

Investigating the Effectiveness, the Intelligent Speech Recognition System in Correctly Retrieving Narrative Literature Information

Oranus Tajedini¹  | Nasrin Fallah²  | Arezoo Momeni³ 

1. Corresponding author, Associate Professor, Department of Knowledge and Information Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: tajedini@uk.ac.ir
2. Assistant Professor, Department of Persian Literature and Language, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: n.fallah@uk.ac.ir
3. M.A., Department of Knowledge and Information Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: aarezoo.mni@gmail.com

Article Info

Article type:

Research Article

Article history:

Received 10 March 2023

Received in revised form 21 April 2023

Accepted 29 April 2023

Published online 10 June 2023

Keywords:

Intelligent speech recognition system,
Effectiveness,
Correct information retrieval,
Narrative literature,
Storytelling software,
Elementary school children.

ABSTRACT

Objective: This research has been carried out with the aim of investigating the effectiveness of using the intelligent speech recognition system in correctly retrieving the information of narrative literature for elementary school children.

Methodology: This research is practical in terms of purpose and was done with a semi-experimental method. The statistical population of the research consists of first to third grade elementary school children, 36 of whom were selected as a sample using stratified sampling method. In order to conduct this study, two intelligent speech recognition system tools and a storytelling application called "Audio Story Book" were used. To collect the data, a questionnaire and a checklist designed according to the research questions. The questionnaire was used to complete the survey and evaluate the determined criteria. SPSS27 statistical software and one-way analysis of variance statistical test were used to analyze the data.

Results: The findings of this research showed that the average number of errors in the story information retrieved by the children of the intelligent speech recognition system group with an average of 0.7 is more than the human voice group and less than the storytelling application group. The average length of story information retrieved by the children of the intelligent speech recognition system group with an average of 3.966 is less than the human voice group and more than the storytelling application group. The average speed of retrieving story information by children in the intelligent speech system group with an average of 2.583 is lower than the human voice group and higher than the storytelling application group. The average accuracy of story information retrieved by the speech recognition system group children is lower than the human voice group and higher than the storytelling application group with an average of 3.4.

Conclusions: The results of this research showed that the use of intelligent speech recognition system is not significantly effective in correctly retrieving the information of narrative literature of primary school children. Children who were told stories by a human had higher accuracy and concentration, and this could be due to physical, emotional, eye, face-to-face communication and environmental conditions that affected their psyche and increased memory and learning performance, increased attention and concentration, quick processing of information, perceptual organization and recall of information, which led to their complete understanding and comprehension.

Cite this article: Tajedini, O., Fallah, N., & Momeni, A. (2023). Investigating the effectiveness, the intelligent speech recognition system in correctly retrieving narrative literature information. *Academic Librarianship and Information Research*, 57 (2), 41-54. <http://doi.org/10.22059/jlib.2023.363017.1703>



© The Author(s).

DOI: <http://doi.org/10.22059/jlib.2023.363017.1703>

Publisher: University of Tehran.

بررسی اثربخشی سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات ادبیات روایی

اورانوس تاج‌الدینی^۱ | نسربین فلاح^۲ | آرزو مؤمنی^۳

۱. نویسنده مسئول، دانشیار، گروه علم اطلاعات و دانش‌شناسی، دانشگاه شهید باهنر کرمان، کرمان، ایران. رایانامه: tajedini@uk.ac.ir

۲. استادیار، گروه ادبیات، دانشگاه شهید باهنر کرمان، کرمان، ایران. رایانامه: n.fallah@uk.ac.ir

۳. کارشناس ارشد، گروه علم اطلاعات و دانش‌شناسی، دانشگاه شهید باهنر کرمان، کرمان، ایران. رایانامه: aarezo.mni@gmail.com

چکیده

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

هدف: این پژوهش با هدف کشف میزان اثربخشی استفاده از سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات ادبیات روایی در کودکان دبستانی و مقایسه آن با تأثیر استفاده از نرم‌افزار قصه‌گویی و قصه‌گوی انسانی در این مورد صورت گرفته است.

روش‌شناسی: این پژوهش از نظر هدف کاربردی بوده و با روش نیمه‌تجربی انجام شده است. جامعه آماری این پژوهش را کودکان دبستانی دختر و پسر پایه اول تا سوم دبستان تشکیل می‌دهند که ۳۶ نفر از آنها با روش نمونه‌گیری طبقه‌ای در دسترس به عنوان نمونه انتخاب شدند. برای انجام این مطالعه از دو ابزار سیستم تشخیص گفتار هوشمند و برنامه کاربردی (اپلیکیشن) قصه‌گویی به نام «کتاب قصه صوتی» استفاده شد. جهت گردآوری داده‌ها از یک پرسش‌نامه و همچنین یک سیاهه واریسی که مطابق با سؤالات پرسش‌نامه طراحی گردیده به منظور تکمیل بررسی و ارزیابی معیارهای تعیین شده، استفاده گردید. جهت تجزیه و تحلیل داده‌ها نرم‌افزار آماری SPSS27 و آزمون آماری تحلیل واریانس یک‌طرفه به کار برده شده است.

تاریخ دریافت: ۱۴۰۱/۱۲/۲۰

تاریخ بازنگری: ۱۴۰۲/۰۲/۰۱

تاریخ پذیرش: ۱۴۰۲/۰۲/۱۰

تاریخ انتشار: ۱۴۰۲/۰۳/۲۰

کلیدواژه‌ها:

سیستم تشخیص گفتار هوشمند، اثربخشی، بازیابی صحیح اطلاعات، ادبیات روایی، نرم‌افزار قصه‌گویی، کودکان دبستانی.

یافته‌ها: یافته‌ها نشان داد که متوسط تعداد خطاها در اطلاعات بازیابی شده از قصه‌ها توسط کودکان هنگام استفاده از سیستم تشخیص گفتار هوشمند با میانگین ۰/۷، نسبت به هنگام استفاده از قصه‌گوی انسانی بیشتر و نسبت به هنگام استفاده از برنامه کاربردی (اپلیکیشن) قصه‌گویی کمتر است. متوسط طول اطلاعات بازیابی شده توسط کودکان در گروهی که از سیستم تشخیص گفتار هوشمند استفاده می‌کردند با میانگین ۳/۹۶۶، از گروه صدای انسان کمتر و از گروه برنامه کاربردی قصه‌گویی بیشتر است. متوسط سرعت بازیابی اطلاعات قصه توسط کودکان گروهی که از سیستم گفتار هوشمند استفاده می‌کردند با میانگین ۲/۵۸۳، از گروه صدای انسان کمتر و از گروه برنامه کاربردی قصه‌گویی بیشتر است. متوسط صحت اطلاعات بازیابی شده توسط کودکان در گروهی که از سیستم تشخیص گفتار استفاده می‌کردند با میانگین ۴/۳، از گروه صدای انسان کمتر و از گروه برنامه کاربردی قصه‌گویی بیشتر است.

نتیجه‌گیری: نتایج این پژوهش نشان داد استفاده از سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات ادبیات روایی کودکان دبستانی به طور محسوس اثربخش نیست. کودکانی که داستان‌ها را یک انسان برایشان تعریف کردند دقت و تمرکز بالاتری داشتند و این موضوع می‌تواند ناشی از برقراری ارتباط فیزیکی، عاطفی، چشمی، چهره‌به‌چهره و شرایط محیطی باشد که بر روی روان آنها تأثیر گذاشته و سبب افزایش عملکرد حافظه و یادگیری، افزایش میزان توجه و تمرکز، پردازش سریع اطلاعات، سازمان‌دهی ادراکی و یادآوری اطلاعات شده که این عوامل درک و فهم کامل آنها را به دنبال داشته است.

