوی شبکهٔ عصبی، نمرهٔ درس ریاضی بیشترین تأثیر را در ارزیابی عملکرد داوطلبان گروه ریاضی و فنی داشته و از سوی دیگر فاصلهٔ دریافت دیپلم تا آزمون سراسری و جنسیت کم‌ترین تأثیر را در ارزیابی عملکرد داشته، به طوری که میزان اهمیت جنسیت در نتیجهٔ عملکرد داوطلب ناچیز است.

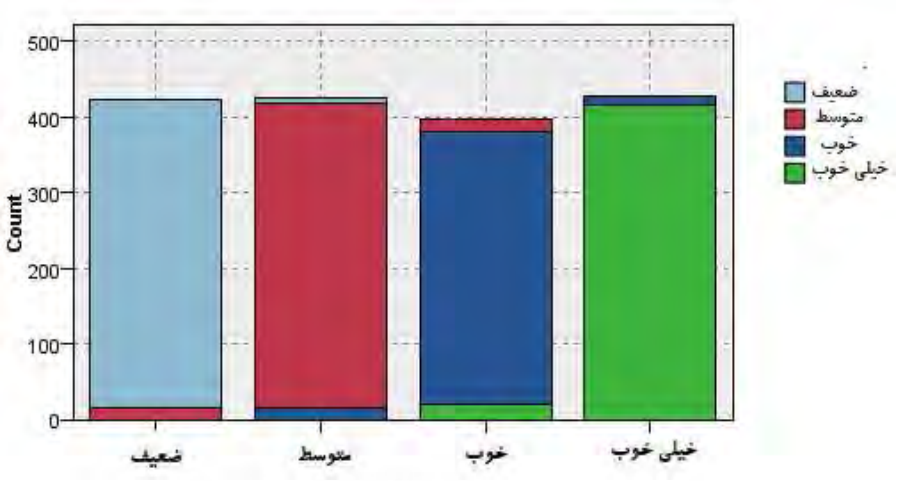
عملکرد شبکهٔ عصبی، توسط ماتریس اغتشاش[1] قابل ارزیابی است. در این ماتریس، سطرها و ستون‌ها به ترتیب بیانگر مقادیر واقعی و مقادیر پیش‌بینی شـده هستند، بدین معنی که هنگامی که عملکرد داوطلب به درسـتی تعیین شده اسـت مقدار یک در سطر و سـتون یکسان مربوط به سـطح عملکرد داوطلب قرار می‌گیرد. به عنوان نمونه، ماتریس اغتشـاش برای شـبکهٔ عصبی که بیانگر میزان پیش‌بینی درست الگوی شـبکهٔ عصبی اسـت در جدول (۶) آورده شده اسـت. با توجه به جدول (۶) مشاهده می‌شود که به عنوان نمونه این شـبکه ۹۷/۴ درصد از داوطلبانی را که در طبقهٔ خیلی خوب قرار داشـتند به درستی پیش‌بینی کرده و تنها ۲/۶ درصد از افراد طبقهٔ خیلی خوب را به اشتباه در طبقهٔ خوب قرار داده است.

جدول ٦. دقت شبکهٔ عصبی پرسپترون در هر یک از طبقه‌ها برای همهٔ متغیرها با احتساب جنسیت

پیش‌بینی شده / مشاهده شده	خیلی خوب	خوب	متوسط	ضعیف
خیلی خوب	٪۹۷/۴	٪۲/۶	٪۰/۰	٪۰/۰
خوب	٪۲/۴	٪۹۳/۹	٪۳/۷	٪۰/۰
متوسط	٪۰/۰	٪۳	٪۹۲/۹	٪۴
ضعیف	٪۰/۰	٪۰/۰	٪۲/۶	٪۹۷/۴

