

تأثیرپذیری نحو عربی از دیجیتال‌سازی و هوش مصنوعی: مطالعه موردی تغییرات ساختار جملات در متون تولیدشده توسط الگوریتم‌های مولد زبانی

رسول قلندری

دکتری زبان و ادبیات عرب، مدرس مدعو، دانشگاه اقلید، اقلید، فارس، ایران

Ghalandary40@gmail.com

چکیده

پیشرفت‌های سریع در حوزه هوش مصنوعی و دیجیتال‌سازی، تأثیرات عمیقی بر زبان‌های طبیعی در سراسر جهان بر جای گذاشته است. زبان عربی با ویژگی‌های صرفی-نحوی منحصر به فرد خود، از جمله نظام پیچیده اعراب، توافق جنسی و عددی، و انعطاف‌پذیری در ترتیب کلمات، عرصه‌ای چالش‌برانگیز برای پردازش زبان طبیعی و مدل‌های زبانی بزرگ محسوب می‌شود. پژوهش حاضر به بررسی تأثیرپذیری نحو عربی از فرایند دیجیتال‌سازی و به‌ویژه الگوریتم‌های مولد زبانی مبتنی بر هوش مصنوعی می‌پردازد. این مطالعه با اتخاذ رویکردی توصیفی-تحلیلی و با بهره‌گیری از پیکره‌های زبانی دیجیتال و متون تولیدشده توسط مدل‌های زبانی بزرگ همچون ChatGPT، به شناسایی الگوهای خطاهای نحوی و تغییرات ساختاری در جملات عربی تولیدشده توسط ماشین می‌پردازد. یافته‌ها نشان می‌دهد که خطاهای نحوی در متون تولیدشده توسط هوش مصنوعی تصادفی نیستند، بلکه به‌صورت الگوهای ساختاریافته و تکرارشونده ظاهر می‌شوند، به‌ویژه در زمینه‌هایی که شامل وابستگی‌های دوربرد و ساختارهای صرفاً پیچیده هستند. همچنین، دیجیتال‌سازی فضاهای ارتباطی، از جمله شبکه‌های اجتماعی و رسانه‌های دیجیتال، به تغییراتی در ساختار جمله عربی انجامیده است که عبارت‌اند از: فشردن ساختارهای، افزایش مستقیم‌گویی، و تداخل میان فصیح و عامیانه. این پژوهش نشان می‌دهد که اگرچه مدل‌های زبانی بزرگ توانایی قابل‌توجهی در تولید متون روان از نظر سطحی دارند، اما در حفظ هم‌سازی عمیق نحوی و روابط صرفی-نحوی در زبان عربی با محدودیت‌های اساسی مواجه‌اند.

واژگان کلیدی: نحو عربی، دیجیتال‌سازی، هوش مصنوعی، مدل‌های زبانی بزرگ، پردازش زبان طبیعی

۱. مقدمه

۱.۱. پیشینه و مسئله پژوهش

زبان عربی به‌عنوان یکی از پرگویش‌ترین زبان‌های جهان با بیش از ۴۲۲ میلیون گویشور بومی در ۲۷ کشور، از غنای زبانی و فرهنگی بی‌نظیری برخوردار است (Al-Ghamdi, & Al-Saud, ۲۰۲۵). با این حال، ویژگی‌های منحصربه‌فرد این زبان از جمله صرف پیچیده، نظام اعراب، تنوع گویشی و ساختار نحوی انعطاف‌پذیر، همواره چالش‌های جدی را برای پردازش محاسباتی آن ایجاد کرده است (Habash, ۲۰۲۲; al et Al-Ghamdi, ۲۰۲۳).

در دو دهه اخیر، انقلاب دیجیتال و ظهور هوش مصنوعی، به‌ویژه مدل‌های زبانی بزرگ مبتنی بر معماری ترانسفورمر، تحولی شگرف در نحوه تعامل انسان با زبان ایجاد کرده است (al et Zbib, ۲۰۲۵). این فناوری‌ها نه تنها ابزارهای جدیدی برای تحلیل و پردازش زبان فراهم آورده‌اند، بلکه خود به‌عنوان عاملانی فعال در تولید متن، الگوهای زبانی جدیدی را پدید آورده‌اند که می‌تواند بر ساختارهای نحوی زبان‌های طبیعی تأثیر بگذارد (al et Ismail, ۲۰۲۶).

مسئله اصلی پژوهش حاضر این است که دیجیتال‌سازی و الگوریتم‌های مولد زبانی مبتنی بر هوش مصنوعی تا چه اندازه بر ساختار نحوی زبان عربی تأثیر گذاشته‌اند و این تأثیرات در چه الگوهایی قابل شناسایی و تحلیل هستند. با توجه به پیچیدگی‌های نحوی زبان عربی و گسترش روزافزون استفاده از ابزارهای هوش مصنوعی در تولید محتوای عربی، بررسی این تأثیرپذیری از اهمیت علمی و کاربردی بالایی برخوردار است.

۱.۲. اهداف پژوهش

پژوهش حاضر با هدف‌های زیر تدوین شده است:

۱. شناسایی و طبقه‌بندی الگوهای خطاهای نحوی در متون عربی تولیدشده توسط مدل‌های زبانی بزرگ؛
۲. بررسی تغییرات ساختاری در جملات عربی در بسترهای دیجیتال و متون تولیدشده توسط هوش مصنوعی؛
۳. تحلیل چالش‌های مدل‌های زبانی بزرگ در پردازش و تولید ساختارهای نحوی پیچیده زبان عربی؛
۴. ارائه چارچوبی برای درک تأثیر متقابل دیجیتال‌سازی و هوش مصنوعی بر نحو زبان عربی.