استناد: تاج‌الدینی، اورانوس؛ فلاح، نسربین، و مؤمنی، آرزو (۱۴۰۲). بررسی اثربخشی سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات ادبیات روایی. تحقیقات

کتابداری و اطلاع‌رسانی دانشگاهی، ۵۷ (۲)، ۵۴-۴۱. <http://doi.org/10.22059/jlib.2023.363017.1703>



در دهه اخیر بیشتر مردم خواهان این هستند که همه اقدامات و فعالیت‌ها را با روشی بسیار آسان‌تر و سریع‌تر انجام دهند. برای تحقق این امر، بیشتر امور باید مدرن و رایانه‌ای شود تا بتوان عملکرد را به صورت کارآمد انجام داد، به این ترتیب بسیاری از فناوری‌های جدید در زمینه رایانه ظهور کرده‌اند (الوارازی و سوسندران^۱، ۲۰۲۰). یکی از این فناوری‌های جدید تشخیص گفتار است که فرایند تبدیل سیگنال صوتی ضبط شده توسط میکروفون یا تلفن، به مجموعه‌ای از کلمات تعریف می‌شود (زراری^۲ و همکاران، ۲۰۱۶). با توجه به رشد سریع فناوری‌ها و تأثیر اثبات شده آنها در آموزش و یادگیری، استفاده از انواع مختلف فناوری در سطوح مختلف سنی در جوامع مختلف رواج یافته است. نتایج پژوهش‌های علمی نشان داده که حواس مختلف در یادگیری نقش یکسانی ندارند. مهمترین دلیل استفاده از فناوری آموزشی، نیز همین است زیرا می‌تواند چند حس را با هم درگیر و تأثیرگذاری را بیشتر کند. یکی از فناوری‌های فعلی که جایگاه خود را در میان ابزارهای آموزشی و پژوهشی باز کرده، انواع مختلف سیستم تشخیص گفتار هوشمند است. امروزه فناوری‌های آموزشی راه خود را به سستی‌ترین رشته‌ها مانند ادبیات هم باز کرده است.

ادبیات^۳ هر دوره، نماینده روحیات و فراز و نشیب‌های آن دوره بوده و همراه با تحولات اجتماعی در زمان‌های مختلف تغییر می‌یابد و در هر عصر معنا و مفهوم خاص خود را پیدا می‌کند (پوردربایی و سعیدی، ۱۳۹۲). ادبیات کودکان که از بنیادی‌ترین و مهمترین مباحث آموزش و پرورش محسوب می‌شود، عبارت از هرگونه متنی که با چارچوب و طرح مشخصی به طور عمد و یا غیرعمدی برای کودکان آفریده شده، است. ادبیات روایی کودک، نوشته‌ها و سروده‌هایی هستند که ارزش ادبی یا هنری دارند و مخاطب آن کودکان هستند (حاجی نصرالله، ۱۳۹۶). از گونه‌های پرکاربرد ادبیات روایی برای کودکان، داستان یا قصه است که روایتی است تخیلی به نثر یا به شعر که تأکیدش بیشتر بر حوادث خارق‌العاده، شگفت‌انگیز و نامعمول است (شهبازی، ۱۳۹۱، ۷). کودک از لحاظ روان‌شناسی یعنی فردی که هنوز هوش هیجانی کافی را برای زندگی به دست نیاورده است (ایرانمنش، ۱۳۹۸). با توجه به ماهیت کودک و آموزش و سرگرمی برای او یکی از بهترین روش‌ها در این امر استفاده از قصه‌های آموزنده و شیرین است. این امر علاوه بر جنبه‌های سرگرم‌کننده می‌تواند در تقویت مهارت خواندن، آشنایی با اصول زندگی، آموزش مهارت‌های اخلاقی و اجتماعی، ایجاد مهارت شنیدن و به‌خاطر سپردن و... نقش به‌سزایی داشته باشد. داستان خوب در انواع گونه‌ها نقش ویژه‌ای در رشد و پرورش زبانی، رشد و پرورش شناختی، رشد و پرورش عاطفی-شخصیتی، رشد و پرورش اجتماعی کودکان دارد. از گونه‌های داستانی، داستان واقع‌گرا و داستان فانتزی در ادبیات کودک بیشتر دیده می‌شود (حاجی نصرالله، ۱۳۹۶). البته شایان ذکر است که دریافت پیام و بازیابی اطلاعاتی^۴ که در هر یک از این داستان‌ها نهفته است یکی از اصلی‌ترین رسالت‌هاست.

بازیابی اطلاعات در پاسخ به چالش‌های مختلف دسترسی به اطلاعات، برای ارائه رویکردهای اصولی در جستجوی اشکال مختلف محتوا تکامل یافته است. با افزایش روزافزون حجم اطلاعات ذخیره شده در منابع قابل‌دسترس و گوناگون، فرایند بازیابی و استخراج اطلاعات اهمیت ویژه‌ای دارد (منینگ^۵ و همکاران، ۲۰۰۹). بازیابی صحیح اطلاعات فرایندی است که در آن مجموعه‌ای از سوابق یا اسناد جستجو می‌شوند تا مواردی را پیدا کنند که به تأمین نیاز اطلاعاتی یا علاقه یک فرد یا گروه کمک می‌کند (کاگولوسکی و موهر^۶، ۲۰۰۳). اگرچه امروزه استفاده از سیستم‌های رایانه‌ای، جستجو و بازیابی اطلاعات را تسهیل نموده است ولی جستجو و بازیابی تاکنون به طور عمده در محتوای متنی انجام می‌پذیرفته است؛ در حالی که نیاز به استخراج اطلاعات از محتوای غیر متنی مانند صدا نیز احساس می‌شود. بنابراین به دنبال روش یا روش‌هایی هستیم که تنها با استفاده از محتوای صدا، جستجو و بازیابی اطلاعات را در آن انجام دهد. به این روش‌ها بازیابی مبتنی بر محتوا گفته می‌شود که بازیابی مبتنی بر محتوای صوتی یکی از این موارد محسوب می‌شود. بازیابی مبتنی بر محتوای صوتی را می‌توان به این صورت تعریف کرد: یک مجموعه از مستندات صوتی در اختیار است. از بین این مجموعه مستندات، اطلاعاتی را باید پیدا کرد که درجه شباهت و یا میزان ارتباط و نزدیکی آن با موضوع مورد نظر بیشتر باشد. یکی از روش‌های بازیابی اطلاعات گفتاری (صوتی) ترکیبی از فناوری‌های

1. Elavarasi & Suseendran
2. Zerari
3. Literature
4. Information retrieval
5. Maning
6. Kagolovsky & Moehr

بازشناسی گفتار و بازیابی اطلاعات است که اطلاعات متنی ابتدا به اطلاعات گفتاری (صوتی) تبدیل شده و سپس عملیات بازیابی بر روی محتوای صوتی انجام می‌شود. با توجه به مواردی که اشاره شد، این پژوهش درصدد است تا با روش نیمه تجربی به بررسی اثربخشی استفاده از سیستم تشخیص گفتار هوشمند در مهارت بازیابی صحیح اطلاعات موجود در داستان‌های حوزه ادبیات روایی برای کودکان دبستانی بپردازد تا افقی جهت همگامی ادبیات روایی برای آموزش و سرگرمی کودکان در عصر نوین و پیشرفته امروزی باشد.

علاوه بر مطالب بالا، بررسی پژوهشگران نشان داد که به نظر می‌رسد با وجود انجام تحقیقات متعدد در زمینه فناوری گفتار هوشمند در خصوص استفاده از این سیستم در آموزش و بازیابی اطلاعات ادبیات روایی کودکان، پژوهشی صورت نگرفته است. بنابراین، با توجه به زمینه‌های کاربردی این موضوع توجه به آن ضروری به نظر می‌رسد. بنابراین، این پژوهش درصدد است با روش نیمه تجربی به بررسی اثربخشی استفاده از سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات موجود در داستان‌های حوزه ادبیات روایی برای کودکان دبستانی بپردازد تا افقی جهت همگامی ادبیات روایی برای آموزش و سرگرمی کودکان در عصر نوین و پیشرفته امروزی باشد. با توجه به مواردی که ذکر شد و نیل به اهداف موردنظر این پژوهش درصدد پاسخ به پرسش‌های زیر است:

۱. تعداد خطاهای موجود در اطلاعات بازیابی شده توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند چقدر است؟
۲. طول اطلاعات بازیابی شده توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند چقدر است؟
۳. سرعت بازیابی اطلاعات توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند چقدر است؟
۴. صحت اطلاعات بازیابی شده توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند چقدر است؟
۵. میزان اطلاعات صحیح بازیابی شده توسط کودکان دبستانی با استفاده از کدام یک از سه روش (سیستم تشخیص گفتار هوشمند، نرم‌افزار قصه‌گویی و قصه‌گوی انسانی) بیشتر است؟

پیشینه پژوهش

بررسی پایگاه‌های اطلاعاتی معتبر داخلی و خارجی نشان داد پژوهشگران مختلف به بررسی جنبه‌های مختلف به کارگیری سیستم گفتار هوشمند پرداخته‌اند اما تاکنون پژوهشی که به بررسی تأثیر استفاده از سیستم گفتار هوشمند در بازیابی اطلاعات توجه کند، انجام نشده است.