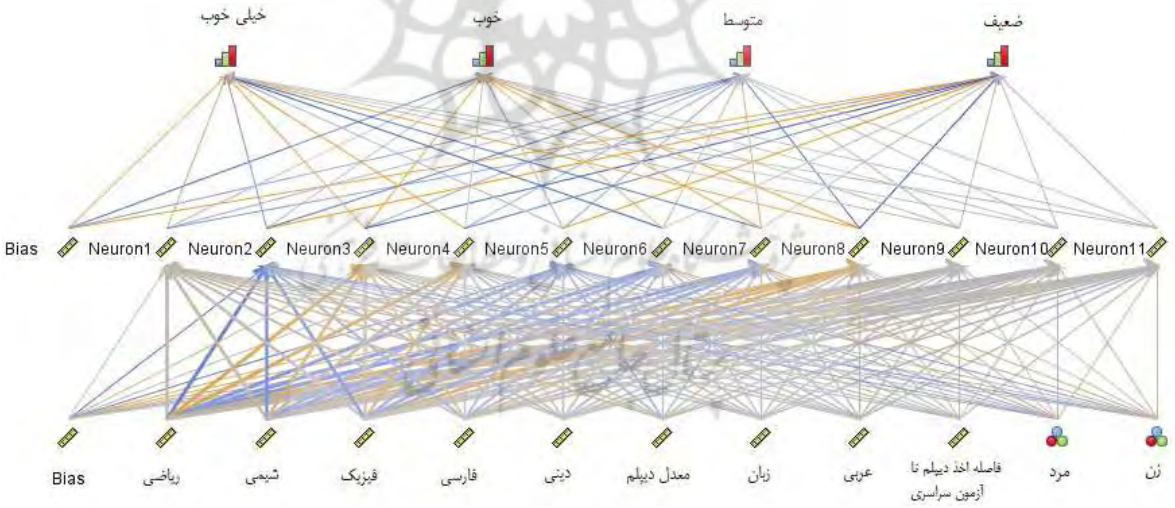
1. Confusion matrix

شکل (۵) نیز تأییدی بر نتایج جدول (۶) است.



شکل ۵. وضعیت پیش‌بینی شبکهٔ عصبی پرسپترون

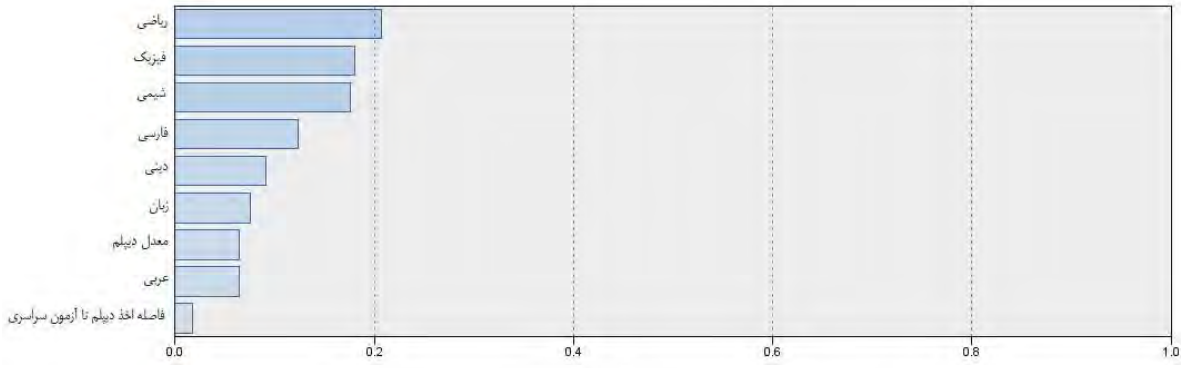
نمودار شبکهٔ عصبی پرسپترون با ۱۱ نورون در شکل (۶) آورده شده است.



شکل ٦. نمودار شبکهٔ عصبی پرسپترون با یازده نورون میانی با احتساب جنسیت

در صورتی که متغیر جنسیت در نظر گرفته نشود تعداد نورون‌های لایهٔ پنهان به نُه نورون کاهش می‌یابد. تنها ترتیب میزان اهمیت متغیرها تغییر می‌کند که این تغییر در واقعیت تا حدی منطقی‌تر است. نتیجهٔ این تغییر در شکل (۷) آورده شده است. می‌توان این گونه نتیجه گرفت که جنسیت در روایی نتیجهٔ آزمون

تأثیر می‌گذارد به طوری که درس‌های تخصصی بیشترین اهمیت را دارند و تأثیر معدل دیپلم از نمره کسب شده عربی در آزمون بیشتر شده است.



