Educational Measurement and Evaluation Studies

National Organization for Educational Testing

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

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

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

1. Assistant Professor, Department of Statistics, University of Sistan and Baluchestan,Sistan and Baluchestan, Zahedan, Iran, Email: h.h.naderi@gmail.com
2. Associate Professor, Faculty of Psychology and Education, University of Tehran, Tehran, Iran ; (Corresponding author), Email: keyvansalehi@ut.ac.ir
3. Ph.D. Candidate, Faculty of Psychology and Education, University of Tehran, Tehran, Iran, Email: maryam.parsaeian@ut.ac.ir

Article Info

Article Type:
Research Article

Received: 2022/01/02
Received in revised form: 2023/05/17
Accepted: 2023/05/22
Published online: 2023/05/23

ABSTRACT

Objective: Comparison of 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 national exam candidates.

Methods: It is based on a quantitative approach and a survey method, and the variables of gender, 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 the general courses of Persian and religion respectively, are the most important in the classification. The distance between the diploma and the national exam had the least effect. It was also found that if only variables related to 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 has a higher performance than the decision tree algorithms. On the other hand, the inclusion of educational background variables has had a favorable effect on the accuracy of the neural network algorithm.

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

Cite this article: Naderi, Habib; Salehi, Keyvan; Parsaeian, Maryam (2023). Comparing the Performance of Decision Tree and Artificial Neural Network in the Classification of National Exam Candidates. *Educational Measurement and Studies*, 13 (41): 31-50 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

مقایسه عملکرد درخت تصمیم و شبکه 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: ۱۴۰۰/۱۰/۱۲
اصلاح: ۱۴۰۲/۰۲/۲۷
پذیرش: ۱۴۰۲/۰۳/۰۱
انتشار : ۱۴۰۲/۰۳/۰۲

چکیده

هدف: مقایسه قابلیت‌های سه الگوریتم مختلف درخت تصمیم و مدل شبکه عصبی به منظور انتخاب مدل 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 دروس به عنوان متغیرهای موثر در طبقه‌بندی استفاده گردید.

یافته‌ها: با در نظر گرفتن همه متغیرها بدون جنسیت براساس مدل شبکه عصبی، مشخص شد که دروس تخصصی ریاضی، فیزیک و شیمی سپس دروس عمومی فارسی و دینی به ترتیب بیشترین اهمیت را در طبقه‌بندی دارند. فاصله دیپلم تا آزمون سراسری کمترین تاثیر را داشت. همچنین مشخص شد در صورتی که تنها متغیرهای مربوط به دروس آزمون در نظر گرفته شود ترتیب میزان اهمیت دروس شیمی و فیزیک جابه‌جا می‌شود.

نتیجه‌گیری: با استناد به شاخص دقت کلی، الگوریتم مدل شبکه عصبی با دقت بیشتر از ۰/۹۵ از عملکرد بالاتری نسبت به الگوریتم‌های درخت تصمیم برخوردار است. از طرفی ورود متغیرهای سوابق تحصیلی در دقت الگوریتم شبکه عصبی تاثیر مطلوبی داشته است.

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

استناد: نادری، حبیب؛ صالحی، کیوان؛ پارساییان، مریم (۱۴۰۲). مقایسه عملکرد درخت تصمیم و شبکه عصبی مصنوعی در طبقه‌بندی داوطلبان آزمون سراسری. *مطالعات اندازه‌گیری و ارزشیابی آموزشی*، ۱۳ (۴۱)، ۳۱-۵۰ صفحه.

DOI: 10.22034/EMES.2023.545992.2335

ناشر: سازمان سنجش آموزش کشور

حق‌مؤلف © نویسندگان.

مقدمه

عملکرد تحصیلی یک جنبه حیاتی از تحقیقات آموزشی است و الگوریتم‌های مختلف یادگیری ماشین برای مدل‌سازی و پیش‌بینی عملکرد دانش‌آموزان استفاده شده است. یادگیری ماشین زیرشاخه‌ای از هوش مصنوعی است که شامل توسعه الگوریتم‌ها و مدل‌های آماری است که می‌توانند از داده‌ها یاد بگیرند و پیش‌بینی یا تصمیم بگیرند. در سال های اخیر، یادگیری ماشین به دلیل توانایی آن در پردازش سریع و دقیق حجم عظیمی از داده‌ها، به طور فزاینده‌ای در حوزه‌های مختلفی از جمله مالی، مراقبت های 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 ماشین یعنی درخت تصمیم و شبکه عصبی مصنوعی در طبقه‌بندی داوطلبان آزمون سراسری با یکدیگر مقایسه شده‌اند. درخت تصمیم برای مدل‌سازی تصمیمات و پیامدهای احتمالی آنها در زمینه پیش‌بینی عملکرد تحصیلی استفاده می‌شود. به علاوه 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 اینکه دانش‌آموز در یک درس قبول می‌شود یا رد می‌شود) و هم برای کارهای رگرسیون (مانند پیش‌بینی نمره نهایی دانش‌آموز) استفاده کرد.

هدف این مقاله، مقایسه ارزیابی عملکرد داوطلبان آزمون سراسری گرو علوم ریاضی و فنی با استفاده از روش‌های داده‌کاوی و با توجه به متغیرهای مختلف تاثیرگذار از جمله جنسیت، سوابق تحصیلی (معدل کتبی دیپلم و فاصله اخذ دیپلم و شرکت در آزمون سراسری) و نمره‌های تراز هر یک از دروس 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 یک آزمون یکسان به رقابت بپردازند، تاثیر متغیر سهمیه نهایی (شامل مناطق ۱ و۲ و۳، شاهد) نیز در نمره داوطلبان در نظر گرفته شد (صادقی جعفری، روشن و شکوری گنجوی، ۱۳۹۰). بدین‌منظور به مقایسه قابلیت‌های سه الگوریتم مختلف درخت تصمیم (C5، CHAID و CART) و مدل شبکه عصبی پرداخته شده است.

نتایج این مطالعه بینش ارزشمندی در مورد نقاط قوت و ضعف الگوریتم درخت تصمیم و شبکه عصبی مصنوعی در پیش‌بینی عملکرد داوطلبان آزمون سراسری ارائه می‌کند. این اطلاعات می‌تواند توسط موسسات آموزشی برای بهبود مدل‌های پیش‌بینی خود و درک بهتر عوامل مؤثر بر عملکرد دانش‌آموزان استفاده شود.

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 کردند. در واقع اگر عملکرد تحصیلی به صورت یک پیوستار در نظر گرفته شود، در یک سر آن پیشرفت تحصیلی و سر دیگر آن افت تحصیلی قرار می‌گیرد. گاهی مشاهده می‌شود که محققان در انجام تحقیقات خود 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 در تحقیق خود مترادف در نظر گرفتند. بهترین شاخص برای نشان دادن افت یا پیشرفت تحصیلی از نظر رحیمی(۱۳۸۷)، مقایسه 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 یک حوزه دانشگاهی به دست آمده باشند. 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). به این معنا که فراگیر از یک سطح بالاتر به سطحی پایین‌تر در تحصیل و آموزش نزول کرده است. پس

[1]. Banik & Kumar
[2]. Mau
[3]. Lawrence & Deepa
[4]. Abdulrazzaq, Kamal, Muhsen, Tareq, Al Zubaidi & Al Mousawi

می‌توان آن را کاهش عملکرد فراگیر دانست (رحیمی، ۱۳۸۷). افت تحصیلی هنگامی اتفاق می‌افتد که فراگیری که از نظر نمره ضعیف است، به انتظار خود نرسد و ملاک‌های آموزشی را برآورده نکند (ویدیاستوتی، کورنیاوان و چاندرا[۱]، ۲۰۱۷).

مرور منظم مطالعات مرتبط نشان داد که معدل کل و ارزیابی داخلی[۲]، دو ویژگی اصلی است که اغلب برای ارزیابی عملکرد تحصیلی مورد استفاده قرار می‌گیرد (آنجلین[۳]، ۲۰۱۳؛ کریستین و ایوب[۴] ، ۲۰۱۴؛ لی، راسک و سانگ[۵]، ۲۰۱۳؛ جیشان، راشو، هاک و رحمان[۶]، ۲۰۱۵). از ۳۰ مقاله مورد بررسی در این زمینه، ۱۳ مقاله از این دو ویژگی استفاده کرده‌اند. دلیل اصلی اکثر محققان برای استفاده از معدل کل این است که ارزش قابل ملاحظه‌ای برای پویایی تحصیلی و شغلی آینده دارد. همچنین می‌توان آن را نشانه‌ای از ظرفیت علمی تحقق یافته دانست (بین‌مات، بونیامین، آرسد و کاسیم[۷]، ۲۰۱۳). ایبراهیم وروسلی[۸] (۲۰۰۷) با استفاده از مقدار ضریب همبستگی (۰/۸۷)، نتیجه گرفتند که مهم‌ترین متغیر ورودی برای ارزیابی عملکرد تحصیلی، معدل کل است. همچنین طبق مطالعه کریستین و ایوب (۲۰۱۴) معدل کل بیشترین تأثیر را در تعیین بقای دانش‌آموزان دارد، خواه بتوانند تحصیل خود را به پایان برسانند یا نه. ویژگی مهم دیگر شامل ارزیابی‌های جمعیت‌شناختی[۹] و ارزیابی خارجی[۱۰] دانش‌آموزان است. مشخصات جمعیت‌شناختی دانش‌آموزان شامل جنسیت، سن، سابقه خانوادگی و معلولیت است (ایبراهیم و روسلی، ۲۰۰۷؛ بین‌مات، بونیامین، آرسد و کاسیم، ۲۰۱۳؛ کریستین و ایوب، ۲۰۱۴). میت، بورگس، کوبیک و سیبل[۱۱] (۲۰۰۴) نشان دادند که اکثر دانش‌آموزان دختر در مقایسه با دانش‌آموزان پسر دارای سبک‌ها و رفتارهای یادگیری مثبت مختلفی هستند (میت و همکاران، ۲۰۰۴). دانش‌آموزان دختر خودنظم دهی و دقت بیشتری در مطالعات خود دارند. از سوی دیگر، دانش‌آموزان دختر دارای استراتژی‌های یادگیری موثری در مطالعات خود هستند (سیم‌سک و بالابان[۱۲]، ۲۰۱۰). آنها خود انگیزشی، سازمان‌دهی و مرور ذهنی دارند و به طور موثر از آنها استفاده می‌کنند. بنابراین نقش جنسیت به عنوان یکی از ویژگی‌های مهم تأثیرگذار بر عملکرد دانش‌آموزان مورد تایید قرار گرفته شده است (عثمان‌بگانویک و سال‌جیک[۱۳]، ۲۰۱۲). به عبارت دیگر دلیل این که اکثر محققان از جنسیت دانش‌آموزان به عنوان یک متغیر استفاده می‌کنند این است که دخترها و پسرها در فرآیند یادگیری خود سبک‌های متفاوتی دارند (بین‌مات، بونیامین، آرسد و کاسیم، ۲۰۱۳). همچنین سابقه تحصیلی هم در پیش‌بینی عملکرد دانش‌آموز نقش موثری دارد (آبوتیر و ایل‌هالیز[۱۴]، ۲۰۱۲؛ اولادوکان، آدبانجو و چارلز اوابا[۱۵]، ۲۰۰۸؛ رامش، پرکاوی و رامر[۱۶]، ۲۰۱۳).