لیو^۱ و همکارانش (۲۰۲۳) در پژوهش خود با توجه به پیشرفت‌های سریع در فناوری سنجش و استفاده از هوش مصنوعی برای کمک به افراد معلول در برقراری ارتباط به بررسی تشخیص گفتار هوشمند پرداختند. این پژوهش با روش تجربی انجام شد و در نتایج آن به تأثیر مثبت استفاده از هوش مصنوعی در این امر اشاره گردید. بای^۲ و همکاران (۲۰۲۳) با انجام یک پژوهش به بررسی یک فراسطح هوشمند برای برنامه‌ریزی دستکاری‌های الکترومغناطیسی بر اساس تشخیص گفتار انسان پرداخته و در نتایج خود پلتفرمی فراسطحی از یک فراسطح کدگذاری دیجیتال، یک ماژول تشخیص گفتار، یک رایانه تک‌تراشه و یک مدار مبدل دیجیتال به آنالوگ بر اساس دستورات صوتی از پیش ذخیره شده، پیشنهاد دادند.

ژیا^۳ و همکارانش (۲۰۲۲) در پژوهش خود سیستم‌های تشخیص گفتار خودکار LeVoice را برای استفاده در کابین خلبان هوشمند ۲۰۲ بررسی کرده است. نکات اصلی بررسی در این پژوهش شامل تقویت گفتار مبتنی بر یادگیری عمیق، تولید گفتار مبتنی بر متن به گفتار، تقویت داده‌های آموزشی از طریق فنون و روش‌های مختلف و ادغام مدل تشخیص گفتار است. این پژوهش به صورت داده‌محور و تجربی انجام گرفت. محمود و کوز^۴ (۲۰۲۱) در پژوهشی با عنوان «تشخیص گفتار بر اساس شبکه‌های عصبی کانولوشن و الگوریتم MFCC» دریافتند که استفاده از MFCC به عنوان استخراج ویژگی و اعمال این ویژگی‌ها برای مدل

1. Liu
2. Bai
3 Jia
4. Mahmood & Kose

CNN جهت استخراج ویژگی‌های سرعتی بیشتر باعث بهبود دقت و کاهش پیچیدگی مدل می‌شود. پژوهشگران بعد از اجرای این طرح در یک نمونه کوچک در نتایج خود به تأثیرات مثبت ویژگی‌های MFCC در این باره اشاره کردند.

الوارازی و سوسندران (۲۰۲۰) در پژوهشی با عنوان «پردازش خودکار ربات با استفاده از سیستم تشخیص گفتار» به این نتیجه رسیدند که تشخیص گفتار فرایند تبدیل سیگنال گفتار به دنباله‌ای از کلمات با استفاده از الگوریتم مناسب است که این یک روش جایگزین و کارآمد برای افرادی که تحصیلات کافی ندارند و یا دانش کافی در رایانه ندارند برای دسترسی به سیستم‌ها و جایی که تایپ کردن مشکل باشد، فراهم می‌کند.

گودارا^۱ (۲۰۱۹) در پژوهشی با عنوان «تشخیص گفتار با استفاده از یادگیری ماشین» بیان کرد که گفتار یک موج فشار است که از طریق هوای ایجاد شده توسط لرزش حنجره با باز و بسته شدن دهان در برنامه‌های تشخیص گفتار با استفاده از فنون و روش‌های یادگیری ماشین حرکت می‌کند و همچنین این که این فناوری در حال رشد است و تأثیر زیادی در آینده دارد.

اراتاهی^۲ و همکاران (۲۰۱۸) در پژوهشی با عنوان «تشخیص و تصحیح خطاهای تشخیص صدا به صورت خودکار» به این نتیجه رسیدند که در مورد اصلاح خودکار خطاهای سامانه بازشناسی خودکار گفتار، نیاز به تحقیقات بیشتری وجود دارد و باید به مواردی از جمله بهره‌وری، قابلیت استفاده و استحکام روش‌های توسعه‌یافته توجه شود.

جهانگشته و همکاران (۱۴۰۰) در پژوهشی با عنوان «تشخیص احساس با روش پردازش گفتار و انتخاب ویژگی» دریافتند که از آنجا که ساختار موجک بیونیک منطبق بر ساختار گوش انسان است و چون انسان درک خوبی از احساسات گفتار دارد، استفاده از موجک بیونیک برای تشخیص خودکار احساسات از گفتار مفید است.

در پژوهش حسنی (۱۳۹۸) بیان شد که نتایج حاصله از این آزمایش به طور مؤثرتری نرخ‌های خطای کلمه پایین‌تری را نسبت به شرایط تحت نویز تطبیق یافته نشان می‌دهد و در اصل زمانی که الگوریتم کاهش نویز به دیتای نویزی آموزش اعمال می‌شود و متعاقب آن تشخیص‌دهنده مجدداً آموزش می‌بیند، نرخ‌های خطای پایین‌تری حاصل می‌شود.

کمالی و شیخ‌طاهری (۱۳۹۷) در پژوهشی با عنوان «مستندسازی گزارش‌های پرستاری با استفاده از فناوری تشخیص گفتار: مزایا، موانع، چالش‌ها و تسهیلگرها» به این نتیجه رسیدند که مدیران بیمارستان‌ها، مدیران پرستاری و مسئولان فناوری اطلاعات بیمارستان‌ها با کمک نتایج این مطالعه می‌توانند رفتار آگاهانه‌تری در انتخاب و پیاده‌سازی سیستم تشخیص گفتار برای مستندسازی گزارش‌های پرستاری داشته باشند.

با توجه به نتایج پژوهش‌های انجام شده در حیطه‌های مرتبط با موضوع پژوهش حاضر، می‌توان گفت که تشخیص گفتار یکی از مهمترین زمینه‌های هوش ماشینی بوده و از آنجا که فعالیت‌های روزمره انسان‌ها پیچیده، زمان‌بر و پرهزینه شده است فنون و روش‌های تشخیص گفتار می‌تواند کمک شایانی در تسهیل امور انسان‌ها در همه زمینه‌ها داشته باشد. طراحی سیستم تشخیص گفتار نیازمند شناسایی انواع مختلف خوشه‌های گفتار، فنون و روش‌های استخراج ویژگی، طبقه‌بندی‌کننده گفتار، پایگاه داده و ارزیابی عملکرد است. در واقع، استفاده از سیستم تشخیص گفتار می‌تواند یک روش جایگزین و کارآمد برای افرادی که تحصیلات کافی ندارند و یا برای دسترسی به سیستم‌ها و جایی که تایپ کردن مشکل باشد و همچنین افزایش بازدهی اقتصادی در سازمان‌ها و ادارات، کاربردهای مربوط به معلولین، نابینایان و ناشنویان در جهت آموزش، یادگیری و تسریع امور آنها، افزایش امنیت در سیستم‌های نظامی، سهولت و تسریع ارتباطات، بهبود کیفیت ارائه خدمات و گردش کار و امکان خلق فناوری‌های جدید به‌واسطه سیستم تشخیص گفتار، امکان جستجو در کل فرمت‌ها از قبیل متن، ویدئو، موسیقی و ... تهیه گزارش وضعیت بیماران در بیمارستان‌ها، کاربردهای تلفنی (از قبیل سیستم‌های تلفن بانک، اطلاع‌رسانی تلفنی، سیستم‌های خودکار رزرو بلیط)، کاربردهای تشخیص فرامین و دستورات صوتی (از قبیل اجرای برنامه‌های رایانه با گفتن نام آنها، کنترل لوازم خانگی با دستورات صوتی، فرمان دادن به ربات‌ها و وسایل صنعتی، فرمان‌های صوتی در اتومبیل‌ها، دستگاه‌های خودپرداز بانک‌ها و موارد مشابه دیگر)، استفاده در رایانه‌های جیبی، تلفن همراه، جستجوگر واژه‌های کلیدی در گفتار، تصدیق و یا تشخیص هویت و غیره، جنبه مهم استفاده از

این سیستم نیز در حوزه آموزش بیشترین کارایی را می‌تواند داشته باشد. همچنین با توجه به نتایج حاصله از پژوهش‌های صورت گرفته می‌توان به موارد زیر اشاره کرد:

- در خصوص روش‌های مورد استفاده در سیستم تشخیص گفتار، روش استخراج ویژگی و تشخیص الگو بیشترین کارایی را داشته است.

- از بین الگوریتم‌ها، الگوریتم MFCC بازده بیشتری را در نتایج عملکرد تشخیص نشان داده است.

- در زمینه تشخیص در محیط نویزی با استفاده از الگوریتم کاهش نویز نرخ خطای پایین‌تری حاصل می‌شود در حالی که در محیط ساکت و بدون نویز نرخ تشخیص به صورت مؤثری بالا می‌رود.

- در زمینه دقت تشخیص احساس برای گوینده مرد و زن، ماشین بردار پشتیبان و شبکه عصبی بیشترین بهره‌وری را نشان داده است. همچنین باید گفت تفکیک جنسیتی و لهجه نیز نقش مهمی در افزایش میزان دقت نتایج تشخیص گفتار را دارد. شاید چون زنان بیشتر از مردان احساساتی هستند و این موضوع بیشتر در صدای آنان نمایان می‌شود تشخیص احساس از صدای آنان آسان‌تر است؛ اما برای مردان این موضوع برعکس است به دلیل این که می‌توانند خود را کنترل کنند کمتر می‌شود از روی صدای آنان احساس‌شان را تشخیص داد.

- در مورد اصلاح خطاهای تشخیص خودکار گفتار، نیاز به تحقیقات بیشتری وجود دارد و باید به مواردی از جمله بهره‌وری، قابلیت استفاده و استحکام روش‌های توسعه یافته توجه شود و از همه توانمندی‌های آن در جهت افزایش سرعت و بهبود کیفیت و ساده‌تر کردن نحوه عملکرد آن و خلق فناوری‌های جدید از طریق این سیستم استفاده شود.

با توجه به موارد ذکر شده و خالی بودن جای توجه به مسئله بازیابی صحیح اطلاعات می‌توان اذعان کرد انجام پژوهشی جهت بررسی اثربخشی سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات و نیل به اهداف تعیین شده در این پژوهش می‌تواند در این حیطه پنجره‌ای روبه آینده باشد.

روش‌شناسی پژوهش

پژوهش حاضر از نظر هدف کاربردی بوده و با روش، نیمه‌تجربی انجام شده است. جامعه آماری این پژوهش شامل کودکان دبستانی دختر و پسر پایه اول تا سوم مدارس دولتی و غیردولتی شهر کرمان در سال تحصیلی ۱۴۰۱-۱۴۰۲ بود. نمونه پژوهش حاضر را ۳۶ نفر از کودکان دبستانی دختر و پسر پایه اول تا سوم تشکیل دادند که با روش نمونه‌گیری طبقه‌ای در دسترس انتخاب شدند که ۱۷ نفر آنها پسر و ۱۹ نفر آنها دختر بودند و همچنین ۲۰ نفر از آنها از مدارس دولتی و ۱۶ نفر از مدارس غیردولتی بودند.

برای انجام این پژوهش ۱۵ قصه کودکان انتخاب شد. مبنای انتخاب قصه‌ها این بود که جدید بوده و کودکان آنها را نشنیده باشند تا بررسی میزان بازیابی صحیح اطلاعات بر اساس شنیده‌ها و محفوظات قبلی کودکان صورت نگیرد. این ۱۵ قصه به سه دسته ۵ تایی تقسیم شده و هر دسته به وسیله یک روش برای همه کودکان بازگو شدند. روش اول؛ با استفاده از سیستم تشخیص گفتار هوشمند که با تبدیل صدا به متن، متن یک قصه را دریافت و سپس پردازش می‌کند و در قالب صدا به عنوان خروجی ارائه می‌دهد. از این ابزار به منظور ضبط و پخش ۵ داستان استفاده شد. روش دوم؛ استفاده از یک برنامه کاربردی (اپلیکیشن) قصه‌گویی به نام «کتاب قصه صوتی» است که مجموعه‌ای از داستان‌های صوتی و متنی را شامل می‌شود که در این پژوهش تعداد ۵ داستان از داستان‌های صوتی آن مورد استفاده قرار گرفت. بدیهی است که هنگام استفاده از این نرم‌افزار قصه‌گویی، قصه‌ها همراه با موسیقی زمینه، صداهای مناسب و احساسات گوینده همراه بود. از این ابزار نیز به منظور ضبط و پخش ۵ داستان استفاده شد. ۵ قصه باقی‌مانده نیز قصه توسط یکی از پژوهشگران برای کودکان به صورت مستقیم تعریف شد.

برای گردآوری داده‌ها در این پژوهش، یک پرسش‌نامه طراحی شد که شامل ۵ سؤال کلیدی چهار گزینه‌ای از محتوای هر قصه بود که بعد از گوش دادن به هر قصه به وسیله کودکان (و با کمک پژوهشگران با توجه به سن نمونه پژوهش) پاسخ داده شد و همچنین یک سیاهه و آرسی مبتنی بر سؤالات پرسشنامه به منظور تکمیل و ارزیابی معیارهای تعیین شده، طراحی گردید که بر مبنای پاسخ‌های دریافتی از نمونه پژوهش توسط پژوهشگران تکمیل می‌شد. این سیاهه و آرسی شامل این موارد بود:

- بررسی سرعت بازیابی اطلاعات توسط کودکان با استفاده بررسی مدت زمان پاسخگویی به سؤالات پرسش‌نامه مربوط به هر قصه؛

- بررسی درستی و نادرستی اطلاعات بازیابی‌شده توسط کودکان با استفاده از یک جدول با آیت‌های کنترلی (تشخیص شخصیت اصلی داستان، تشخیص مکان داستان، تشخیص زمان داستان، تشخیص حادثه اصلی داستان، تشخیص درون‌مایه داستان) با گزینه‌های تأیید و عدم تأیید؛ و

- بررسی طول اطلاعات بازیابی‌شده توسط کودکان با استفاده از یک جدول با آیت‌های کنترلی جدول قبلی به همراه تعداد واقعی کلمات (با توجه به گزینه‌های صحیح پاسخ‌های پرسشنامه) با گزینه‌های کامل و ناقص.

شایان ذکر است که جهت سنجش روایی این ابزارها از چهار متخصص در حوزه ادبیات کودک و دو متخصص در حوزه بازیابی اطلاعات کمک گرفته‌شد و روایی این ابزارها مورد بررسی و تأیید قرار گرفت. جهت تعیین پایایی این ابزارها نیز از آلفای کرونباخ استفاده‌شد که ضریب آلفا به ترتیب ۰/۸۹ (پرسشنامه) و ۰/۸۲ (سیاهه واری) تعیین گردید.

جهت تجزیه و تحلیل داده‌های پژوهش از آمار استنباطی مشتمل بر آزمون تحلیل واریانس یک‌طرفه با به کارگیری نرم‌افزار SPSS21 استفاده شد.

یافته‌های پژوهش

در این پژوهش میزان اثربخشی از جنبه تعداد خطاها، طول، سرعت و صحت اطلاعات بازیابی‌شده توسط کودکان در سه گروه صدای انسان، سیستم تشخیص گفتار هوشمند و برنامه کاربردی (اپلیکیشن) قصه‌گویی مورد بررسی قرار گرفت.

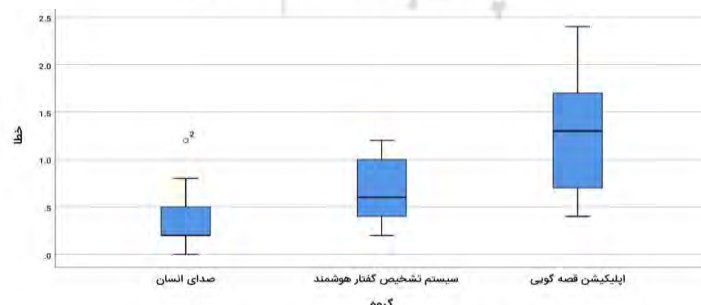
پرسش ۱. تعداد خطاهای موجود در اطلاعات بازیابی‌شده

بر اساس یافته‌های جدول ۱، P -مقدار کمتر از ۰/۰۰۱ است بنابراین فرض صفر در سطح ۰/۰۵ رد می‌شود و وضعیت تعداد خطاهای موجود در اطلاعات بازیابی‌شده توسط نمونه پژوهش بالاست و در ضمن متوسط تعداد خطاهای موجود در اطلاعات بازیابی‌شده قصه‌ها در سه روش مورد بررسی یکسان نیست.

جدول ۱. آزمون تحلیل واریانس یک‌طرفه برای بررسی متوسط تعداد خطاهای موجود در اطلاعات بازیابی‌شده قصه‌ها

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	P-مقدار
درون گروهی	۴/۶۹۶	۲	۲/۴۸۴	۱۱/۴۲۹	* کمتر از ۰/۰۰۱
بین گروهی	۷/۱۷۳	۳۳	۰/۲۱۷		
کل	۱۲/۱۴۲	۳۵			

* معنادار در سطح ۰/۰۵



نمودار ۱. نمودار جعبه‌ای متوسط تعداد خطاهای موجود در اطلاعات بازیابی‌شده قصه‌ها با استفاده از سه روش

در نمودار ۱ متوسط تعداد خطاهای موجود در اطلاعات بازیابی‌شده در سه روش استفاده شده قابل مشاهده است. با توجه به معناداری تحلیل واریانس و نمودار ۱ مشخص می‌شود که در روش استفاده از نرم‌افزار قصه‌گویی بیشترین میزان خطا و در روش استفاده از قصه‌گویی انسانی کمترین خطا مشاهده می‌شود.

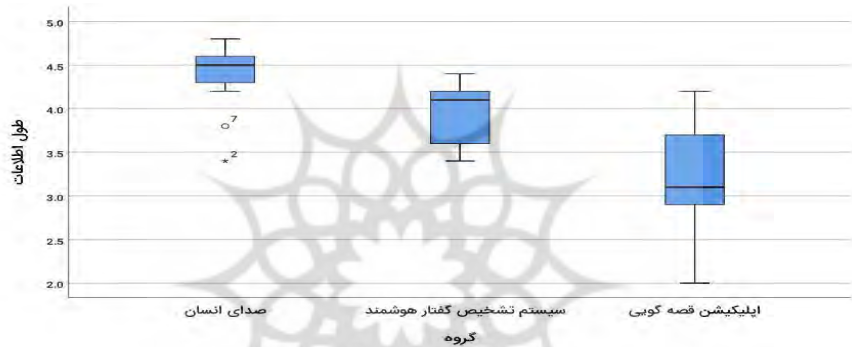
پرسش ۲. طول اطلاعات بازیابی شده

جدول ۲. آزمون تحلیل واریانس یک طرفه برای بررسی متوسط طول اطلاعات بازیابی شده قصه‌ها

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	P-مقدار
درون گروهی	۸/۱۸	۲	۴/۰۹۰	۱۴/۱۹۴	*کمتر از ۰/۰۰۱
بین گروهی	۷/۸۵	۳۳	۰/۲۳۸		
کل	۱۶/۰۳	۳۵			

*معنی‌دار در سطح ۰/۰۵

بر اساس یافته‌های جدول ۲، P-مقدار کمتر از ۰/۰۰۱ است بنابراین فرض صفر در سطح ۰/۰۵ رد می‌شود پس متوسط طول اطلاعات بازیابی شده از قصه‌ها در سه روش مورد استفاده یکسان نیست. در نمودار ۲ متوسط طول اطلاعات بازیابی شده قصه‌ها توسط کودکان در سه روش مورد استفاده قابل مشاهده است. با توجه به معناداری تحلیل واریانس و نمودار بالا مشخص می‌شود که روش استفاده از قصه‌گوی انسانی دارای بیشترین طول اطلاعات بازیابی شده و روش استفاده از نرم‌افزار قصه‌گویی دارای کمترین طول اطلاعات بازیابی شده است.



نمودار ۲. نمودار جعبه‌ای متوسط طول اطلاعات بازیابی شده قصه‌ها با استفاده از سه روش

پرسش ۳. سرعت بازیابی اطلاعات

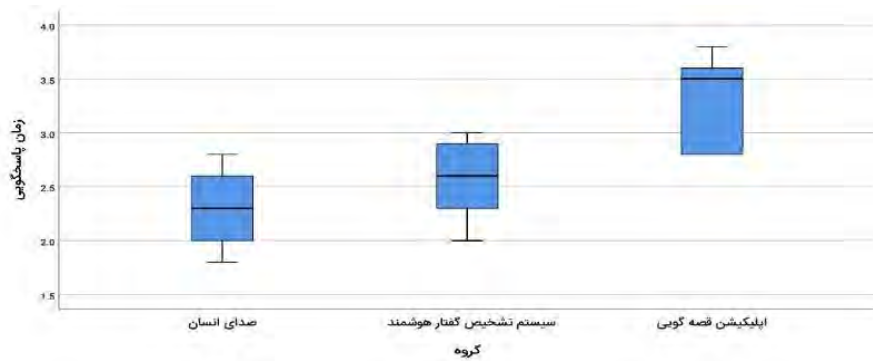
بر اساس یافته‌های جدول ۳، P-مقدار کمتر از ۰/۰۰۱ است. بنابراین، فرض صفر در سطح ۰/۰۵ رد می‌شود. پس، متوسط سرعت بازیابی اطلاعات بازیابی شده در سه روش مورد استفاده یکسان نیست.

جدول ۳. آزمون تحلیل واریانس یک طرفه برای بررسی متوسط سرعت بازیابی اطلاعات قصه‌ها

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	P-مقدار
درون گروهی	۵/۸۲	۲	۲/۹۱۰	۲۱/۷۲۶	*کمتر از ۰/۰۰۱
بین گروهی	۴/۴۲	۳۳	۰/۱۳۴		
کل	۱۰/۲۴	۳۵			

*معنادار در سطح ۰/۰۵

در نمودار ۳ متوسط سرعت اطلاعات بازیابی شده قصه‌ها توسط کودکان در سه روش مورد استفاده قابل مشاهده است. با توجه به معناداری تحلیل واریانس و نمودار بالا مشخص می‌شود که روش استفاده از قصه‌گوی انسانی دارای بیشترین سرعت بازیابی (پاسخگویی در زمان کمتر) و روش استفاده از نرم‌افزار قصه‌گویی دارای کمترین سرعت بازیابی (پاسخگویی در زمان بیشتر) است.



نمودار ۳. نمودار جعبه‌ای متوسط سرعت بازیابی اطلاعات قصه‌ها با استفاده از سه روش

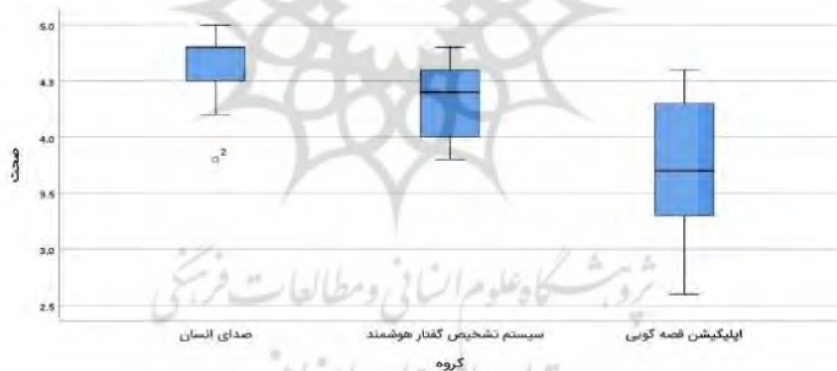
پرسش ۴. صحت اطلاعات بازیابی شده

بر اساس یافته‌های جدول ۴، p -مقدار کمتر از $0/001$ است. بنابراین، فرض صفر در سطح $0/05$ رد می‌شود پس متوسط صحت اطلاعات بازیابی شده در سه روش مورد استفاده یکسان نیست.

جدول ۴. آزمون تحلیل واریانس یک‌طرفه برای بررسی متوسط صحت بازیابی اطلاعات قصه‌ها

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	P-مقدار
درون گروهی	۴/۹۶۹	۲	۲/۴۸۴	۱۱/۴۲۹	* کمتر از $0/001$
بین گروهی	۷/۱۷۳	۳۳	$0/۲۱۷$		
کل	۱۲/۱۴۲	۳۵			

* معنادار در سطح $0/05$



نمودار ۴. نمودار جعبه‌ای متوسط صحت بازیابی اطلاعات قصه‌ها با استفاده از سه روش

در نمودار ۴ متوسط صحت اطلاعات بازیابی شده قصه‌ها توسط کودکان در سه روش مورد استفاده قابل مشاهده است. با توجه به معناداری تحلیل واریانس و نمودار بالا مشخص می‌شود که روش استفاده از قصه‌گوی انسانی دارای بیشترین میزان صحت اطلاعات بازیابی شده و روش استفاده از نرم‌افزار قصه‌گویی دارای کمترین میزان صحت اطلاعات بازیابی شده است.

جدول ۵. متوسط تعداد خطاهای موجود در اطلاعات بازیابی شده قصه‌ها در سه روش استفاده شده

تعداد	انحراف معیار	میانگین	گروه
۱۲	$0/۳۴۹$	$0/۳۶۶۷$	صدای انسان
۱۲	$0/۳۳۵$	$0/۷$	سیستم تشخیص گفتار هوشمند
۱۲	$0/۶۴۵$	$۱/۲۶۶۷$	برنامه کاربردی (اپلیکیشن) قصه‌گویی
۳۶	$0/۵۸۹$	$0/۷۷۷$	کل

پرسش ۵. مقایسه میزان اطلاعات صحیح بازیابی شده با سه روش استفاده شده

در جدول‌های زیر متوسط متغیرهای تعداد خطاها، طول، سرعت و صحت اطلاعات بازیابی شده از قصه‌ها توسط کودکان با استفاده از سه روش استفاده شده مشاهده می‌شود.

بر اساس یافته‌های جدول ۵، متوسط تعداد خطاهای موجود در اطلاعات بازیابی شده قصه‌ها توسط کودکان در روش استفاده از سیستم تشخیص گفتار هوشمند ۰/۷ است که از روش استفاده از قصه‌گوی انسانی بیشتر و از روش استفاده از نرم‌افزار قصه‌گویی کمتر است.

جدول ۶. متوسط طول اطلاعات بازیابی شده قصه‌ها در سه روش استفاده شده

تعداد	انحراف معیار	میانگین	گروه
۱۲	۰/۳۹۸	۴/۳۶۶	صدای انسان
۱۲	۰/۳۲۸	۲/۹۶۶	سیستم تشخیص گفتار هوشمند
۱۲	۰/۶۶۸	۳/۲۱۶	برنامه کاربردی (اپلیکیشن) قصه‌گویی
۳۶	۰/۶۷۶	۳/۸۵	کل

جدول ۷. متوسط سرعت بازیابی اطلاعات قصه‌ها در سه روش استفاده شده

تعداد	انحراف معیار	میانگین	گروه
۱۲	۰/۳۴۴	۲/۳۳۳	صدای انسان
۱۲	۰/۳۴۵	۲/۵۸۳	سیستم تشخیص گفتار هوشمند
۱۲	۰/۴۰۴	۳/۲۸۳	برنامه کاربردی (اپلیکیشن) قصه‌گویی
۳۶	۰/۵۴	۲/۷۳۳	کل

بر اساس یافته‌های جدول ۶، متوسط طول اطلاعات بازیابی شده قصه‌ها توسط کودکان در روش استفاده از سیستم تشخیص گفتار هوشمند ۳/۹۶۶ است که از روش استفاده از قصه‌گوی انسانی کمتر و از روش استفاده از نرم‌افزار قصه‌گویی بیشتر است. جدول ۷ نشان می‌دهد که متوسط طول سرعت بازیابی اطلاعات قصه‌ها توسط کودکان در روش استفاده از سیستم تشخیص گفتار هوشمند ۲/۵۸۳ است که از روش استفاده از قصه‌گوی انسانی بیشتر و از روش استفاده از نرم‌افزار قصه‌گویی کمتر است. یافته‌های جدول ۸ نمایان می‌سازد که متوسط طول صحت بازیابی اطلاعات قصه‌ها توسط کودکان در روش استفاده از سیستم تشخیص گفتار هوشمند ۴/۳ است که از روش استفاده از قصه‌گوی انسانی کمتر و از روش استفاده از نرم‌افزار قصه‌گویی بیشتر است.

جدول ۸. متوسط صحت بازیابی اطلاعات قصه‌ها در سه روش استفاده شده

تعداد	انحراف معیار	میانگین	گروه
۱۲	۰/۳۴۹	۴/۶۳۳	صدای انسان
۱۲	۰/۳۳۵	۴/۳	سیستم تشخیص گفتار هوشمند
۱۲	۰/۶۴۵	۳/۷۳۳	برنامه کاربردی (اپلیکیشن) قصه‌گویی
۳۶	۰/۵۸۹	۴/۲۲۲	کل

بحث و نتیجه‌گیری

سلامت و بقای فرهنگی و تربیتی هر جامعه‌ای، مدیون مدیریت فرهنگی باکفایت و برنامه‌ریزی اصولی، برای گروه سنی کودک و نوجوان است؛ زیرا این گروه سنی به عنوان شاخصه‌هایی جهت ارزیابی ارتقاء و توسعه فرهنگی جامعه بشمار می‌روند و اینها ثمره نوع تربیت و تغذیه فرهنگی در سنین کودکی و نوجوانی آنهاست. مسئله گفتار هوشمند و یا تبدیل گفتار به متن از مسائل مهم پژوهشی و کاربردی است. در واقع، هدف نهایی گفتار هوشمند، ساخت ماشین‌هایی است که بتوانند مانند انسان بشنوند و عکس‌العمل

نشان دهند (بحرانی و صامتی، ۱۳۸۹). بازیابی اطلاعات به سرعت در حال تبدیل شدن به شکل غالب دسترسی به اطلاعات در همهٔ حیطه‌ها از جمله ادبیات روایی و سبقت از جستجوی سبک سنتی است (مینینگ و همکاران، ۲۰۰۹).

نتایج این پژوهش نشان داد که تعداد خطاهای موجود در اطلاعات بازیابی شده قصه توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند نسبتاً زیاد است. این اطلاعات غلط بازیابی شده ممکن است ناشی از عدم وضوح گفتار گوینده در هنگام ضبط صدا باشد که در برقراری ارتباط اختلال ایجاد کرده و بر شنونده تأثیر نامطلوب گذاشته است که در این خصوص باید عواملی مانند روان بودن کلام، کیفیت و شدت صدا که بر وضوح گفتار گوینده و درک شنونده مؤثر است، مورد توجه قرار گیرد. از جهت دیگر، نویزها و مداخله‌گرهای محیطی که عبارتند از سروصدای پس‌زمینه، تداخل گفتار و بازتاب صدا، افت سیگنال، اغتشاش در سیستم انتقال یا میکروفون؛ از دیگر عواملی هستند که می‌توانند در ایجاد این اطلاعات غلط بازیابی شده تأثیر داشته باشند (پورا بر اهییم، رزازی و صامتی، ۱۴۰۰). از سوی دیگر، ممکن است محیط ماشینی و نامأنوس بودن استفاده از سیستم گفتار هوشمند در بعد شنیداری نیز در افزایش میزان خطا مؤثر بوده باشد. این بخش از نتایج با نتیجه پژوهش اراتاهی و همکاران (۲۰۱۸) مرتبط است که بیانگر این است که اصلاح خطاهای تشخیص گفتار نیاز به تحقیقات بیشتری دارد.

با توجه به نتایج طول اطلاعات بازیابی شده قصه توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند نسبتاً کم است. این اطلاعات ناقص بازیابی شده می‌تواند به دلیل لحن کلام گوینده، تن صدای ضعیف یا قوی گوینده باشد که سبب عملکرد ضعیف سیستم شده و فرایند انتقال اطلاعات را نیز ضعیف کرده و از جهت دیگر، ممکن است به دلیل عدم تعامل شنونده در برقراری ارتباط با سیستم باشد و همچنین عوامل زبانی و ارتباطی مؤثر بر درک شنونده از گفتار دیگران، استرس و اضطراب شنونده در برقراری مکالمه، ضعیف بودن شنونده در درک مطلب (که عوامل ژنتیکی و محیطی در آن نقش دارد)، عدم توانایی در مهارت‌های استدلال و تفکر و مشکل در پیدا کردن کلمات توسط شنونده؛ از جمله مواردی هستند که احتمالاً بر بازیابی ناقص اطلاعات تأثیر گذاشته‌اند. در این راستا، مرور پژوهش‌های گذشته نشان داد که تاکنون در این رابطه پژوهشی صورت نگرفته است و لازم است که متخصصان گفتار، زبان‌شناسان، متخصصان رایانه با همکاری یکدیگر فنون و روش‌های جدیدی را اعمال کنند تا دقت و کیفیت تشخیص گفتار در ارائه اطلاعات حاصل از بازشناسی بالا رود. از جهتی نیز می‌توان گفت که اضطراب ناشی از ابزار جدید می‌تواند در عدم تمایل کودکان برای پاسخ‌های طولانی‌تر مؤثر باشد. این قسمت از نتایج با یافته‌های حاصل از پژوهش الاوارازی و سوسندران، ۲۰۲۰ هم‌سو می‌باشد که طول بازیابی اطلاعات را با مؤلفه‌های جانبی مرتبط می‌دانند.

نتایج پژوهش حاضر نشان داد که سرعت کم بازیابی اطلاعات قصه توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند ممکن است به دلیل ناکافی بودن دقت شنونده، عدم تمرکز کافی در هنگام دریافت اطلاعات و تجزیه و تحلیل مطالب و شرایط محیطی که در یادگیری و یادآوری مؤثر است و همچنین عملکرد ضعیف سیستم در هنگام پخش صدا که همه این موارد منجر به کاهش سرعت، اختلال و وقفه در مدت زمان بازیابی شده، باشد. این بخش از نتایج پژوهش با یافته‌های پژوهش مطالبی و باستان فرد (۱۳۹۶) مرتبط است که بیانگر فنون و روش‌هایی جهت افزایش سرعت و بهبود کیفیت تشخیص به کار گرفته شود.

با توجه به نتایج صحت اطلاعات بازیابی شده قصه توسط کودکان دبستانی از سیستم تشخیص گفتار هوشمند نسبتاً کم است. دلیل این امر می‌تواند ناشی از تلفظ ناصحیح کلمات، اختلال در تکلم روان و شرایط غیرمعمول مانند استرس در گوینده باشد که سبب کیفیت نامطلوب صدای ضبط شده و همچنین کاهش دقت سیستم تشخیص گفتار در بازشناسی صدا نیز عامل دیگری که منجر به کاهش صحت اطلاعات بازیابی شده، گردیده است. این بخش از نتایج با نتیجه پژوهش صرافان و همکاران (۱۳۹۰) مرتبط و بیانگر آن است که استفاده از طبقه‌بندی‌کننده، درصد صحت مناسبی را دارا بوده و استفاده از سیستم‌های شناخت با منطق فازی و شبکه‌های عصبی دینامیک برای بهبود صحت تشخیص را توصیه کردند.

در واقع، به طور کلی نتایج این پژوهش مبین این است که کودکانی با داستان‌های شنیده شده از انسان تمرکز و دقت بالاتری نسبت به کودکانی که داستان‌ها را از طریق سیستم تشخیص گفتار هوشمند و برنامه کاربردی (اپلیکیشن) قصه‌گویی شنیده بودند، داشتند. دلیل این امر می‌تواند به علت فراهم بودن شرایط محیطی مناسب، برقراری ارتباط کلامی و غیرکلامی گوینده با شنونده، رسا بودن صدا و... باشد که ظاهراً در دو روش دیگر این دست موارد به صورت مطلوب وجود نداشته است. بنابراین، نتایج به دست

آمده نشان داد که استفاده از سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات ادبیات روایی کودکان دبستانی به طور کامل اثربخش نبوده است. این یافته با نتیجه پژوهش‌های (مهرداد و کلینی، ۱۳۸۶) و دیانت و همکاران (۱۳۹۵) در بُعد جستجوی اطلاعات مرتبط است. همچنین نتیجه پژوهش حاضر با یافته‌های پژوهش‌های حسنی، ۱۳۹۸؛ کمالی و شیخ‌طاهری، ۱۳۹۷؛ گودارا، ۲۰۱۹؛ بوکیا، ۲۰۱۸؛ اراتاهی و همکاران، ۲۰۱۸؛ پرودئوس و کوخاریچوا، ۲۰۱۷؛ زراری و همکاران، ۲۰۱۶؛ کاتیل و همکاران، ۲۰۱۴؛ ویمالا و رادا، ۲۰۱۲ همخوانی دارد و بیانگر آن است که سیستم‌های تشخیص گفتار کلمات گفتاری و صداها موجود را به سیگنال‌هایی تبدیل می‌کنند که قابلیت پردازش دارند و این یک روش جایگزین برای افرادی است که تحصیلات کافی ندارند و یا جایی که دسترسی به سیستم‌ها مشکل بوده و یا فرصت کافی برای انجام فعالیت‌های روزمره وجود ندارد؛ اما در زمینه آموزش و یادگیری مانند آموزش زبان، چگونگی تلفظ کلمات و... می‌توان از این سیستم‌ها به صورت کارآمد استفاده کرد.

با توجه به مطالب بالا می‌توان نتیجه گرفت که استفاده از سیستم تشخیص گفتار هوشمند در بازیابی صحیح اطلاعات ادبیات کودکان دبستانی به صورت شایسته و مورد انتظار اثربخش نیست و هنوز ارتباط چهره‌به‌چهره و شرایط نزدیکی محیطی به علت تأثیر مستقیم با کودک سبب افزایش عملکرد حافظه و یادگیری، افزایش میزان توجه و تمرکز، پردازش سریع اطلاعات، سازمان‌دهی ادراکی و یادآوری اطلاعات شده که این عوامل درک و فهم کامل آنها را به دنبال داشته است. البته شایان ذکر است که کودکان می‌توانند با فراگیری همزمان اطلاعات دیداری و شنیداری عملکرد بهتری داشته باشند. ناگفته نماند که استفاده از این سیستم‌ها سرعت روند انجام کارها را تا حد باورنکردنی بالا ببرند. البته در پایان باید گفت از آنجا که این سیستم‌ها بسیار کاربردی هستند و اهداف زیادی را دنبال می‌کنند؛ تاکنون از تمامی توانمندی‌هایشان استفاده نشده و همواره در حال رشد بوده و تأثیر زیادی در آینده دارند پس انتظار می‌رود که با کار بر روی آنها و ارائه فنون و روش‌های جدید در جهت بهبود کیفیت یا ساده‌تر کردن نحوه عملکرد و ارزیابی عملکرد، اصلاح خودکار خطاهای تشخیص صدا، نوع جنسیت و لهجه برای افزایش دقت تشخیص گفتار و افزایش دقت تشخیص در محیط‌های دارای نویز و پُر سروصدا و همچنین ادغام چندحالتی سمعی و بصری به دستاوردهای خوبی رسید که جبران نقاط ضعف و از میان برداشتن موانع اثربخشی آنها در همه ابعاد آموزشی و فرهنگی باشد که این امر رؤیای شهرهای الکترونیکی در دنیای الکترونیکی را تسریع و تسهیل می‌نماید.

پیشنادهای پژوهش

از آنجا که تعداد خطاهای موجود در اطلاعات بازیابی شده قصه توسط کودکان دبستانی با استفاده از سیستم تشخیص گفتار هوشمند زیاد است، پیشنهاد می‌شود که قصه‌گو با فراهم کردن شرایط محیطی مناسب واضح و شمرده و بدون تغییر سبک صحبت کند تا از این طریق کیفیت ضبط صدا بالا رود تا در هنگام پخش برای کودک (شنونده) تا حد امکان بتوان میزان خطاها را کاهش داد. با توجه به این که طول اطلاعات بازیابی شده قصه توسط کودکان با استفاده از سیستم تشخیص گفتار هوشمند کامل نیست، پیشنهاد می‌شود که قصه‌گو کلمات و جملات را به صورت واضح و شمرده بیان کند تا سیستم تشخیص گفتار، صدا را با دقت بالاتری تشخیص داده و سبب سهولت انتقال مفاهیم و معانی شود که این امر منجر به دقت شنونده و درک و فهم کامل محتوا در زمان پخش می‌شود.

از آنجا که سرعت بازیابی اطلاعات قصه توسط کودکان با استفاده از سیستم تشخیص گفتار هوشمند کم است، پیشنهاد می‌شود که قصه‌گو در هنگام ضبط صدا علاوه بر این بیان جملات را به صورت واضح، سعی کند لحن صحبت خود را طنین‌انداز نموده و همچنین حالات هیجانی متناسب با محتوای قصه را به صدای خود اضافه کند که از این طریق توجه و تمرکز و حس کنجکاوی کودک (شنونده) را برانگیزد که همین عامل سبب درک بهتر و به خاطر سپاری و یادآوری سریع‌تر کودک می‌شود.

با توجه به اینکه صحت اطلاعات بازیابی شده توسط کودکان با استفاده از سیستم تشخیص گفتار هوشمند کم است، پیشنهاد می‌شود که هنگام ضبط صدا قصه‌گو در محیط بدون سر و صدا با صدای واضح و رسا صحبت کند تا سیستم تشخیص، صدا را با

دقت بالایی تشخیص داده و از جهت دیگر پخش محتوای صوتی نیز در محیطی بدون سر و صدا صورت گیرد تا تمام حواس و تمرکز کودک متوجه صدا و محتوای آن گردیده و دقت کودک را در فهم محتوا به دنبال داشته باشد.

تعارض منافع

بنابر اظهار نویسندگان این مقاله تعارض منافع ندارد.

سپاسگزاری

نگارندگان بر خود لازم می‌دانند از داوران محترم به خاطر ارائه نظرهای ساختاری و علمی ارزشمند سپاسگزاری نمایند.

منابع

- ایرانمنش، زهرا (۱۳۹۸). تأثیر ادبیات داستانی بر رشد و تفکر کودک. *اومزرد*، ۴۷، ۲۸-۴۱.
- بحرانی، محمد، و صامتی، حسین (۱۳۸۹). به کارگیری اطلاعات زبانی در یک سیستم بازشناسی گفتار پیوسته فارسی. *زبان و زبان‌شناسی*، ۱۱، ۸۷-۱۱۲.
- پورابراهیم، یوسف؛ رزازی، فرید، و صامتی، حسین (۱۴۰۰). بازشناسی احساسات از روی گفتار با استفاده از ترکیب شبکه‌های عصبی ترنسفورمر و کانولوشن. *روش‌های هوشمند در صنعت برق*، ۱۳(۵۲)، ۷۹-۹۸.
- پوردریایی، علی، و سعیدی، سهراب (۱۳۹۲). ادبیات چیست. هفتمین همایش پژوهش‌های زبان و ادبیات فارسی، ۲۸۳-۲۹۱.
- جهانگشته، اسماعیل؛ دوستی، عبدالرحمن، و دهقانی، مجتبی (۱۴۰۰). تشخیص احساس با روش پردازش گفتار و انتخاب ویژگی. نهمین کنفرانس ملی پژوهش‌های کاربردی در علوم برق، کامپیوتر و مهندسی پزشکی.
- حاجی نصرالله، شکوه (۱۳۹۶). *شناخت ادبیات کودک*. تهران: وزارت فرهنگ و ارشاد اسلامی.
- حسنی، ملیحه (۱۳۹۸). یک روش جدید برای کاهش نویز گفتار و تشخیص دقیق گفتار با استفاده از مدل‌های احتمالی برای گفتار تمیز و نویز. *پایان‌نامه کارشناسی ارشد، آموزشکده فنی و حرفه‌ای سما، دانشگاه آزاد اسلامی، بابل، ایران*.
- دیانت، روح‌الله؛ علی احمدی، مرتضی؛ اخلاقی، محمد یحیی، و باباعلی، یاقر (۱۳۹۵). ارائه یک روش جدید بازیابی اطلاعات مناسب برای متون حاصل از بازشناسی گفتار. *پردازش علائم و داده‌ها*، ۳۰، ۹۳-۱۰۸.
- شهبازی، مهری (۱۳۹۱). *مواد و خدمات کتابخانه برای کودکان و نوجوانان*. تهران: دانشگاه پیام‌نور.
- کمالی، مهدیه، و شیخ طاهری، عباس (۱۳۹۷). مستندسازی گزارش‌های پرستاری با استفاده از فناوری تشخیص گفتار (مزایا، موانع و چالش‌ها و تسهیلگرها). *مجله انفورماتیک سلامت و زیست پزشکی*، ۵ (۱)، ۷۰-۸۲.
- مطالبی، مژده، و باستان‌فرد، اعظم (۱۳۹۶). مروری بر سیستم‌های تشخیص صوت، مفاهیم و روش‌ها. *سومین کنفرانس بین‌المللی فناوری اطلاعات، مهندسی کامپیوتر و مخابرات*.
- مهرداد، جعفر، و کلینی، سارا (۱۳۸۶). بررسی مدل فضای برداری در بازیابی اطلاعات. *کتابداری و اطلاع‌رسانی*، ۱۰(۲)، ۱۹۸-۲۱۰.

References

- Bahrani, M., & Sameti, H. (2010). Using linguistic information in a Persian continuous speech recognition system. *Language and Linguistics*, 11, 87-112. (In Persian)
- Bai, L., Liu, Y. K., Xu, L., Zhang, Z., Wang, Q., Jiang, W. X. & Cui, T. J. (2023). A Smart Metasurface for Electromagnetic Manipulation Based on Speech Recognition. *Engineering*, 22, 185-190. <https://doi.org/10.1016/j.eng.2022.06.026>
- Bhukya, S. (2018). Effect of gender on improving speech recognition system. *International Journal of Computer Applications*, 179 (14), 22-30.
- Diyanat, R., Ali Ahmadi, M., Akhlaghi, M. Y., & Baba Ali, B. (2016). Presenting a new method of retrieving suitable information for texts obtained from speech recognition. *Signal and Data Processing Quarterly*, 30, 93-108. (In Persian)

- Elavarasi, S., & Suseendran, G. (2020). Automatic robot processing using speech recognition system. *Department of Information and Technology, School of Computing Sciences, Vels Institute of Science Technology & Advanced Studies*, 185-195. https://doi.org/10.1007/978-981-32-9949-8_14
- Erratahi, R., El Hannani, A., & Ouahmane, H. (2018). Automatic speech recognition errors detection and correction: A review. *Procedia Computer Science*, 128, 32-37. <https://doi.org/10.1016/j.procs.2018.03.005>
- Godara, S. (2019). Speech recognition using machine learning: a review. *International Journal of Electronics Engineering*, 11 (1), 971-976.
- Haji Nasrallah, S. (2016). *Understanding children's literature*. Tehran: Ministry of Culture and Islamic Guidance. (In Persian)
- Hasani, M. (2018). *A new method for speech noise reduction and accurate speech recognition using probabilistic models for clean speech and noise*. Sama Technical and Vocational School, Master thesis, Islamic Azad University, Babol, Iran. (In Persian)
- Iranmanesh, Z. (2018). The effect of fiction on children's development and thinking. *Urmazd Quarterly*, 47, 28-41. (In Persian)
- Iranmanesh, Z. (2019). The effect of fiction on children's development and thinking. *Ourmazd*, 47, 31-47. (In Persian)
- Jahangeshte, I., Dusti, A. R., & Dehghani, M. (2021). Emotion recognition by speech processing and feature selection. *The 9th National Conference of Applied Researches in Electrical Sciences, Computers and Medical Engineering*. (In Persian)
- Jia, Y., Hong, M., Hou, J., Ren, K., Ma, S., Wang, J., & Wang, J. (2022, December). LeVoice ASR systems for the ISCSLP 2022 intelligent cockpit speech recognition challenge. In *2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP)* (pp. 517-521). IEEE. <https://doi.org/10.1109/ISCSLP57327.2022.10038155>
- Kagolovsky, Y., & Moehr, J. R. (2003). Terminological problems in information retrieval. *Journal of Medical Systems*, 27(5), 399-408. <https://doi.org/10.1023/A:1025687220609>
- Kamali, M., & Sheikh Taheri, A. (2018). Documentation of nursing reports using speech recognition technology (advantages, obstacles and challenges and facilitators). *Journal of Health and Biomedical Informatics*, 5(1), 72-80. (In Persian)
- Katyal, A., Kaur, A., & Gill, J. (2014). Automatic speech recognition: a review. *International Journal of Engineering and Advanced Technology*, 3(3), 71-74. https://doi.org/10.1007/1-4020-2673-0_3
- Liu, Y., Li, H., Liang, X., Deng, H., Zhang, X., Heidari, H., & Zhang, X. (2023). Speech Recognition Using Intelligent Piezoresistive Sensor Based on Polystyrene Sphere Microstructures. *Advanced Intelligent Systems*, 2200427. <https://doi.org/10.1002/aisy.202200427>
- Mahmood, A., & Kose, U. (2021). Speech recognition based on convolutional neural networks and MFCC algorithm. *Advances in Artificial Intelligence Research*, 1(1), 6-12. <https://doi.org/10.3390/bdcc7030132>
- Mahrad, J., & Coleinney, S. (2007). Investigating the vector space model in information retrieval. *Library and Information Quarterly*, 10(2), 198-210.
- Maning, C., Raghavan, P., & Schutze, H. (2009). *An introduction to information retrieval*. Cambridge University Press, p. 569.
- Matalebi, M., & Bastan Fard, A. (2017). Overview of voice recognition systems, concepts and methods. *The third international conference on information technology, computer engineering and telecommunications*. (In Persian)
- Prodeus, A., & Kukharicheva, K. (2017). Automatic speech recognition performance for training on noised speech. *International Conference Advanced Information and Communication Technologies*, pp. 1-4. <https://doi.org/10.1109/AIACT.2017.8020068>
- Purdaryai, A., & Saidi, S. (2012). What is literature? *The 7th Persian Language and Literature Research Conference*, pp.283-291. (In Persian)

- Purebrahim, Y., Razazi, F., & Sameti, H. (2021). Recognition of emotions from speech using a combination of transformer and convolutional neural networks. *Smart methods in the electricity industry*, 13(52), 79-98. (In Persian)
- Shahbazi, M. (2012). *Library materials and services for children and teenagers*. Tehran: Payam Noor University. (In Persian)
- Vimala, C., & Radha, D. V. (2012). A review on speech recognition challenges and approaches. *World of Computer Science and Information Technology Journal*, 2 (1), 1-7. <http://dx.doi.org/10.5120/20284-2839>
- Zerari, N., Yousfi, B., & Abdelhamid, S. (2016). Automatic speech recognition: a review. *International Academic Research Journal of Business and Technology*, 2(2), 63-68. https://doi.org/10.1007/1-4020-2673-0_3