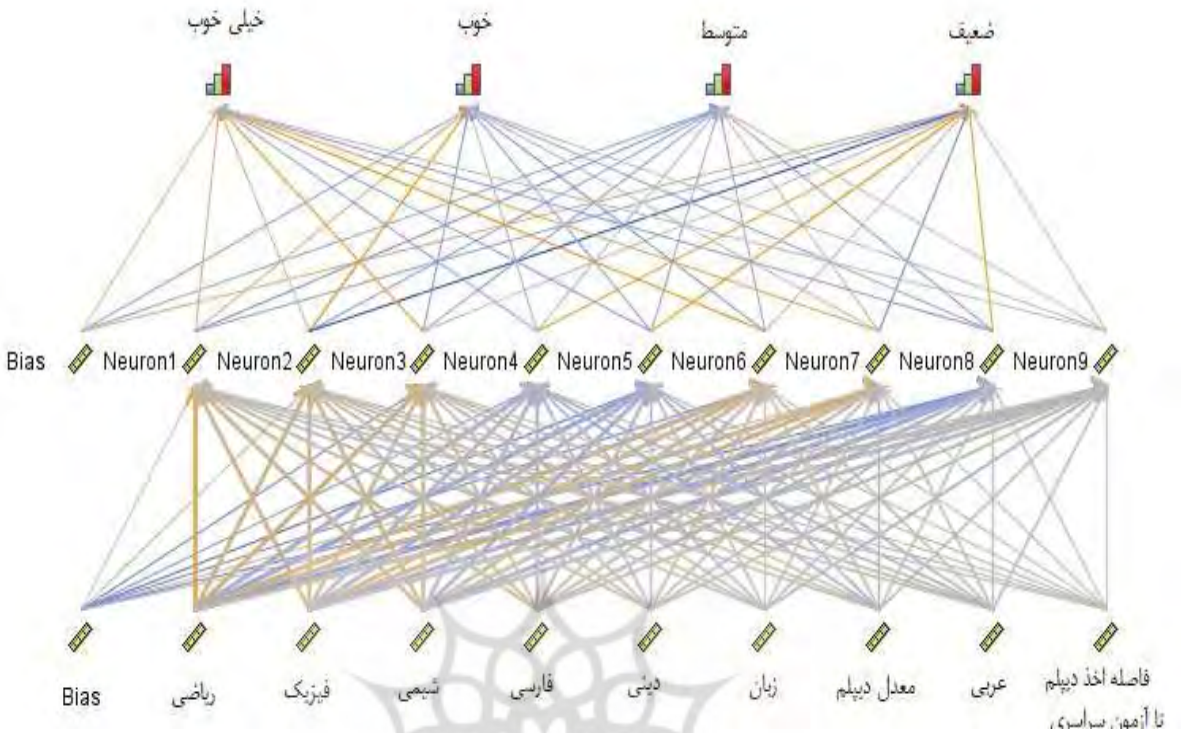
شکل ۷. رتبه‌بندی همهٔ متغیرها اعم از درسی و غیردرسی (سوال اول پژوهش) در شبکهٔ عصبی بدون احتساب جنسیت

ماتریس اغتشاش مربوط به این شبکه عصبی در جدول (۷) آورده شده است.

جدول ۷. دقت شبکهٔ عصبی پرسپترون در هر یک از طبقه‌ها برای همهٔ متغیرها اعم از درسی و غیردرسی بدون جنسیت

پیش‌بینی شده / مشاهده شده	خیلی خوب	خوب	متوسط	ضعیف
خیلی خوب	٪۹۷/۵	٪۲/۵	٪۰/۰	٪۰/۰
خوب	٪۱/۹	٪۹۴/۷	٪۳/۵	٪۰/۰
متوسط	٪۰/۰	٪۲/۹	٪۹۳/۴	٪۳/۷
ضعیف	٪۰/۰	٪۰/۰	٪۳/۲	٪۹۶/۸

با مقایسهٔ جدول‌های (۶) و (۷) به سادگی می‌توان دریافت که دقت شبکهٔ عصبی بدون در نظر گرفتن جنسیت تا حدودی بهبود یافته است.



شکل ۸. نمودار شبکهٔ عصبی پرسپترون با نُه نورون میانی بدون احتساب جنسیت

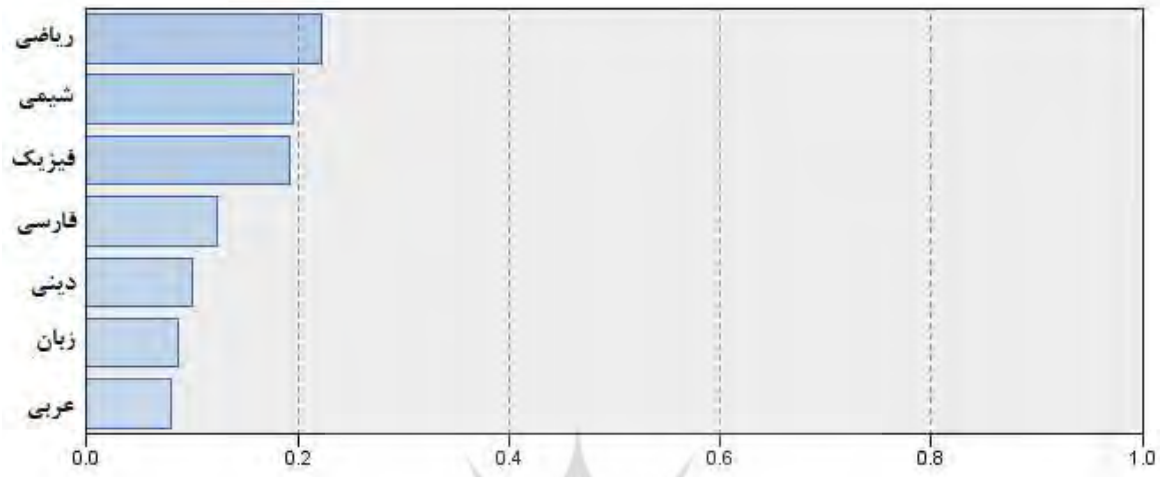
در مورد سـؤال دوم پژوهـش که تنها نمره‌های درس‌هـای عمومی و تخصصی در نظر گرفته شـده‌اند یک الگوی شـبکهٔ عصبی از نوع پرسـپترون با یازده نورون در لایهٔ پنهان حاصل شد. دقت الگوی شبکهٔ عصبی پرسـپترون بر حسـب فراوانی در هر یک از مجموعه داده‌های آموزشـی و آزمون در جدول (۸) آورده شـده است.