بیش از چهار دهه است که ارزیابی عملکرد دانش‌آموزان دوره متوسطه به منظور ورود به دانشگاه‌های ایران از طریق آزمون سراسری به طور سالانه برای پنج گروه آزمایشی علوم ریاضی و فنی، علوم تجربی، علوم انسانی، زبان‌های خارجی و هنر به‌طور جداگانه برگزار می‌شود که با توجه به اهمیت داشتن نتایج آن برای شرکت‌کنندگان، آزمون‌های سرنوشت‌ساز محسوب می‌شوند. بنابراین نحوه گزینش دقیق داوطلبان در این آزمون از اهمیت ویژه‌ای برخوردار است. در واقع با توجه به مواد درسی مورد آزمون که دروس آموخته‌شده دوره متوسطه دوم در دبیرستان است می‌توان این آزمون را به عنوان ابزاری برای سنجش پیشرفت تحصیلی دانش‌آموزان و نوع عملکرد آنها در دوره متوسطه دوم در دبیرستان در نظر گرفت.

سازمان سنجش به‌منظور سنجش دقیقی از داوطلبان، ملاک‌های مختلفی از جمله نمره علمی کسب‌شده در آزمون را در نظر می‌گیرد. داوطلبان با پاسخ به سوال‌های 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 یک نمره واحد به نام شاخص نمره کل تعریف شود. نمره کل داوطلبان در سازمان سنجش آموزش

---

1. Widyastuti, Kurniawan & Chandra
2. internal assessment
3. Angeline
4. Christian & Ayub
5. Li, Rusk, & Song
6. Jishan, Rashu, Haque, & Rahman
7. Bin Mat, Buniyamin, Arsad, & Kassim
8. Ibrahim & Rusli
9. demographic
10. external assessment
11. Meit, Borges, Cubic & Seibel
12. Simsek & Balaban
13. Osmanbegovic & Suljic
14. Abu Tair & El-Halees
15. Oladokun, Adebanjo, & Charles-Owaba
16. Ramesh, Parkavi, & Ramar

کشور طی یک فرآیند سه مرحله‌ای تعیین نمره خام و نمره تراز تمام دروس و در نهایت محاسبه نمره کل انجام می‌شود. در محاسبه نمره خام، تعداد پاسخ‌های غلط از سه برابر تعداد پاسخ‌های صحیح کم شده و حاصل در سه برابر تعداد سوال‌های آن درس تقسیم می‌شود و در نهایت نتیجه در ۱۰۰ ضرب می‌شود. برای تعیین نمره 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 شوند. لازم به ذکر است که با توجه به ویژگی چندک‌ها، در همه درس‌ها تعییر نمره‌های Z_{ij} ها حداکثر در بازه $(+5 , -5)$ قرار می‌گیرند. یعنی نمره همه داوطلبان در محدوده 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 روابط بالا اندیس‌های i و j به ترتیب بیانگر شماره آزمودنی i در درس j است. در انتها نمره کل از طریق گرفتن میانگین وزنی نمره‌های تراز شده 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 می‌شوند که مشخص می‌نمایند هر مورد به کدام گروه تعلق دارد. این الگو هم می‌تواند برای فهم داده‌های موجود و هم برای پیش بینی اینکه هر نمونه جدید چگونه کار می‌کند استفاده شود. طبقه‌بندی در واقع ارزیابی ویژگی‌های مجموعه‌ای از داده‌ها و سپس اختصاص دادن آنها به مجموعه‌ای از گروه‌های از پیش تعریف شده است. هریک از الگوریتم‌های موجود برای طبقه‌بندی دارای رویکردها و تکنیک‌های متفاوتی برای بررسی داده‌ها و یافتن قوانین طبقه‌بندی هستند. این روش‌ها را می‌توان به سه نوع الگوریتم خطی، غیرخطی و مبتنی بر قاعده تقسیم کرد. مدل‌های خطی و غیرخطی عمدتا از رویکرد ریاضی پیروی می‌کنند و توابعی را برای امتیازدهی و جداسازی داده‌ها و کلاس‌ها ایجاد می‌کنند. با این تفاوت که روش‌های خطی، کلاس‌های مختلف را با توابع خطی از هم مشخص می‌کنند ولی در روش‌های غیرخطی از توابع پیچیده‌تری استفاده می‌شود. برخلاف رویکردهای ریاضی، خروجی مدل‌های مبتنی بر قانون به صورت قوانین "اگر .... آنگاه" است. الگوریتم‌های مهم طبقه‌بندی را می‌توان به صورت جدول ۱ فهرست کرد (شاهیری و حوسین[2]، ۲۰۱۵).

جدول ۱. الگوریتم‌های رایج طبقه‌بندی

مبتنی بر قانون	غیرخطی	خطی
درخت تصمیم	ماشین بردار پشتیبان[3]	رگرسیون لوجستیک
CART	شبکه‌های عصبی	تحلیل تشخیص خطی
C4.5, C5	شبکه‌های بیزی	–
CHAID	نزدیک‌ترین همسایه[4] (KNN)	–
QUEST	–	–

از آنجایی که تعداد متغیرهای مورد استفاده در این پژوهش زیاد است امکان طبقه‌بندی صحیح داوطلبان با استفاده روش‌های خطی مانند رگرسیون لوجستیک ناممکن است لذا روش‌های غیرخطی منجر به دقت بالاتری در طبقه‌بندی داوطلبان می‌شود. مطالعات انجام شده بیانگر پتانسیل درخت تصمیم و الگوریتم‌های شبکه عصبی برای پیش‌بینی عملکرد تحصیلی هستند. لذا در این مطالعه از این الگوریتم‌ها برای طبقه‌بندی داوطلبان کنکورسراسری استفاده شده است که سه الگوریتم مهم درخت تصمیم و مدل شبکه عصبی پرسپترون برای این پژوهش در نظر گرفته شدند که در ادامه به تفکیک مبانی نظری روش‌های مربوطه مطرح می‌شوند.

درخت تصمیم

استفاده از درخت تصمیم بدلیل سادگی در استفاده و فهم، قابلیت آن برای کار با داده‌های بزرگ، قابلیت ترکیب با روش‌های دیگر و همچنین استفاده مجدد آسان از درخت تصمیم ساخته‌شده مورد استقبال پژوهشگران بسیاری قرار گرفته است. از فن درخت تصمیم هم می‌توان برای کشف و استخراج دانش از یک پایگاه داده و هم برای ایجاد مدل‌های پیش‌بینی استفاده نمود (کاظمی‌نژاد، اعتمادشهیدی و موسوی، ۲۰۰۵). به

[1]. Classification
[2]. Shahiri & Husain
[3]. Support Vector Machine
[4]. K-Nearest Neighbor

عبارت دیگر درخت تصمیم یکی از ابزارهای متداول برای دسته‌بندی و پیش‌بینی است که قابلیت تولید قانون را نیز دارد (شهرابی و ذوالقدرشجاعی، ۱۳۹۰). درخت تصمیم بر اساس اینکه متغیر پاسخ پیوسته یا گسسته باشد به دو دسته رگرسیونی و طبقه‌بندی تقسیم می‌شوند (طالبی و اکبری، ۱۳۹۲).

