

# ردپای هوش مصنوعی در آموزش آمار مدرسه‌ای

نرگس یافتیان<sup>۱\*</sup>، آناهیتا کمیجانی<sup>۲</sup> و ریحانه نیکنام اسکویی<sup>۲</sup>

<sup>۱،۲،۳</sup> گروه ریاضی، دانشکده علوم، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران

تاریخ دریافت: ۱۴۰۴/۰۷/۱۸

تاریخ پذیرش: ۱۴۰۵/۰۱/۲۴

## چکیده

این مطالعه، با هدف بررسی و مقایسه کیفی عملکرد چند ابزار مبتنی بر هوش مصنوعی در آموزش مفاهیم واریانس و انحراف معیار با رویکرد کیفی و به روش مطالعه موردی انجام شد. جامعه آماری، ابزارهای مبتنی بر هوش مصنوعی با کاربرد در آموزش آمار بود. نمونه‌گیری به صورت هدفمند و بر اساس دسترسی‌پذیری و توانایی تولید محتوای آماری انجام و پنج ابزار Gemini، ChatGPT، MathGPT، MagicSchool و DeepSeek انتخاب شد. داده‌ها نیز با چهار پرامپت جمع‌آوری شد و برای ارزیابی پاسخ‌های هوش مصنوعی به این چهار پرامپت، از معیارهایی مانند وضوح، جذابیت، ساختار، تعامل، ارتباط با دنیای واقعی و تناسب با دانش‌آموزان استفاده شد. تحلیل داده‌ها با استفاده از کدگذاری به کمک نرم‌افزار MAXQDA انجام شد و پایایی تحلیل‌ها با ضریب کاپا (۰/۸) تأیید گردید. یافته‌ها نشان داد که ابزارهای هوش مصنوعی می‌توانند با ارائه محتوای جذاب و مرتبط با دنیای واقعی، یادگیری آمار را بهبود بخشند، اما محدودیت‌هایی نظیر کمبود خلاقیت یا دقت ناکافی دارند. به طور کلی، در این زمینه، ابزارهای مختلف عملکرد متفاوتی از خود نشان دادند. Gemini و DeepSeek در ارائه طرح درس‌های دقیق و منسجم، ChatGPT در طراحی سؤالات تحلیلی و تصاویر بصری جذاب، MagicSchool در تسهیل فعالیت‌های گروهی و اصلاح خطاهای رایج، و MathGPT در ارائه محتوای قابل درک در ارتباط با مفاهیم واریانس و انحراف معیار، عملکرد نسبتاً بهتری داشته‌اند. ترکیب این ابزارها با نظارت معلم که برای حفظ سلامت آکادمیک ضروری است، می‌تواند سواد آماری و تفکر انتقادی را تقویت کرده و آنها را برای مواجهه با چالش‌های دنیای داده‌محور آماده کند.

**واژه‌های کلیدی:** آموزش آمار، واریانس و انحراف معیار، هوش مصنوعی، مقایسه قابلیت‌های ابزارهای مبتنی بر هوش مصنوعی.

## ۱ مقدمه و بیان مسئله

به دانش و مهارت‌های آماری نیاز دارند [۹]. آمار چیزی فراتر از مجموعه‌ای از رویه‌ها و تکنیک‌هایی است که باید آموخته شوند [۱۳]. برنامه‌های آموزشی مبتنی بر داده به دانش‌آموزان این امکان را می‌دهند که مهارت‌های لازم برای تحلیل داده‌های دنیای واقعی را کسب کنند [۱۴]. برای اینکه افراد بتوانند اطلاعات را به درستی تفسیر کرده و تصمیم‌گیری‌های آگاهانه انجام دهند، تقویت مهارت‌های آماری از سنین پایین ضروری است. به همین دلیل، نیاز به سواد داده‌ها، استدلال آماری و تفکر آماری، که شامل توانایی تفکر انتقادی و درک نحوه استفاده از

آمار یکی از موضوعات کلیدی گنجانده‌شده در برنامه‌های آموزشی مدارس است و نقش مهمی در توسعه مهارت‌ها و شایستگی‌های دانش‌آموزان برای مدیریت زندگی در عصر دیجیتال ایفا می‌کند [۶، ۴۶]. آموزش آمار باید بیش از هر چیز، بر پرورش درک عمیق مفاهیم آماری تمرکز داشته باشد [۴۳]، زیرا رویدادهای متعددی در دنیای واقعی رخ می‌دهند که برای تفسیر و تجزیه و تحلیل آنها، افراد

غیرقابل انکار است [۲۴]. با این حال، فرصت‌های یادگیری آمار در برنامه‌های درسی ریاضی اغلب بر جنبه‌های ریاضیاتی این رشته تأکید دارند و کمتر به مواردی مانند ارزیابی کیفیت سؤالات آماری، استفاده از نرم‌افزارهای آماری، مدیریت مجموعه‌های بزرگ و نامرتب داده‌ها و پذیرش عدم قطعیت در تفسیر نتایج تحقیقات آماری پرداخته می‌شود [۴۲]. فرصت‌ها و تدابیری مانند تقویت تجارب یادگیری عملی در دانش‌آموزان، آموزش با الهام از ردپای علم در محیط اطراف، بهره‌گیری از روش‌های ارزشیابی مؤثرتر و افزایش تدریس عملی می‌توانند توانایی دانش‌آموزان را در حل مسائل عملی آماری پرورش داده و کیفیت تدریس را ارتقا دهد [۴۳]. فرایند یاددهی-یادگیری مفاهیم آماری زمانی ثمربخش خواهد بود که دانش‌آموزان، علاوه بر یادگیری مفاهیم نظری اولیه، چگونگی به‌کارگیری این مفاهیم در مجموعه داده‌های واقعی را بیاموزند. به همین دلیل، استفاده از داده‌های واقعی در آموزش آمار برای توسعه درک مفهومی و داده‌های مفهومی مرتبط ضروری است [۴۳]. یکی از راهبردهای آموزشی مؤثر برای ایجاد انگیزه در دانش‌آموزان، تفسیر معنادار داده‌ها و فراهم کردن فرصت‌هایی است که در آنها دانش‌آموزان یاد بگیرند چگونه با استفاده هوشمندانه از داده‌های موجود، به حل مسائل پیش روی خود بپردازند و تصمیمات درست و مؤثری اتخاذ کنند [۴۲].

به‌کارگیری فناوری در ایجاد نگرش مثبت نسبت به یادگیری ریاضیات از دیرباز مورد توجه بوده است [۴۶]. این موضوع برای آمار نیز از نظر سند راهنمای GAISE II<sup>۱</sup> (چارچوبی برای آموزش و ارزیابی آمار در مقاطع پیش‌دبستانی تا دوازدهم) صادق است و هر جا که امکان‌پذیر باشد، باید از فناوری برای تقویت درک مفاهیم آماری و اجرای تحقیقات آماری بهره گرفت و تدریس آمار به سبک نوین بدون فناوری قابل تصور نیست [۲۱]. پیشرفت فناوری‌های کامپیوتری، به‌ویژه هوش مصنوعی، فرصتی فراهم می‌کند تا با تشخیص مشکلات یادگیری فردی دانش‌آموزان و ارائه حمایت‌های شخصی‌سازی‌شده، عملکرد یادگیری آنها را در دروس ریاضیات به‌ویژه آمار به حداکثر برسانند [۲۷]. تعاریف متعددی برای هوش مصنوعی وجود دارد که هر یک جنبه‌های مختلفی از این حوزه را مورد توجه قرار می‌دهند. یکی از تعاریف رایج و پذیرفته‌شده بیان می‌کند که هوش مصنوعی (AI) به فناوری‌هایی اشاره دارد که دارای توانایی مواجهه با موقعیت‌های جدید، حل مسائل، پاسخ به سؤالات، تدوین برنامه‌ها و سایر کارکردهایی هستند که با بهره‌گیری از روش‌هایی مبتنی بر رفتار هوشمند انسان‌ها و سایر موجودات برای حل مسائل پیچیده طراحی شده‌اند [۱۷]. این

داده‌ها برای اطلاع‌رسانی است، بیش از پیش احساس می‌شود [۴۲]. مهم‌ترین مهارت آماری، استدلال آماری است که به‌عنوان توانایی تفسیر مفاهیم، فرایندها و نتایج آماری، شناسایی روابط بین آنها و ارائه دلایل منطقی برای استفاده از آنها تعریف می‌شود [۴]. این مهارت به طرز فکر دانش‌آموز درباره ایده‌های آماری و درک اطلاعات آماری مرتبط است [۴۱].

نتایج پژوهش‌ها نشان داده‌اند، دانش‌آموزانی که از مهارت‌های استدلال آماری قوی‌تری برخوردارند، با سرعت و کارایی بیشتری به اهداف یادگیری خود دست می‌یابند. با این حال، بسیاری از دانش‌آموزان، با وجود کسب نمرات بالا در دروس آماری، در استدلال آماری عملکرد ضعیفی دارند و علاقه چندانی به این موضوع نشان نمی‌دهند. دشواری قوانین آماری و احساس اضطراب ناشی از مواجهه با حجم زیاد داده‌هایی که نیاز به تفسیرهای متفاوت و مهارت‌های تخصصی دارند، از جمله مشکلات شناخته‌شده در آموزش آمار هستند [۴۶]. بزرگ‌ترین مانع یادگیری آمار از جنبه‌های روان‌شناختی، مانند انگیزه پایین و نگرش منفی نسبت به یادگیری این موضوع، نشأت می‌گیرد. بنابراین، ایجاد نگرش مثبت در دانش‌آموزان به‌عنوان یک هدف ضمنی در آموزش مباحث آماری مطرح است [۱۱]. به‌کارگیری داده‌های واقعی در آموزش، علاوه بر تقویت کنجکاوی و انگیزه یادگیری در دانش‌آموزان، زمینه‌ای برای فهم عمیق‌تر مفاهیم و اصول آماری فراهم می‌کند [۱۲]. از سوی دیگر، استفاده از داده‌های واقعی در تهیه منابع آموزشی، توانایی تفکر انتقادی دانش‌آموزان را هنگام مواجهه با اطلاعات آماری در زندگی روزمره تقویت می‌کند [۳۳]. آموزش روش‌های مختلف تحلیل، مانند روش‌های آماری، گرافیکی و کیفی و گنجانیدن آموزش استفاده از نرم‌افزارهای آماری برای انجام تحلیل‌های پیچیده‌تر و استفاده از داده‌های واقعی و مرتبط با زندگی روزمره برای انجام تمرین‌ها و پروژه‌ها در کتاب درسی، تشویق دانش‌آموزان به پرسشگری در مورد نتایج و یافتن پاسخ‌های احتمالی می‌تواند در بهبود سواد آماری دانش‌آموزان مؤثر باشد [۲].

معلمان، به‌عنوان یکی از ارکان اصلی تربیت شهروندان آگاه، با چالش‌های متعددی برای ارتقای درک آماری دانش‌آموزان مواجه هستند تا بتوانند شهروندانی روشن‌فکر پرورش دهند که توانایی مدیریت، تجزیه و تحلیل و تفکر انتقادی درباره داده‌ها را داشته باشند. این رویکرد بهترین راه‌حل برای جهانی است که با اخبار جعلی محاصره شده یا نسبت به حقایق بی‌توجه است [۱۹]. اهمیت نقش معلم در بهبود مهارت‌های تفکر آنها برای استفاده از دانش آماری در حل مسائل روزمره

<sup>1</sup>Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II)

این مؤلفه‌ها، تعامل کاربران با مدل‌های هوش مصنوعی را مؤثرتر و هدفمندتر می‌سازد. به بیان دیگر، در مهندسی پرامپت، تکنیک‌های مؤثری مانند تخصیص نقش به مدل (مانند «فرض کن که تو متخصص ... هستی»)، ساختاربندی مرحله به مرحله پرامپت‌ها با رعایت قواعد نگارشی، افزودن عبارات تأییدی (مانند «آیا پرامپت کافی بود؟ سؤالی داری؟»)، تأکید بر شفافیت و دقت در نوشتن پرامپت‌های بدون ابهام، و اعمال محدودیت‌های صریح (مانند «نام شخص خاصی را ذکر نکن» یا «پاسخ را در چهار خط بنویس») کیفیت و کنترل‌پذیری خروجی‌های هوش مصنوعی مولد را به طور قابل توجهی ارتقا می‌دهند [۱]. در نوشتن پرامپت‌های درست و دقیق برای هوش مصنوعی توجه به تفکر الگوریتمی<sup>۲</sup> ضروری است.

تفکر الگوریتمی فرآیندی نظام‌مند برای حل مسئله است که بر توسعه گام‌های متوالی و منطقی جهت دستیابی به نتایج مطلوب تأکید دارد و هدف اصلی آن، فرموله‌سازی دقیق مراحل ضروری برای حل بهینه مسئله است. این نوع تفکر مستلزم مجموعه‌ای از توانایی‌ها از جمله تحلیل و تعریف دقیق مسئله، شناسایی اقدامات پایه برای طراحی الگوریتم صحیح، و بهبود مستمر کارایی آن است [۲۹]. همچنین، می‌توان تفکر الگوریتمی را به عنوان روشی رسمی‌شده از تفکر عملیاتی توصیف کرد که دارای جنبه‌های فکری (دانش الگوریتمی) و ابزاری (مهارت‌های اجرایی) است و پایه‌ای برای سواد دیجیتال و تعامل با سیستم‌های هوشمند مانند هوش مصنوعی فراهم می‌کند [۴۵]. تفکر الگوریتمی به معلمان کمک می‌کند تا پرامپت‌های داده شده به ابزارهای هوش مصنوعی را به گونه‌ای طراحی کنند که خروجی‌ها دقیق‌تر و مرتبط با مفاهیم واقعی باشند.