جدول ۸. نتایج حاصل از شبکهٔ عصبی برای تمام درس‌ها

	داده‌های آموزش	داده‌های آزمون
درست	۶۱۴۸ (٪۹۲/۶۷)	۱۵۵۱ (٪۹۲/۸۲)
غلط	۴۸۶ (٪۷/۳۳)	۱۲۰ (٪۷/۱۸)

با مقایسـهٔ جدول‌های (۵) و (۸) این گونه استنباط می‌شود که متغیرهای سوابق تحصیلی بدون احتساب جنسیت که در سؤال اول پژوهش حضور دارند در دقت الگو تأثیر به‌سزایی داشته و نیز باعث ساده‌تر شدن

الگو شـده‌اند. میزان اهمیت متغیرهای استفاده شـده (سؤال دوم پژوهش) در شبکهٔ عصبی در شکل (۹) مشخص شده است.



شکل ۹. رتبه‌بندی همهٔ متغیرهای درسی (سؤال دوم پژوهش) در شبکهٔ عصبی

با مقایسـهٔ شـکل‌های (۹) و (۷) می‌توان نتیجه گرفت که متغیرهایی غیـر از نمره‌های درس‌های آزمون (یعنی سـوابق تحصیلی) در ترتیب میزان اهمیت درس‌های فیزیک و شـیمی تأثیر گذاشته‌اند. اما از طرفی میزان تأثیر هر یک از درس‌ها به نسـبت شـکل (۷) افزایش پیدا کرده اسـت. میزان پیش‌بینی درست این الگوی شبکهٔ عصبی در هر یک از طبقه‌های مشخص شده در جدول (۹) آورده شده است.

جدول ۹. دقت شبکهٔ عصبی پرسپترون در هر یک از طبقه‌ها برای تنها متغیرهای درسی (سؤال دوم پژوهش)

پیش‌بینی شده / مشاهده شده	خیلی خوب	خوب	متوسط	ضعیف
خیلی خوب	٪۹۷	٪۳	٪۰/۰	٪۰/۰
خوب	٪۳/۱	٪۸۹/۹	٪۷	٪۰/۰
متوسط	٪۰/۰	٪۴/۲	٪۸۹/۱	٪۶/۷
ضعیف	٪۰/۰	٪۰/۰	٪۵/۳	٪۹۴/۷

از جـدول (۹) می‌تـوان نتیجه گرفت که ۹۷ درصـد از داوطلبانی که به طور واقعـی عملکرد خیلی خوبی داشته‌اند به درستی در طبقهٔ خیلی خوب پیش‌بینی شده‌اند و تنها سه درصد از این داوطلبان، به اشتباه،

در طبقهٔ خوب قرار گرفته‌اند. به طور کلی در این جدول مشاهده می‌شود که بیشتر افراد در محل تقاطع سطر و ستون یکسان (مانند (متوسط ، متوسط)) قرار گرفته‌اند که به معنی پیش‌بینی درست این افراد و عملکرد خوب شبکهٔ عصبی است. با مقایسهٔ جدول‌های (۷) و (۹) می‌توان دریافت که حضور متغیرهای مربوط به سوابق تحصیلی در کنار نمره‌های درس‌های امتحانی منجر به افزایش دقت شبکهٔ عصبی شده است.