درخت تصمیم اغلب از چندین گره تشکیل می‌شود که با نام گره‌های ورودی و خروجی شناخته می‌شوند که از سه بخش اصلی برگ[۱]، ریشه[۲] و شاخه[۳] تشکیل شده است. به عبارت دقیق‌تر برگ‌ها گره‌هایی هستند که تقسیم‌های متوالی در آنجا پایان می‌یابد و بیانگر یک کلاس هستند. به گره اول هر درخت تصمیم ریشه گفته می‌شود. در هر گره داخلی به تعداد جواب‌های ممکن تقسیم صورت می‌گیرد و در نتیجه این تصمیمات در بخش میانی درخت شاخه ایجاد می‌شود که هر شاخه بیانگر ارزش ویژگی مورد نظر است (میچل[۴]، ۱۹۹۷). قوانین ایجاد شده در درخت تصمیم به صورت اگر و آنگاه بیان می‌شوند. همچنین در یک درخت تصمیم، مهم است که کدام یک از ویژگی‌ها (یا همان ابعاد) در سطوح بالاتری از درخت قرار گیرند تا طبقه‌بندی بهتری انجام شود. لذا ویژگی که توزیع دسته‌ها در گره‌های حاصل از آن همگن باشد نسبت به سایر ویژگی‌ها برتری دارد. منظور از همگن بودن گره این است که همه رکوردهای موجود در آن متعلق به یکدسته خاص باشند و چون در آن صورت آن گره به برگ تبدیل می‌شود بنابراین گره‌ای به عنوان گره همگن در نظر گرفته می‌شود که کمترین میزان ناخالصی را داشته باشد. دو روش مهم برای محاسبه ناخالصی گره شاخص جینی و آنتروپی است (صنیعی‌آباده، محمودی، طاهرپور، ۱۳۹۱). الگوریتم‌های مختلفی برای ساخت درخت تصمیم توسعه یافته شده‌اند که از آن جمله می‌توان به ID3 (کوییلن[۵]، ۱۹۸۶)، C4.5 (کوییلن، ۱۹۹۳)، CART[۶] (بریمان[۷] و همکاران، ۱۹۸۴)، CHAID[۸] (کاس[۹]، ۱۹۸۱) و QUEST[۱۰] (لوح و شیح[۱۱]، ۱۹۹۷) اشاره کرد (شهرابی و ذوالقدرشجاعی، ۱۳۹۰).

درخت تصمیم در الگوریتم‌های ID3، C4.5 و C5 براساس مفهوم بی‌نظمی ساخته می‌شود. به این مفهوم که الگوریتم درصدد آن است که میزان بی‌نظمی در گره‌های بالایی درخت حداقل باشد. بنابراین برای تمام ویژگی‌های داده‌های اولیه، بی‌نظمی با استفاده از آنتروپی شانون محاسبه شده و سپس آن ویژگی که بیشترین سودمندی را دارد به عنوان ریشه انتخاب می‌شود (شهرابی و ذوالقدرشجاعی، ۱۳۹۰). تنها تفاوتی که این الگوریتم‌ها با هم دارند این است که در الگوریتم ID3 فقط قانون تقسیم وجود دارد ولی الگوریتم C4.5، قانون هرس نیز دارد. بدین‌معنی که بعد از ساخت درخت، تعدادی از شاخه‌ها به دلیل وجود نقاط دورافتاده در داده‌های آموزشی ناهنجاری‌هایی ایجاد می‌کنند. روش هرس کردن با به کارگیری از معیارهای آماری، شاخه‌هایی که کمتر مورد اطمینان هستند را حذف می‌کند که این امر منجر به رده‌بندی سریع‌تر و بهبود در توانایی درخت برای رده‌بندی صحیح داده‌های آزمون می‌شود (مشکانی و ناظمی، ۱۳۸۸). با الگوریتم C5 می‌توان مجموعه داده‌ها را به بیش از دو قسمت تقسیم کرد و هرس را پس از رشد درخت فراهم می‌کند و قوانین تقسیم گره‌ها از طریق اندازه‌گیری ناخالصی انتخاب می‌شود (وندلر و گراتآپ[۱۲]، ۲۰۲۱).

در هر سه الگوریتم ID3، C4.5 و C5 برای انتخاب مهم‌ترین خصیصه از شاخص آنتروپی استفاده می‌شود که معیاری عددی از میزان اطلاعات یا میزان تصادفی بودن یک متغیر تصادفی است که به هر یک از ملاک‌ها وزن منحصر به فردی می‌دهد و به صورت زیر محاسبه می‌شود:

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

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
12. Wendler, T., & Gröttrup

که در آن $p(i|n)$ کسری از نمونه‌های متعلق به کلاس i است که در گره n واقع هستند. بنابراین طبق فرمول آنتروپی می‌توان نتیجه گرفت که ویژگی در بالای درخت قرار می‌گیرد که بیشترین بار اطلاعاتی را داشته باشد و باعث کاهش آنتروپی یا بی‌نظمی شود که بدین‌منظور بهره اطلاعات[1] محاسبه می‌شود.

منظور از بهره اطلاعات یک ویژگی، مقدار کاهش آنتروپی است که به واسطه جداسازی داده‌ها از طریق این ویژگی حاصل می‌شود. به عبارت دیگر بهره اطلاعات برای یک ویژگی نظیر A نسبت به مجموعه داده‌های S از طریق فرمول زیر محاسبه می‌شود:

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

که در آن؛

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

بنابراین برای یافتن مهم‌ترین ویژگی در مجموعه داده‌ها ابتدا براساس هدف طبقه‌بندی باید آنتروپی کل را حساب کرده و سپس برای تک‌تک متغیرها، بهره اطلاعات محاسبه شود و متغیری که بهره اطلاعات آن از بقیه متغیرها بیشتر است در ریشه قرار داده شود (شهرابی و ذوالقدرشجاعی، ۱۳۹۰).

CART یک درخت باینری است این بدان معناست که هر گره بدون برگ دقیقا دو شاخه دارد. در این الگوریتم برای ساخت درخت از ضریب جینی[2] استفاده می‌شود که یک معیار ناخالصی است و پراکندگی تقسیم را توصیف می‌کند. به عبارت دیگر ضریب جینی برابر با واریانس متغیر تصادفی برنولی حاصل از استخراج یک نمونه تصادفی (با جایگذاری) از این گره و مشاهده کلاس آن می‌باشد. بنابراین می‌توان کاهش ناخالصی را به عنوان کاهش وایانس در نظر گرفت و هدف به دست آوردن گره‌هایی با واریانس صفر در برچسب کلاس می‌باشد (مشکانی و ناظمی، ۱۳۸۸). به زبان ریاضی ضریب جینی در گره σ 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 طبقه j گره σ قرار دارند 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 بهره ضریب جینی آن ماکسیمم باشد انتخاب می‌شود.

در الگوریتم CHAID مجموعه داده‌ها را می‌توان به بیش از دو گروه تقسیم کرد. در این الگوریتم هرس انجام نمی‌شود. در عوض از آزمون استقلال خی‌دو برای تصمیم‌گیری در مورد قانون تقسیم برای هر گره استفاده می‌کند. برای هر متغیر ورودی، کلاس‌ها بر اساس شباهت آماری خود در یک ابرکلاس[3] 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 عبارتند از: ۱) نسبت به نقاط پرت نیرومند[4] است. ۲) نمایش راه‌حل در قالب فلوچارت است و بنابراین درک مدل 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، بنابراین آماده‌سازی داده‌ها ساده‌تر است. ۶) می‌تواند مقادیر گمشده را مدیریت کند (وندلر و گرات‌آپ، ۲۰۲۱؛ وانگ و لی[5]، ۲۰۰۶).

---

[1] . Information gain

[2] . Gini

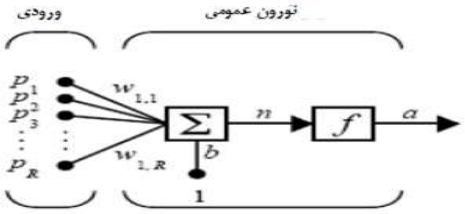
[3] . super-class

[4] . Robust

[5] . Wang & Lee

شبکه عصبی

شبکه عصبی مصنوعی شاخه‌ای از هوش مصنوعی و روشی مناسب برای تشخیص الگوهای ناشناخته در داده‌ها است. منشا الهام این ابزارها، سیستم‌های بیولوژیکی و سیستم‌های طبیعی بدن انسان هستند. خروجی مطلوب در شبکه عصبی مصنوعی نیز مانند شبکه عصبی انسان طی دو مرحله آموزش و یادگیری حاصل می‌شود. شبکه‌های عصبی از یک سری لایه‌هایی شامل نورون‌ها تشکیل شده‌اند که به صورت موازی با هم عمل می‌کنند و اساس عملکرد شبکه‌های عصبی را تشکیل می‌دهند. به عبارت دقیق‌تر شبکه عصبی شامل سه لایه ورودی، پنهان یا میانی و خروجی می‌باشد. لایه ورودی جهت معرفی پارامترها به شبکه است و یک لایه انتقال‌دهنده محسوب می‌شود. لایه پنهان بین لایه ورودی و خروجی است و نقش پردازشگری اطلاعات را برعهده دارد. لایه خروجی محل قرارگیری مقادیر خروجی شبکه است. در شکل ۱ یک نورون ساده با R ورودی نشان داده شده است.



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

خروجی نورون به وسیله معادله زیر تعیین می‌شود:

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

که در این رابطه b مقدار ثابت است و به عنوان مولفه مرز تشخیص استفاده می‌شود و می‌تواند نوع طبقه‌بندی داده‌ها را تغییر دهد. بنابراین وجود مقدار b در مدل ضروری است. به f تابع محرک یا تابع فعال‌سازی می‌گویند که معمولا یک تابع غیرخطی است. لازم به ذکر است پارامترهای b و w قابل تنظیم هستند و تابع f هم توسط پژوهشگر انتخاب می‌شود. به طور کلی به این سه مرحله در شکل بالا یعنی ورودی، وزن‌ها و تابع محرک به اصطلاح یک لایه گفته می‌شود که باید به این لایه آموزش[2] داده شود. لازم به ذکر است که آموزش به معنی پیدا کردن وزن‌هایی است که ورودی را به خروجی مورد نظر تبدیل کند (هاسون[3]، ۱۹۹۵). بنابراین شبکه 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. بدین‌منظور فرآیند یادگیری را لایه به لایه انجام می‌دهد تا در نهایت اختلاف بین مقادیر مطلوب و مقادیر پیش‌بینی شده شبکه به حداقل ممکن برسد (هان و همکاران، ۲۰۱۱).

از طرفی نتیجه (خروجی) یک شبکه عصبی به صورت باینری ارائه می‌شود، بنابراین اساس کار یک نورون در شبکه‌های عصبی، ترکیب گزاره‌های منطقی و ایجاد یک گزاره 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 می‌توانند طبقه‌بندی را انجام دهند. در جدول ۲ ترکیب‌های عبارات به صورت ریاضی نشان داده شده است. به عنوان توضیح بیشتر در مورد گزاره‌های منطقی، فرض کنید دو متغیر ورودی X و Y وجود دارد که هر یک شامل ویژگی‌هایی هستند در آن صورت نتیجه حاصل از ترکیب عطفی درست خواهد بود اگر و فقط اگر هر دو عبارت ساده تشکیل‌دهنده آن درست باشند و نتیجه حاصل از ترکیب فصلی، درست خواهد بود اگر و فقط اگر حداقل یکی از عبارت‌های تشکیل‌دهنده آن درست باشد و در آخر نتیجه حاصل از ترکیب فصلی ضمنی درست خواهد بود اگر و تنها اگر دو عبارت ساده تشکیل‌دهنده آن مثل هم نباشند.

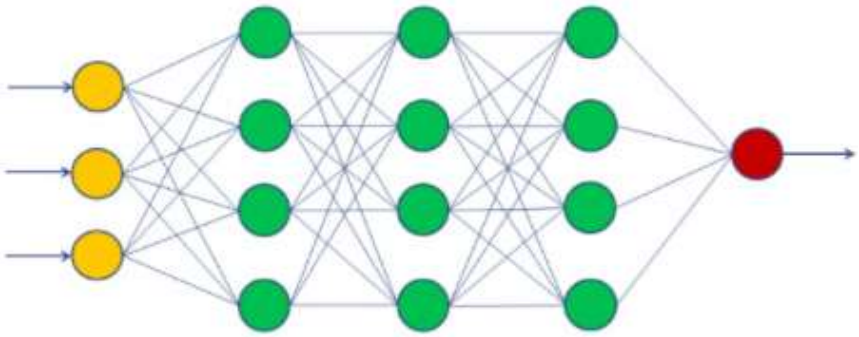
[1] . Demuth , & Beale
[2] . train
[3]. Hassoun

جدول ۲. گزاره‌های منطقی در ریاضیات

عبارت ساده اول	عبارت ساده دوم	ترکیب عطفی	ترکیب فصلی	ترکیب فصلی ضمنی
X	Y	X∧Y	X ∨ Y	XOR
۰	۰	۰	۰	۰
۰	۱	۰	۱	۱
۱	۰	۰	۱	۱
۱	۱	۱	۱	۰

از آنجایی که همیشه با یک نورون نمی‌توان به جواب رسید شبکه‌های عصبی پیشرفته‌تری مطرح شدند. به طور کلی دو نوع شبکه عصبی پیشخور[۱] (FFN) و بازگشتی[۲] (RNN) وجود دارد که تفاوت آنها در این است که در شبکه‌های عصبی بازگشتی، حداقل یک سیگنال برگشتی از یک نورون به همان نورون یا نورون‌های همان لایه و یا لایه قبل وجود دارد (منهاج، ۱۳۹۷).

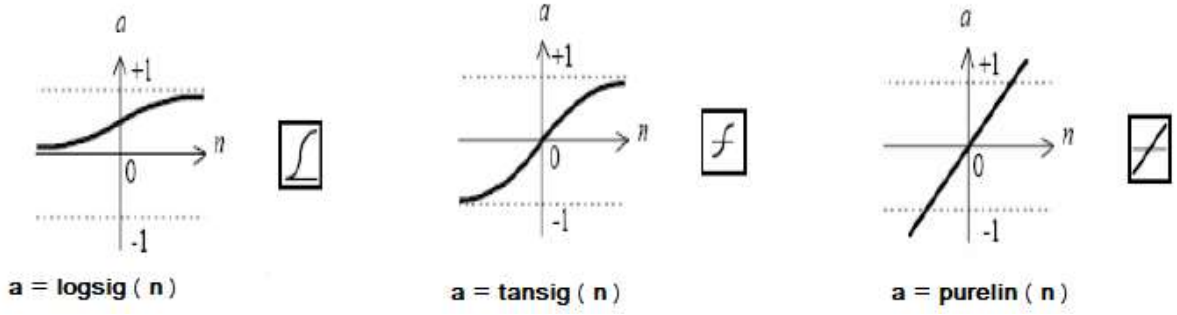
یکی از شبکه‌های عصبی پیشخور، شبکه عصبی پرسپترون[۳] است که قادر هستند با انتخاب مناسب تعداد لایه‌ها و سلول‌های عصبی، یک نگاشت غیرخطی با دقت دلخواه ایجاد کنند. شبکه‌های پرسپترون به صورت تک‌لایه و چندلایه ارائه می‌شوند. شبکه‌های پرسپترون چندلایه شبکه‌هایی کاملا مرتبط هستند بدین‌معنی که هر نورون در هر لایه به تمام نورون‌های لایه قبل متصل است (منهاج، ۱۳۹۷). شکل ۲ مثالی از یک شبکه پرسپترون با سه لایه پنهان است که در هر یک از این لایه‌ها چهار نورون وجود دارد. در لایه ورودی آن، سه نورون و یک نورون در لایه 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 شبکه عصبی پرسپترون چندلایه می‌توان به توابع لگاریتم سیگموئیدی، تانژانت سیگموئیدی و تابع محرک خطی اشاره کرد (دموت و بیل[۴]، ۲۰۰۰).

[1]. Feed Forward Network
[2]. Recurrent Neural Network
[3]. Perceptron
[4]. Demuth & Beale



شکل ۳ . توابع محرک رایج در شبکه عصبی پرسپترون چندلایه

در همه شبکه‌های عصبی دو مساله انتخاب یکی در مورد معماری شبکه و دیگری در مورد الگوریتم آموزشی مناسب مطرح است. منظور از معماری شبکه، انتخاب بهینه 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.

مهم‌ترین انتخاب ویژگی هر شبکه عصبی، توانایی یادگیری است. یعنی بر اساس یک الگوریتم یادگیری مشخص، وزن‌ها تغییر کنند تا حدی که میزان اتلاف شبکه مینیمم شود. برای یادگیری در شبکه پرسپترون از قاعده پس انتشار خطا[1] استفاده می‌شود. در واقع پس انتشار خطا با مقایسه بین خروجی مشاهده شده (واقعی) و خروجی مطلوب ضرایب وزنی شبکه را تغییر می‌دهد به طوری که در دفعات بعدی، خروجی صحیح‌تری حاصل شود. در این روش برای هر الگو، مجموع مربعات خطا طبق رابطه زیر محاسبه می‌شود:

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

که در آن t_{pk} و o_{pk} به ترتیب بیانگر مقادیر مطلوب و مقادیر مشاهده شده هستند (مشکانی و ناظمی، ۱۳۸۸). پس از محاسبه خطا، مقدار آن به صورت معکوس (از لایه خروجی به سمت لایه اول) در شبکه انتشار می‌یابد. سپس 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 درباره عملکرد داوطلبان آزمون سراسری، در این مطالعه به مقایسه قابلیت‌های سه الگوریتم مختلف درخت تصمیم (C5، CHAID و CART) و مدل شبکه عصبی پرداخته می‌شود. بدین منظور براساس مبانی نظری مطرح شده در پیشینه پژوهش از متغیرهای جنسیت، سوابق تحصیلی (معدل کتبی دیپلم و فاصله اخذ دیپلم و شرکت در آزمون سراسری) و نمره‌های تراز هر یک از دروس به عنوان متغیرهای موثر در طبقه‌بندی و پیش‌بینی سطح عملکرد داوطلبان استفاده گردید و بر این اساس به سوال‌های زیر پاسخ داده می‌شود:

۱. مناسب‌ترین مدل طبقه‌بندی داوطلبان آزمون سراسری براساس کلیه دروس با احتساب متغیرهای سوابق تحصیلی کدام است؟

۲. مناسب‌ترین مدل طبقه‌بندی داوطلبان آزمون سراسری براساس کلیه دروس با کنترل سایر متغیرها کدام است؟

لازم به ذکر است ایده نویسندگان برای افزودن متغیر فاصله اخذ دیپلم تا کنکور سراسری به فرآیند 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، لذا این پژوهش از نظر گردآوری داده‌ها جزء مطالعات توصیفی و از نظر تحلیلی جزء تحلیل ثانویه به شمار می‌آید. داده‌های این پژوهش شامل همه فارغ‌التحصیلان نظام جدید متوسطه ساکن استان البرز است که در گروه آزمایشی علوم ریاضی و فنی در آزمون سراسری شرکت کرده‌اند که ۸۳۰۵ نفر هستند.

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 شبکه عصبی به منظور طبقه‌بندی داوطلبان گروه ریاضی و فنی استفاده می‌شود. به منظور استفاده و تحلیل بهتر از

[1] backpropagation

درخت تصمیم لازم است که همه متغیرهای کمّی به متغیرهای طبقه‌بندی شده با طبقه‌های چهارتایی براساس چارک اول و دوم و سوم تبدیل شدند. همچنین برای به کارگیری درست شبکه عصبی لازم است که مقیاس داده‌ها یکسان شوند لذا ابتدا متغیرهای کمّی استاندارد و سپس مورد استفاده قرار گرفتند.

براساس مبانی نظری مطرح شده در پیشینه پژوهش از متغیرهای جنسیت، سوابق تحصیلی (معدل کتبی دیپلم و فاصله اخذ دیپلم و شرکت در آزمون سراسری)، سهمیه نهایی (شامل مناطق ۱ و ۲ و ۳ و شاهد) و نمره‌های تراز هر یک از دروس به عنوان متغیرهای موثر در طبقه‌بندی و پیش‌بینی سطح عملکرد داوطلبان استفاده گردید. قبل از استفاده از متغیرهای جنسیت و سهمیه که کیفی بودند تاثیر آنها در نمره کل داوطلبان با استفاده از آزمون t مستقل و تحلیل واریانس بررسی شد. نتایج نشان داد که دخترها و پسرها در نمره 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 مدل شد. هدف تحقیق حاضر طبقه‌بندی نامزدها بر اساس سطح عملکرد آنهاست. برای دستیابی به این هدف، نمره کل به عنوان معیاری برای ارزیابی سطح عملکرد عمل می‌کند و بنابراین 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 مرحله آموزش و آزمون انجام می‌شود در این پژوهش 80 درصد مجموعه داده‌ها به عنوان داده آموزشی و ۲۰ درصد داده‌ها به عنوان داده آزمون در نظر گرفته شد تا قابلیت اعتبار مدل سنجیده شود. در واقع با استفاده از داده‌های آزمون، دقت مدلی که آموزش داده شده است سنجیده می‌شود.

جدول ۳. وضعیت نمره‌های کل بر اساس 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 به ذکر است که تعداد لایه‌ها و تعداد نورون‌ها در هر لایه پنهان در شبکه عصبی به وسیله آزمون و خطا مشخص می‌شود (پاتل و میستری[1]، ۲۰۱۵).

یافته‌ها

با استفاده از نرم‌افزار SPSS Modeler نسخه ۱۸، به سوال‌های پژوهش در مورد انتخاب مدل مناسب طبقه‌بندی پاسخ داده شد. علی‌رغم اینکه جنسیت تاثیری در نتیجه آزمون ندارد ولی چون بر اساس آزمون t بین عملکرد دخترها و پسرها تفاوت معنی‌داری وجود داشت نتایج الگوریتم‌های استفاده شده با جنسیت و بدون جنسیت مورد بحث و بررسی قرار می‌گیرد. نتیجه بررسی یافته‌ها در سوال اول پژوهش مبنی بر تعیین دقت کلی مدل‌های شبکه عصبی از نوع پرسپترون چندلایه و درخت تصمیم با الگوریتم‌های C5، CHAID و CART در طبقه‌بندی داوطلبان آزمون سراسری، در جدول ۴ آورده شده است. با توجه به نتایج موجود در جدول ۴ می‌توان نتیجه گرفت که در بین همه انواع مدل‌های مورد مطالعه، مدل شبکه عصبی از نوع پرسپترون چندلایه، بیشترین دقت (۹۵/۱۵۳ و ۹۵/۶۳۱ و ۹۲/۸۱۹) را در طبقه‌بندی صحیح داوطلبان در هر دو حالت مقایسه تمام دروس تخصصی و عمومی با احتساب متغیرهای دیگر با جنسیت و بدون جنسیت و مقایسه تمام دروس تخصصی و عمومی دارد و الگوریتم C5 هم بیشترین دقت (۸۰/۰۱۲ و ۸۰/۴۳۱ و ۷۸/۵۱۶) را در بین الگوریتم‌های درخت تصمیم دارد. از طرفی دقت شبکه عصبی بدون احتساب جنسیت افزایش می‌یابد بدین معنی که در صورت وجود متغیرهای مربوط به سوابق تحصیلی بدون حضور جنسیت، شبکه عصبی با دقت بیشتری سطح عملکرد داوطلبان را تعیین می‌کند.

جدول ٤. دقت کلی برآورد شده هریک از الگوریتم‌ها بر اساس سوال‌های پژوهش

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

[1]. Patel & Mistry

در ادامه با توجه به دقت بالای شبکه عصبی در مقایسه با درخت تصمیم تنها به ارائه نتایج مربوط به شبکه عصبی اکتفا می‌شود.

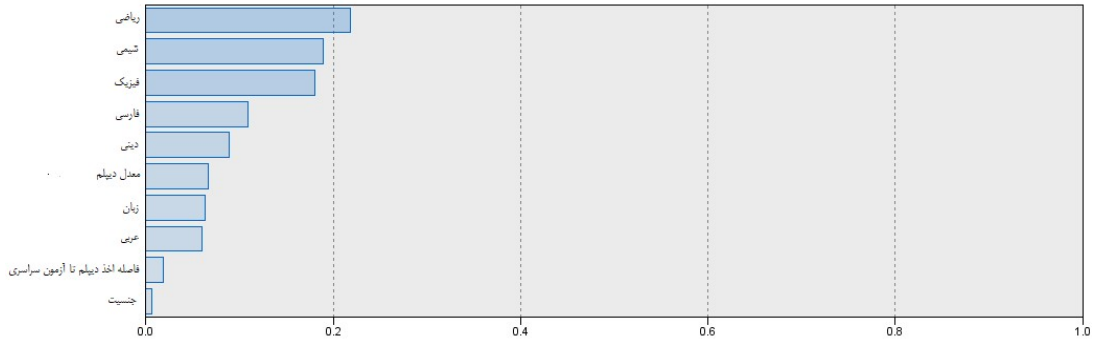
دقت مدل شبکه عصبی پرسپترون بر حسب فراوانی بدون در نظر گرفتن جنسیت در هر یک از مجموعه داده‌های آموزشی و آزمون در جدول ۵ آورده شده است.

جدول ۵ . نتایج حاصل ماتریس اغتشاش شبکه عصبی برای تمام دروس به انضمام متغیرهای دیگر بدون جنسیت

	داده‌های آموزش	داده‌های آزمون
درست	٦٣٨٤ (٩٦٪/٢٣)	١٦٠٧ (٩٦٪/١٦)
غلط	٢٥٠ (٣٪/٧٧)	٦٤ (٣٪/٨٤)

با استفاد از جدول ۵ نتیجه گرفته می‌شود که یادگیری شبکه به خوبی انجام گرفته است به طوری که وقتی داده‌های آزمون به مدل داده شد تنها کمتر از ۴ درصد از داده‌های آزمون به غلط پیش‌بینی شده و ۹۶ درصد به درستی در طبقه مناسب قرار داده شده‌اند.

در مورد سوال ۱ پژوهش که تمام دروس تخصصی و عمومی به انضمام متغیرهای جنسیت، سوابق تحصیلی و فاصله اخذ دیپلم تا زمان آزمون سراسری در نظر گرفته شده‌اند یک مدل شبکه عصبی از نوع پرسپترون با یازده نورون در لایه پنهان حاصل شد. میزان اهمیت متغیرهای استفاده شده در شبکه عصبی با احتساب جنسیت در شکل ۴ مشخص شده است.



شکل ٤. رتبه‌بندی همه متغیرها اعم از درسی و غیردرسی (سوال ۱ پژوهش) در شبکه عصبی با احتساب جنسیت

همان‌گونه که مشاهده می‌شود با استفاده از مدل شبکه عصبی، نمره درس ریاضی بیشترین تاثیر را در ارزیابی عملکرد داوطلبان گروه ریاضی و فنی داشته و از سوی دیگر فاصله اخذ دیپلم تا آزمون سراسری و جنسیت کمترین تاثیر را در ارزیابی عملکرد داشته‌اند به طوری که میزان اهمیت جنسیت در نتیجه عملکرد داوطلب ناچیز است.

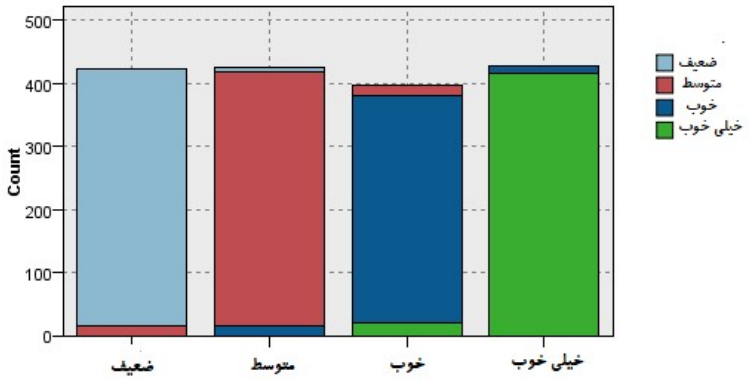
عملکرد شبکه عصبی توسط ماتریس اغتشاش[1] قابل ارزیابی است. در این ماتریس، سطرها و ستون‌ها به ترتیب بیانگر مقادیر واقعی و مقادیر پیش‌بینی شده هستند بدین معنی که هنگامی که عملکرد داوطلب به درستی تعیین شده است مقدار یک در سطر و ستون یکسان مربوط به سطح عملکرد داوطلب قرار می‌گیرد. به عنوان نمونه ماتریس اغتشاش برای شبکه عصبی که بیانگر میزان پیش‌بینی درست مدل شبکه عصبی است در جدول ۶ آورده شده است. با توجه به جدول ۶ مشاهده می‌شود که به عنوان مثال این شبکه ۹۷/۴٪ از داوطلبانی که در طبقه خیلی خوب قرار داشتند را به درستی پیش‌بینی کرده و تنها ۲/۶٪ از افرادی که در طبقه خیلی خوب بودند را به اشتباه در طبقه خوب قرار داده است.