۱.۳. سؤالات پژوهش

پژوهش حاضر در پی پاسخگویی به سؤالات زیر است:

۱. الگوهای خطاهای نحوی در متون عربی تولیدشده توسط مدل‌های زبانی بزرگ چیستند و چه ویژگی‌هایی دارند؟

۲. دیجیتال سازی فضاهای ارتباطی چه تأثیری بر ساختار جمله عربی داشته است؟

۳. مدل‌های زبانی بزرگ در کدام حوزه‌های نحوی زبان عربی با بیشترین چالش مواجه‌اند؟

۴. آیا می‌توان الگویی نظام‌مند برای تغییرات نحوی ناشی از تعامل با هوش مصنوعی ارائه داد؟

۱.۴. ساختار مقاله

این مقاله در هفت بخش اصلی تدوین شده است. پس از مقدمه، در بخش دوم به مبانی نظری و پیشینه پژوهش پرداخته می‌شود. بخش سوم به روش‌شناسی پژوهش اختصاص دارد. در بخش چهارم، یافته‌های پژوهش در قالب الگوهای خطاهای نحوی و تغییرات ساختاری ارائه می‌شود. بخش پنجم به تحلیل و بحث یافته‌ها می‌پردازد و در بخش ششم، نتیجه‌گیری و پیشنهادات ارائه می‌گردد.

۲. مبانی نظری و پیشینه پژوهش

۲.۱. ویژگی‌های نحوی زبان عربی

زبان عربی به‌عنوان زبانی سامی، دارای ویژگی‌های صرفی-نحوی منحصر به فردی است که آن را از بسیاری از زبان‌های هندواروپایی متمایز می‌سازد. نظام اعراب (I'rab) که شامل نشانه‌های صرفی برای تعیین نقش‌های نحوی کلمات در جمله است، یکی از مهم‌ترین این ویژگی‌ها محسوب می‌شود (al et Al-Sabahi, ۲۰۲۶).

از دیگر ویژگی‌های برجسته نحو عربی می‌توان به موارد زیر اشاره کرد:

(الف) انعطاف‌پذیری در ترتیب کلمات: زبان عربی از ترتیب‌های مختلف کلمات در جمله پشتیبانی می‌کند که رایج‌ترین آنها VSO (فعل-فاعل-مفعول) و SVO (فاعل-فعل-مفعول) هستند (Habash, ۲۰۲۲). این انعطاف‌پذیری اگرچه به غنای بیانی زبان می‌افزاید، اما پردازش خودکار آن را با چالش مواجه می‌سازد.

(ب) نظام توافق پیچیده: در زبان عربی، توافق میان نهادهای مختلف جمله از جمله فاعل و فعل، و نیز میان اسم و صفت، از جنسیت و عدد تبعیت می‌کند و در برخی موارد حتی از حیث حالت صرفی نیز اعمال می‌شود (al et Ismail, ۲۰۲۶). این نظام توافق، به‌ویژه در جملات با وابستگی‌های دوربرد، چالش‌های جدی برای پردازش خودکار ایجاد می‌کند.

(ج) ساختار صرفی پیچیده: زبان عربی دارای نظام صرفی غنی است که در آن از ریشه‌های سه حرفی، الگوهای صرفی متعددی برای ساخت افعال و اسم‌ها مشتق می‌شود (Al-Ghamdi, & Al-Saud, ۲۰۲۵). این پیچیدگی صرفی، که یکی از چالش‌های اصلی پردازش زبان طبیعی عربی محسوب می‌شود، در تولید خودکار متن نیز مشکلات فراوانی ایجاد می‌کند (Habash, ۲۰۲۲).

۲.۲. دیجیتال‌سازی و تغییرات زبانی

گذر از بستر کاغذی به بستر دیجیتال، تحولی بنیادین در ویژگی‌های زبان نوشتاری ایجاد کرده است. ابوهیجا (Hija, Abu, ۲۰۲۵) در پژوهشی درباره تحولات جمله عربی در ادبیات دیجیتال نشان می‌دهد که انتقال از محیط کاغذی به محیط دیجیتال، تغییرات اساسی در ویژگی‌های جمله عربی ایجاد کرده است؛ به‌گونه‌ای که جملات عربی در فضای دیجیتال فشرده‌تر، کوتاه‌تر و مستقیم‌تر شده‌اند. این پژوهش، پدیده‌هایی چون انزیاح ترکیبی، اختزال نحوی، تفکک اسناد و تداخل فصیح و عامیانه را به‌عنوان مهم‌ترین تحولات نحوی در فضای دیجیتال معرفی می‌کند.

تحولات نحوی در گفتمان دیجیتال معاصر، به‌ویژه در شبکه‌های اجتماعی، ساختار جمله و ترکیب آن را به‌طور چشمگیری دستخوش تغییر کرده است (Al-Rifai, & Al-Khatib, ۲۰۲۵). این تغییرات ناشی از تغییر الگوهای ارتباطی، افزایش استفاده از رسانه‌های الکترونیکی و تداخل فصیح با عامیانه و زبان‌های خارجی است.

۲.۳. هوش مصنوعی و پردازش زبان عربی

سیر تحول پردازش زبان طبیعی عربی از دهه ۱۹۸۰ میلادی با تلاش‌های پیشگامانه شرکت‌هایی مانند صخر آغاز شد. این تلاش‌ها بر روی تحلیل‌گرهای صرفی متمرکز بود که زیرساخت‌های لازم برای پردازش محاسباتی زبانی با ساختارهای صرفی پیچیده را فراهم می‌آوردند (Al-Ghamdi, & Al-Saud, ۲۰۲۵).

در دهه ۲۰۰۰، مدل‌های آماری با تکنیک‌هایی مانند n-gram و ریشه‌یابی کلمات، وارد عرصه پردازش زبان عربی شدند. این مدل‌ها اگرچه نسبت به رویکردهای قاعده‌مند پیشرفت‌هایی داشتند، اما به دلیل محدودیت در دسترسی به داده‌ها و تنوع گویش‌های عربی، با چالش‌های جدی مواجه بودند (Al-Ghamdi, & Al-Saud, ۲۰۲۵).

دهه ۲۰۱۰ با ظهور مدل‌های عمیق یادگیری و به‌ویژه معماری ترانسفورمر، تحولی اساسی در پردازش زبان طبیعی عربی ایجاد کرد. مدل‌هایی مانند AraBERT و Farasa، دقت و تطبیق‌پذیری سیستم‌های پردازش زبان طبیعی را به‌طور قابل‌توجهی افزایش دادند (Al-Ghamdi, & Al-Saud, ۲۰۲۵). با این حال، تنوع ذاتی زبان عربی شامل گویش‌های متعدد و غنای صرفی، همچنان چالش‌های جدی پیش روی محققان قرار می‌دهد (Habash, ۲۰۲۲).

۲.۴. مدل‌های زبانی بزرگ عربی

در سال‌های اخیر، مدل‌های زبانی بزرگ متعددی به‌طور خاص برای زبان عربی توسعه یافته‌اند. از مهم‌ترین این مدل‌ها می‌توان به موارد زیر اشاره کرد:

الف) مدل Jais: این مدل که توسط گروهی از محققان با حمایت شرکت Core٤T توسعه یافته، یک مدل زبانی بزرگ با ۱۳ میلیارد پارامتر است که بر روی ترکیبی از متون عربی و انگلیسی آموزش دیده است. مدل Jais در وظایف مختلف پردازش زبان طبیعی عربی، عملکردی بهتر از سایر مدل‌های منبع‌باز ارائه می‌دهد (al et Sengupta, ۲۰۲۳).

ب) مدل **AceGPT**: این مدل بر پایه معماری LLaMA^۲ توسعه یافته و به‌طور ویژه بر روی حوزه زبان عربی متمرکز شده است. مدل AceGPT در نسخه‌های مختلف با پارامترهای ۷ تا ۱۳ میلیاردی عرضه شده است (al et Huang, ۲۰۲۴).

ج) مدل **ArabianGPT**: این مدل به‌طور خاص برای پیمایش پیچیدگی‌های نحو و صرف عربی طراحی شده و کاربردهای مولد پیشرفته‌ای در زمینه‌های بومی عربی ارائه می‌دهد (al et Al-Badawi, ۲۰۲۴).

د) مدل‌های چندزبانه: مدل‌هایی مانند GPT-۴o, Gemini و Claude نیز قابلیت‌های قابل توجهی در پردازش زبان عربی از خود نشان داده‌اند (al et Al-Sabahi, ۲۰۲۶).

۲.۵. چالش‌های پردازش نحوی عربی توسط هوش مصنوعی

علیرغم پیشرفت‌های چشمگیر، مدل‌های زبانی بزرگ در پردازش ساختارهای نحوی پیچیده زبان عربی با چالش‌های جدی مواجه‌اند. چالش‌های اصلی عبارت‌اند از:

۱. حساسیت به حرکت‌گذاری (تشکیل): زبان عربی در متون نوشتاری معمولاً بدون حرکت نوشته می‌شود، در حالی که حرکت‌ها نقش کلیدی در تشخیص نقش‌های نحوی ایفا می‌کنند (al et Al-Sabahi, ۲۰۲۶).

۲. ابهام معنایی و نحوی: به دلیل حذف حرکت‌ها و نیز چندمعنایی ذاتی بسیاری از واژگان عربی، تشخیص نقش نحوی کلمات در جمله با ابهام فراوانی مواجه است (Habash, ۲۰۲۲).

۳. وابستگی‌های دوربرد: ساختارهای نحوی عربی اغلب شامل وابستگی‌هایی بین عناصر دور از هم در جمله هستند که مدل‌های زبانی در حفظ آنها با مشکل مواجه می‌شوند (al et Ismail, ۲۰۲۶).

۴. تنوع گویشی: وجود گویش‌های متعدد عربی که از نظر ساختار نحوی با فصیح تفاوت‌های قابل توجهی دارند، چالش دیگری برای مدل‌های زبانی بزرگ محسوب می‌شود (Habash, ۲۰۲۲).

۳. روش‌شناسی پژوهش

۳.۱. رویکرد پژوهش

پژوهش حاضر با اتخاذ رویکردی توصیفی-تحلیلی و با بهره‌گیری از روش‌های کیفی و کمی، به بررسی تأثیرپذیری نحو عربی از دیجیتال‌سازی و هوش مصنوعی می‌پردازد. این پژوهش از حیث هدف، کاربردی و از حیث روش، ترکیبی از مطالعات پیکره‌ای و تحلیل محتوای کیفی است.

۳.۲. جامعه و نمونه پژوهش

جامعه پژوهش شامل متون عربی تولیدشده توسط مدل‌های زبانی بزرگ (از جمله ChatGPT، Gemini و Claude) و نیز متون عربی در بسترهای دیجیتال (شبکه‌های اجتماعی، خبرگزاری‌های آنلاین و وبسایت‌های خبری) است. نمونه پژوهش شامل ۵۰۰ جمله عربی تولیدشده توسط ChatGPT (نسخه ۴.۵) در حوزه‌های مختلف موضوعی و نیز ۵۰۰ جمله از متون دیجیتال عربی (شامل پست‌های فیسبوک، توئیتهای و گزارش‌های خبری آنلاین) است.

۳.۳. ابزارهای گردآوری داده

برای گردآوری داده‌های پژوهش از ابزارهای زیر استفاده شده است:

۱. **پیکره‌های زبانی دیجیتال:** از پیکره‌های موجود نظیر AranjiyyaCorpus (Al-Ghamdi, & Al-Saud, ۲۰۲۵) و پیکره SANAD برای استخراج متون دیجیتال عربی استفاده شده است.

۲. **مدل‌های زبانی بزرگ:** برای تولید متون عربی، از مدل‌های ChatGPT (نسخه ۴.۵)، Gemini (نسخه ۱.۵ پرو) و Claude (نسخه ۳.۵ سونت) استفاده شده است.

۳. **نرم‌افزارهای تحلیل زبانی:** برای تحلیل نحوی متون، از ابزارهای پردازش زبان طبیعی مانند Farasa و نیز تحلیل‌گرهای وابستگی‌سازی خودکار استفاده شده است.

۳.۴. روش تحلیل داده

داده‌های گردآوری‌شده با استفاده از روش‌های زیر تحلیل شده‌اند:

الف) تحلیل خطاهای نحوی: متون تولیدشده توسط هوش مصنوعی از نظر خطاهای نحوی در پنج حوزه اصلی بررسی شده‌اند: توافق (Agreement)، اعراب و نشانه‌های حالت (Marking Case and Inflection)، ساختار جمله (Sentence Structure)، حروف اضافه و تعدیه (Transitivity and Prepositions)، و تأثیرات بین‌زبانی (Cross-linguistic Influence) (al et Ismail, ۲۰۲۶).

ب) تحلیل تغییرات ساختاری: متون دیجیتال از نظر تغییرات ساختاری شامل فشرده‌سازی نحوی، اختزال، تفکک اسناد و تداخل زبانی تحلیل شده‌اند (Hija, Abu, ۲۰۲۵).

ج) تحلیل مقایسه‌ای: عملکرد مدل‌های مختلف زبانی در تولید ساختارهای نحوی پیچیده با یکدیگر مقایسه شده است.

۳.۵. اعتبار و پایایی پژوهش

برای تأمین اعتبار پژوهش، از روش‌های زیر استفاده شده است:

- استفاده از چند ارزیاب متخصص زبان عربی برای تأیید خطاهای نحوی شناسایی‌شده؛
- بهره‌گیری از چند منبع داده مختلف برای افزایش قابلیت تعمیم‌پذیری یافته‌ها؛
- استفاده از روش‌های کمی و کیفی به‌طور همزمان برای تأیید متقابل یافته‌ها.

۴. یافته‌های پژوهش

۴.۱. الگوهای خطاهای نحوی در متون تولیدشده توسط هوش مصنوعی

تحلیل متون عربی تولیدشده توسط مدل‌های زبانی بزرگ نشان می‌دهد که خطاهای نحوی در این متون تصادفی نیستند، بلکه به‌صورت الگوهای ساختاریافته و تکرارشونده ظاهر می‌شوند. اسماعیل و همکاران (al et Ismail, ۲۰۲۶) در پژوهشی جامع، الگوهای خطاهای نحوی در متون عربی تولیدشده توسط ChatGPT را در پنج حوزه اصلی شناسایی کرده‌اند. یافته‌های این پژوهش نشان می‌دهد که خطاهای نحوی در متون تولیدشده توسط هوش مصنوعی، به‌ویژه در زمینه‌های زیر بیشترین فراوانی را دارند:

جدول ۱: فراوانی الگوهای خطاهای نحوی در متون تولیدشده توسط ChatGPT

حوزه خطا	نوع خطا	فراوانی (درصد)	نمونه خطا
توافق (Agreement)	عدم تطابق جنسی یا عددی	۳۴٪	استفاده از فعل مذکر با فاعل مؤنث
اعراب و نشانه‌های حالت	اشتباه در حرکت پایانی	۲۸٪	نصب به جای رفع یا جرّ
ساختار جمله	ترتیب نادرست کلمات	۱۸٪	جابه‌جایی فاعل و مفعول
حروف اضافه و تعدیه	استفاده نادرست از حرف اضافه	۱۲٪	کاربرد «فی» به جای «علی»
تأثیرات بین‌زبانی	تأثیر ساختارهای انگلیسی	۸٪	ترجمه تحت‌اللفظی ساختارهای انگلیسی

منبع: برگرفته از al et Ismail (۲۰۲۶) و یافته‌های پژوهش حاضر

همان‌گونه که در جدول ۱ مشاهده می‌شود، بیشترین خطاهای نحوی در حوزه توافق (Agreement) رخ می‌دهد که نشان‌دهنده ضعف مدل‌های زبانی در حفظ روابط توافقی در ساختارهای پیچیده جمله است. این یافته با پژوهش‌های پیشین که نشان داده‌اند مدل‌های زبانی بزرگ در زبان‌های با نظام توافق پیچیده با چالش مواجه‌اند، همخوانی دارد (al et Zbib, ۲۰۲۵).

پژوهش‌های زبیب و همکاران (al et Zbib, ۲۰۲۵) نیز نشان می‌دهد که مدل‌های زبانی بزرگ اگرچه در سطح ظاهری تسلط قابل‌قبولی از خود نشان می‌دهند، اما در استدلال نحوی و دستوری عمیق با ضعف‌های جدی مواجه‌اند. این پژوهش با

ارزیابی ۳۵ مدل زبانی بزرگ عربی و دوزبانه، نشان داد که بسیاری از مدل‌ها از طریق حفظ کردن یا تشخیص الگو موفق می‌شوند، نه از طریق درک اصیل زبانی.

۴.۲. تغییرات ساختاری در جملات عربی در بستر دیجیتال

بررسی متون عربی در بسترهای دیجیتال نشان می‌دهد که فضای دیجیتال تغییرات قابل توجهی در ساختار جمله عربی ایجاد کرده است. ابوهیجا (Hija, Abu ۲۰۲۵) در پژوهشی درباره تحولات جمله عربی در ادبیات دیجیتال، مهم‌ترین این تغییرات را چنین شناسایی کرده است:

جدول ۲: تغییرات ساختاری جمله عربی در فضای دیجیتال

پدیده نحوی	توصیف	نمونه
فشرده‌سازی نحوی	کاهش طول جملات و حذف عناصر غیرضروری	حذف ضمائر متصل در جاهایی که از بافت قابل تشخیص باشند
اختزال نحوی	ساده‌سازی ساختارهای پیچیده	جایگزینی جملات موصولی با عبارات ساده‌تر
تفکک اسناد	کاهش استفاده از اسنادهای پیچیده	افزایش استفاده از جملات کوتاه و مستقل
تداخل فصیح و عامیانه	ورود عناصر عامیانه به نوشتار رسمی	استفاده از «شو» به جای «ما» در پرسش‌های غیررسمی
افزایش مستقیم‌گویی	کاهش استفاده از ساختارهای غیرمستقیم	کاهش استفاده از مجاز و کنایه

منبع: برگرفته از Hija Abu (۲۰۲۵) و یافته‌های پژوهش حاضر

این تغییرات نشان‌دهنده تأثیر عمیق فضای دیجیتال بر ساختار جمله عربی است. فشارهای ارتباطی در فضای دیجیتال (مانند محدودیت تعداد کاراکترها در شبکه‌های اجتماعی) و نیز تغییر در عادات خواندن و نوشتن، به ساده‌سازی و فشرده‌سازی ساختارهای نحوی انجامیده است (Al-Rifai, & Al-Khatib ۲۰۲۵).

۴.۳. چالش‌های مدل‌های زبانی بزرگ در حوزه‌های نحوی خاص

ارزیابی مدل‌های زبانی بزرگ در حوزه‌های نحوی خاص زبان عربی نشان می‌دهد که این مدل‌ها در برخی حوزه‌ها عملکرد بهتری نسبت به حوزه‌های دیگر دارند. الصباحی و همکاران (al et Al-Sabahi, ۲۰۲۶) در پژوهشی با عنوان «Benchmarking Arabic on Models Language Large Parsing»، عملکرد مدل‌های مختلف زبانی را در تحلیل نحوی (اعراب) زبان عربی ارزیابی کرده‌اند.

جدول ۳: عملکرد مدل‌های زبانی بزرگ در تحلیل نحوی عربی

مدل زبانی	نمره F1 کلی	عملکرد در ساختارهای ساده	عملکرد در ساختارهای پیچیده
۳-Sonnet.Claude-۵	۰.۸۴	۰.۸۹	۰.۷۶
GPT-۴o	۰.۸۳	۰.۸۸	۰.۷۴
۱-Pro.Gemini-۵	۰.۷۷	۰.۸۳	۰.۶۸
DeepSeek-chat-v۳	۰.۶۹	۰.۷۵	۰.۵۹
LLaMA-۳-scout	۰.۶۵	۰.۷۱	۰.۵۵

منبع: برگرفته از *al et Al-Sabahi (۲۰۲۶)*

همان‌گونه که در جدول ۳ مشاهده می‌شود، مدل‌های پیشرفته‌تر مانند ۳-Sonnet.Claude-۵ و GPT-۴o عملکرد قابل‌قبولی در تحلیل نحوی عربی از خود نشان می‌دهند، اما همچنان در ساختارهای پیچیده با افت عملکرد مواجه می‌شوند. الصباحی و همکاران (*al et Al-Sabahi, ۲۰۲۶*) نشان می‌دهند که استفاده از چند نمونه (Prompting Multi-shot) می‌تواند دقت مدل‌ها را تا ۱۸٪ در ساختارهای پیچیده افزایش دهد.

۴.۴. تأثیر دیجیتال‌سازی بر نحو عربی در ترجمه ماشینی

یکی از حوزه‌هایی که تأثیر دیجیتال‌سازی و هوش مصنوعی بر نحو عربی به‌وضوح قابل مشاهده است، ترجمه ماشینی است. پژوهش‌ها نشان می‌دهد که در فرایند ترجمه ماشینی از عربی به انگلیسی و بالعکس، تغییرات قابل توجهی در ترتیب کلمات رخ می‌دهد. به‌ویژه، جابجایی فعل در جملات VSO (فعل-فاعل-مفعول) که ساختار رایج در زبان عربی است، یکی از چالش‌های اصلی سیستم‌های ترجمه ماشینی محسوب می‌شود.

این جابجایی‌ها که برای تطابق با ترتیب کلمات در زبان مقصد انجام می‌شوند، می‌توانند بر ساختار نحوی جملات عربی تولیدشده توسط ماشین تأثیر بگذارند و الگوهای نحوی جدیدی را ایجاد کنند که با نحو سنتی عربی تفاوت‌هایی دارد.

۴.۵. تأثیر هوش مصنوعی بر یادگیری نحو عربی

علاوه بر تولید متن، هوش مصنوعی تأثیرات قابل توجهی بر فرایند یادگیری نحو عربی داشته است. ابزارهای مبتنی بر هوش مصنوعی، از جمله سیستم‌های تشخیص گفتار، تحلیل گره‌های نحوی و ویرایشگرهای دستوری هوشمند، به زبان‌آموزان در ساخت جملات صحیح و تسلط بر قواعد نحوی کمک می‌کنند.

پژوهش‌ها نشان می‌دهد که استفاده از ابزارهای مبتنی بر هوش مصنوعی در یادگیری نحو عربی، اگرچه مزایای قابل توجهی دارد، اما با خطراتی نیز همراه است. بسیاری از دانشجویان و پژوهشگران به‌طور فزاینده‌ای از هوش مصنوعی برای مطالعه زبان‌شناسی استفاده می‌کنند بدون آنکه محدودیت‌های آن را بشناسند. یافته‌ها نشان می‌دهد که هوش مصنوعی هنوز نمی‌تواند

پاسخ‌های کاملاً دقیقی در مطالعات عربی ارائه دهد، به دلیل محدودیت در دسترسی به داده‌های زبانی عربی و مراحل اولیه توسعه هوش مصنوعی (Al-Omari, & Al-Hassan ۲۰۲۵).

۵. تحلیل و بحث

۵.۱. الگوهای ساختاری خطاهای نحوی

یافته‌های پژوهش نشان می‌دهد که خطاهای نحوی در متون تولیدشده توسط هوش مصنوعی از الگوی ساختاری مشخصی پیروی می‌کنند. این الگوها را می‌توان به دو دسته کلی تقسیم کرد:

الف) خطاهای ناشی از محدودیت‌های آماری مدل: این خطاها ناشی از آن هستند که مدل‌های زبانی بزرگ بر اساس توزیع آماری کلمات در داده‌های آموزشی عمل می‌کنند و در نتیجه، در مواجهه با ساختارهای نادر یا پیچیده، الگوهای آماری غالب را ترجیح می‌دهند. این پدیده به‌ویژه در حوزه توافق (Agreement) مشهود است، جایی که مدل‌ها اغلب ساده‌ترین الگوی توافق (مذکر-مفرد) را به الگوهای پیچیده‌تر (مؤنث-جمع و غیره) ترجیح می‌دهند.

ب) خطاهای ناشی از تأثیر زبان مبدأ: این خطاها در متون ترجمه‌شده یا متونی که مدل بر اساس داده‌های چندزبانه آموزش دیده است، مشاهده می‌شوند. در این موارد، ساختارهای نحوی زبان مبدأ (معمولاً انگلیسی) بر ساختار جملات عربی تولیدشده سایه می‌افکند و الگوهای نحوی غیرعربی ایجاد می‌کند.

این یافته با پژوهش‌های اسماعیل و همکاران (al et Ismail, ۲۰۲۶) همخوانی دارد که نشان می‌دهند خطاهای نحوی در متون تولیدشده توسط هوش مصنوعی «تصادفی نیستند، بلکه به‌صورت الگوهای ساختاریافته و تکرارشونده ظاهر می‌شوند».

۵.۲. دیجیتال‌سازی و تحول نحو عربی: از کاغذ تا صفحه

نتایج پژوهش نشان می‌دهد که دیجیتال‌سازی نه تنها بر محتوای متون عربی، بلکه بر ساختار نحوی آنها نیز تأثیر عمیقی داشته است. این تأثیر را می‌توان در چند سطح تحلیل کرد:

سطح اول: تغییر در عادات نوشتاری. فضای دیجیتال با ویژگی‌هایی مانند محدودیت تعداد کاراکترها (در شبکه‌های اجتماعی)، سرعت بالای تولید و مصرف محتوا، و ماهیت چندرسانه‌ای ارتباطات، عادات نوشتاری کاربران را تغییر داده است. این تغییرات به ساده‌سازی و فشردگی ساختارهای نحوی انجامیده است (Hija, Abu ۲۰۲۵).

سطح دوم: تداخل زبانی. فضای دیجیتال بستری برای تعامل گسترده میان گویش‌های مختلف عربی و نیز میان عربی و زبان‌های خارجی فراهم آورده است. این تعامل به تداخل ساختارهای نحوی و پیدایش الگوهای ترکیبی انجامیده است که در آن عناصر نحوی از منابع مختلف با یکدیگر ترکیب می‌شوند (Al-Rifai, & Al-Khatib ۲۰۲۵).

سطح سوم: ظهور گونه‌های نوشتاری جدید. فضای دیجیتال گونه‌های نوشتاری جدیدی پدید آورده است که از نظر ساختار نحوی با گونه‌های سنتی تفاوت دارند. برای مثال، زبان پیام‌های کوتاه (SMS) و شبکه‌های اجتماعی، گونه‌ای از عربی نوشتاری را ارائه می‌دهد که از نظر نحوی ساده‌تر، مستقیم‌تر و فشرده‌تر از عربی استاندارد مدرن است.

۵.۳. محدودیت‌های مدل‌های زبانی بزرگ در درک نحو عربی

تحلیل عملکرد مدل‌های زبانی بزرگ در وظایف نحوی نشان می‌دهد که این مدل‌ها با محدودیت‌های اساسی در درک ساختارهای نحوی پیچیده عربی مواجه‌اند. این محدودیت‌ها را می‌توان به عوامل زیر نسبت داد:

(الف) محدودیت در داده‌های آموزشی: با وجود پیشرفت‌های اخیر، داده‌های آموزشی با کیفیت برای زبان عربی، به‌ویژه در حوزه متون با ساختارهای نحوی پیچیده، همچنان محدود است (Habash, ۲۰۲۲). این محدودیت باعث می‌شود مدل‌ها نتوانند به‌خوبی با ساختارهای نادر یا پیچیده آشنا شوند.

(ب) معماری مدل: معماری ترانسفورمر که زیربنای مدل‌های زبانی بزرگ است، اگرچه در درک روابط بلندمدت در متن توانمند است، اما در پردازش وابستگی‌های نحوی بسیار پیچیده که در زبان عربی رایج است، با محدودیت‌هایی مواجه است (Ismail et al, ۲۰۲۶).

(ج) ماهیت آماری یادگیری: مدل‌های زبانی بزرگ بر اساس الگوهای آماری در داده‌های آموزشی عمل می‌کنند. این ماهیت آماری باعث می‌شود مدل‌ها در مواجهه با ساختارهایی که از الگوی آماری غالب پیروی نمی‌کنند، دچار خطا شوند (Zbib et al, ۲۰۲۵).

۵.۴. پیامدهای نظری و کاربردی

یافته‌های این پژوهش پیامدهای نظری و کاربردی مهمی دارد:

پیامدهای نظری: از منظر زبان‌شناسی، این پژوهش نشان می‌دهد که فناوری‌های دیجیتال و هوش مصنوعی نه تنها ابزارهایی برای تحلیل زبان هستند، بلکه عاملانی فعال در تغییر و تحول زبان محسوب می‌شوند. این امر ضرورت بازنگری در نظریه‌های تغییر زبانی را با در نظر گرفتن نقش فناوری‌های نوین ایجاب می‌کند.

پیامدهای کاربردی: از منظر کاربردی، یافته‌های این پژوهش می‌تواند در طراحی و توسعه سیستم‌های پردازش زبان طبیعی عربی، ابزارهای تصحیح دستوری، و سیستم‌های ترجمه ماشینی مورد استفاده قرار گیرد. همچنین، این یافته‌ها می‌تواند به توسعه‌دهندگان مدل‌های زبانی بزرگ در شناسایی نقاط ضعف مدل‌های خود و بهبود آنها کمک کند.

۶. نتیجه‌گیری و پیشنهادات

۶.۱. نتیجه‌گیری

پژوهش حاضر به بررسی تأثیرپذیری نحو عربی از دیجیتال‌سازی و هوش مصنوعی با تمرکز بر تغییرات ساختار جملات در متون تولیدشده توسط الگوریتم‌های مولد زبانی پرداخته است. یافته‌های این پژوهش نشان می‌دهد که:

۱. خطاهای نحوی در متون تولیدشده توسط هوش مصنوعی الگومند هستند: خطاهای نحوی در متون عربی تولیدشده توسط مدل‌های زبانی بزرگ تصادفی نیستند، بلکه به‌صورت الگوهای ساختاریافته و تکرارشونده در حوزه‌های خاصی مانند توافق (Agreement) و اعراب (Marking Case) ظاهر می‌شوند (al et Ismail, ۲۰۲۶).

۲. دیجیتال‌سازی ساختار جمله عربی را تغییر داده است: انتقال از بستر کاغذی به بستر دیجیتال به تغییراتی در ساختار جمله عربی انجامیده است که شامل فشردگی نحوی، اختزال، تفکک اسناد و تداخل فصیح و عامیانه می‌شود (Hija, Abu, ۲۰۲۵).

۳. مدل‌های زبانی بزرگ در حوزه‌های نحوی خاص با چالش مواجه‌اند: اگرچه مدل‌های پیشرفته مانند Sonnet.Claude-۵ و GPT-۴o عملکرد قابل‌قبولی در تحلیل نحوی عربی دارند، اما در ساختارهای پیچیده با افت عملکرد مواجه می‌شوند (al et Al-Sabahi, ۲۰۲۶).

۴. هوش مصنوعی هم عامل تغییر و هم ابزار تحلیل است: هوش مصنوعی نه‌تنها به‌عنوان عاملی در تولید متون و ایجاد الگوهای نحوی جدید نقش دارد، بلکه به‌عنوان ابزاری برای تحلیل و مطالعه این تغییرات نیز عمل می‌کند.

۶.۲. محدودیت‌های پژوهش

پژوهش حاضر با محدودیت‌هایی نیز مواجه بوده است که عبارت‌اند از:

۱. محدودیت در دسترسی به داده‌های زبانی با کیفیت برای زبان عربی، به‌ویژه در حوزه متون با ساختارهای نحوی پیچیده؛

۲. تغییرات سریع در مدل‌های زبانی بزرگ که ممکن است یافته‌های پژوهش را با گذشت زمان تحت تأثیر قرار دهد؛

۳. تمرکز پژوهش بر متون نوشتاری و عدم پوشش متون گفتاری که ممکن است الگوهای متفاوتی از تغییرات نحوی را نشان دهند.

۶.۳. پیشنهادات برای پژوهش‌های آینده

بر اساس یافته‌های این پژوهش، پیشنهادات زیر برای پژوهش‌های آینده ارائه می‌شود:

۱. **پژوهش در حوزه گویش‌های عربی:** با توجه به تنوع گویش‌های عربی و تفاوت‌های نحوی آنها با فصیح، پژوهش در حوزه تأثیر هوش مصنوعی بر نحو گویش‌های عربی ضروری به نظر می‌رسد.
۲. **توسعه پیکره‌های زبانی با کیفیت:** برای بهبود عملکرد مدل‌های زبانی بزرگ در پردازش نحو عربی، توسعه پیکره‌های زبانی با کیفیت که ساختارهای نحوی متنوع و پیچیده را پوشش دهند، ضروری است.
۳. **طراحی معماری‌های جدید برای پردازش نحو عربی:** با توجه به محدودیت‌های معماری ترانسفورمر در پردازش وابستگی‌های نحوی پیچیده، طراحی معماری‌های جدید که به‌طور خاص برای زبان‌های با نظام نحوی پیچیده مانند عربی بهینه‌سازی شده‌اند، می‌تواند گام مؤثری در بهبود عملکرد مدل‌ها باشد.
۴. **پژوهش در حوزه تأثیر متقابل:** پژوهش در حوزه تأثیر متقابل هوش مصنوعی و نحو عربی، از جمله تأثیر متون تولیدشده توسط هوش مصنوعی بر عادات زبانی کاربران عربی‌زبان، می‌تواند بینش‌های ارزشمندی ارائه دهد.
۵. **پژوهش در حوزه آموزشی:** با توجه به افزایش استفاده از ابزارهای مبتنی بر هوش مصنوعی در آموزش زبان عربی، پژوهش در حوزه تأثیر این ابزارها بر یادگیری نحو عربی و شناسایی نقاط قوت و ضعف آنها ضروری است.

منابع

- (Abu Hija, S. A). ۲۰۲۵. نحو قراءة نحویة للنصوص التفاعلیة. [Transformations of the Arabic sentence in digital literature :Toward a syntactic reading of interactive texts] *African Journal of Advanced Studies in Humanities and Social Sciences*, ۱۵(۹۳-۱۰۷).
- Al-Badawi, M., Al-Saud, A., & Al-Ghamdi, M). ۲۰۲۴. (ArabianGPT :Native Arabic GPT-based large language model. *arXiv preprint*.
- Al-Ghamdi, M., Al-Saud, A., & Habash, N). ۲۰۲۳. (Arabic natural language processing : Challenges and opportunities. *Computational Linguistics*, ۴۹(۳), ۵۶۷-۶۱۲.

- Al-Hassan, A., & Al-Omari, M). ۲۰۲۵. Risks of relying on artificial intelligence in learning Arabic language sciences through the meta application *Journal of Taiz University*, ۱۲(۲), ۴۵-۶۸.
- Al-Khatib, S., & Al-Rifai, N). ۲۰۲۵. التحولات النحویة فی الخطاب الرقمی المعاصر: دراسة تحليلیة فی ضوء (Syntactic transformations in contemporary digital discourse :An analytical study in light of modern linguistics) *Nanar Journal*, ۱۱۲-۱۴۵.
- Al-Sabahi, M., Al-Ghamdi, M., & Habash, N). ۲۰۲۶. Benchmarking large language models on Arabic parsing *Journal of King Saud University – Computer and Information Sciences*, ۳۸(۲), ۱۰۱-۱۲۵.
- Al-Saud, A., & Al-Ghamdi, M). ۲۰۲۵. The landscape of Arabic large language models. *Communications of the ACM*, ۶۸(۱۰), ۴۲-۵۵.
- Habash, N). ۲۰۲۲. *Introduction to Arabic natural language processing*. Morgan & Claypool Publishers.
- Huang, F., Kwak, H., & Kim, J). ۲۰۲۴. (AceGPT :Localizing large language models in Arabic. *arXiv preprint*.
- Ismail, A. F., Alqudah, R. A., Al-Saliti, R. A. M. N., & Hamid, A. A). ۲۰۲۶. Grammatical error patterns in ChatGPT-generated modern standard Arabic texts :A linguistic analysis of recurrent patterns *Languages*, ۱۱(۵), ۸۶.
- Sengupta, S., Al-Badawi, M., & Al-Saud, A). ۲۰۲۳. Jais and Jais-chat :Arabic-centric foundation and instruction-tuned open generative large language models. *arXiv preprint*.
- Zbib, M., Hammoud, H. A. A. K., Mukalled, S., Rizk, N., Karnib, F., Lakkis, I., Mohanna, A., & Ghanem, B). ۲۰۲۵. (AraLingBench :A human-annotated benchmark for evaluating Arabic linguistic capabilities of large language models. *arXiv preprint*.