بحث و نتیجه‌گیری

آزمون‌ها به‌خصوص آزمون‌های سرنوشت‌ساز مانند آزمون سراسری تأثیر مهمی در سلامت روان جامعهٔ دانش‌آموزان دارند. بنابر این تا حد امکان لازم است هر چه بهتر و دقیق‌تر نتایج این آزمون‌ها ارزیابی شوند تا هر فرد در جایگاه متناسب با توانایی خود قرار گیرد. با توجه به تنوع دانشگاه‌ها (دولتی، غیرانتفاعی، پیام‌نور, علمی کاربردی، آزاد) و همچنین رشته‌های بسیاری که برای ادامهٔ تحصیل دایر شده است، داوطلبان بسیاری به دانشگاه راه می‌یابند. لذا در این پژوهش صرف نظر از قبولی یا عدم قبولی داوطلب با توجه به نمره کل احتسابی، عملکرد داوطلب در یکی از دسته‌های ضعیف، متوسط، خوب و خیلی خوب قرار می‌گیرد و بدین‌ترتیب می‌توان بررسی کرد که چند درصد از داوطلبان هر یک از این دسته‌ها در کدام دانشگاه‌ها و چه رشته‌هایی پذیرفته شده‌اند. بدین منظور لازم است از روش‌های آماری یا داده‌کاوی استفاده شود. یکی از مهم‌ترین کاربردهای روش‌های آماری و داده‌کاوی که در علوم مختلف استفاده می‌شود، شناسایی الگوها و طبقه‌بندی‌هایی است که براساس داده‌های واقعی موجود انجام می‌شود. از جمله روش‌های رایج در آمار برای طبقه‌بندی و پیش‌بینی، رگرسیون لوجستیک است. برای اجرای رگرسیون لوجستیک لازم است پیش‌فرض‌هایی مانند نبودن هم‌خطی چندگانه کامل برقرار باشد که در صورت برقرار نبودن این مفروضه‌ها، استفاده از این روش‌ها امکان‌پذیر نیست و یا با خطای قابل توجهی انجام می‌شود که نمی‌توان به نتایج آن اعتماد کرد. از دیگر محدودیت‌های روش‌های آماری این است که به دادهٔ پرت و داده گمشده حساس هستند. به طور مسلّم با توجه به متغیرهای زیادی که برای ارزیابی داوطلبان در نظر گرفته می‌شود و همچنین تنوع زیاد دانشگاه‌ها و رشته‌ها، تعیین دقیق نتیجهٔ عملکرد داوطلبان با روش‌های آماری مانند رگرسیون لوجستیک قابل اعتماد نیست. به عبارت دیگر خطای طبقه‌بندی در روش‌های آماری مانند رگرسیون لوجستیک بیشتر است. از سوی دیگر در صورتی رگرسیون لوجستیک روش مناسبی به شمار می‌آید که سطوح متغیرهای مستقل، گسسته محدود باشند و بین متغیرها رابطهٔ خطی برقرار باشد. بر این اساس برای اجرای رگرسیون لوجستیک در این پژوهش، لازم بود که متغیرهای مستقل به متغیرهای گسسته تبدیل شوند، که در این صورت اطلاعات زیادی از دست

می‌رفت و نتایج قابل اعتمادی حاصل نمی‌شد (اسکندری، ۱۳۹۵). در جدول (۱) برگرفته از پژوهش شاهیری و حوسین (۲۰۱۵) مؤید آن است که رگرسیون لوجستیک در ردهٔ الگوریتم‌های خطی طبقه‌بندی قرار می‌گیرد. این در حالی است که برای اجرای روش‌های داده‌کاوی نیازی به بررسی توزیع داده‌ها و هیچ پیش فرض دیگری نیست. این روش‌ها همچنین قادر هستند که هر گونه رابطهٔ خطی و غیرخطی را کشف کنند. از آنجا که متغیرهای به کار رفته در این پژوهش زیاد است احتمال اینکه داوطلبان به صورت خطی طبقه‌بندی شوند بسیار کم و یا حتی ناممکن است به همین دلایل در این پژوهش از روش‌های داده‌کاوی از جمله شبکهٔ عصبی مصنوعی و درخت تصمیم برای طبقه‌بندی داوطلبان آزمون سراسری گروه ریاضی و فنی استفاده شد و این داوطلبان از نظر نوع عملکرد طبقه‌بندی شدند.

در این مطالعه، از اطلاعاتی نظیر جنسیت، سوابق تحصیلی داوطلبان، سهمیهٔ نهایی و نمره‌های علمی کسب شده در هر یک از درس‌ها استفاده شد. با استفاده از آزمون‌های آماری مشخص شد که دخترها و پسرها در نمرهٔ کل تفاوت معناداری با یکدیگر دارند که با مبانی نظری مطرح شده مطابقت دارد، اما بین داوطلبان سهمیه‌های مختلف. تفاوت معناداری وجود نداشت. لذا با وجود بی‌تأثیری جنسیت بر نتیجهٔ آزمون سراسری، نتایج الگوریتم‌ها با جنسیت و بدون جنسیت بررسی شد. از طرفی در هر مسئله علم داده‌ای، قبل از شروع کار لازم است داده‌ها پیش‌پردازش شود که یکی از عملیات‌های مهم این فرآیند، هم‌مقیاس کردن داده‌ها است تا همگی تأثیر یکسانی در تحلیل داشته باشند. روش‌های مختلفی برای هم‌مقیاس کردن داده‌ها وجود دارد که در این مقاله از روش استانداردسازی استفاده شد. با توجه به اینکه عملکرد داوطلبان با نمرهٔ کل نهایی مشخص می‌شود لذا چارک‌های نمره‌های کل نهایی داوطلبان مطابق جدول (۳) تعیین و به داده‌های مربوط به نمرهٔ کل، چهار برچسب ضعیف، متوسط، خوب و خیلی خوب داده شد. داده‌های مورد استفاده در این پژوهش، اطلاعات مربوط ۸۳۰۵ نفر از داوطلبان استان البرز بود که در گروه ریاضی و فیزیک شرکت کرده بودند که هشتاد درصد این مجموعه داده به عنوان دادهٔ آموزشی و بیست درصد به عنوان داده برای آزمون در نظر گرفته شد تا قابلیت اعتماد الگو سنجیده شود.