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

پیش‌بینی شده / مشاهده شده	خیلی خوب	خوب	متوسط	ضعیف
خیلی خوب	٩٧٪/٤	٢٪/٦	٠٪/٠	٠٪/٠
خوب	٢٪/٤	٩٣٪/٩	٣٪/٧	٠٪/٠
متوسط	٠٪/٠	٣٪	٩٢٪/٩	٤٪
ضعیف	٠٪/٠	٠٪/٠	٢٪/٦	٩٧٪/٤

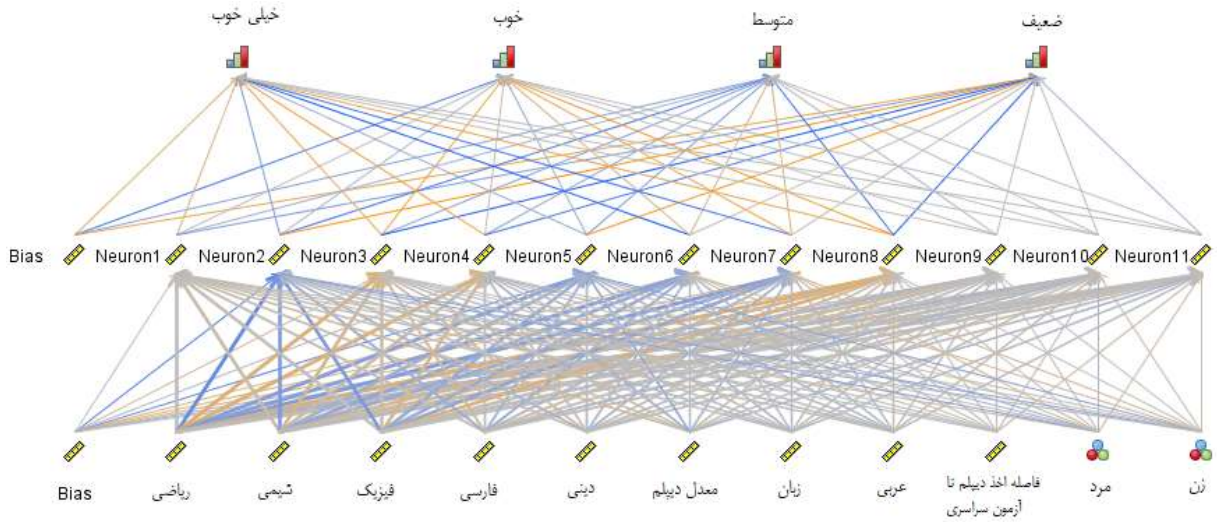
[1]. Confusion matrix

که شکل ۵ نیز تاییدی بر نتایج جدول ۶ است.



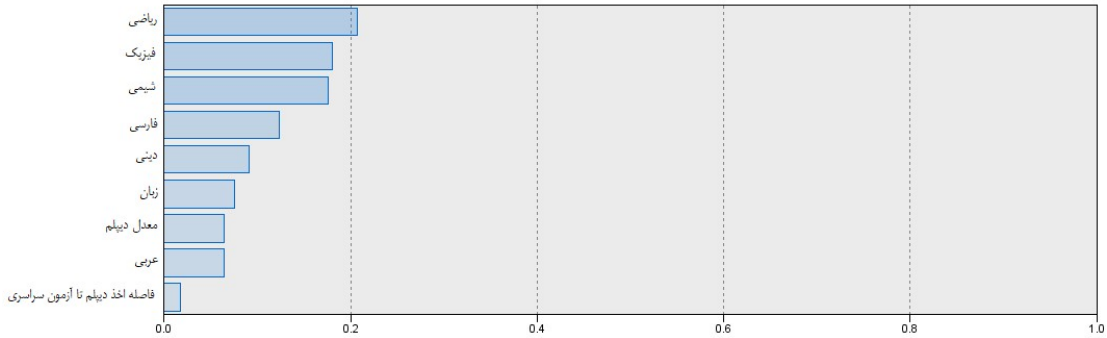
شکل ۵. وضعیت پیش‌بینی شبکه عصبی پرسپترون

نمودار شبکه عصبی پرسپترون با ۱۱ نورون در شکل ۶ آورده شده است.



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

در صورتی که متغیر جنسیت در نظر گرفته نشود تعداد نورون‌های لایه پنهان به ۹ نورون کاهش پیدا می‌کند. تنها ترتیب میزان اهمیت متغیرها تغییر می‌کند که این تغییر در واقعیت تا حدی منطقی‌تر است که نتیجه این تغییر در شکل ۷ آورده شده است. می‌توان اینگونه نتیجه گرفت که جنسیت در روایی نتیجه آزمون تاثیر می‌گذارد به طوری که دروس تخصصی بیشترین اهمیت را دارند و تاثیر معدل دیپلم از نمره کسب شده عربی در آزمون بیشتر شده است.



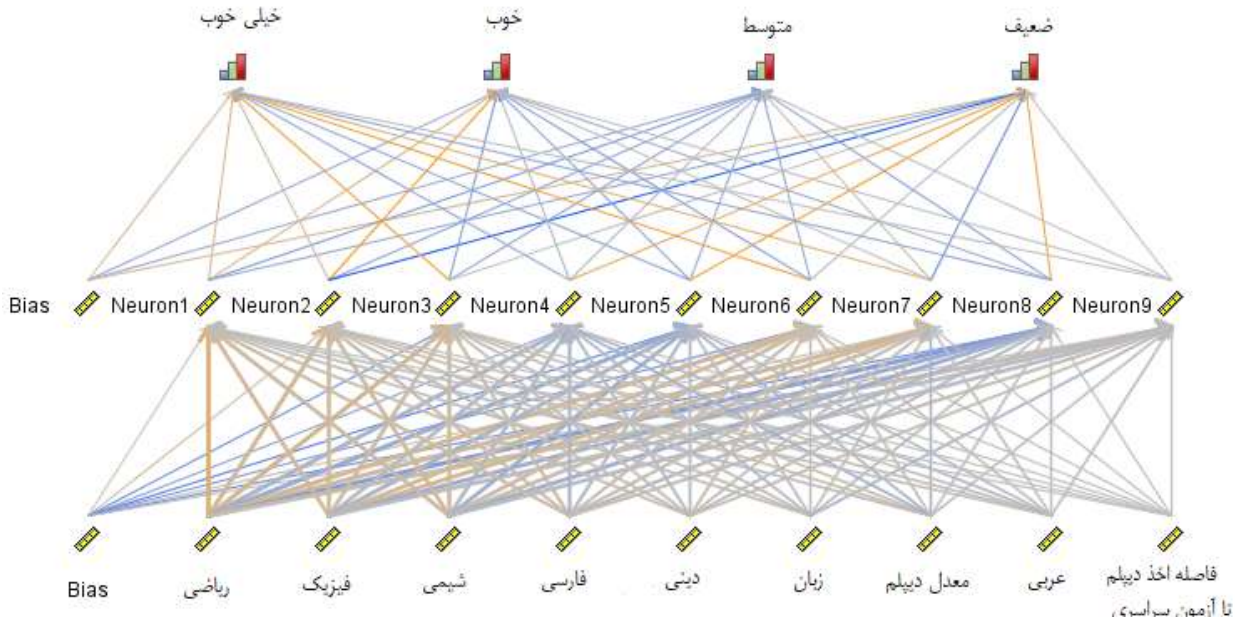
شکل۷. رتبه‌بندی همه متغیرها اعم از درسی و غیردرسی (سوال ۱ پژوهش) در شبکه عصبی بدون احتساب جنسیت

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

جدول ۷. دقت شبکه عصبی پرسپترون در هر یک از طبقه‌ها برای همه 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	۹٤٪/۷	۳٪/۵	۰٪/۰
متوسط	۰٪/۰	۲٪/۹	۹۳٪/٤	۳٪/۷
ضعیف	۰٪/۰	۰٪/۰	۳٪/۲	۹٦٪/۸

با مقایسه جدول‌های ۶ و ۷ به سادگی می‌توان دریافت که دقت شبکه عصبی بدون در نظر گرفتن جنسیت تا حدودی بهبود یافته است.



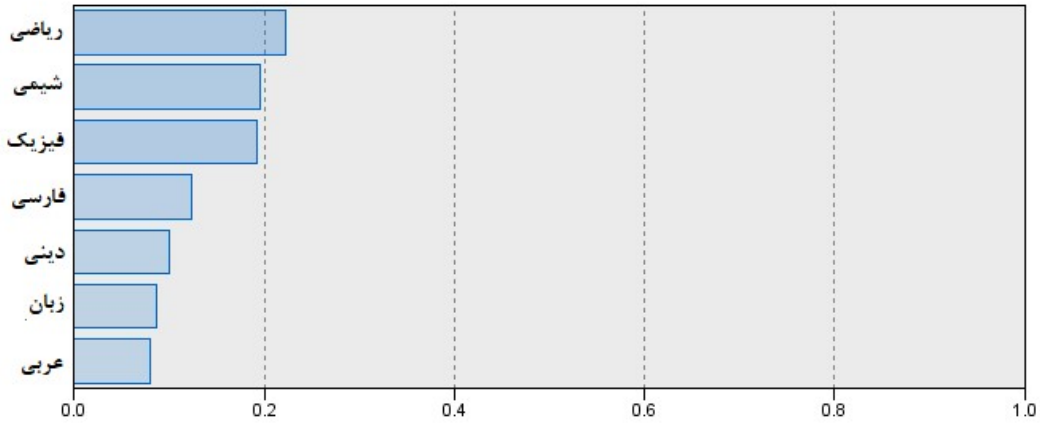
شکل ۸. نمودار شبکه عصبی پرسپترون با نه نورون میانی بدون احتساب جنسیت

در مورد سوال ۲ پژوهش که تنها نمره‌های دروس عمومی و تخصصی در نظر گرفته شده‌اند یک مدل شبکه عصبی از نوع پرسپترون با یازده نورون در لایه پنهان حاصل شد. دقت مدل شبکه عصبی پرسپترون بر حسب فراوانی در هر یک از مجموعه داده‌های آموزشی و آزمون در جدول ۸ آورده شده است.

جدول ۸ . نتایج حاصل از شبکه عصبی برای تمام دروس

	داده‌های آموزش	داده‌های آزمون
درست	6148 (۹۲٪/٦۷)	۱۵۵۱ (۹۲٪/۸۲)
غلط	٤۸٦ (۷٪/۳۳)	۱۲۰ (۷٪/18)

با مقایسه جدول‌های ۵ و ۸ اینگونه استنباط می‌شود که متغیرهای شامل سوابق تحصیلی بدون احتساب جنسیت که در سوال ۱ پژوهش حضور دارند در دقت مدل تاثیر بسزایی داشته‌اند و نیز باعث ساده‌تر شدن مدل شده‌اند. میزان اهمیت متغیرهای استفاده شده (سوال ۲ پژوهش) در شبکه عصبی در شکل ۹ مشخص شده است.



شکل 9. رتبه‌بندی همه متغیرهای درسی (سوال ۲ پژوهش) در شبکه عصبی

با مقایسه شکل‌های ۹ و ۷ نتیجه گرفته می‌شود که متغیرهایی غیر از نمره‌های دروس آزمون (یعنی سوابق تحصیلی) در ترتیب میزان اهمیت دروس فیزیک و شیمی تاثیر گذاشته‌اند. ولی از طرفی میزان تاثیر هر یک از دروس به نسبت شکل ۷ افزایش پیدا کرده است. میزان پیش‌بینی درست این مدل شبکه عصبی در هر یک از طبقه‌های مشخص شده در جدول ۹ آورده شده است.

جدول ۹. دقت شبکه عصبی پرسپترون در هر یک از طبقه‌ها برای تنها متغیرهای درسی (سوال ۲ 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	٤٪/۲	۸۹٪/۱	٦٪/۷
ضعیف	۰٪/۰	۰٪/۰	۵٪/۳	۹٤٪/۷

از جدول ۹ می‌توان نتیجه گرفت که ۹۷ درصد از داوطلبانی که به طور واقعی عملکرد خیلی خوبی داشته‌اند به درستی در طبقه خیلی خوب پیش‌بینی شده‌اند و تنها ۳ درصد از این داوطلبان در طبقه خوب به اشتباه 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 شبکه عصبی است. با مقایسه جدول‌های ۷ و ۹ می‌توان دریافت که حضور متغیرهای مربوط به سوابق تحصیلی در کنار نمره‌های دروس امتحانی منجر به افزایش دقت شبکه عصبی شده‌اند.

بحث و نتیجه‌گیری

آزمون‌ها مخصوصا آزمون‌های سرنوشت‌ساز مانند آزمون سراسری نقشی مهم در سلامت روان جامعه دانش‌آموزان دارند. بنابراین تا حد امکان لازم است هر چه بهتر و دقیق‌تر نتایج این آزمون‌ها ارزیابی شوند تا هر کسی 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 ادامه تحصیل به وجود آمده، داوطلبان بسیاری به دانشگاه راه می‌یابند. لذا در این پژوهش علیرغم قبولی و 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. بدین‌منظور بایستی از روش‌های آماری یا داده‌کاوی استفاده نمود. یکی از مهم‌ترین کاربردهای روش‌های آماری و داده‌کاوی که در علوم مختلف استفاده می‌شود شناسایی الگو و طبقه‌بندی است که بر اساس 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 عدم وجود هم‌خطی چندگانه کامل برقرار باشد که در صورت عدم برقراری این مفروضه‌ها، استفاده از این

روش‌ها امکان‌پذیر نیست و یا با خطای قابل‌توجه‌ای انجام می‌شوند که نمی‌توان به نتایج حاصله اعتماد کرد. از دیگر محدودیت‌های روش‌های آماری این است که به داده پرت و داده گمشده حساس هستند. مسلما با توجه به متغیرهای زیادی که برای ارزیابی داوطلبان در نظر گرفته می‌شود و همچنین تنوع زیاد دانشگاه‌ها و رشته‌ها، تعیین دقیق نتیجه 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 می‌شود که سطوح متغیرهای مستقل، گسسته محدود باشند و بین متغیرها رابطه خطی برقرار باشد. بدین معنی به منظور اجرای رگرسیون لوجستیک در این پژوهش لازم بود که متغیرهای مستقل به متغیرهای گسسته‌ای تبدیل شوند که بدین‌ترتیب اطلاعات زیادی از دست می‌رفت و نتایج قابل اعتمادی حاصل نمی‌شد (اسکندری، ۱۳۹۵) که در جدول ۱ که برگرفته از منبع شاهیری و حوسین (۲۰۱۵) است نیز تایید می‌شود که رگرسیون لوجستیک در رده الگوریتم‌های خطی طبقه‌بندی قرار می‌گیرد. در حالی که به منظور اجرای روش‌های داده‌کاوی نیاز به بررسی توزیع داده‌ها و هیچ پیش فرض دیگری نیست و همچنین این روش‌ها قادر هستند هر گونه رابطه خطی و غیرخطی را کشف کنند. از آنجایی 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 شبکه عصبی مصنوعی و درخت تصمیم برای طبقه‌بندی داوطلبان آزمون سراسری گروه ریاضی و فنی استفاده گردید و این داوطلبان از نظر نوع عملکرد طبقه‌بندی شدند.

در این مطالعه، از اطلاعاتی نظیر جنسیت، سوابق تحصیلی داوطلبان، سهمیه نهایی و نمره‌های علمی کسب شده در هر یک از دروس استفاده گردید. با استفاده از آزمون‌های آماری مشخص شد که دخترها و پسرها در نمره کل تفاوت معناداری با یکدیگر دارند که با مبانی نظری مطرح شده مطابقت دارد ولی بین داوطلبان سهمیه‌های مختلف. تفاوت معناداری وجود ندارد. لذا علیرغم بی‌تاثیر بودن جنسیت در نتیجه آزمون سراسری نتایج الگوریتم‌ها با جنسیت و بدون جنسیت بررسی شدند. از طرفی در هر مساله علم داده‌ای، قبل از شروع کار لازم است پیش‌پردازشی بر روی داده‌ها انجام شود که یکی از عملیات‌های مهم این فرآیند، هم‌مقیاس کردن داده‌ها است تا همگی تاثیر یکسانی در تحلیل داشته باشند. روش‌های مختلفی برای هم‌مقیاس کردن داده‌ها وجود دارد که در این مقاله از روش استانداردسازی استفاده شده‌است. با توجه به اینکه عملکرد داوطلبان با نمره کل نهایی مشخص می‌شود لذا چارک‌های نمره‌های کل نهایی داوطلبان مطابق جدول ۳ تعیین و به داده‌های مربوط به نمره کل چهار برچسب ضعیف، متوسط، خوب و خیلی خوب داده‌شد. داده‌های مورد استفاده در این پژوهش، اطلاعات مربوط ۸۳۰۵ نفر از داوطلبان استان البرز است که در گروه ریاضی و فیزیک شرکت کرده‌اند که 80 درصد این مجموعه داده به عنوان داده آموزشی و ۲۰ درصد به عنوان داده برای آزمون در نظر گرفته شد تا قابلیت اعتماد مدل سنجیده شود.

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

مشاهده شد که با حذف جنسیت و افزودن سایر متغیرها که مربوط به سوابق تحصیلی است دقت شبکه عصبی و همچنین الگوریتم c5 بهبود می‌یابد. به علاوه مشخص شد که در صورت به کارگیری تنها نمره‌های دروس، دقت شبکه 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، شبکه عصبی با دقت بیشتری سطح عملکرد داوطلبان را تعیین می‌کند و می‌توان نتیجه گرفت که جنسیت در سطح عملکرد داوطلب تاثیری ندارد. در دو الگوریتم c5 و CART، بهره اطلاعات درس فیزیک و در الگوریتم CHAID بهره اطلاعات درس ریاضی بیشتر شد که در نتیجه دراین الگوریتم‌ها فیزیک و ریاضی در ریشه 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 تاثیری نداشت و از طرفی دقت الگوریتم CART بدون حضور متغیرهای سوابق تحصیلی کاهش یافت ولی حضور یا عدم حضور متغیرهای سوابق تحصیلی تاثیری در دقت الگوریتم CHAID نداشت. به طور کلی دقت مدل‌های مربوط به درخت تصمیم در حضور متغیرهای دیگر از جمله جنسیت، معدل و فاصله اخذ دیپلم تا آزمون سراسری بیشتر از مدل دوم است.

در الگوریتم c5 با احتساب جنسیت، دروس ریاضی و فیزیک بیشترین تاثیر را داشتند و سپس به ترتیب عربی و معدل دیپلم و شیمی بر سطح عملکرد تاثیر می‌گذارند در این الگوریتم اهمیت جنسیت بیشتر از فاصله اخذ دیپلم تا آزمون سراسری است ولی به طور کلی میزان تاثیر هر دوی این متغیرها ناچیز است. به علاوه در صورتی که تنها نمره‌های دروس به عنوان متغیرهای تاثیرگذار در سطح عملکرد در نظر گرفته شود بیشترین درسی که در سطح عملکرد داوطلبان گروه ریاضی و فنی تاثیر 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 شبکه عصبی کمتر است. دلیل کم شدن دقت درخت تصمیم به این علت است که متغیرهای کمی به کیفی تبدیل شدند و بدین‌ترتیب اطلاعات زیادی از دست رفته است در صورتی که اگر متغیرها ذاتا کیفی بودند از درخت تصمیم نتایج رضایتبخشی به دست می‌آمد.

براساس شبکه عصبی نتیجه گرفته شد در صورتی که همه متغیرهای پژوهش اعم از درسی و غیردرسی بدون جنسیت در نظر گرفته شوند ترتیب اهمیت دروس تخصصی همانند مقدار تعیین شده توسط سازمان سنجش آموزش کشور است و تنها میزان اهمیت معدل دیپلم بیشتر از درس عربی و فاصله اخذ دیپلم تا آزمون سراسری است. این بدین معنی است که معدل دیپلم در طبقه‌بندی داوطلبان بر اساس سطح عملکرد بیشتر از درس عربی موثر است که نتیجه حاصل شده می‌تواند به روایی آزمون مرتبط باشد. در صورتی که طبقه‌بندی تنها بر اساس متغیرهای درسی صورت گیرد به ترتیب دروس ریاضی، شیمی، فیزیک بیشترین اهمیت را دارند و درس عربی کمترین تاثیر را دارد. نکته جالب توجه اینجاست که ترتیب اهمیت دروس عمومی در این دو مدل (بر اساس سوال اول و دوم پژوهش) تغییری نکرده است و تنها در مدل دوم، اهمیت درس شیمی بیشتر از فیزیک شده است.

همانگونه که مشاهده می‌شود متغیرهای سوابق تحصیلی و فاصله اخذ دیپلم تا آزمون سراسری در مدل شبکه عصبی تاثیر کمتری دارند. همچنین طبق تحقیق‌های انجام شده مشخص شد که در صورت نادیده گرفتن متغیر فاصله اخذ دیپلم تا آزمون سراسری، این تناقص 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 مقایسه مقادیرحاصل شده در جدول‌های ۶ و ۷ و ۹ می‌توان نتیجه گرفت که حضور متغیرهایی غیر از دروس آزمون باعث بالا رفتن دقت شبکه عصبی می‌شود و این به معنای قدرت شبکه عصبی است که با وجود متغیرهای مختلف می‌تواند بهترین طبقه‌بندی و با کمترین خطا را انجام دهد. همچنین این نتیجه در مورد درخت تصمیم نیز صادق است ولی از آنجایی که متغیرهای به کار رفته در این طبقه‌بندی، کمی 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 ادامه به منظور توسعه این پژوهش پنج پیشنهاد کاربردی و پژوهشی ارائه شده است:

۱) می‌توان مدل طبقه‌بندی مناسب را برای دختران و پسران به طورجداگانه بررسی نمود و نتیجه گرفت که آیا جنسیت در انتخاب مدل طبقه‌بندی تاثیر دارد یا خیر.

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

۳) مدل طبقه‌بندی مناسبی براساس متغیرهای اسمی مانند جنسیت و سهمیه نهایی انتخاب شود که در این حالت بهتر است از نمره‌های کل اصلی و نه نمره‌های کل طبقه‌بندی شده استفاده شود.

۴) مدل طبقه‌بندی مناسبی بر اساس قبولی و رد داوطلبان با استفاده از شبکه عصبی و رگرسیون لوجستیک انتخاب شود و نتایج با یکدیگر مقایسه شود.

۵) می‌توان از روش‌های دیگری برای تبدیل متغیرهای کمی به کیفی استفاده کرد و سپس مدل شبکه عصبی و درخت تصمیم را به کار گرفت.

References

Abdulrazzaq, N., Kamal, M., Muhsen, A., Tareq, A., Al Zubaidi, R., & Al Mousawi, A. (۲۰۱۷). Academic Failure And Student's Viewpoint: the Influence of Individual, Internal and External Organizational Factors .

Abu Tair, M. M., & El-Halees, A. M. (۲۰۱۲). Mining educational data to improve students' performance: a case study. *International Journal of Information,* ۲(2).

Afroz, G. (۱۳۷۵). *Educational Psychology*. Retrieved from

Angeline, D. M. D. (۲۰۱۳). Association rule generation for student performance analysis using apriori algorithm. *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA),* ۱(۱), ۱۶-۱۲.

Bakhshayesh, A. (۱۳۹۳). Investigating the relationship between thinking styles and learning strategies with academic performance in students. *Research Quarterly in Curriculum Planning,* ۱۱*(۱٤)*, ۱۳۵- ۱٤٦.

Banik, P., & Kumar, B. (۲۰۱۹). Impact of information literacy skill on students' academic performance in Bangladesh. *International Journal of European Studies,* ۳(۱), ۳۳-۲۷.

Bin Mat, U., Buniyamin, N., Arsad, P. M., & Kassim, R. (۲۰۱۳). *An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention.* Paper presented at the ۲۰۱۳IEEE ۵th conference on engineering education (ICEED).

Christian, T. M., & Ayub, M. (۲۰۱٤). *Exploration of classification using NBTree for predicting students' performance.* Paper presented at the ۲۰۱٤international conference on data and software engineering (ICODSE).

Demuth, H., & Beale, M .(۲۰۰۰) .Neural network toolbox user's guide .

Eskandari, F. (۱۳۹۵). *Categorical data analysis*. Tehran: Allameh Tabatabai University.

Fataa, L., Azari, S., Baradaran, H., & Atlasi, R. (۱۳۹۲). A systematic review of the causes of academic failure of medical students. *Journal of development steps in medical education,* ۱۰*(۲)*, ۳۱- .۳۸

Golshani-Foomani, M. (۱۳۷۵). *Sociology of Education*: shifteh.

Applied Statistics Research Group. (۱۳۹٤). *Constructing the total score of the national exam in the experimental group of mathematical and technical sciences in* ۲۰۱۳*based on the actual distribution of scores and comparing it with the current method.* Retrieved from Tehran

Han, J., Pei, J., & Kamber, M. (۲۰۱۱). *Data mining: concepts and techniques*: Elsevier.

Hassoun, M. H. (۱۹۹۵). *Fundamentals of artificial neural networks*: MIT press.

Ibrahim, Z., & Rusli, D. (۲۰۰۷). *Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression.* Paper presented at the ۲۱st Annual SAS Malaysia Forum, ۵th September.

Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (۲۰۱۵). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. *Decision Analytics,* ۲(۱), .۲۵-۱

Kazeminezhad, M., Etemad-Shahidi, A., & Mousavi, S. (۲۰۰۵). Application of fuzzy inference system in the prediction of wave parameters. *Ocean Engineering,* ۳۲(۱۵-۱٤), .۱۷۲۵-۱۷۰۹

Lawrence, A., & Deepa, T. (۲۰۱۳). Emotional Intelligence and Academic Achievement of High School Students in Kanyakumari District. *Online submission,* ۳(۲), .۱۰۷-۱۰۱

Li, K. F., Rusk, D., & Song, F. (۲۰۱۳). *Predicting student academic performance* .Paper presented at the ۲۰۱۳Seventh International Conference on Complex, Intelligent, and Software Intensive Systems.

Mau, W. C. (۱۹۹۷). Parental influences on the high school students' academic achievement: A comparison of Asian immigrants, Asian Americans, and White Americans. *Psychology in the Schools,* ۳٤(۳), ۲۷۷-۲٦۷.

Meit, S. S., Borges, N. J., Cubic, B. A., & Seibel, H. R. (۲۰۰٤). Personality Differences in Incoming Male and Female Medical Students. *Online submission* .

Menhaj, M. (۱۳۹۷). *Basics of Neural Networks - Volume One: Computational Intelligence*. Tehran: Amirkabir University of Technology.

Meshkani, A., & Naazemi, A. (۱۳۸۸). *An introduction to data mining*. Mashhad: Ali Meshkani Publications in cooperation with Islamic Azad University, Neyshabur branch.

Mitchell, T. M. (۱۹۹۷). Does machine learning really work? *AI magazine,* ۱۸(۳), ۱۱-۱۱.

Navidi, A. (۱۳۸۲). 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,* ۱۹*(٤)*, ۹۷- .۱۲۹

Oladokun, V., Adebanjo, A., & Charles-Owaba, O. (۲۰۰۸). Predicting students academic performance using artificial neural network: A case study of an engineering course .

Osmanbegovic, E., & Suljic, M. (۲۰۱۲). Data mining approach for predicting student performance. *Economic Review: Journal of Economics and Business,* ۱۰(۱), .۱۲-۳

Patel, P., & Mistry, K. (۲۰۱۵). A review: Text classification on social media data. *IOSR Journal of Computer Engineering,* ۱۷(۱), .۸٤-۸۰

Rahimi, M. (۱۳۸۷). *Investigating the factors related to the academic decline of female high school students in Region* ۱٤. (Master's thesis), University of Tehran ,.

Ramesh, V., Parkavi, P., & Ramar, K. (۲۰۱۳). Predicting student performance: a statistical and data mining approach. *International journal of computer applications,* ٦۳(8).

Sadeghi-Jafari, J., Roshan, M., & Shakoori-Ganjavi, H. (۱۳۹۰). 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,* ۳*(۱۰)*, ٥۱- .۷٥

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

Shahiri, A. M., & Husain, W. (۲۰۱٥). A review on predicting student's performance using data mining techniques. *Procedia Computer Science,* ۷۲, .٤۲۲-٤۱٤

Shahrabi, J., & ZolghadreShojae, A. (۱۳۹۰). *Advanced data mining concepts and algorithms*: Jihad University Publications of Amir Kabir University.

Simsek, A., & Balaban, J. (۲۰۱۰). Learning strategies of successful and unsuccessful university students. *Contemporary Educational Technology,* ۱(۱), .٤٥-۳٦

Talebi, A., & Akbari, Z. (۱۳۹۲). 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,* ٦۳, ۱۰۹- .۱۲۱

WANG, & LEE. (۲۰۰٦). Constructing a fuzzy decision tree by integrating fuzzy sets and entropy .*WSEAS transactions on information science and applications* .

Wendler, T., & Gröttrup, S. (۲۰۲۱). *Data mining with SPSS modeler: theory, exercises and solutions*: Springer.

Widyastuti, T., Kurniawan, A., & Chandra, N. P. (۲۰۱۷). Coping Strategies on Students After Experiencing Academic Failure: An Indigenous Study in Javanese Context. *Work Pap Ser,* ۳, -۲۲ .۲٦