اکنون که هوش مصنوعی به‌عنوان یک فناوری نوظهور و در دسترس با تمام جنبه‌های زندگی ما ادغام شده است [۴۶]، می‌توان از آن برای تجزیه و تحلیل داده‌ها و آماده‌سازی دانش‌آموزان برای مشاغل آینده بهره گرفت [۴۳]. بر اساس دستورالعمل‌های سند راهنمای GAISE II با فراوانی داده‌هایی که روزانه از طریق اینترنت و سایر رسانه‌ها جمع‌آوری می‌شوند، یادگیری ماشین، یادگیری عمیق و هوش مصنوعی حوزه‌های روبه‌رشدی هستند که در آنها از مجموعه داده‌های بزرگ استفاده می‌شود [۲۱]. مدارس باید با گنجاندن و تشویق استفاده از هوش مصنوعی، راهبردهای آموزشی طراحی کنند که در آنها دانش‌آموزان بتوانند فرایندهای آموزشی مبتنی بر هوش مصنوعی را تجربه کنند و به یادگیرندگانی مادام‌العمر تبدیل شوند [۳۶]. هوش مصنوعی می‌تواند

تعریف، هوش مصنوعی را به‌عنوان ابزاری پویا و انعطاف‌پذیر معرفی می‌کند که می‌تواند فراتر از الگوهای ثابت و به‌گونه‌ای عمل کند که گویی از یادگیری و استدلال انسانی الهام گرفته است. یکی از انواع هوش مصنوعی که در سال‌های اخیر توجه زیادی را به خود جلب کرده‌اند، مدل‌های زبانی بزرگ هستند. این مدل‌ها با بهره‌گیری از مجموعه‌های عظیم داده‌های متنی و از طریق فرآیند آموزش گسترده، به‌گونه‌ای توسعه یافته‌اند که توانایی پردازش و درک زبان انسانی را با دقت و کارایی بالایی ارائه دهند [۴۹]. در واقع هوش مصنوعی به‌عنوان یک فناوری پیشرفته، برای ایجاد محتوای جدید، مانند تصاویر، موسیقی و متن، طراحی شده است که با استفاده از تکنیک‌های یادگیری عمیق برای تحلیل الگوها و روابط در مجموعه داده‌های بزرگ، خروجی‌های نوآورانه و خلاقانه در پاسخ به پرامپت‌های درخواست شده توسط کاربر، تولید می‌کنند.

گرویمینکو [۲۲] معتقد است پرامپت‌دهی به هوش مصنوعی مولد، برای جهت‌دهی دقیق به خروجی آن ضروری است و نوشتن حرفه‌ای آن نیازمند توجه به مولفه‌هایی مانند شفافیت، دقت، ایجاز و جلوگیری از هرگونه ابهام است. این مهارت، که به عنوان مهندسی پرامپت<sup>۳</sup> شناخته می‌شود، شامل نوشتن پرامپت‌های واضح و ساختارمند برای مدل‌های زبانی بزرگ است تا بدون آزمون و خطای زیاد، خروجی‌های باکیفیت تولید کنند. همچنین، پرامپت‌نویسی حرفه‌ای کمک می‌کند تا از مشکلات رایج مانند خروجی‌های نادرست<sup>۴</sup> جلوگیری شود و تعامل انسان با هوش مصنوعی را به سطحی هدفمند و کارآمد برساند [۳۰]. کاربر هوش مصنوعی باید مسئولیت کامل پیامدهای ناشی از استفاده از خروجی‌های تولیدشده توسط مدل را بر عهده گیرد. این اصل نه تنها از منظر اخلاقی مورد تأکید است، بلکه در عمل نیز ضروری به نظر می‌رسد و در مواردی که خروجی‌های هوش مصنوعی برای ارائه‌های عمومی، انتشار علمی یا اتخاذ تصمیمات مهم مورد استفاده قرار می‌گیرد، مسئولیت بررسی صحت، دقت و پیامدهای احتمالی آن صرفاً بر عهده کاربر است [۱]. در چارچوب پیشنهادی [۱۸]، مهندسی پرامپت بر شش عنصر کلیدی استوار است که با ایجاد ساختاری سیستماتیک، کیفیت و دقت خروجی‌های هوش مصنوعی مولد را ارتقا می‌دهند. این عناصر شامل فعل (مشخص‌کننده عمل مورد نظر)، تمرکز (تعریف محصول یا نتیجه انتظار)، زمینه (ارائه دامنه کلی)، تمرکز و شرط (تعیین موضوع دقیق و هدف)، هم‌راستایی (هماهنگی با اهداف خاص)، و محدودیت‌ها (اعمال قیود مانند طول متن یا منابع معتبر) می‌شوند. بهره‌گیری از

<sup>2</sup>Prompt engineering

<sup>3</sup>Hallucinations

<sup>4</sup>Algorithmic Thinking

شخصی‌سازی‌شده‌تر و جذاب‌تر کنند. با تکامل فناوری، روش‌های آموزش و یادگیری نیز تکامل می‌یابند [۳] و در بسیاری از موارد بهینه می‌شوند، به این ترتیب که با استفاده از فناوری برخی از فعالیت‌های زمان‌بر را در زمان کمتری و به شیوه بهتری انجام دهند، به عنوان مثال، استفاده از تخته‌ها و ابزارهای هوشمند به جای نوشتن بر روی تخته سیاه می‌تواند وقت ارزشمند کلاس را برای فعالیت‌های اثرگذارتری ذخیره کند. ابزارهای مبتنی بر هوش مصنوعی می‌توانند بازخوردهای لحظه‌ای، تصاویر کاربردی و توضیحات تعاملی متناسب با نیازهای هر دانش‌آموز ارائه دهند که یادگیری آمار را اثربخش‌تر می‌سازد [۲۵]. از سوی دیگر، تحلیل‌های اخیر حاکی از آن است که با توجه به تکامل سریع هوش مصنوعی و بهبود عملکرد آن در بسیاری از مهارت‌های انسانی، آشنایی معلمان با این فناوری و استفاده هوشمندانه از آن برای ارتقای کیفیت آموزش، به‌ویژه در دروسی مانند آمار که مبتنی بر استدلال و تحلیل هستند، ضروری است [۲۶]. بنابراین، چرخه موفق و پیشرفته آموزش آمار نیازمند پیوندی عمیق با فناوری‌های هوشمند است و معلمان باید مهارت‌های خود را در استفاده از هوش مصنوعی تقویت کنند تا فرایند یاددهی را جذاب‌تر و ماندگارتر سازند [۲۵].

پیشرفت‌های اخیر در حوزه هوش مصنوعی، ابزارهای قدرتمندی را برای بهبود آموزش آمار معرفی کرده است که می‌توانند فرایند یاددهی-یادگیری را جذاب‌تر و مؤثرتر کنند، برای مثال، Gemini، Perplexity، DeepSeek، MagicSchool، MathGPT، ChatGPT، Grok و Claude، Kimi، Scispace از جمله ابزارهای مبتنی بر هوش مصنوعی هستند که در بسیاری از زمینه‌ها می‌توانند مانند یک همکار در کنار معلم حضور داشته باشند و تلفیق تخصص و مهارت انسانی معلم و راهکارهای ارائه شده توسط این ابزارهای مبتنی بر هوش مصنوعی می‌تواند زمینه یک فرایند تدریس حرفه‌ای را فراهم آورد. در پژوهش‌های اخیر درباره تأثیرگذاری این ابزارهای مبتنی بر هوش مصنوعی در حوزه آموزش ریاضی و به ویژه آموزش آمار تأکید داشته‌اند. برای مثال، چونگ‌چیتان و همکاران [۱۶] معتقدند ChatGPT چشم‌انداز آموزشی را یک‌باره و به طور بنیادی تغییر داد و دانش‌آموزان تقریباً ناگهانی و به صورت رایگان، توانستند پاسخ‌های سوالات ریاضی خود را به کمک این ابزار دریافت کنند که بسیار فراتر از توانایی‌های دستیارهای هوش مصنوعی یا ابزارهای جستجویی مانند Google یا Siri بود. الیس و اسلید [۷] نیز، به بررسی چگونگی استفاده از ChatGPT برای تولید محتوای درسی و طراحی پرسش‌های امتحانی در آمار پرداختند. همچنین شینگ [۵۰] به بررسی استفاده از ChatGPT برای پاسخ‌دهی به سوالات مفهومی در آمار پرداخت و به

با ارائه مثال‌های کاربردی، تحلیل داده‌های پیچیده و ایجاد محیط‌های یادگیری تعاملی، درک دانش‌آموزان را از مفاهیم آماری تسهیل کند [۴۳] و به پرورش تفکر آماری عمیق‌تر آنها کمک نماید [۲۱]. دانش‌آموزان می‌توانند با استفاده از ابزارهای مبتنی بر هوش مصنوعی، اطلاعاتی که می‌تواند برای حل مسائل ریاضی و به‌ویژه موضوعات آماری، کمک‌کننده باشد را جست‌وجو کنند [۳۶]. برای مثال، این ابزارها می‌توانند داده‌های واقعی مانند آمارهای محلی یا جهانی را تحلیل کرده، الگوها را شناسایی کنند و به دانش‌آموزان کمک کنند تا نتایج را در زمینه‌های واقعی مانند پیش‌بینی روندهای اجتماعی یا ارزیابی داده‌های اقتصادی تفسیر کنند [۳۶]. همچنین، دانش‌آموزان با استفاده از مثال‌هایی که هوش مصنوعی ارائه می‌دهد، می‌توانند نظریه‌های آماری مانند توزیع‌ها یا استنباط‌های آماری را بهتر درک کنند [۴۴]، البته اگر استفاده دانش‌آموزان از ابزارهای مبتنی بر هوش مصنوعی به شیوه‌ای مدیریت شده نباشد، می‌تواند تأثیراتی منفی بر فرایند یادگیری آنها بگذارد، اما در صورتی که دانش‌آموزان به‌صورت فعال در فرایند تدریس، طراحی تمرین‌ها و سؤالات آزمون‌ها مشارکت داشته باشند و از ابتکار ذهنی خود بهره بگیرند، می‌توانند میثاق آمار را به‌گونه‌ای عمیق‌تر فرا بگیرند، در واقع، باید تلاش شود تا نقش دانش‌آموزان از صرف گوش دادن و خواندن به مشارکت فعال در فرایند آموزش تغییر کند [۴۴]. هدف از مقاله حاضر، معرفی تعدادی از ابزارهای مبتنی بر هوش مصنوعی است که می‌توانند در کنار معلم، برای انجام اقدامات مؤثرتر برای شکل‌گیری و تقویت مهارت‌های آماری دانش‌آموزان، مفید واقع شوند.

## ۲ اهمیت هوش مصنوعی در آموزش آمار مدرسه‌ای

پیشرفت‌های هوش مصنوعی فرصت‌های بی‌سابقه‌ای را برای تحول بنیادین در آموزش فراهم کرده است، به‌گونه‌ای که این فناوری نه تنها می‌تواند جایگزینی برای انجام برخی از وظایف معلمان باشد، بلکه قادر است شناخت و یادگیری عمیق‌تر دانش‌آموزان را تقویت نماید. هوش مصنوعی باید فراتر از سرعت بخشیدن به برخی از وظایف معلمان عمل کند و به‌عنوان ابزاری برای توانمندسازی آنها در راستای تدریس مؤثرتر، به‌ویژه در موضوعاتی اساسی مانند آمار، به کار گرفته شود [۲۵]. گزارش‌ها نشان می‌دهند که معلمان با بهره‌گیری از قابلیت‌های هوشمند هوش مصنوعی می‌توانند فرایند یاددهی-یادگیری آمار را متنوع‌تر،

۲. یک سوال یا فعالیت برای تدریس و ارزیابی مبحث واریانس و انحراف معیار طراحی کن که مبتنی بر زندگی واقعی دانش‌آموزان باشد. سعی کن در طراحی این سوالات دانش‌آموزان را به چالش بکشی تا بتوانند اطلاعات را از متن سوال استخراج کنند.

۳. یک ماجرا یا فعالیت کلاسی و فیزیکی طراحی کن که در طول آن دانش‌آموزان فرایندی را برای جمع‌آوری اطلاعات طی کنند و با استفاده از داده‌هایی که خودشان گردآوری کرده‌اند، واریانس و انحراف معیار را محاسبه کنند.

۴. در رابطه با ماجرا یا فعالیتی که ارائه دادی، تصویری طراحی کن، در طراحی این تصویر، فرهنگ ایرانی را در نظر بگیر.

برای تدوین و طراحی پرامپت‌ها، از نقطه‌نظرات سه متخصص (یک متخصص آموزش ریاضی، یک متخصص آموزش آمار و یک معلم آمار با تجربه) بهره گرفته شد تا پرامپت‌ها از نظر تناسب با سطح دانش‌آموزان پایه یازدهم، پوشش سطوح مختلف و تمرکز بر مفاهیم واریانس و انحراف معیار، معتبر باشند. برای تحلیل پاسخ‌ها نیز از معیارهای ارزیابی، وضوح و سادگی توضیحات، ایجاد انگیزه و جذابیت، ساختار و سازمان‌دهی، تعامل و فعالیت عملی، ارتباط با دنیای واقعی، و تناسب با دانش‌آموزان بدون پیش‌زمینه، با اقتباس از سیستم ارزیابی کیفیت منابع آموزشی دیجیتال تولیدشده توسط هوش مصنوعی، پیشنهادی هوانگ و همکاران [۲۷] بهره گرفته شد. محتوای به‌دست‌آمده از پاسخ‌های این پنج هوش مصنوعی به چهار پرامپت، با استفاده از نرم‌افزار MAXQDA کدگذاری، تحلیل و سپس مقایسه شد (در پیوست مقاله، تصویری از محیط کدگذاری در MAXQDA نیز ارائه شده است تا فرایند کدگذاری برای خوانندگان قابل مشاهده باشد)؛ ارزیابان در تحلیل پاسخ‌ها از رویکرد طراحی شده بر اساس رویکرد پژوهش حمود و همکاران [۲۴] استفاده نمودند. پایایی تحلیل‌ها نیز با محاسبه ضریب کاپای کوهن برای توافق بین کدگذاری‌های متخصصان تأیید شد که مقدار  $\frac{0}{8}$  به‌دست‌آمده، نشان‌دهنده سطح توافق مناسب بود.

## ۴ یافته‌ها

با توجه به اینکه هدف این پژوهش، مقایسه عملکرد پنج هوش مصنوعی است، ابتدا به معرفی محیط آنها پرداخته می‌شود: (۱) Gemini، دستیار هوش مصنوعی گوگل، امکان گفت‌وگوی متنی با کاربران را فراهم می‌کند و از جمله ابزارهای مفید در آموزش به شمار

این نتیجه رسید که این ابزار می‌تواند به برخی از این سوالات پاسخ دهد و توضیحات مناسبی نیز برای آنها ارائه کند. او معتقد است با توجه به دانش گسترده ChatGPT در آمار سنتی و علم داده معاصر، می‌توان آن را به‌عنوان ابزاری مؤثر در آموزش آمار مدرن ادغام کرد. با این حال، در پژوهش‌های داخلی، مطالعه‌ای که به کاربرد ابزارهای هوش مصنوعی بر فرایند آموزش و یادگیری در حوزه آموزش آمار پرداخته باشد، مشاهده نشد که با توجه به این خلأ و همچنین اهمیت هوش مصنوعی در این زمینه، هدف پژوهش حاضر، بررسی و مقایسه کیفی عملکرد پنج ابزار هوش مصنوعی (MagicSchool، MathGPT، ChatGPT، Gemini) و DeepSeek) بر اساس چهار پرامپت برای تولید محتوای آموزشی در ارتباط با تدریس مفاهیم واریانس و انحراف معیار به دانش‌آموزان پایه یازدهم است.

## ۳ روش پژوهش

این پژوهش با هدف بررسی عملکرد ابزارهای هوش مصنوعی منتخب در آموزش مفاهیم آماری واریانس و انحراف معیار با رویکرد کیفی و با روش مطالعه موردی انجام شد. جامعه آماری شامل ابزارهای هوش مصنوعی با کاربرد آموزشی بود و نمونه‌گیری به‌صورت هدفمند معیارمحور و بر مبنای چهار معیار همزمان دسترسی‌پذیری پایدار در ایران، توانایی تولید محتوای آموزشی روان به زبان فارسی، سابقه استفاده در پژوهش‌های معتبر آموزش ریاضی و آمار و بهره‌گیری از آخرین نسخه به‌روز هر ابزار انجام شد. برای این منظور از مدل‌های به‌روز هوش‌های مصنوعی (Math-، MagicSchool (1.4.2)، GPT (2025)، Gemini (5)، ChatGPT (5) و DeepSeek (3.1)) استفاده شد تا عملکرد آنها در آموزش مفاهیم آماری واریانس و انحراف معیار ارزیابی شود. انتخاب این پنج هوش مصنوعی از این جهت بود که پژوهش‌های متعددی در آموزش ریاضی و به ویژه آموزش آمار، از این هوش‌های مصنوعی استفاده کرده‌اند. برای جمع‌آوری داده‌ها، پاسخ‌های پنج ابزار مبتنی بر هوش مصنوعی به چهار پرامپت زیر، به‌صورت متنی و تصویری ثبت شد:

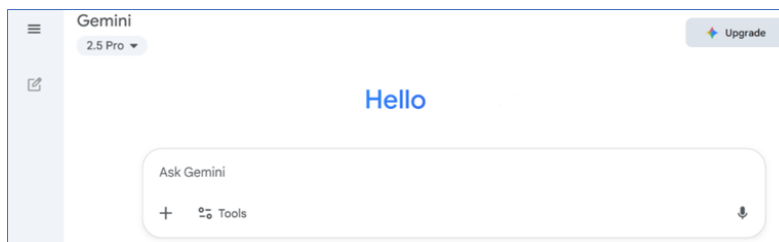
۱. من می‌خواهم مبحث واریانس و انحراف معیار را برای دانش‌آموزان پایه یازدهم رشته ریاضی-فیزیک تدریس کنم. طرح درسی برای ۳۰ دقیقه طراحی کن. این طرح درس طوری باشد که دانش‌آموزی که هیچ آشنایی با این بحث ندارد، بتواند به خوبی با جریان تدریس همراه شود.

می‌آید. این ابزار می‌تواند بر اساس اهداف آموزشی معلمان، طرح درس طراحی کند، سؤالات آموزشی مناسب تولید کند [۲۷] و تصاویر مرتبط با موضوعات آماری ارائه دهد تا یادگیری را تسهیل کند. شکل ۱ پیش‌نمایشی از محیط Gemini را نشان می‌دهد.

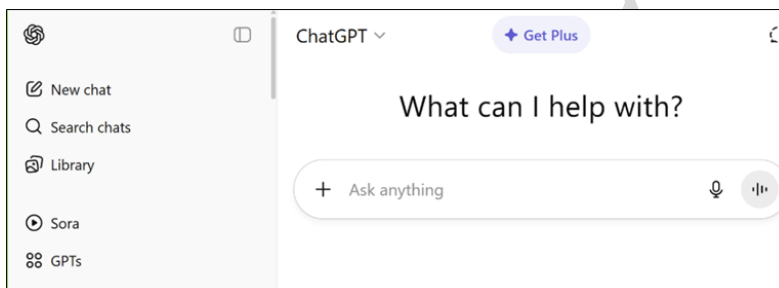
ChatGPT بستری را برای تعامل آسان با هوش مصنوعی فراهم

می‌آید. این ابزار می‌تواند بر اساس اهداف آموزشی معلمان، طرح درس طراحی کند، سؤالات آموزشی مناسب تولید کند [۲۷] و تصاویر مرتبط با موضوعات آماری ارائه دهد تا یادگیری را تسهیل کند. شکل ۱ پیش‌نمایشی از محیط Gemini را نشان می‌دهد.

ChatGPT بستری را برای تعامل آسان با هوش مصنوعی فراهم



شکل ۱: پیش‌نمایشی از محیط Gemini



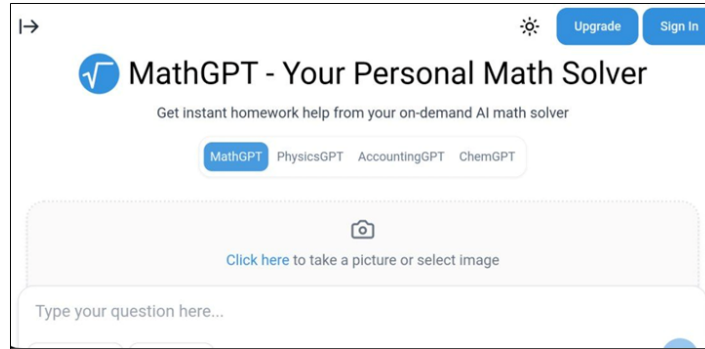
شکل ۲: پیش‌نمایشی از محیط ChatGPT

از محیط DeepSeek را نشان می‌دهد. یافته‌های پژوهش از تحلیل پاسخ‌های ابزارهای هوش مصنوعی به پرامپت‌های طراحی شده برای آموزش مفاهیم آماری استخراج شد. عملکرد این ابزارها با معیارهایی که در بخش روش پژوهش اشاره شد، مقایسه گردید و نقاط قوت و محدودیت‌های آن‌ها در بهبود فرایند یادگیری مفاهیم آماری واریانس و انحراف معیار مشخص شد. در ادامه، پرامپت‌ها همراه با پاسخ‌های هر ابزار ارائه شده است. شکل ۵، پرامپت ۱ را در محیط ChatGPT نشان می‌دهد.

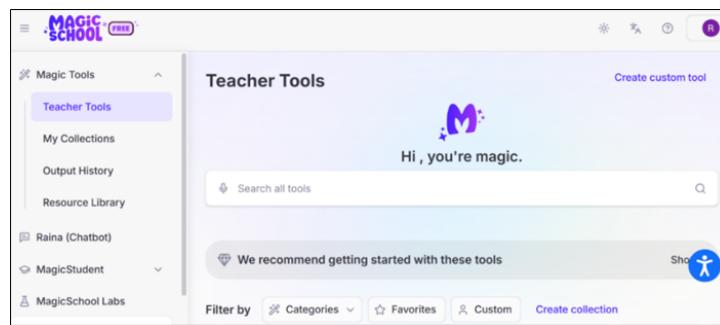
از محیط DeepSeek را نشان می‌دهد. یافته‌های پژوهش از تحلیل پاسخ‌های ابزارهای هوش مصنوعی به پرامپت‌های طراحی شده برای آموزش مفاهیم آماری استخراج شد. عملکرد این ابزارها با معیارهایی که در بخش روش پژوهش اشاره شد، مقایسه گردید و نقاط قوت و محدودیت‌های آن‌ها در بهبود فرایند یادگیری مفاهیم آماری واریانس و انحراف معیار مشخص شد. در ادامه، پرامپت‌ها همراه با پاسخ‌های هر ابزار ارائه شده است. شکل ۵، پرامپت ۱ را در محیط ChatGPT نشان می‌دهد.

از محیط DeepSeek را نشان می‌دهد. یافته‌های پژوهش از تحلیل پاسخ‌های ابزارهای هوش مصنوعی به پرامپت‌های طراحی شده برای آموزش مفاهیم آماری استخراج شد. عملکرد این ابزارها با معیارهایی که در بخش روش پژوهش اشاره شد، مقایسه گردید و نقاط قوت و محدودیت‌های آن‌ها در بهبود فرایند یادگیری مفاهیم آماری واریانس و انحراف معیار مشخص شد. در ادامه، پرامپت‌ها همراه با پاسخ‌های هر ابزار ارائه شده است. شکل ۵، پرامپت ۱ را در محیط ChatGPT نشان می‌دهد.

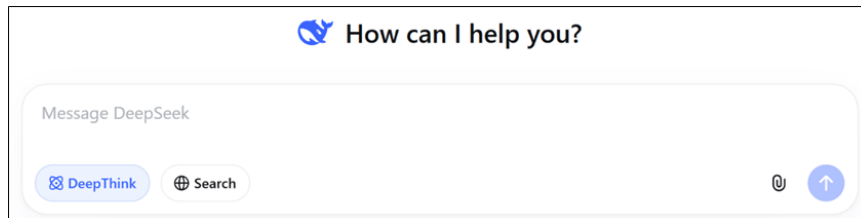
از محیط DeepSeek را نشان می‌دهد. یافته‌های پژوهش از تحلیل پاسخ‌های ابزارهای هوش مصنوعی به پرامپت‌های طراحی شده برای آموزش مفاهیم آماری استخراج شد. عملکرد این ابزارها با معیارهایی که در بخش روش پژوهش اشاره شد، مقایسه گردید و نقاط قوت و محدودیت‌های آن‌ها در بهبود فرایند یادگیری مفاهیم آماری واریانس و انحراف معیار مشخص شد. در ادامه، پرامپت‌ها همراه با پاسخ‌های هر ابزار ارائه شده است. شکل ۵، پرامپت ۱ را در محیط ChatGPT نشان می‌دهد.



شکل ۳: پیش‌نمایشی از محیط MathGPT



شکل ۴: پیش‌نمایشی از محیط MagicSchool



شکل ۵: پیش‌نمایشی از محیط DeepSeek



شکل ۶: پرامپت ۱ در محیط ChatGPT

## پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۱:

Gemini .۱

مرحله اول: مقدمه و ایجاد انگیزه (۵ دقیقه)

پرسش و پاسخ: تا به حال به این فکر کرده‌اید که چطور می‌توانیم پراکندگی نمرات یک کلاس را با یک عدد توصیف کنیم؟ مثلاً اگر میانگین نمرات دو کلاس یکسان باشد، آیا این به معنای یکسان بودن سطح هر دو کلاس است؟

فرض کنید دو بازیکن بسکتبال داریم که هر دو در ۵ بازی اخیر خود، میانگین ۲۰ امتیاز کسب کرده‌اند. بازیکن اول در بازی‌ها امتیازات ۱۸، ۲۲، ۱۹، ۲۱، ۲۰ را کسب کرده و بازیکن دوم ۵، ۳۵، ۱۰، ۳۰، ۲۰ را. به نظر شما کدام بازیکن عملکرد ثابت‌تری داشته؟ (اثبات بودن را با پراکندگی کمتر مرتبط کنید).

مرحله دوم: معرفی واریانس (۱۰ دقیقه)

یادآوری میانگین: برای اینکه بدانیم هر داده چقدر از میانگین دور است، ابتدا باید میانگین را محاسبه کنیم.

انحراف از میانگین: حالا می‌خواهیم ببینیم هر کدام از امتیازات چقدر از میانگین ۲۰ فاصله دارند. این فاصله را انحراف از میانگین می‌نامیم.

مشکل مجموع انحرافات و راه‌حل (واریانس): اگر این انحرافات را با هم جمع کنیم، حاصل صفر می‌شود، پس این روش به تنهایی نمی‌تواند پراکندگی را نشان دهد و برای غلبه بر این مشکل، هر یک از این انحرافات را به توان ۲ می‌رسانیم تا علامت منفی از بین برود. حالا این مربع انحرافات را با هم جمع می‌کنیم و بر تعداد داده‌ها (یا تعداد داده‌ها منهای یک برای نمونه) تقسیم می‌کنیم. این مقدار را واریانس می‌نامیم.

مرحله سوم: معرفی انحراف معیار (۱۰ دقیقه)

مشکل واحد واریانس: مثلاً اگر امتیازات را با واحد امتیاز اندازه گرفته باشیم، واریانس با واحد امتیاز مربع خواهد بود که امتیاز مربع برای ما قابل درک نیست.

معرفی انحراف معیار: برای اینکه به واحد اولیه برگردیم و یک معیار پراکندگی قابل تفسیرتر داشته باشیم، از واریانس جذر می‌گیریم. به این مقدار انحراف معیار می‌گوییم.

انحراف معیار نشان می‌دهد که داده‌ها به طور متوسط چقدر از میانگین فاصله دارند. هرچه انحراف معیار کمتر باشد، داده‌ها به میانگین نزدیک‌تر و پراکندگی کمتر است و بالعکس. حالاکه انحراف معیار هر دو بازیکن را داریم، می‌فهمیم بازیکن اول با انحراف معیار کمتر، ثبات بیشتری دارد زیرا داده‌هایش کمتر از میانگین فاصله دارند.

مرحله چهارم: جمع‌بندی و پرسش و پاسخ (۵ دقیقه)

ChatGP .۲

شروع انگیزشی (۵ دقیقه): پرسش از دانش‌آموزان: «اگر معدل دو کلاس یکی باشد، آیا سطح آنها یکسان است؟»

ارائه مفاهیم (۱۰ دقیقه) تعریف «پراکندگی»، «واریانس» و «انحراف معیار» با مثال‌های روزمره

حل تمرین (۱۰ دقیقه) حل یک مثال عددی ساده همراه با دانش‌آموزان (مثلاً نمرات ۵ دانش‌آموز)

جمع‌بندی و مشارکت (۵ دقیقه) پرسش و پاسخ، مرور مفاهیم، تعیین تکلیف کوتاه خانگی

- شروع با موقعیت ملموس: با طرح سوالی جالب شروع کنید: فرض کنید معدل دو کلاس ۱۶ است، ولی در یکی همه نمره نزدیک ۱۶ گرفته‌اند و در دیگری نمره‌ها بین ۱۲ و ۲۰ پراکنده است. کدام کلاس سطح علمی یکدست‌تری دارد؟
- معرفی مفاهیم میانگین، واریانس و انحراف معیار.
- تمرین تعاملی: مسئله‌ای ساده همراه با دانش‌آموزان حل کنید: اگر نمرات ریاضی ۵ دانش‌آموز به ترتیب ۱۲، ۱۴، ۱۶، ۱۸، ۲۰ باشد، انحراف معیار را حساب کنیم.
- به کاربردهای واقعی اشاره کنید: تحلیل نمرات دانش‌آموزان، نوسان قیمت‌ها در بازار، کنترل کیفیت در صنایع
- تکلیف پیشنهادی: از دانش‌آموزان بخواهید نمرات ۵ نفر از دوستان یا اعضای خانواده خود را جمع‌آوری و واریانس و انحراف معیار آن را محاسبه کنند.

MathGPT .۳

- مقدمه (۵ دقیقه): شروع با یک مثال ساده از زندگی روزمره (مثلاً قد دانش‌آموزان کلاس یا نمرات یک آزمون) و معرفی مفهوم پراکندگی داده‌ها به زبان ساده پرسش: «آیا همه داده‌ها دقیقاً برابرند؟ اگر نه، چطور می‌توانیم میزان تفاوت آنها را بسنجیم؟»
- معرفی واریانس (۱۰ دقیقه): تعریف واریانس به عنوان میانگین مربعات اختلاف داده‌ها از میانگین و توضیح گام به گام فرمول واریانس، بیان یک مثال عددی و تأکید بر اینکه واریانس عددی غیرمنفی است
- معرفی انحراف معیار (۵ دقیقه): تعریف انحراف معیار به عنوان جذر واریانس، بیان اینکه انحراف معیار واحد داده‌ها را حفظ می‌کند و راحت‌تر قابل فهم است
- تمرین کوتاه و مشارکت (۷ دقیقه): ارائه یک فعالیت که دانش‌آموزان به صورت گروهی یا فردی انجام دهند، بررسی پاسخ‌ها و رفع اشکال.
- جمع‌بندی و نتیجه‌گیری (۳ دقیقه): مرور نکات کلیدی: مفهوم پراکندگی، واریانس، انحراف معیار، بیان کاربردهایی در زندگی و علوم مختلف و پاسخ به سؤالات احتمالی دانش‌آموزان

(۵-۰ دقیقه) مقدمه و ایجاد انگیزه (پراکندگی چیست؟)

سوال: «میانگین نمرات هر دو کلاس ۱۹ شده. آیا این یعنی وضعیت نمرات در دو کلاس کاملاً یکسانه؟ تفاوت اصلی کجاست؟» علاوه بر میانگین، نیاز داریم بدانیم داده‌ها چقدر دور یا نزدیک به میانگین هستند (پراکندگی داده‌ها). دو روش کمی برای اندازه‌گیری این پراکندگی یاد می‌گیریم: واریانس و انحراف معیار.»

(۵-۱۵ دقیقه) مفهوم انحراف از میانگین و چالش آن

(۳ دقیقه) انحراف از میانگین: «برای فهم پراکندگی، اولین ایده اینه: ببینیم هر داده چقدر از میانگین فاصله داره. به این فاصله می‌گن انحراف از میانگین.»

سوال: «چرا جمع انحرافات همیشه صفر میشه؟» (چون داده‌های بالای میانگین و پایین میانگین همدیگر را خنثی می‌کنند).

جمع انحرافات به ما اطلاعات مفیدی درباره پراکندگی نمیده، چون همیشه صفر میشه. باید راه‌حل دیگه‌ای پیدا کنیم.

(۲ دقیقه) راه‌حل: مجذور انحرافات: «چطور می‌تونیم از شر علامت منفی خلاص بشیم و همزمان انحرافات بزرگ‌تر رو پررنگ‌تر کنیم؟»

(۱۵-۲۵ دقیقه) معرفی واریانس و انحراف معیار، نوشتن فرمول و حل یک مثال عددی ساده

(۳ دقیقه) مشکل واحد واریانس: سوال: «واحد واریانس چیه؟» (واحد مجذور واحد اصلی است که معنی فیزیکی واضحی ندارد).

(۲ دقیقه) انحراف معیار «برای حل مشکل واحد، کافیه جذر واریانس رو بگیریم. به این مقدار جدید می‌گن انحراف معیار (فرمول را بنویسید و انحراف معیار را برای کلاس‌ها محاسبه کنید)

(۳ دقیقه) تفسیر انحراف معیار: «انحراف معیار واحدش همون واحد داده‌های اصلیه (مثلاً نمره). انحراف معیار به ما می‌گه به طور میانگین، هر داده حدوداً چقدر از میانگین فاصله داره.»

ارتباط با دامنه تغییرات: «انحراف معیار معیار بهتری از دامنه تغییرات چون از همه داده‌ها استفاده می‌کنه، نه فقط ماکسیمم و مینیمم.»

(۲۵-۳۰ دقیقه) جمع‌بندی، مقایسه و نتیجه‌گیری

(۳ دقیقه) خلاصه مفاهیم کلیدی

(۲ دقیقه) تفاوت واریانس و انحراف معیار: تأکید بر واحد و این‌که انحراف معیار تفسیرپذیری مستقیم‌تری دارد.

نکات مهم برای اجرای موفق: تمرکز روی مفهوم و منطق پشت واریانس و انحراف معیار نه روی فرمول‌های پیچیده، تعامل بین اعضای کلاس، ارائه مثال‌های عینی، تخته‌نویسی منظم و برقراری ارتباط با دنیای واقعی

نکات کلیدی:

• انحراف معیار: معیاری از پراکندگی داده‌ها حول میانگین است.

• واریانس: مربع انحراف معیار و نشان‌دهنده میزان پراکندگی داده‌هاست.

• نکته رایج: دانش‌آموزان ممکن است تصور کنند که انحراف معیار همواره یک عدد بزرگ است، در حالی که می‌تواند کوچک باشد.

مقدمه: جلسه با یک سوال آغاز می‌شود: "چه چیزی داده‌ها را پراکنده می‌کند؟" و سپس، یک مثال از داده‌های واقعی (مانند نمرات امتحانات) برای جلب توجه دانش‌آموزان ارائه می‌شود.

معرفی مطالب جدید:

• توضیح مفهوم واریانس و انحراف معیار به صورت گام به گام (استفاده از نمودارها و جداول برای توضیح مفاهیم).

• یک تصور غلط رایج این است که انحراف معیار همیشه بزرگ است، در حالی که ممکن است در برخی داده‌ها کوچک باشد.

تمرین هدایت شده: دانش‌آموزان به گروه‌های کوچک تقسیم می‌شوند و با استفاده از داده‌های مشخص، واریانس و انحراف معیار را در سؤالاتی که به تدریج از آسان به سخت مطرح می‌شوند، محاسبه می‌کنند و معلم هم بر کار دانش‌آموزان نظارت می‌کند.

تمرین مستقل: دانش‌آموزان یک تکلیف انفرادی شامل محاسبه واریانس و انحراف معیار برای یک مجموعه داده جدید دریافت می‌کنند.

جمع‌بندی: جلسه با مرور نقاط کلیدی و پاسخ به سوالات باقی‌مانده به پایان می‌رسد و یک سوال نهایی برای بررسی درک دانش‌آموزان مطرح می‌شود.

فعالیت توسعه‌ای: دانش‌آموزانی که زودتر کار خود را تمام کردند، می‌توانند یک پروژه کوچک برای جمع‌آوری داده‌ها و محاسبه واریانس و انحراف معیار آنها انجام دهند.

تکلیف منزل: دانش‌آموزان باید یک مجموعه داده جدید جمع‌آوری کرده و واریانس و انحراف معیار آن را محاسبه کنند.

۱، پاسخ‌های بدست آمده از این ابزارها را در معیارهایی مانند وضوح و سادگی توضیحات، ایجاد انگیزه و جذابیت، ساختار و سازمان‌دهی، تعامل و فعالیت عملی، ارتباط با دنیای واقعی و مناسب بودن برای

مقایسه پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۱:

برای مقایسه و ارزیابی عملکرد پنج ابزار معرفی شده برای پرامپت

دانش‌آموزان بدون پیش‌زمینه بررسی می‌کنیم:

شده است. طرح درس ارائه شده توسط MathGPT ساختار زمانی مشخصی دارد و ترتیب مراحل تدریس نسبتاً مناسب طراحی شده است، اما زمان اختصاص یافته برای مرحله معرفی انحراف معیار، ممکن است کافی نباشد و روند تدریس را با چالش مواجه سازد. با وجود اینکه MagicSchool توانسته طرح درسی با ساختار مشخص ارائه کند، اما مشخص نبودن زمان‌بندی دقیقی برای مراحل تدریس، تا حدودی از کیفیت این طرح درس می‌کاهد.

• در بررسی عملکرد ابزارهای مبتنی بر هوش مصنوعی از حیث تعامل و فعالیت عملی می‌توان دریافت که تعامل در طرح درس ارائه شده توسط Gemini عمدتاً برپایه پرسش و پاسخ است و فعالیت عملی یا کار مشارکتی واضحی ذکر نشده است. اما ChatGPT توانسته از طریق تشویق معلمان به پشتیبانی دانش‌آموزان و راهنمایی آنها برای حل تمرین‌های عددی و همچنین ارائه فعالیت‌ها و تکالیفی برای محیط خارج از کلاس، بستر خوبی را برای تعامل فراهم آورد و بدین ترتیب کیفیت فرایند تدریس را ارتقا بخشد. به طور مشابه MathGPT نیز پیشنهاد انجام یک فعالیت گروهی یا فردی را ارائه داده است، اما جزئیات تمرین مشخص نیست. MagicSchool توانسته با تقسیم‌بندی دانش‌آموزان به گروه‌های کوچک، ارائه تمرینات هدایت‌شده و مستقل، زمینه تعاملی نسبتاً علمی و قوی را فراهم آورد، DeepSeek نیز ضمن حل مثال عددی، بستری را برای تعامل و پرسش و پاسخ فراهم می‌آورد اما کمتر به تمرینات عملی مستقل پرداخته است.

• در معیار ارتباط با دنیای واقعی، همانند مثال بازیکنان بسکتبال ارائه شده توسط Gemini که مفاهیم را تقریباً با زندگی واقعی مرتبط می‌کند، ChatGPT نیز به کاربردهایی از میحث واریانس و انحراف معیار مانند تحلیل نمرات دانش‌آموزان، نوسان قیمت‌ها در بازار و کنترل کیفیت در صنایع اشاره کرده است که علاوه بر ارتباط با دنیای واقعی، موضوعاتی نو و بدیع نیز هستند. همچنین، موضوع نمرات امتحان ارائه شده توسط MagicSchool ملموس و عینی است و پیشنهاد واگذاری پروژه‌های کوچک به دانش‌آموزانی با عملکرد قوی‌تر، جذاب به نظر می‌رسد و می‌تواند سایر دانش‌آموزان را هم برای داشتن عملکردی بهتر برانگیزد. اگرچه MathGPT به شکلی کلی به کاربردهای مفاهیم در علوم و زندگی اشاره کرده است. اما مثال‌های مشخص را در تعداد کمتری ارائه داده است.

• در معیار وضوح و سادگی توضیحات، به نظر می‌رسد ابزار Gemini با ارائه توضیحات گام‌به‌گام و بهره‌گیری از مثال‌های واقعی مانند مقایسه دو بازیکن بسکتبال برای نشان دادن تفاوت پراکندگی، عملکردی نسبتاً واضح و قابل فهم داشته است. ابزار ChatGPT مفاهیم را به صورت کلی و با مثال‌های روزمره توصیف کرده است، هرچند جزئیات محاسباتی آن در مقایسه با Gemini نسبتاً محدودتر است. توضیحات MagicSchool نیز تقریباً گام‌به‌گام و روشن است و با پرداختن به تصورات غلط رایج (مانند تفسیر نادرست بزرگی انحراف معیار)، جنبه آموزشی بیشتری دارد. DeepSeek توضیحاتی نسبتاً دقیق و منطقی ارائه داده است که بر حل مسئله صفر شدن جمع انحرافات، واحد واریانس و مثال‌های عددی ساده تمرکز دارد. در مقابل، MathGPT اگرچه توضیحاتی ساده با مثال‌های روزمره (مانند قد یا نمرات) عرضه کرده است، اما عمق آن در مقایسه با Gemini یا DeepSeek کمتر است.

• در معیار ایجاد انگیزه و جذابیت، Gemini توانسته است تا با ارائه مثال‌های با موضوع بازیکنان بسکتبال که مرتبط با زندگی واقعی نیز هست و با فراهم کردن زمینه‌ای برای بحث و گفت‌وگو و همچنین پرسش و پاسخ، دانش‌آموزان را درگیر کند. سؤالاتی با موضوع معدل نمرات دانش‌آموزان دو کلاس که توسط ChatGPT و DeepSeek مطرح شده‌اند، اگرچه ساده و روان هستند، اما به نظر می‌رسد که نسبت به مثال ارائه شده توسط Gemini از جذابیت کمتری برخوردار هستند. MathGPT و MagicSchool نیز مثال‌هایی مثل قد دانش‌آموزان یک کلاس یا نمرات یک آزمون ارائه داده‌اند که برگرفته از واقعیت‌های موجود در محیط اطراف هستند، اما از لحاظ جذابیت محتوا و نو بودن موضوعات، نسبت به Gemini ضعیف‌تر عمل کرده‌اند.

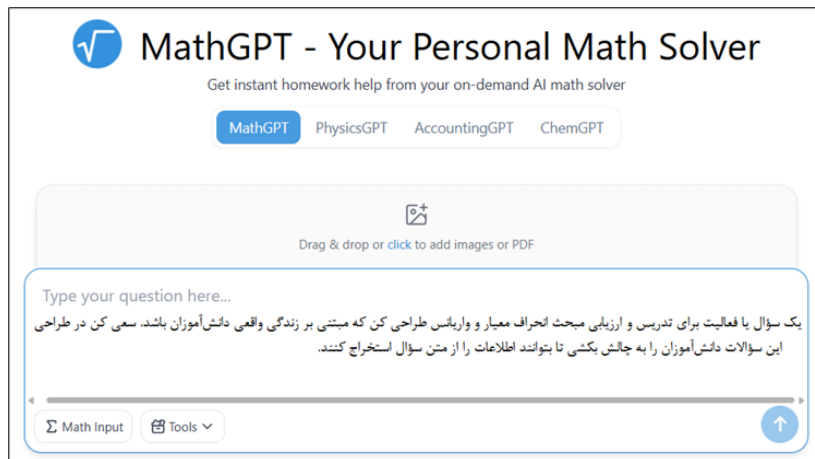
• در معیار ساختار و سازماندهی، Gemini و DeepSeek توانسته‌اند طرح درس‌هایی تقریباً دقیق ارائه دهند به طوری که مراحل انجام فرایند تدریس مبحث را به صورتی نسبتاً واضح و ریزبینانه تقسیم‌بندی کرده و توضیح داده‌اند، زمان‌بندی معقول و منطقی ارائه شده در این طرح درس‌ها می‌تواند به معلمان در اجرای آنها یاری رساند. اگرچه دقت ChatGPT در توضیح مراحل تدریس به خوبی دو ابزار قبلی نیست، اما ساختار مشخص و زمان‌بندی مناسبی در طرح درس ارائه شده، گنجانده

گام به گام، طرح مثال‌هایی ساده و حل تعاملی آنها برای آموزش مبتدیان تقریباً مفید واقع شوند، البته توضیحات ارائه شده توسط ChatGPT کمی کلی‌تر است و نیاز است تا خود معلم از دانش خویش در ارائه محتوا استفاده کند.

DeepSeek نیز با وجود اینکه توانسته ارتباط انحراف معیار با دامنه تغییرات و مثال نمرات را تا حد خوبی بیان کند، اما کمتر به کاربردهای گسترده پرداخته است.

- هر پنج ابزار مبتنی بر هوش مصنوعی به کار گرفته شده توانسته‌اند در ارائه محتوایی برای دانش‌آموزان بدون پیش‌زمینه قبلی، به خوبی عمل کنند و با ارائه توضیحات ساده، ملموس و

شکل ۷، پرامپت ۲ را در محیط MathGPT نشان می‌دهد.



شکل ۷: پرامپت ۲ در محیط MathGPT

پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۲:

۱. Gemini

یک تولیدکننده لامپ ادعا می‌کند که عمر مفید لامپ‌های تولیدی‌اش بسیار یکنواخت است. برای بررسی این ادعا، شما از دو خط تولید الف و ب به صورت تصادفی ۵ لامپ را انتخاب و عمر مفید آن را بر حسب ساعت ثبت می‌کنید:

خط تولید الف: ۹۹۵، ۱۰۰۵، ۹۹۰، ۱۰۱۰، ۱۰۰۰

خط تولید ب: ۱۰۵۰، ۹۵۰، ۱۰۰۰، ۱۱۰۰، ۹۰۰

اگر بخواهید ادعای تولیدکننده را از نظر یکنواختی بررسی کنید، کدام خط تولید ادعای تولیدکننده را بهتر پشتیبانی می‌کند؟ (میانگین عمر لامپ‌ها در هر دو خط تولید ۱۰۰۰ ساعت است).

۲. ChatGPT

بررسی هزینه روزانه خرید تنقلات

رضا در طول ۷ روز گذشته از بوفه مدرسه خرید کرده. مبلغ‌های خرج‌شده به صورت زیره:

روز اول: ۱۵ هزار تومان، روز دوم: ۲۰ هزار تومان، روز سوم: ۱۰ هزار تومان، روز چهارم: ۲۵ هزار تومان، روز پنجم: ۱۵ هزار تومان، روز ششم: ۳۰ هزار تومان، روز هفتم: ۲۰ هزار تومان

- داده‌ها را از متن استخراج کن و میانگین، واریانس و انحراف معیار مبلغ‌های خرج‌شده را محاسبه کن.

- تحلیل کن که آیا خرج‌کرد رضا در این هفته دارای پراکندگی زیادی بوده یا خیر؟

## MathGPT .۳

پنج دانش‌آموز تعداد قدم‌هایی که در یک روز برداشته‌اند را به این صورت گزارش کرده‌اند: امیر: ۷۵۰۰ قدم، فاطمه: ۸۲۰۰ قدم، رضا: ۶۹۰۰ قدم، نگار: ۷۸۰۰ قدم، سحر: ۸۱۰۰ قدم.  
 الف) میانگین تعداد قدم‌ها را به دست آورید.  
 ب) واریانس و انحراف معیار تعداد قدم‌ها را محاسبه کنید.  
 ج) اگر قرار باشد یک جایزه به کسی داده شود که تعداد قدم‌هایش به میانگین نزدیک‌تر است، چه کسی برنده می‌شود؟

## MagicSchool .۴

## فعالیت تعداد پیامک‌های روزانه

در یک هفته، تعداد پیامک‌هایی که هر روز برای دوستان خود می‌فرستید به این صورت است: ۱۰، ۱۲، ۸، ۱۵، ۹، ۶، ۱۴.  
 الف) میانگین، واریانس و انحراف معیار تعداد پیامک‌های روزانه خود را محاسبه کنید.  
 ب) فکر کنید اگر هر روز دقیقاً ۱۰ پیامک ارسال می‌کردید، انحراف معیار چه تغییری می‌کرد؟ چرا؟  
 پ) این اطلاعات چه کمکی به شما در مدیریت زمان و ارتباطاتتان می‌کند؟

## DeepSeek .۵

مطالعه برای امتحان و تأثیر آن بر نمرات  
 یک کلاس ۵ نفره، ساعات مطالعه هفتگی (بر حسب ساعت) و نمره امتحان ریاضی (از ۲۰) خود را گزارش کرده‌اند.

نمره امتحان	میزان ساعت مطالعه	دانش‌آموز
۱۶	۴	علی
۱۸	۷	مریم
۱۵	۵	رضا
۱۷	۹	زهرا
۱۴	۳	سامان

الف) انحراف معیار ساعات مطالعه و انحراف معیار نمرات را جداگانه محاسبه کنید.  
 ب) کدام مجموعه (ساعات مطالعه یا نمرات) پراکندگی نسبی بیشتری دارد؟  
 ج) آیا به نظر می‌رسد افزایش ساعات مطالعه همواره منجر به افزایش نمره شده است؟ خیر.  
 واریانس چه اطلاعاتی درباره رابطه این دو متغیر به ما نمی‌دهد؟ (پاسخ: واریانس ارتباط دو متغیر را نشان نمی‌دهد!).  
 د) اگر معلم بگوید: "پراکندگی نمرات این کلاس پایین است"، آیا با داده‌ها موافقت می‌کند؟ چرا؟ (از انحراف معیار نمرات استفاده کنید)

ولی ممکن است این موضوع برای دانش‌آموزان کمی انتزاعی باشد زیرا مستقیماً با تجربیات روزمره آنها مرتبط نیست اما می‌تواند زمینه یک پژوهش حرفه‌ای و کار عملی جذاب را برای دانش‌آموزان فراهم آورد. ابزارهای ChatGPT، MathGPT، MagicSchool و DeepSeek توانسته‌اند با ارائه سؤالاتی با موضوعاتی چون خرید تنقلات، تعداد قدم‌ها، فرستادن پیامک، رابطه بین میزان مطالعه و نمرات کسب شده در امتحان، دانش‌آموزان را با فعالیت‌هایی مبتنی بر پدیده‌های اطرافشان به چالش بکشند.

• از لحاظ چالش‌برانگیز بودن، تمام سؤالات ارائه شده توسط

مقایسه پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۲:

برای مقایسه و ارزیابی عملکرد پنج ابزار معرفی شده برای پرامپت ۲، پاسخ‌های بدست آمده از این ابزارها را در معیارهایی مانند ارتباط با زندگی واقعی، چالش‌برانگیز بودن و استخراج اطلاعات از متن بررسی می‌کنیم:

• در بررسی عملکرد ابزارهای مبتنی بر هوش مصنوعی ارائه شده در معیار ارتباط با زندگی واقعی، اینگونه برداشت می‌شود که اگرچه Gemini توانسته با موضوع طول عمر لامپ، پیوند نسبتاً خوبی را بین سؤال و دنیای اطراف دانش‌آموزان برقرار سازد،

مبتنی بر هوش مصنوعی در معیار استخراج داده‌ها از متن، اینگونه دریافت می‌شود که Gemini داده‌ها را نسبتاً واضح در متن ارائه داده است که استخراج آنها ساده است و دانش‌آموز نیاز به پردازش پیچیده‌ای ندارد. در مقابل، اگرچه داده‌های سؤال ارائه شده توسط ChatGPT تقریباً ساده است اما استخراج مبالغ و تبدیل آنها به اعداد (از «هزارتومان» به عدد)، شاید کمی مهارت خواندن و سازمان‌دهی دانش‌آموز را بطلبد. به طور مشابه، استخراج داده‌ها از متن سؤال‌اتی که MatGPT و MagicSchool ارائه داده‌اند، ساده است اما پاسخگویی به بخش‌های مختلف این سؤالات نیازمند تفکر خلاقانه است. DeepSeek عملکردی متمایز داشته است و با طراحی یک جدول، استخراج داده‌ها را ساده‌تر کرده است اما تحلیل رابطه بین دو متغیر و پاسخ به پرسش‌های تحلیلی نیاز به مهارت بالایی دارد.

ابزارهای معرفی شده فرصت مناسبی را برای به چالش کشیدن دانش‌آموزان در محاسبات ایجاد می‌کنند. اما در حالی که سؤال ارائه شده توسط Gemini تنها به مقایسه دو مجموعه محدود شده و تحلیل عمیق‌تری (مثل تفسیر نتایج) را درخواست نمی‌کند، ChatGPT با طرح پرسشی تحلیلی درباره پراکندگی هزینه‌ها، دانش‌آموزان را به تفکر انتقادی واریانس و پرسش درباره نزدیک‌ترین فرد به میانگین، ترکیبی از محاسبات و تفکر تحلیلی را می‌طلبد. MagicSchool با ارائه پرسشی با فرض اینکه «هر روز دقیقاً ده پیامک ارسال شود» و طرح مسئله مدیریت زمان، ترکیبی از محاسبات و تفکر نقادانه را خلق کرده است، به طور مشابه DeepSeek نیز با تشویق دانش‌آموزان به محاسبه انحراف معیار دو مجموعه داده ساعات مطالعه و نمرات، مقایسه پراکندگی نسبی آنها و تحلیل رابطه بین متغیرها، چالش‌های محاسباتی و تحلیلی مفیدی را خلق می‌کند.

- در بررسی کیفیت سؤالات ارائه شده توسط ابزارهای مختلف شکل ۸، پرامپت ۳ را در محیط DeepSeek نشان می‌دهد.



شکل ۸: پرامپت ۳ در محیط DeepSeek

پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۳:

۱. Gemini

آزمایشگاه واکنش سریع: ماجرای کشف پراکندگی داده‌ها  
شما و هم‌گروهی‌هایتان یک تیم از دانشمندان علوم اعصاب هستید که در "آزمایشگاه عملکرد انسانی" کار می‌کنید. یک سازمان معتبر فضایی از شما خواسته تا برای انتخاب فضانوردان آینده، یک آزمون ساده برای سنجش زمان واکنش طراحی و اجرا کنید. برای این سازمان، فقط "میانگین" زمان واکنش یک فرد مهم نیست، بلکه پایداری و ثبات عملکرد او نیز اهمیت حیاتی دارد. یک فضانورد باید واکنش‌های همواره سریعی داشته باشد، نه اینکه یک بار فوق‌العاده سریع و بار دیگر بسیار کند عمل کند. برای فهمیدن اینکه هر فرد به طور میانگین چقدر سریع است و عملکردش چقدر پایدار و قابل پیش‌بینی است، باید زمان واکنش اعضای تیم خودتان را اندازه بگیرید. ابزار شما برای سنجش این پایداری، واریانس و انحراف معیار خواهد بود.  
مرحله اول: گروه‌بندی (۱۰ دقیقه): کلاس به گروه‌های ۳ یا ۴ نفره تقسیم می‌شود و نقش‌های فرد آزمایش‌شونده، آزمایشگر و ثبت‌کننده مشخص می‌شوند.  
مرحله دوم: جمع‌آوری داده‌ها (۱۵ دقیقه): روش آزمایش: فرد آزمایش‌شونده دست خود را طوری لبه می‌زیر قرار می‌دهد که انگشت شست و اشاره‌اش کمی از لبه میز بیرون زده و از هم باز باشند. آزمایشگر خطکش را به صورت عمودی نگه می‌دارد، به طوری که عدد صفر خطکش دقیقاً بین دو انگشت فرد آزمایش‌شونده قرار بگیرد (بدون اینکه آن را لمس کند). سپس آزمایشگر بدون اطلاع قبلی، خطکش را رها می‌کند. فرد آزمایش‌شونده باید به محض دیدن رها شدن خطکش، با گرفتن آن واکنش نشان دهد. ثبات عددی را که انگشتان فرد روی آن قرار گرفته است (بر حسب سانتی‌متر) یادداشت می‌کند. این عدد نماینده زمان واکنش است (هرچه عدد کمتر، واکنش سریع‌تر).  
ثبت داده‌ها: هر عضو گروه باید ۱۰ بار این آزمایش را به عنوان "فرد آزمایش‌شونده" تکرار کند. نتایج هر فرد باید در یک جدول جداگانه برای او ثبت شود.  
مرحله سوم: محاسبات (۲۰ دقیقه): هر گروه برای هر یک از اعضای خود محاسبات مربوط به میانگین، واریانس و انحراف معیار را انجام می‌دهد.

## ۲. ChatGPT

ماموریت قدم‌شمار

مراحل فعالیت:

- مرحله ۱: کلاس به گروه‌های ۳ تا ۴ نفره تقسیم می‌شود و هر گروه یک قدم‌شمار موبایلی یا اپلیکیشن شمارش گام‌ها یا هر ابزار مشابه در اختیار دارد.
- مرحله ۲: مسیر مشخصی در حیاط مدرسه یا اطراف آن انتخاب می‌شود (مثلاً یک دور کامل ۲۰۰ متری).
- مرحله ۳: هر عضو گروه مسیر را طی می‌کند و تعداد گام‌های ثبت‌شده توسط اپلیکیشن را یادداشت می‌کند.
- مرحله ۴: گروه داده‌های مربوط به تعداد قدم‌های هر عضو را گردآوری و ثبت می‌کند، سپس میانگین تعداد قدم‌ها، واریانس و انحراف معیار را محاسبه می‌کند.
- مرحله ۵: گروه‌ها تفاوت‌ها را بررسی می‌کنند (مثلاً چرا افراد مختلف در مسیر یکسان تعداد قدم‌های متفاوتی دارند؟)

## ۳. MathGPT

«مسابقه پرتاب کاغذی»

مراحل اجرا:

۱. تشکیل گروه‌ها: دانش‌آموزان به گروه‌های ۴ یا ۵ نفره تقسیم می‌شوند.
۲. ساخت هواپیمای کاغذی: هر دانش‌آموز یک هواپیمای کاغذی می‌سازد.
۳. پرتاب و ثبت داده‌ها: هر دانش‌آموز هواپیمای خود را سه بار پرتاب می‌کند و هر بار فاصله پرتاب (بر حسب متر یا سانتی‌متر) توسط اعضای گروه اندازه‌گیری و ثبت می‌شود.
۴. جمع‌آوری داده‌ها: هر گروه جدولی تهیه می‌کند و داده‌های مربوط به هر پرتاب را در آن یادداشت می‌کند.
۵. محاسبات آماری: هر گروه برای داده‌های هر دانش‌آموز (سه پرتاب) میانگین، واریانس و انحراف معیار را محاسبه می‌کند.
۶. تحلیل و مقایسه: گروه‌ها نتایج خود را با هم مقایسه می‌کنند و درباره این‌که کدام دانش‌آموز یا گروه پرتاب‌های یکنواخت‌تر (انحراف معیار کمتر) یا متغیرتری داشته‌اند، بحث می‌کنند.
۷. جمع‌بندی معلم: معلم با کمک دانش‌آموزان توضیح می‌دهد که واریانس و انحراف معیار چه اطلاعاتی درباره پایداری یا پراکندگی داده‌ها به ما می‌دهند.

## ۴. MagicSchool

«مسابقه زمان‌سنجی: از داده تا واریانس و انحراف معیار»

وسایل مورد نیاز: کروномتر، برگه ثبت داده برای هر گروه و یک برگه بزرگ کلاسی، تخته سفید یا وایت‌برد + ماژیک، وسایل مربوط به فعالیت فیزیکی انتخاب شده (خودکار و کاغذ برای نوشتن، توپ و سطل برای پرتاب، کاغذ ۴A برای هواپیما)

مراحل اجرای فعالیت:

۱. آماده‌سازی و توضیح مفاهیم (۱۰ دقیقه): کلاس را به گروه‌های ۳-۴ نفره تقسیم کنید.
۲. جمع‌آوری داده‌ها (۱۵ دقیقه): جمع‌آوری درون گروهی: در هر گروه، اعضا یکی یکی فعالیت را انجام می‌دهند، یک نفر زمان را با کروномتر اندازه می‌گیرد و ثبت می‌کند.
۳. محاسبات گروهی (۱۵ دقیقه): به هر گروه بگویید روی داده‌های خودشان، میانگین، انحراف از میانگین برای هر داده، واریانس و انحراف معیار را محاسبه کنند.
۴. تحلیل و مقایسه کلاسی (۱۵ دقیقه): به داده‌ها و نتایج روی تخته اشاره کنید. «کدام گروه انحراف معیار (s) کوچکتری دارد؟» توضیح دهید که انحراف معیار کوچکتر یعنی زمان‌های اعضای آن گروه به هم شبیه‌تر بودند (پراکندگی کمتر). انحراف معیار بزرگتر یعنی تفاوت زمان‌های اعضا بیشتر بود (پراکندگی بیشتر). «آیا ممکن است دو گروه میانگین یکسان داشته باشند اما انحراف معیارهای متفاوتی داشته باشند؟ چه مفهومی دارد؟» (پاسخ: بله! میانگین مرکز را نشان می‌دهد، انحراف معیار میزان پراکندگی حول آن مرکز را. دو گروه با میانگین یکسان اما انحراف معیار متفاوت، یکنواختی متفاوتی دارند)
۵. جمع‌بندی و نتیجه‌گیری (۵ دقیقه): مرور مفاهیم و تأکید بر معنای عملی: انحراف معیار کوچک = داده‌ها نزدیک به هم و به میانگین، انحراف معیار بزرگ = داده‌ها پراکندتر.

## فعالیت کلاسی قد دانش‌آموزان کلاس

مراحل فعالیت:

۱. جمع‌آوری داده‌ها: هر دانش‌آموز قد خود را به سانتی‌متر اندازه می‌گیرد و یادداشت می‌کند، سپس داده‌ها به صورت یک لیست کامل روی تخته یا کاغذ بزرگ نوشته می‌شود.
۲. محاسبه میانگین: با کمک معلم یا به صورت گروهی، میانگین قد‌ها را محاسبه کنند.
۳. محاسبه واریانس: برای هر قد، مقدار تفاوت آن با میانگین را پیدا کرده و مربع آن را محاسبه کنند، سپس مجموع مربع تفاوت‌ها را محاسبه و بر تعداد داده‌ها تقسیم کنند.
۴. محاسبه انحراف معیار: جذر واریانس را بگیرند تا انحراف معیار به دست آید.
۵. بحث و تفسیر: دانش‌آموزان درباره معنی واریانس و انحراف معیار در این داده‌ها بحث کنند: آیا قد‌ها خیلی به هم نزدیک هستند یا پراکندگی زیادی دارند؟ انحراف معیار چقدر است و چه چیزی درباره تنوع قد‌ها می‌گوید؟

### مقایسه پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۳:

برای مقایسه و ارزیابی عملکرد پنج ابزار معرفی شده برای پرامپت ۳، پاسخ‌های بدست آمده از این ابزارها را در معیارهایی مانند خلاقیت و جذابیت، وضوح و ساختار مراحل، ارتباط با مفاهیم آماری، عملی بودن و تعامل و مشارکت بررسی می‌کنیم:

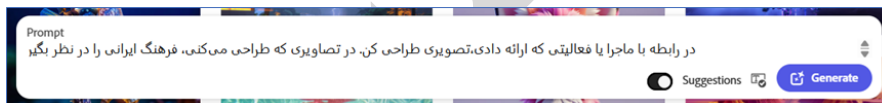
- در معیار خلاقیت و جذابیت، Gemini توانسته با ارائه داستانی با موضوع فضانوردی و علوم اعصاب، فعالیتی هیجان‌انگیز و مرتبط با دنیای واقعی طراحی کند، همچنین استفاده از خط‌کش برای سنجش زمان واکنش، ایده‌ای خلاقانه و در عین حال ساده است. ایده ChatGPT در طراحی فعالیت «قدم‌شمار» و MagicSchool در طراحی فعالیت «اندازه‌گیری قد»، اگرچه گویای مبحث در حال تدریس، ساده و سرگرم‌کننده است، اما عنصر داستان‌سرایی یا زمینه‌ای جذاب در آن کم‌رنگ است. Gemini با ارائه فعالیت ساخت و پرتاب هواپیمای کاغذی، فعالیتی سرگرم‌کننده و انگیزه‌بخش ارائه داده است که خلاقیت دانش‌آموزان را نیز درگیر می‌کند. به همین صورت، DeepSeek با توضیح در مورد «مسابقه زمان‌سنجی» ایده‌ای انعطاف‌پذیر ارائه داده است که می‌توان فعالیت‌های مختلفی مانند پرتاب توپ یا فعالیت‌های فیزیکی دیگر را انتخاب کرد، اما داستان یا زمینه جذابی برای این ایده ارائه نشده است.
- Gemini و DeepSeek با ارائه توضیحات و زمان‌بندی‌هایی مشخص برای هر مرحله، به برنامه‌ریزی دقیق‌تر معلم کمک می‌کنند و توانسته‌اند در معیار وضوح و ساختار مراحل نسبتاً خوب عمل کنند. اگرچه ChatGPT مراحل را به صورتی ساده و مشخص ارائه داده است اما در توضیح زمان‌بندی و

جزئیات، نسبت به دو ابزار قبلی ضعیف‌تر عمل کرده است. MathGPT مراحل انجام فعالیت را تقریباً کامل تعریف کرده است و توضیحاتی که درباره مقایسه پرتاب‌های یکنواخت‌تر ارائه شده، به درک بهتر مفاهیم آماری کمک می‌کند. DeepSeek نیز توانسته فعالیتی را طراحی کند که در آن، مراحل بسیار واضح و ساده بیان شده‌اند، همچنین توضیح گام‌به‌گام محاسبات، یادگیری را آسان می‌کند.

- در معیار ارتباط با مفاهیم آماری، Gemini با فعالیت جذابی که طراحی کرده است و تلاش می‌کند تا از طریق تشویق دانش‌آموزان به پیدا کردن سریع‌ترین و باثبات‌ترین فرد، آنها را به تحلیل عمیق‌تر وادارد و از این طریق بین فعالیت طرح شده و واریانس و انحراف معیار ارتباطی قوی برقرار سازد، در مقابل، در فعالیتی که ChatGPT طراحی کرده است اگرچه به تحلیل پراکندگی داده‌ها اشاره می‌شود اما ارتباط با مبحث واریانس و انحراف معیار ضعیف است و چگونگی ارتباط مفاهیم و تحلیل داده‌ها به خوبی توضیح داده نمی‌شود. MathGPT نیز توانسته با طراحی «مسابقه پرتاب کاغذی»، دانش‌آموزان را تشویق به تحلیل ارتباط قوی موجود بین واریانس و انحراف معیار و پراکندگی داده‌ها کند، بحث درباره یکنواختی پرتاب‌ها به دانش‌آموزان کمک می‌کند تا معنای عملی این مفاهیم را نیز درک کنند. MagicSchool با ارائه فعالیتی با موضوع بررسی تنوع قد‌ها، ارتباط بین واریانس و انحراف معیار و پراکندگی قد‌ها را تحلیل می‌کند. DeepSeek بر مقایسه انحراف معیار بین گروه‌ها و توضیح معنای پراکندگی تأکید ورزیده و به درک عمیق مفاهیم کمک می‌کند. سؤال مربوط به میانگین یکسان با انحراف معیار متفاوت، یادگیری را تعمیق می‌بخشد.

• در تحلیل عملکرد ابزارهای مختلف مبتنی بر هوش مصنوعی از حیث تعامل و مشارکت، تقسیم نقش‌هایی (آزمایش‌شونده، آزمایشگر، ثبت‌کننده) که Gemini در نظر گرفته است، مشارکت تقریبی دانش‌آموزان را تضمین می‌کند. به صورتی مشابه، فعالیت ارائه شده توسط DeepSeek نیازمند تقسیم گروهی و تعیین نقش‌های مشخص (زمان‌سنج، ثبت‌کننده) است که مشارکت را تقریباً بالا می‌برد. فعالیتی که ChatGPT ارائه داده است، فعالیتی گروهی است، اما نقش‌های مشخصی برای دانش‌آموزان تعریف نشده، که ممکن است مشارکت همه را تضمین نکند. MathGPT نیز فعالیتی گروهی طراحی کرده است، همچنین ساخت هواپیما مشارکت را بالا می‌برد، اما نقش‌های خاص و مشخصی تعریف نشده است. در فعالیت ارائه شده توسط MagicSchool مشارکت محدود است، زیرا فعالیت به صورت فردی (اندازه‌گیری قد) انجام می‌شود و کار گروهی کمتر است.

شکل ۹، پرامپت ۴ را در محیط MagicSchool نشان می‌دهد.



شکل ۹: پرامپت ۴ در محیط MagicSchool

پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۴:

۱. Gemini: شکل ۱۰ را مشاهده کنید.

۲. ChatGPT: شکل ۱۱ را مشاهده کنید.

۳. MathGPT

من به صورت مستقیم امکان تولید فایل تصویری ندارم، اما می‌توانم طرح نهایی را به صورت یک اسلاید یا پوستر آماده کنم تا شما در PowerPoint، Canva یا هر نرم‌افزار طراحی دیگر، به سرعت آن را پیاده‌سازی کنید.

۴. MagicSchool: شکل ۱۲ را مشاهده کنید.

۵. DeepSeek

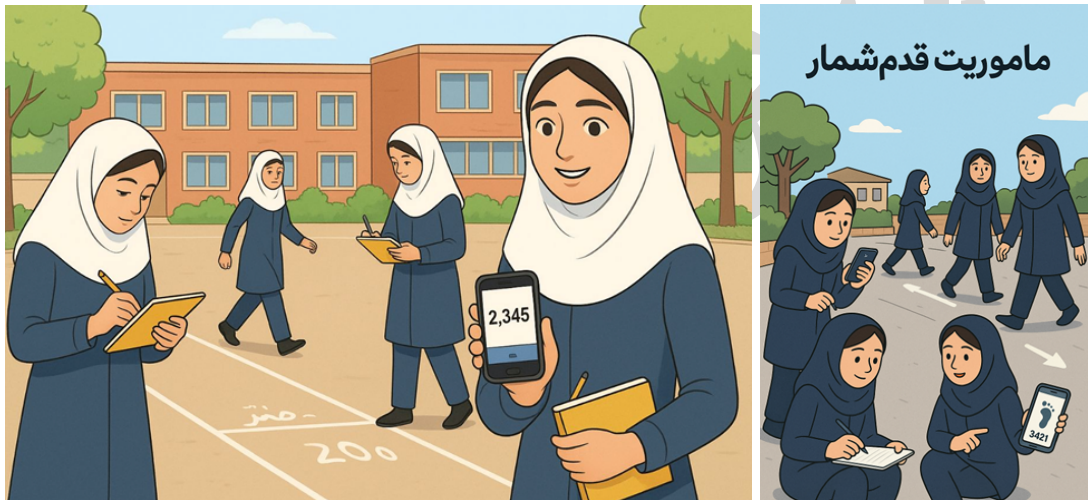
در حال حاضر، من قابلیت طراحی یا تولید تصاویر را ندارم.

در حالی که DeepSeek و MathGPT نتوانسته‌اند تصاویری مرتبط با فعالیت ارائه شده، طراحی کنند، سایر ابزارها از این قابلیت برخوردار

مقایسه پاسخ‌های ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی برای پرامپت ۴:



شکل ۱۰: تصویر خروجی پرامپت ۴ در محیط Gemini



شکل ۱۱: تصاویر خروجی پرامپت ۴ در محیط ChatGPT



شکل ۱۲: تصاویر خروجی پرامپت ۴ در محیط MagicSchool

نیز محتوایی ساده اما کم‌عمق ارائه کرد. برای پرامپت ۲ (طراحی سؤال یا فعالیت)، ChatGPT، MagicSchool و DeepSeek با سؤالات تحلیلی و مرتبط با زندگی واقعی مانند تحلیل هزینه‌ها یا ساعات مطالعه، در ایجاد چالش‌های فکری موفق‌تر بودند، در حالی که Gemini با سؤال طول عمر لامپ، کمتر با تجربیات روزمره دانش‌آموزان مرتبط بود. برای پرامپت ۳ (فعالیت کلاسی)، Gemini و MathGPT با فعالیت‌های تقریباً خلاقانه‌ای مانند آزمایش واکنش سریع و پرتاب هواپیمای کاغذی، جذابیت و تعامل بالایی ایجاد کردند، فعالیت‌های ChatGPT و MagicSchool ساده‌تر و DeepSeek انعطاف‌پذیر اما بدون داستان‌سرایی بود. برای پرامپت ۴ (طراحی تصویر)، Gemini، ChatGPT و MagicSchool تصاویری مرتبط و باکیفیت تولید کردند، اما DeepSeek و MathGPT قابلیت ارائه تصویر نداشتند. به طور کلی می‌توان گفت که ابزارهای هوش مصنوعی می‌توانند به‌عنوان مکمل‌هایی مؤثر در آموزش آمار عمل کنند، به‌ویژه در ارائه محتوای ساده، جذاب و مرتبط با دنیای واقعی به دانش‌آموزان، مناسب است. در ارتباط با چهار پرامپت، Gemini و DeepSeek در ارائه محتوای ساختاریافته و نسبتاً دقیق برجسته‌تر بودند. ChatGPT و MagicSchool با تمرکز بر فعالیت‌های تحلیلی و تعاملی، تفکر نقادانه را تقویت کردند که با هدف پرورش سواد آماری و تفکر انتقادی هم‌راستا است. MathGPT با وجود سادگی، در عمق محتوا و ارائه تصاویر ضعیف بود. تفاوت در عملکرد ابزارها نشان‌دهنده اهمیت انتخاب ابزار مناسب بر اساس نیازهای آموزشی خاص، مانند تمرکز بر فعالیت عملی، توضیح سؤالات و تصورات رایج یا ایجاد انگیزه است. این نتایج نشان می‌دهد که ترکیب نقاط قوت این ابزارها با نظارت معلم می‌تواند یادگیری را بهینه کند.

یافته‌های به‌دست‌آمده از طریق تعامل با ابزارهای مختلف مبتنی بر هوش مصنوعی، با مطالعات پیشین در حوزه کاربرد ابزارهای هوش مصنوعی در آموزش، همخوانی و همپوشانی دارد. پژوهش وانگ و همکاران [۴۷] بر توانایی ChatGPT در ارائه توضیحات ساده و روان و تشویق دانش‌آموزان به تفکر نقادانه از طریق طراحی سؤالات مبتنی بر زندگی روزمره تأکید کرده‌اند و به محدودیت‌های این ابزار در ارائه محتوای دقیق و همراه با جزئیات اشاره داشتند. پاسخ‌های دریافت شده از ابزارهای مبتنی بر هوش مصنوعی در پژوهش حاضر نیز حاکی از آن است که استفاده از هوش مصنوعی در آموزش آمار می‌تواند تا حد قابل توجهی نیازهای آموزشی را تأمین کند، بدین ترتیب که ChatGPT توضیحاتی کلی اما قابل فهم ارائه می‌دهد و در طراحی تصاویر و فعالیت‌های تحلیلی عملکرد مناسبی دارد، هرچند در مقایسه با ابزارهایی مانند Gemini و DeepSeek از دقت کمتری برخوردار

بودند. عملکرد این سه ابزار را در معیارهای میزان ارتباط با موضوع و کیفیت هنری مقایسه می‌کنیم:

- در معیار میزان ارتباط با موضوع، Gemini نسبتاً بیشترین وفاداری را به موضوع داشته است و تصویری که ارائه کرده است دقیقاً عمل‌رها کردن خط‌کش را نشان می‌دهد. در واقع، تصویر ارائه شده توسط Gemini تقریباً روش آزمایش را آموزش می‌دهد و برای شفاف‌سازی نحوه انجام آزمایش و مراحل آن بسیار مفید است. در تصویری که ChatGPT طراحی کرده است، گروه‌هایی از دانش‌آموزان را معرفی می‌کند که در حال قدم زدن هستند، یادداشت برداری و ثبت تعداد قدم‌ها نیز نشان‌دهنده دقت نسبی این ابزار در به تصویر کشیدن جزئیات است. تصویری که MagicSchool در توصیف فعالیت اندازه‌گیری قد دانش‌آموزان ارائه داده است، اگرچه نشان‌دهنده اندازه‌گیری با خط‌کش و همکاری بین دانش‌آموزان است، اما بیشتر به محاسبات اشاره دارد تا اندازه‌گیری مستقیم قد دانش‌آموزان.

- در معیار کیفیت هنری، Gemini تصویری با سبک نسبتاً حرفه‌ای اما کمی رسمی‌تر و کمتر رنگارنگ ارائه داده است، در حالی که ChatGPT تصویری با سبک شاد، رنگ‌های روشن و ترکیبی پویا از کلاس درس، طراحی کرده است. MagicSchool در تصویری که طراحی کرده است از رنگ‌های زنده استفاده کرده است، همچنین نمایش جزئیات تخته و یادداشت‌ها نیز دقت این ابزار را نشان می‌دهد.

## ۵ بحث و نتیجه‌گیری

بررسی عملکرد پنج ابزار مبتنی بر هوش مصنوعی (Gemini، ChatGPT، MagicSchool، MathGPT و DeepSeek) در پاسخ به چهار پرامپت مرتبط با آموزش «واریانس و انحراف معیار» نشان داد که هر ابزار، ویژگی‌ها و محدودیت‌های خاص خود را دارد. برای پرامپت ۱ (طراحی طرح درس)، Gemini و DeepSeek با ارائه طرح درس‌های نسبتاً دقیق، گام‌به‌گام و مرتبط با دنیای واقعی در انسجام و جذابیت ساختار، عملکرد بهتری داشتند. اما Gemini در فعالیت‌های عملی و توضیح سؤالات رایج و DeepSeek در معرفی کاربردهای گسترده‌تر، ضعیف‌تر عمل کردند. ChatGPT توضیحات ساده اما کلی ارائه داد، MagicSchool با تمرکز بر تصورات غلط رایج و فعالیت‌های گروهی به شیوه‌ای متمایز عمل کرد، اما زمان‌بندی مبهمی داشت. MathGPT

عملکرد تحصیلی و انگیزه دانش‌آموزان را در آموزش آمار بررسی کنند. همچنین، تحلیل پرامپت‌ها با تعداد بیشتر و متنوع‌تر و ارزیابی عملکرد ابزارها در دیگر مفاهیم آماری و به ویژه موضوعات آماری پیچیده‌تر (مانند مدل‌سازی پیشرفته یا تحلیل‌های چندمتغیره) پیشنهاد می‌شود. علاوه بر این، مناسب است مطالعاتی تجربی با گروه نمونه از معلمان و دانش‌آموزان همراه با تحلیل‌های آماری مناسب به این حوزه افزوده شود تا نتایج قابل تعمیمی به دست آید. مطالعه تأثیر این ابزارها بر نگرش دانش‌آموزان نسبت به آمار و توانایی آنها در تحلیل داده‌های واقعی نیز می‌تواند ارزشمند باشد. ادغام هوشمندانه ابزارهای هوش مصنوعی در آموزش آمار، با تقویت سواد آماری و تفکر نقادانه، دانش‌آموزان را برای مواجهه با چالش‌های دنیای داده‌محور آماده می‌کند و آینده‌ای روشن برای آموزش نوین رقم می‌زند.

در مجموع می‌توان گفت که ادغام هوش مصنوعی در آموزش آمار می‌تواند با ایجاد محیط‌های یادگیری پویا، تقویت مهارت‌های آماری و آماده‌سازی دانش‌آموزان برای مواجهه با دنیای داده‌محور، کیفیت آموزش را به‌طور قابل‌توجهی ارتقا دهد، اما نباید از چالش‌هایی که در آموزش ایجاد می‌کند غافل شد [۴۳]. برای مثال، با اینکه هوش مصنوعی می‌تواند فرایند آموزش آمار را بهبود بخشد، اما با این حال، درباره خطر جایگزینی کامل تلاش‌های معلمان با ابزارهای مبتنی بر هوش مصنوعی و نیاز به استفاده از این ابزارها به‌عنوان مکمل باید هوشیار بود و بر ضرورت نظارت معلم برای جلوگیری از کمرنگ شدن فرایند تفکر دانش‌آموزان در موقعیت‌هایی که هوش مصنوعی پاسخ‌ها را به‌طور کامل در اختیار آنها می‌گذارد و یا تفسیری نادرست ارائه می‌دهد، تأکید ورزید [۳۴]. در واقع نقش و مسئولیت معلم در استفاده از هوش مصنوعی به‌عنوان مکمل آموزشی این است که بر اقدامات دانش‌آموزان نظارت دقیق داشته باشد [۴]. در حوزه آموزش، معلمان به‌عنوان کاربران کلیدی ابزارهای هوش مصنوعی، مسئولیت اخلاقی و حرفه‌ای بر عهده دارند تا نه تنها خروجی‌های تولیدشده توسط آن را ارزیابی کنند، بلکه مسئولیت نهایی این خروجی‌ها را بپذیرند. این مسئولیت شامل بررسی اعتبار و صحت اطلاعات تولیدشده و پذیرش آگاهانه نتایج به‌عنوان بخشی از فرایند آموزشی است [۱].

همچنین، هوش مصنوعی می‌تواند خلاقیت کاربران را محدود سازد، در برخی مواقع اطلاعات قدیمی‌تر ارائه دهد و در به‌روزرسانی اطلاعات ضعیف عمل کند و یا پاسخ‌هایی نادرست ارائه دهد، لذا نظارت معلم و تفکر نقادانه در استفاده از ابزارهای هوش مصنوعی، ضروری است [۳۸] و معلمان، مسئول تشخیص اشتباهات احتمالی هوش مصنوعی و درک این خطاها هستند [۴]. به عبارت دیگر، در حالی که هوش مصنوعی

است. مطالعه بساس و همکاران [۱۰] توانایی Gemini در ارائه طرح درس‌های جامع، گام‌به‌گام و جذاب با اهداف یادگیری روشن را تأیید کرده، اما به ضعف آن در پاسخ به سؤالات متداول دانش‌آموزان و ارائه ابزارهای بصری اشاره کرده‌اند. پژوهش فوستر و همکاران [۲۰] و والینگتون و همکاران [۲۸] بر نقش MagicSchool به‌عنوان ابزاری چندمنظوره در طراحی محتوای آموزشی و برطرف نمودن تصورات غلط تأکید دارند. این یافته‌ها بر نتایج بدست آمده از تعامل با ابزارهای مبتنی بر هوش مصنوعی که MagicSchool را به دلیل توضیحات روشن، تمرکز بر تصورات غلط و طراحی فعالیت‌های ساده و قابل اجرا مفید و مؤثر می‌دانند، اما به ضعف آن در ارائه زمان‌بندی دقیق و فعالیت‌های گروهی اشاره دارد، صحنه می‌گذارد. به‌طور مشابه، کوتسیس [۳۲] و نها و بتی [۳۹] بر توانایی DeepSeek در ارائه توضیحات دقیق، منطقی و ساختارمند برای بهبود یادگیری مفاهیم دشوار ریاضی تأکید کرده‌اند، اما به کمبود خلاقیت و جذابیت در محتوای آن اشاره داشتند. این نتایج تأییدی بر بررسی پژوهش حاضر از DeepSeek است، که بر دقت علمی و تمرکز بر حل مسئله و تفکر انتقادی تأکید دارد اما در خلاقیت و ارائه فعالیت‌های جذاب‌تر محدودیت‌هایی نشان می‌دهد. هولمز و همکاران [۲۵] بر قابلیت هوش مصنوعی در تحول آموزش و ایجاد یادگیری تعاملی و شخصی‌سازی شده تأکید کرده‌اند، که با توانایی ابزارهای بررسی‌شده در این پژوهش (مانند Gemini، ChatGPT و MagicSchool) در ارائه محتوای جذاب، تعاملی و مرتبط با دنیای واقعی همخوانی دارد.

به‌طور کلی می‌توان گفت که در مبحث واریانس و انحراف معیار، Gemini و DeepSeek در ارائه طرح درس‌های دقیق و ساختاریافته، ChatGPT در طراحی سؤالات تحلیلی و تصاویر جذاب، MagicSchool در فعالیت‌های گروهی و رفع اشتباهات رایج و MathGPT در ارائه محتوای ساده عملکرد نسبتاً بهتری داشتند. در حوزه آموزش، استفاده از این ابزارها می‌تواند سواد آماری دانش‌آموزان را تقویت کند، که در دنیای داده‌محور امروزی، حیاتی است. همچنین، این ابزارها می‌توانند با شبیه‌سازی سناریوهای واقعی و فعالیت‌های تعاملی، مهارت‌های دانش‌آموزان را در مدیریت عدم قطعیت و ارزیابی انتقادی نتایج آماری تقویت کنند [۴۳]، انگیزه آنها را افزایش دهند و به پرورش شهروندانی آگاه و توانمند کمک کنند. این ابزارها با ارائه محتوای مرتبط با دنیای واقعی و فعالیت‌های تعاملی، آموزش آمار را برای دانش‌آموزان بدون پیش‌زمینه تسهیل می‌کنند، اما نیاز به نظارت معلم برای جلوگیری از ساده‌سازی بیش از حد یا تفسیر نادرست دارند. پژوهش‌های آتی می‌توانند تأثیرات بلندمدت استفاده از این ابزارها بر

آماري (مانند مثال‌های توزیع داده، آزمون‌های فرضیه یا مدل‌های پیش‌بینی) را تحت تأثیر قرار دهد؛ این مشکل با نظارت مستمر معلمان آمار و بهره‌گیری از ابزارهای هوش مصنوعی متنوع و مکمل قابل کاهش است. سلامت آکادمیک در آموزش آمار نیز با طراحی فعالیت‌هایی برای ترویج تفکر نقادانه (مانند ارزیابی متخصصان از نتایج تولید شده توسط هوش مصنوعی) و کاربرد هوش مصنوعی به‌عنوان مکمل و نه جایگزین فرایند یادگیری مفاهیم آماری، حفظ می‌شود.

به‌طور کلی، نتایج ارائه شده توسط ابزارهای مبتنی بر هوش مصنوعی در پژوهش حاضر نشان می‌دهد که هوش مصنوعی می‌تواند با ارائه محتوای ساده، جذاب و تعاملی، آموزش آمار را بهبود بخشد، اما محدودیت‌هایی مانند کمبود خلاقیت، دقت ناکافی در برخی موارد و نیاز به نظارت معلم، ضرورت استفاده هوشمندانه و ترکیبی از این ابزارها را برجسته می‌کند.

## سپاسگزاری

این مطالعه با حمایت دانشگاه تربیت دبیر شهید رجایی طبق ابلاغ گزنت شماره ۳۷۴۰۰۱/۱۴۰۴ انجام گردیده است، بدین وسیله از این دانشگاه تشکر و قدردانی می‌گردد.

می‌تواند در ارتقای درک مفاهیم آماری و مشارکت فعال دانش‌آموزان مفید واقع شوند [۳۴، ۴۴]، اما نمی‌توان از نقش بی‌بدیل معلم در ایجاد انگیزه و هدایت فرایند یادگیری غافل شد [۱۹]، زیرا هیچ جایگزینی برای انگیزه و تشویقی که یک معلم می‌تواند ارائه دهد وجود ندارد و فرایند ارزیابی را نمی‌توان به‌طور کامل به فناوری واگذار کرد. به علاوه، اگرچه استفاده از فناوری‌هایی چون هوش مصنوعی در آموزش آمار مؤثر است اما میزان اثرگذاری مواد آموزشی بر کیفیت یادگیری دانش‌آموزان، به انتخاب مناسب ابزار و هدایت معلم بستگی دارد. بنابراین، حتی زمانی که فناوری بار اصلی تدریس را بر عهده دارد، نقش اساسی معلمان انسانی همچنان پابرجاست و نظارت انسانی و دانش آماری برای جلوگیری از تفسیر نادرست نتایج تولیدشده توسط هوش مصنوعی حیاتی است [۴۳].

در استفاده از ابزارهای هوش مصنوعی در آموزش آمار، باید مسائل اخلاقی را در نظر گرفت، از جمله حریم خصوصی داده‌های آماری، سوگیری الگوریتمی در تحلیل‌های آماری و سلامت آکادمیک در یادگیری آمار. حریم خصوصی داده‌های آماری نیازمند بهره‌گیری از ابزارهای هوش مصنوعی سازگار با استانداردهایی مانند مقررات عمومی حفاظت از داده‌ها<sup>۵</sup>، کسب رضایت آگاهانه افراد، و شفافیت کامل در ذخیره‌سازی و پردازش داده‌های آماری است. سوگیری الگوریتمی، که اغلب ناشی از داده‌های آموزشی ناعادلانه یا نامتوازن است، می‌تواند محتوای تولیدشده

## مراجع

- [۱] صادقی بی‌غم، بهرام و ذوالقدر، نغمه. (۱۴۰۴). هوش مصنوعی برای آدم‌های معمولی. تهران: انتشارات گوتنبرگ.
- [۲] کمیجانی، آناهیتا، ریحانی، ابراهیم، رحیمی، زهرا و بهرامی سامانی، احسان. (۱۴۰۲). فراترکیبی از تعاریف و مولفه‌های تفکر آماری. اندیشه آماری، ۲۸(۱)، ۸۵-۱۰۰.
- [3] Adair, A. (2023). Teaching and learning with AI: How artificial intelligence is transforming the future of education. *XRDS: Crossroads, The ACM Magazine for Students*, 29(3), 7-9.
- [4] Akgun, S., and Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.
- [5] Al-ganaam, S. (2018). The level of statistical reasoning among faculty members at the faculty of education. *Mathematics Education Journal*, 21(1), 170-212.
- [6] Al Sarhani, F. (2023). A suggested proposal of a training program to acquire statistical skills for female mathematics teachers in the preparatory stage. *Fayoum University Journal of Educational and Psychological Sciences*, 17(2), 908-957.

<sup>5</sup>General Data Protection Regulation

- [7] Ellis, A. R., and Slade, E. (2023). A new era of learning: Considerations for ChatGPT as a tool to enhance statistics and data science education. *Journal of Statistics and Data Science Education*, **31**(2), 128–133.
- [8] Ashley, M. S. (2026). Empowering STEM education: Harnessing MagicSchool AI for experiential learning. In *Harnessing AI's potential to support student success and teaching excellence*, 51–108. IGI Global Scientific Publishing.
- [9] Ben-Zvi, D., and Makar, K. (2016). *The teaching and learning of statistics*. Springer, Cham.
- [10] Bessas, N., Tzanaki, E., Vavougiou, D., and Plagianakos, V. P. (2025). Comparative analysis of ChatGPT and Gemini; Implications for junior high school physics education: Opportunities and ethical challenges. *International Journal of Advanced Multidisciplinary Research and Studies*, **5**(1), 7–18.
- [11] Bromage, A., Pierce, S., Reader, T., and Compton, L. (2022). Teaching statistics to non-specialists: Challenges and strategies for success. *Journal of Further and Higher Education*, **46**(1), 46–61.
- [12] Budgett, S., and Renelle, A. (2023). Statistical literacy—the golden rules: A review of Tim Harford's *The Data Detective: Ten easy rules to make sense of statistics*. *The Mathematics Enthusiast*, **20**(1), 256–265.
- [13] Burgess, T. (2009). Teacher knowledge and statistics: What types of knowledge are used in the primary classroom? *The Mathematics Enthusiast*, **6**(1), 3–24.
- [14] Burrill, G., and Camden, M. (2005). *Curricular development in statistics education: International Association for Statistical Education 2004 Roundtable*. International Statistical Institute, Voorburg, The Netherlands.
- [15] Bi, X., Chen, D., Chen, G., Chen, S., Dai, D., Deng, C., et al. (2024). DeepSeek LLM: Scaling open-source language models with long-termism. *arXiv preprint arXiv:2401.02954*.
- [16] Chongchitnan, S., Parker, M., Mahal, M., and Petrie, S. (2025). Exploring the use of AI in mathematics and statistics assessments. *MSOR Connections*, **23**(2), 35–48.
- [17] Coppin, B. (2004). *Artificial intelligence illuminated*. Jones and Bartlett Learning.
- [18] Eager, B., and Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching and Learning Practice*, **20**(5), 1–19.
- [19] Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, **16**(1), 44–49.
- [20] Foster, A., Khazanchi, P., and Khazanchi, R. (2024). MagicSchool.ai: A personal assistant. In *Proceedings of the Society for Information Technology and Teacher Education International Conference*, 74–80. Association for the Advancement of Computing in Education (AACE).
- [21] Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., and Spangler, D. A. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data literacy*. American Statistical Association.
- [22] Geroimenko, V. (2025). Key principles of good prompt design. In *The essential guide to prompt engineering: Key principles, techniques, challenges, and security risks*, 17–36. Springer Nature Switzerland, Cham.
- [23] Ghulami, H., Hamid, M., and Zakaria, R. (2015). Comparative analysis of positive and negative attitudes toward statistics. *AIP Conference Proceedings*, **1643**, 212–218.

- [24] Hmoud, M., Swaity, H., Anjass, E., and Aguaded-Ramírez, E. M. (2024). Rubric development and validation for assessing tasks' solving via AI chatbots. *Electronic Journal of e-Learning*, **22**(6), 1–17.
- [25] Holmes, W., Bialik, M., and Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- [26] Holmes, W., and Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, **57**(4), 542–570.
- [27] Huang, Q., Lv, C., Lu, L., and Tu, S. (2025). Evaluating the quality of AI-generated digital educational resources for university teaching and learning. *Systems*, **13**(3), 174.
- [28] Imran, M., and Almusharraf, N. (2024). Google Gemini as a next generation AI educational tool: A review of emerging educational technology. *Smart Learning Environments*, **11**(1), 22.
- [29] Katai, Z. (2014). The challenge of promoting algorithmic thinking of both sciences- and humanities-oriented learners. *Journal of Computer Assisted Learning*, **31**(4), 287–299.
- [30] Knoth, N., Tolzin, A., Janson, A., and Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, **6**, 100225.
- [31] Kollipara, V. K., Chugh, Y. P., and Mondal, K. (2014). Physical, mineralogical and wetting characteristics of dusts from Interior Basin coal mines. *International Journal of Coal Geology*, **127**, 75–87.
- [32] Kotsis, K. T. (2025). ChatGPT and DeepSeek evaluate one another for science education. *EIKI Journal of Effective Teaching Methods*, **3**(1), 98–102.
- [33] Kurnia, A. B., Lowrie, T., and Patahuddin, S. M. (2024). The development of high school students' statistical literacy across grade level. *Mathematics Education Research Journal*, **36**(Suppl 1), 7–35.
- [34] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, **13**(4), 410.
- [35] Masykar, N. T. (2019). *Exploring 8th grade students' statistical competency through the investigation enquiry cycle*. Master's thesis, National Taiwan Normal University.
- [36] Morales-Vargas, S. R., and Benitez-Medina, R. (2024). The use of artificial intelligence as a learning tool: A first focus on solving statistics problems. In *ICERI2024 Proceedings*, 3412–3416.
- [37] Memarian, B., and Doleck, T. (2023). ChatGPT in education: Methods, potentials, and limitations. *Computers in Human Behavior: Artificial Humans*, **1**(2), 100022.
- [38] National Council of Teachers of Mathematics. (2024). *Artificial intelligence and mathematics teaching: A position of the National Council of Teachers of Mathematics*.
- [39] Neha, F., and Bhati, D. (2025). DeepSeek models in STEM education: Capabilities, applications, and challenges. *Authorea Preprints*.
- [40] Nurul, T. (2019). *Exploring 8th grade students' statistical competency through the investigation enquiry cycle*. Doctoral dissertation, National Taiwan Normal University.

- [41] Rohana, R., and Ningsih, Y. L. (2020). Students' statistical reasoning in statistics method course. *Jurnal Pendidikan Matematika*, **14**(1), 81–90.
- [42] Schoen, R. C., Rhoads, C., Perez, A., Jacobbe, T., and Li, L. (2025). Improving the teaching and learning of statistics. *Learning and Instruction*, **95**, 102018.
- [43] Schwarz, J. (2025). The use of generative AI in statistical data analysis and its impact on teaching statistics at universities of applied sciences. *Teaching Statistics*, **47**(2), 118–128.
- [44] Tian, X. (2021). Research on teaching reform of probability theory and mathematical statistics in artificial intelligence major. *Frontiers in Educational Research*, **4**(11), 110–114.
- [45] Tsvetkova, M. S., and Kiryukhin, V. M. (2021). Algorithmic thinking and new digital literacy. *Olympiads in Informatics*, **15**, 105–118.
- [46] Wahba, F., Ajlouni, A. O., and Abumosa, M. A. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education*, **20**(7), em2468.
- [47] Wang, X., Gong, Z., Wang, G., Jia, J., Xu, Y., Zhao, J., et al. (2023). ChatGPT performs on the Chinese national medical licensing examination. *Journal of Medical Systems*, **47**(1), 86.
- [48] Walkington, C., Bernacki, M., and Beauchamp, T. (2024). *Personalizing instruction to students' interests: Foundations and new directions*. Preprint.
- [49] Wolfram, S. (2023). What is ChatGPT doing... and why does it work? *Stephen Wolfram Writings*.
- [50] Xing, Y. (2024). Exploring the use of ChatGPT in learning and instructing statistics and data analytics. *Teaching Statistics*, **46**(2), 95–104.

## پیوست ۱: تصویری از کدگذاری پاسخ‌های ابزارهای هوش مصنوعی در محیط نرم‌افزار MAXQDA

MAXQDA Analytics Pro (24.7.0)

Home Import Codes Memos Variables Analysis Mixed Methods Visual Tools Reports MAXDicto Stats TeamCloud

Documents (13 Paragraphs)

مرحله اول: مقدمه و ایجاد انگیزه (۵ دقیقه)

پرسش و پاسخ: تا به حال به این فکر کرده‌اید که چطور می‌توانیم پراکندگی نمرات یک کلاس را با یک عدد توصیف کنیم؟ مثلاً اگر میانگین نمرات دو کلاس یکسان باشد، آیا این به معنای یکسان بودن سطح هر دو کلاس است؟

فرض کنید دو بازیکن بسکتبال داریم که هر دو در ۵ بازی اخیر خود، میانگین ۲۰ امتیاز کسب کرده‌اند. بازیکن اول در بازی‌ها امتیازات ۲۰، ۲۱، ۱۹، ۲۲، ۱۸ را کسب کرده و بازیکن دوم ۵، ۲۵، ۱۰، ۳۰، ۲۰ را. به نظر شما کدام بازیکن عملکرد بالاتری داشته؟ (اثبات بودن را با پراکندگی کمتر مرتبط کنید).

مرحله دوم: معرفی واریانس (۱۰ دقیقه)

یادآوری میانگین: برای اینکه بدانیم هر داده چقدر از میانگین دور است، ابتدا باید میانگین را محاسبه کنیم.

انحراف از میانگین: حالا می‌خواهیم ببینیم هر کدام از امتیازات چقدر از میانگین ۲۰ فاصله دارند. این فاصله را انحراف از میانگین می‌نامیم.

مشکل مجموع انحرافات و راحل (واریانس): اگر این انحرافات را با هم جمع کنیم، حاصل صفر می‌شود، پس این روش به تنهایی نمی‌تواند پراکندگی را نشان دهد و برای غلبه بر این مشکل، هر یک از این انحرافات را به توان ۲ می‌رسانیم تا علامت منفی از بین برود. حالا این مربع انحرافات را با هم جمع می‌کنیم و بر تعداد داده‌ها (یا تعداد داده‌ها منهای یک برای نمونه) تقسیم می‌کنیم. این مقدار را واریانس می‌نامیم.

مرحله سوم: معرفی انحراف معیار (۱۰ دقیقه)

مشکل واحد واریانس: مثلاً اگر امتیازات را با واحد امتیاز اندازه گرفته باشیم، واریانس با واحد امتیاز مربع خواهد بود که امتیاز مربع برای ما قابل درک نیست.

معرفی انحراف معیار: برای اینکه به واحد اولیه برگردیم و یک معیار پراکندگی قابل تفسیرتر داشته باشیم، از واریانس جذر می‌گیریم. به این مقدار انحراف معیار می‌گوییم.

انحراف معیار نشان می‌دهد که داده‌ها به طور متوسط چقدر از میانگین فاصله دارند. هرچه انحراف معیار کمتر باشد، داده‌ها به میانگین نزدیک‌تر و پراکندگی کمتر

Codes

- زمان بندی معقول و منطقی
- تفسیربندی نسبتاً واضح و زیربنایی
- طرح درس/تغریبات دقیق
- پداس و پاسخ
- فراهم کردن زمینه‌های بزرگ بحث و گفت...
- دوگانه کردن دانش آموزان با مثال واضح
- عملکرد/دانش‌آموختن واضح و قابل فهم
- بهره-گیری از منابع و منابع
- توضیح گام به گام

شکل ۱۳: کدگذاری پاسخ‌های ابزارهای هوش مصنوعی در محیط نرم‌افزار MAXQDA

## The Footprint of Artificial Intelligence in School Statistics Education

Narges Yaftian<sup>1\*</sup>, Anahita Komeijani<sup>2</sup>, Niknam Oskooei<sup>3</sup>

<sup>1,2,3</sup> Department of Mathematics, Faculty of Science, Shahid Rajaee Teacher Training University, Tehran, Iran

Received: 2025/10/10

Accepted: 2026/04/14

### Abstract

This study was conducted with the aim of qualitatively examining and comparing the performance of several AI-based tools in teaching the concepts of variance and standard deviation, using a qualitative approach and case study method. The statistical population consisted of AI-based tools applicable in statistics education. Sampling was purposeful, based on accessibility and the ability to generate statistical content, and five tools—Gemini, ChatGPT, MathGPT, MagicSchool, and DeepSeek—were selected. Data were collected using four input prompts, and to evaluate the AI responses to these prompts, criteria such as clarity, engagement, structure, interactivity, real-world relevance, and appropriateness for students were employed. Data analysis was performed through coding with the assistance of MAXQDA software, and the reliability of the analyses was confirmed with a kappa coefficient of 0.8. The findings indicated that AI tools can enhance statistics learning by providing engaging and real-world-relevant content, though they have limitations such as lack of creativity or insufficient accuracy. Overall, in this domain, the tools exhibited varying performance. Gemini and DeepSeek performed relatively better in delivering precise and coherent lesson plans, ChatGPT in designing analytical questions and attractive visual images, MagicSchool in facilitating group activities and addressing common errors, and MathGPT in providing comprehensible content related to the concepts of variance and standard deviation. Combining these tools with teacher supervision—which is essential for maintaining academic integrity—can strengthen statistical literacy and critical thinking, preparing students to face the challenges of a data-driven world.

**Keywords:** Statistics Education, Variance and Standard Deviation, Artificial Intelligence, Comparison of AI-based Tools' Capabilities.