سه الگوریتم درخت تصمیم (CHAID، C5 و CART) و نیز شبکهٔ عصبی برای طبقه‌بندی داده‌ها با استفاده از نرم‌افزار "SPSS Modeler" نسخهٔ ۱۸ اجرا شد و در نهایت با توجه به دقت کلی که این نرم‌افزار براساس پیش‌بینی درست ارائه می‌دهد نتیجه گرفته شد که شبکهٔ عصبی پرسپترون کم‌ترین خطا را در طبقه‌بندی داوطلبان دارد و بنابر این مورد اعتمادتر است.

در این پژوهش مشاهده شد که با حذف جنسیت و افزودن سایر متغیرهای مربوط به سوابق تحصیلی، دقت شبکهٔ عصبی و همچنین الگوریتم "c5" بهبود می‌یابد. به علاوه مشخص شد که در صورت به‌کارگیری نمره‌های درس‌ها به تنهایی، دقت شبکهٔ عصبی و الگوریتم‌های "c5" و "CART" کاهش می‌یابد ولی دقت

الگوریتم "CHAID" در هر سه حالت یکسان باقی می‌ماند که به ماهیت الگوریتم برمی‌گردد. همچنین در صورت وجود متغیرهای مربوط به سوابق تحصیلی بدون حضور جنسیت، شبکهٔ عصبی با دقت بیشتری سطح عملکرد داوطلبان را تعیین می‌کند و می‌توان نتیجه گرفت که جنسیت در سطح عملکرد داوطلب تأثیری ندارد. در دو الگوریتم "c5" و "CART"، بهرهٔ اطلاعات درس فیزیک و در الگوریتم "CHAID" بهرهٔ اطلاعات درس ریاضی بیشتر شد که در نتیجه در این الگوریتم‌ها فیزیک و ریاضی در ریشهٔ درخت قرار گرفتند. به علاوه دقت الگوریتم "c5" در حضور متغیرهای سوابق تحصیلی بدون جنسیت افزایش یافت ولی جنسیت در دقت الگوریتم‌های "CART" و "CHAID" تأثیری نداشت و از طرفی دقت الگوریتم "CART" بدون حضور متغیرهای سوابق تحصیلی کاهش یافت ولی حضور یا عدم حضور متغیرهای سوابق تحصیلی تأثیری در دقت الگوریتم "CHAID" نداشت. به طور کلی دقت الگوهای مربوط به درخت تصمیم در حضور متغیرهای دیگر از جمله جنسیت، معدل و فاصلهٔ دریافت دیپلم تا آزمون سراسری، بیشتر از الگوی دوم بود.

در الگوریتم "c5" با احتساب جنسیت، درس‌های ریاضی و فیزیک بیشترین تأثیر را داشتند و سپس به ترتیب عربی و معدل دیپلم و شیمی بر سطح عملکرد تأثیر گذار بودند. در این الگوریتم اهمیت جنسیت بیشتر از فاصلهٔ دریافت دیپلم تا آزمون سراسری بود ولی به طور کلی میزان تأثیر هر دوی این متغیرها ناچیز بود. به علاوه در صورتی که تنها نمره‌های درس‌ها به عنوان متغیرهای تأثیرگذار در سطح عملکرد در نظر گرفته شود، بیشترین درسی که در سطح عملکرد داوطلبان گروه ریاضی و فنی تأثیر می‌گذارد به ترتیب شیمی، فیزیک و ریاضی است که این ترتیب در گروه ریاضی و فنی منطقی نیست. همچنین میزان دقت الگوریتم "c5" به نسبت شبکهٔ عصبی کمتر است. دلیل کم شدن دقت درخت تصمیم به این علت است که متغیرهای کمّی به کیفی تبدیل شدند و بدین ترتیب اطلاعات زیادی از دست رفته است در صورتی که اگر متغیرها ذاتاً کیفی بودند از درخت تصمیم نتایج رضایتبخشی به دست می‌آمد.

براساس شبکهٔ عصبی نتیجه گرفته شد در صورتی که همهٔ متغیرهای پژوهش، اعم از درسی و غیردرسی بدون جنسیت در نظر گرفته شود ترتیب اهمیت درس‌های تخصصی همانند مقدار تعیین شده توسط سازمان سنجش آموزش کشور است و تنها میزان اهمیت معدل دیپلم بیشتر از درس عربی و فاصلهٔ دریافت دیپلم تا آزمون سراسری است. این بدین معنی است که معدل دیپلم در طبقه‌بندی داوطلبان براساس سطح عملکرد، بیشتر از درس عربی مؤثر است که نتیجهٔ حاصل شده می‌تواند به روایی آزمون مرتبط باشد. در صورتی که طبقه‌بندی تنها براساس متغیرهای درسی صورت گیرد به ترتیب درس‌های ریاضی، شیمی و فیزیک بیشترین اهمیت را دارند و درس عربی کم‌ترین تأثیر را دارد. نکتهٔ جالب توجه اینجاست که ترتیب اهمیت درس‌های عمومی در این دو الگو (براساس سؤال اول و دوم پژوهش) تغییری نکرده است و تنها در

الگوی دوم، اهمیت درس شیمی بیشتر از فیزیک شده است.

همان گونه که مشاهده می‌شود متغیرهای سوابق تحصیلی و فاصلهٔ دریافت دیپلم تا آزمون سراسری در الگوی شبکهٔ عصبی تأثیر کمتری دارد. همچنین طبق تحقیق‌های انجام شده مشخص شد که در صورت نادیده گرفتن متغیر فاصله دریافت دیپلم تا آزمون سراسری، این تناقض در نتایج از بین می‌رود و ترتیب اهمیت متغیرها به صورت ریاضی، فیزیک، شیمی، فارسی، دینی، عربی، زبان و معدل دیپلم شد که منطقی است. از سوی دیگر با مقایسهٔ مقادیر حاصل شده در جدول‌های (۶)، (۷) و (۹) می‌توان نتیجه گرفت که حضور متغیرهایی غیر از درس‌های آزمون باعث بالا رفتن دقت شبکهٔ عصبی می‌شود و این به معنای قدرت شبکهٔ عصبی است که با وجود متغیرهای مختلف می‌تواند بهترین طبقه‌بندی را با کمترین خطا انجام دهد. همچنین این نتیجه در مورد درخت تصمیم نیز صادق است، اما از آنجا که متغیرهای به کار رفته در این طبقه‌بندی، کمّی بودند و بنا به اقتضای درخت تصمیم به کیفی تبدیل شدند این امر منجر به از دست رفتن اطلاعات زیادی شد که دقت الگوریتم‌های درخت تصمیم را کاهش داد، لذا از نتایج درخت تصمیم برای ارزیابی با اعتماد بالا نمی‌توان استفاده کرد.

در پایان به منظور توسعهٔ این پژوهش، پنج پیشنهاد کاربردی و پژوهشی ارائه می‌شود:

۱) می‌توان الگوی طبقه‌بندی مناسب را برای دختران و پسران به طور جداگانه بررسی کرد و نتیجه گرفت که آیا جنسیت در انتخاب الگوی طبقه‌بندی تأثیر دارد یا خیر.

۲) نتایج الگوریتم‌های هوش مصنوعی مانند شبکهٔ عصبی با رگرسیون لوجستیک مقایسه شود.

۳) الگوی طبقه‌بندی مناسبی براساس متغیرهای اسمی مانند جنسیت و سهمیهٔ نهایی انتخاب شود که در این حالت بهتر است از نمره‌کل‌های اصلی و نه نمره‌کل‌های طبقه‌بندی شده استفاده شود.

۴) با استفاده از شبکهٔ عصبی و رگرسیون لوجستیک و براساس قبولی و رد داوطلبان، الگوی طبقه‌بندی مناسبی انتخاب شود و نتایج با یکدیگر مقایسه شود.

۵) می‌توان از روش‌های دیگری برای تبدیل متغیرهای کمّی به کیفی استفاده کرد و سپس الگوی شبکهٔ عصبی و درخت تصمیم را به کار گرفت.

References

Abdulrazzaq, N., Kamal, M., Muhsen, A., Tareq, A., Al Zubaidi, R., & Al Mousawi, A. (2017). Academic Failure And Student's Viewpoint: the Influence of Individual, Internal and External Organizational Factors.

Abu Tair, M. M., & El-Halees, A. M. (2012). Mining educational data to improve students' performance: a case study. *International Journal of Information,* 2(2).

Afroz, G. (1375). *Educational Psychology*. Retrieved from

Angeline, D. M. D. (2013). Association rule generation for student performance analysis using apriori algorithm. *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA),* 1(1), 12-16 .

Bakhshayesh, A. (1393). Investigating the relationship between thinking styles and learning strategies with academic performance in students. *Research Quarterly in Curriculum Planning,* 11*(*14*)*, 135 - 146 .

Banik, P., & Kumar, B. (2019). Impact of information literacy skill on students' academic performance in Bangladesh. *International Journal of European Studies,* 3(1), 27-33 .

Bin Mat, U., Buniyamin, N., Arsad, P. M., & Kassim, R. (2013). *An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention.* Paper presented at the 2013 IEEE 5th conference on engineering education (ICEED).

Christian, T. M., & Ayub, M. (2014). *Exploration of classification using NBTree for predicting students' performance.* Paper presented at the 2014 international conference on data and software engineering (ICODSE).

Demuth, H., & Beale, M. (2000). Neural network toolbox user's guide.

Eskandari, F. (1395). *Categorical data analysis*. Tehran: Allameh Tabatabai University.

Fataa, L., Azari, S., Baradaran, H., & Atlasi, R. (1392). A systematic review of the causes of academic failure of medical students. *Journal of development steps in medical education,* 10 *(2)*, 31 - 38.

Golshani-Foomani, M. (1375). *Sociology of Education*: shifteh.

Applied Statistics Research Group. (1394). *Constructing the total score of the national exam in the experimental group of mathematical and technical sciences in* 2013 *based on the actual distribution of scores and comparing it with the current method.* Retrieved from Tehran

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*: Elsevier.

Hassoun, M. H. (1995). *Fundamentals of artificial neural networks*: MIT press.

Ibrahim, Z., & Rusli, D. (2007). *Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression.* Paper presented at the 21st Annual SAS Malaysia Forum, 5th September.

Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic

minority over-sampling technique. *Decision Analytics,* 2(1), 1-25.

Kazeminezhad, M., Etemad-Shahidi, A., & Mousavi, S. (2005). Application of fuzzy inference system in the prediction of wave parameters. *Ocean Engineering,* 32(14-15), 1709-1725.

Lawrence, A., & Deepa, T. (2013). Emotional Intelligence and Academic Achievement of High School Students in Kanyakumari District. *Online submission,* 3(2), 101-107.

Li, K. F., Rusk, D., & Song, F. (2013). *Predicting student academic performance*. Paper presented at the 2013 Seventh International Conference on Complex, Intelligent, and Software Intensive Systems.

Mau, W. C. (1997). Parental influences on the high school students' academic achievement: A comparison of Asian immigrants, Asian Americans, and White Americans. *Psychology in the Schools,* 34(3), 267-277 .

Meit, S. S., Borges, N. J., Cubic, B. A., & Seibel, H. R. (2004). Personality Differences in Incoming Male and Female Medical Students. *Online submission*.

Menhaj, M. (1397). *Basics of Neural Networks - Volume One: Computational Intelligence*. Tehran: Amirkabir University of Technology.

Meshkani, A., & Naazemi, A. (1388). *An introduction to data mining*. Mashhad: Ali Meshkani Publications in cooperation with Islamic Azad University, Neyshabur branch.

Mitchell, T. M. (1997). Does machine learning really work? *AI magazine,* 18(3), 11-11 .

Navidi, A. (1382). Investigating the joint and specific contribution of previous academic performance variables, academic self-concept and general intelligence in predicting students' academic progress. *education and training,* 19 *(4)*, 97 - 129.

Oladokun, V., Adebanjo, A., & Charles-Owaba, O. (2008). Predicting students academic performance using artificial neural network: A case study of an engineering course.

Osmanbegovic, E., & Suljic, M. (2012). Data mining approach for predicting student performance. *Economic Review: Journal of Economics and Business,* 10(1), 3-12.

Patel, P., & Mistry, K. (2015). A review: Text classification on social media data. *IOSR Journal of Computer Engineering,* 17(1), 80-84.

Rahimi, M. (1387). *Investigating the factors related to the academic decline of female high school students in Region* 14. (Master's thesis), University of Tehran.,

Ramesh, V., Parkavi, P., & Ramar, K. (2013). Predicting student performance: a statistical and data mining approach. *International journal of computer applications,* 63(8).

Sadeghi-Jafari, J., Roshan, M., & Shakoori-Ganjavi, H. (1390). A comparative study of the academic performance of students of self-sacrifice quotas and regional quotas in daily courses of state universities. *Higher education letter,* 3 *(*10*)*, 51- 75.

Sanie-Abadeh, M., Mahmoodi, S., & Taherpoor, M.(1391). *Applied data mining*. Tehran: need for knowledge.

Shahiri, A. M., & Husain, W. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science,* 72, 414-422.

Shahrabi, J., & ZolghadreShojae, A. (1390). *Advanced data mining concepts and algorithms*: Jihad University Publications of Amir Kabir University.

Simsek, A., & Balaban, J. (2010). Learning strategies of successful and unsuccessful university students. *Contemporary Educational Technology,* 1(1), 36-45.

Talebi, A., & Akbari, Z. (1392). Investigating the effectiveness of the decision tree model in estimating river suspended sediments (Case study: Ilam Dam Basin. *Agricultural sciences and techniques and natural resources,* 63, 109 - 121.

WANG, & LEE. (2006). Constructing a fuzzy decision tree by integrating fuzzy sets and entropy. *WSEAS transactions on information science and applications*.

Wendler, T., & Gröttrup, S. (2021). *Data mining with SPSS modeler: theory, exercises and solutions*: Springer.

Widyastuti, T., Kurniawan, A., & Chandra, N. P. (2017). Coping Strategies on Students After Experiencing Academic Failure: An Indigenous Study in Javanese Context. *Work Pap Ser,* 3, 22-26.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی