



## طراحی و ساخت یک مکانیزم گفت‌وگو مناسب ربات خیمه‌شب‌باز با قابلیت یادگیری عمیق

مسعود امیرخانی، حامد شهبازی\*

گروه مهندسی مکانیک، دانشگاه اصفهان، اصفهان، ایران.

### تاریخچه داوری:

دریافت: ۱۴۰۱/۰۹/۱۸  
بازنگری: ۱۴۰۲/۰۵/۲۸  
پذیرش: ۱۴۰۲/۰۷/۱۱  
ارائه آنلاین: ۱۴۰۲/۰۹/۰۶

### کلمات کلیدی:

ربات خیمه شب باز  
پردازش زبان طبیعی  
هوش مصنوعی  
یادگیری عمیق، شبکه عصبی  
پاسخ هوشمند

**خلاصه:** هدف از این پژوهش، طراحی و ساخت یک مکانیزم گفت‌وگو بر مبنای متن‌کاوی و پردازش زبان طبیعی بر روی یک ربات خیمه‌شب‌باز با قابلیت یادگیری عمیق گفت‌وگو با رویکرد استدلال عمیق است. عروسک‌های خیمه‌شب‌باز یکی از قدیمی‌ترین عروسک‌ها در حوزه تعامل با مخاطبین است که به صورت مستقیم توسط یک گرداننده کنترل می‌شود. با توجه به پیشرفت‌های بیان‌شده در هوش مصنوعی و یادگیری عمیق، می‌توان شرایطی را فراهم کرد که وابستگی عروسک‌های خیمه‌شب‌باز به گرداننده کاهش یابد و با هوشمندی با مخاطبین ارتباط برقرار کند. این ربات، با تشخیص گفتار فارسی مخاطب، یک پاسخ مناسب برای پرسش تعیین می‌کند و به صورت صوتی و با زبان فارسی پخش می‌کند. اهمیت این موضوع در ایجاد مکانیزم گفت‌وگو مناسب است. این مکانیزم، یک الگوریتم یادگیری عمیق است که با تشخیص سوال مطرح‌شده توسط کاربر، یک مجموعه احتمالات از دسته‌بندی‌های قرارگرفته در مجموعه داده ربات ارائه می‌دهد و با در نظر گرفتن بالاترین احتمال، دسته موردنظر که پرسش کاربر در آن قرار می‌گیرد مشخص می‌شود و از بین پاسخ‌هایی که برای آن دسته سوال در نظر گرفته شده‌است، یک پاسخ به صورت تصادفی انتخاب می‌گردد. علاوه‌براین، مکانیزم گفت‌وگو ربات خیمه‌شب‌باز دارای چند بخش ساده شرطی است که می‌تواند به سوالات تکراری یا نامفهوم پاسخ مناسبی ارائه دهد. در نتایج حاصل از آموزش‌های مختلف با تغییر پارامترها در مدل یادگیری عمیق این ربات با مجموعه داده ۶۴ کلاسی مشخص شد که استفاده از لایه‌های مزدحم با نرون‌های زیاد بهتر از چند لایه از آنها عمل می‌کند و میزان دقت آنها تاثیر زیادی در دقت نهایی مدل ندارد.

### ۱- مقدمه

یک بازگفت مشهور است که می‌گوید: «آیا ماشین‌ها می‌توانند فکر کنند؟». این سوال موجب شده‌است تا دانشمندان علوم کامپیوتر، سیستم‌هایی بسازند که نه تنها قدرت تفکر دارند، بلکه می‌توانند به صورت طبیعی نیز با انسان‌ها ارتباط برقرار کنند. امروزه ماشین‌های مکالمه‌محور به صورت تخصصی پیرامون یک موضوع خاص ساخته می‌شوند. از چالش‌های مکالمه موفق یک ماشین، می‌توان به منطقی و اجتماعی بودن آن اشاره کرد که یکی از پارامترهای مهم در آن دریافت معنی و مفهوم جملات است. بیان کلمه‌ها در یک جمله، بسته به موقعیت آن می‌تواند معنای متفاوتی را ایجاد کند. از طرفی، برخی از کلمات برای هر فرد معنای به‌خصوصی دارد که از زندگی اجتماعی او سرچشمه می‌گیرد [۱]. علاقه و نیاز روزافزون به تولید رابط‌های کاربری بین سیستم و کاربر منجر به توسعه ماشین‌های مکالمه‌محور می‌شود.

این ماشین‌ها برای تبدیل اطلاعات استفاده می‌شود. بنابراین، کاربر با زبان طبیعی و حرکات، مکالمه را در هر زمان و مکان آغاز می‌کند و انتظار می‌رود ماشین بتواند مانند یک دستیار صوتی هوشمند نیازهای کاربر خود را عملی کند [۲].

در گذشته، وجود یک دستیار صوتی هوشمند که توانایی درک و توضیح مطالب مختلف را داشته باشد، تنها در فیلم‌های علمی‌تخیلی دیده می‌شد اما امروزه، ساخت سیستم مکالمه خودکار بین انسان و کامپیوتر به عنوان دستیار شخصی دیگر یک توهم نیست. به دلیل وجود داده‌های بزرگ در وب، می‌توان سیستم‌های مکالمه‌محور هوشمند برای ارتباط با انسان طراحی نمود [۳]. تنها کافی است که بر روی دیتاهای موجود پردازش‌های لازم متناسب با هدف مکالمه ربات صورت گیرد. پرسش مطرح‌شده توسط کاربر، با مراحل مثل تجزیه جمله، نشانه‌گذاری، ریشه‌یابی و استخراج کلمات کلیدی که توسط

\* نویسنده عهده‌دار مکاتبات: shahbazi@eng.ui.ac.ir



از بهینه‌ساز ریشه میانگین مربعات انتشار<sup>۱۱</sup> و تابع هزینه آتروپی متقاطع طبقه ای پراکنده<sup>۱۲</sup>، هیچ بهبود عمده‌ای در نتایج نداشته است. در این مقاله مشاهده شد که نتایج پیاده‌سازی شبکه‌های حافظه<sup>۱۳</sup> نسبت به شبکه حافظه طولانی کوتاه‌مدت<sup>۱۴</sup> بهبود یافته‌اند و شبکه‌های حافظه پویا<sup>۱۵</sup> پیشرفت بیشتری نسبت به شبکه‌های حافظه دارند که ثابت شد یک فرضیه درست است. در نتایج این پژوهش، از ۲۰ روش با در نظر گرفتن پارامترهای مختلف در سه شبکه حافظه طولانی کوتاه‌مدت، شبکه حافظه<sup>۱۶</sup> و شبکه حافظه دینامیکی<sup>۱۷</sup> استفاده شده و میانگین دقت در این سه شبکه به ترتیب ۰.۴۹٪، ۰.۷۵٪ و ۰.۹۱٪ رسیده است. مینگ تان<sup>۱۸</sup> و همکاران [۶]، یک مدل یادگیری عمیق برای پاسخ به سوالات پیچیده بر اساس پاسخ‌های متوالی و انتخاب آن از متن پیشنهاد شده است. این مدل، ترکیبی از شبکه‌های عصبی کانولوشن<sup>۱۹</sup> و بازگشتی را برای تطبیق پرسش و پاسخ با حفظ روابط معنایی پیشنهاد می‌کند. نتایج به دست آمده، با مجموعه داده‌های کنفرانس بازیابی متن<sup>۲۰</sup> و پرسش و پاسخ داده‌های بیمه<sup>۲۱</sup> آموزش و ارزیابی شده‌اند. همه مدل‌های پیشنهادی در این مقاله، از معماری شبکه‌های حافظه طولانی کوتاه‌مدت دوطرفه ساخته شده‌اند. داده‌های تجربی برتری عملکرد مدل‌های پیشنهادی بر مدل‌های مبنایی را تأیید می‌کنند. در نتایج این پژوهش، ۱۴ روش با در نظر گرفتن پارامترهای مختلف و در دو مرحله تست انجام شده است. بیشترین دقت به دست آمده مربوط به مدل حافظه طولانی کوتاه‌مدت توجه<sup>۲۲</sup> با حداکثر تجمع<sup>۲۳</sup> ۵۰ است که در مرحله ارزیابی هنگام یادگیری به ۰.۶۸/۹٪ و در دو مرحله تست به ترتیب ۰.۶۹/۰٪ و ۰.۶۴/۸٪ رسیده است. لی<sup>۲۴</sup> و همکاران [۲۱]، در مورد روش‌های طبقه‌بندی سوال در زبان انگلیسی بحث شده است. در این روش‌ها، نوع سوال طبقه‌بندی می‌شود و می‌تواند پاسخ‌هایی را برای سوالات پرسشی<sup>۲۵</sup> از منابع مختلف دانش ارائه کند. نتایج تجربی در این مقاله

پردازش زبان طبیعی<sup>۱</sup> طراحی می‌شود برای ماشین مورد تحلیل قرار می‌گیرد تا بتواند با تطبیق الگو و قوانین از پیش نوشته شده برای آن، مفهوم سوال را درک کرده و پاسخ مناسبی برای آن تولید کند. موارد بیان شده راهکاری برای ساخت سیستم‌های پرسش و پاسخ سنتی است. امروزه، روش‌های یادگیری عمیق، توانایی‌های بهتری برای پاسخ به سوال‌های پیچیده در سیستم‌های پرسش و پاسخ ارائه می‌دهد [۲]. استفاده از این روش‌ها می‌تواند تاثیر بسزایی در آینده ربات‌های انسان‌نما و عروسک‌های هوشمند داشته باشد.

امروزه ربات‌های انسان‌نما و عروسک‌های هوشمند سخن‌گو، نقش زیادی در تعامل با مخاطبین دارند. آمکا<sup>۲</sup> یکی از پیشرفته‌ترین ربات‌های جهان با شکل انسانی است که توسط شرکت هنرهای مهندسی شده<sup>۳</sup> تولید شده است. آمکا می‌تواند توسعه فناوری‌های جدید هوش مصنوعی و یادگیری ماشین ارائه دهد. صورت و دست‌های سیلیکونی با رنگ پوست انسان، ظاهر بی‌نظیری را برای این ربات بیان می‌کند. این ربات دارای هوش ابری است که به آن اجازه می‌دهد با پیشرفت فناوری ارتقا یابد [۴]. کلیه این ربات‌ها، با یک الگوریتم خاصی از گفت‌وگو به یک ماشین مکالمه طبیعی تبدیل شده‌اند. آندری چواندا<sup>۴</sup> و همکاران [۱] با روش‌های یادگیری عمیق با استفاده از شبکه‌های بازگشتی<sup>۵</sup> و حافظه کوتاه‌مدت طولانی<sup>۶</sup>، تلاش کرده‌اند تا یک ماشین با مکالمه طبیعی به زبان آندونزی تولید کنند. بر طبق این پژوهش‌ها، استفاده از مدل‌های یادگیری عمیق نسبت به سایر روش‌های یادگیری ماشینی ارجحیت دارد زیرا روش‌ها در یادگیری عمیق این اجازه را به الگوریتم می‌دهد تا ویژگی‌های مهم در تشخیص توسط آن تهیه شود اما در روش‌های دیگر یادگیری ماشینی، نیاز به محققانی است که ویژگی‌های زیادی را برای یادگیری شبکه تهیه کنند. نتیجه این مدل با معیار خطای سردرگمی<sup>۷</sup> برابر ۱/۱۲ شده است.

شارما<sup>۸</sup> و همکاران [۵]، در مورد توسعه سیستم‌های پرسش و پاسخ با تکنیک‌های یادگیری عمیق جهت بیان رویکردهای مختلف با مبنای پردازش زبان طبیعی<sup>۹</sup> اشاره می‌کند. مدل پیشنهادی با مجموعه داده بابی<sup>۱۰</sup> فیس‌بوک پیاده‌سازی و ارزیابی شده است. نتایج به دست آمده نشان می‌دهد که استفاده

- 11 Root Mean Squared Propagation (RMSProp)
- 12 Sparse Categorical Cross Entropy
- 13 Memory Networks
- 14 Long Short-Term Memory (LSTM)
- 15 Dynamic Memory Networks (DMN)
- 16 Memory Network
- 17 Dynamic Memory Network
- 18 Ming Tan
- 19 Convolutional Neural Network (CNN)
- 20 Text REtrieval Conference (TREC)
- 21 Insurance QA
- 22 Attention LSTM
- 23 Max-Pooling
- 24 Li
- 25 What, When, Where, Who, Whom, Which, Whose, Why and How

- 1 Natural Language Processing (NLP)
- 2 Ameca
- 3 Engineered Arts
- 4 Andry Chowanda
- 5 Recurrent Neural Network (RNN)
- 6 Long-Short Term Memory (LSTM)
- 7 Perplexity
- 8 Sharma
- 9 Natural Language Processing (NLP)
- 10 Babi

مشترک را در بین زبان‌ها تجزیه و تحلیل می‌کند. در نهایت، مدلی که ارائه شده‌است با پشتیبانی از ۵۳ زبان مختلف در زمینه درک گفتار با منابع کم مورد استفاده قرار گرفته و راه را برای سیستم‌های تشخیص گفتار خودکار<sup>۵</sup> چند زبانه هموار می‌کند.

توسعه سریع جهان‌های مجازی مانند متاورس<sup>۶</sup>، در تعامل و ارتباطات انسانی متحول می‌شود. لی و همکاران [۱۰] پتانسیل یکپارچه‌سازی فناوری رابط مغز و رایانه<sup>۷</sup> با متاورس را برای ایجاد الگوی جدیدی از تعامل انسانی از طریق فکر بررسی کردند. در این مقاله یک سیستم مغز به گفتار<sup>۸</sup> را پیشنهاد می‌کند که با استفاده از گفتار خیالی<sup>۹</sup>، یک الگوی شبکه عصبی بصری برای ارتباط هوشمند با دنیای واقعی را با استفاده از سیگنال‌های مغزی فعال می‌کند. این مقاله امکان کنترل خانه هوشمند از طریق ارتباط خیالی مبتنی بر گفتار با یک دستیار مجازی را به نمایش گذاشته‌است که می‌تواند یک برنامه کاربردی آینده در یک سیستم متاورس مغزی باشد. این تحقیق پتانسیل ارتباطات هوشمند مبتنی بر گفتار خیالی را در متاورژن برجسته می‌کند.

توسعه ربات‌های کنترل شده با صدا تا حد زیادی به برچسب‌های صریح برای صداها و تصاویر و همچنین عملکردهای پاداش دستی تنظیم شده بستگی دارد. با این حال، این روش‌ها به طور دقیق توسعه حسی-حرکتی انسان را منعکس نمی‌کنند. چانگ<sup>۱۰</sup> و همکاران [۱۱]، مدلی برای توسعه این مسئله پیشنهاد می‌کنند. این مقاله، تصاویر و دستورات صوتی را با حداقل نظارت مرتبط می‌کند و از آن برای ایجاد یک تابع پاداش ذاتی برای یادگیری تقویتی در وظایف رباتیک استفاده می‌کند. این مقاله بر روی بازوی رباتیک متمرکز بوده‌است و با خط مشی<sup>۱۱</sup> آموخته‌شده از شبیه‌سازی و در دنیای واقعی توانسته است که راهی برای تعمیم بازنمایی‌های بصری-شنیداری، یادگیری خود نظارتی، و استفاده از تفسیر صدا در فراگیری زبان ربات و تعامل انسان و آن ایجاد کند.

در [۱۲] شهبازی، یک بازی رایانه‌ای تعاملی و جدیدی طراحی و ساخته است که با استفاده از سیستم رباتیک هوشمند، قادر به یادگیری و تقلید حرکات و صداها بازی‌کننده است. این بازی جایگزینی نوین برای بازی‌های سنتی خیمه‌شب‌بازی است و می‌تواند در کاربردهای سرگرمی و تبلیغاتی استفاده شود. این ربات انسان‌نما از روش‌های یادگیری الهام‌گرفته از ساختار سیستم

ثابت می‌کند که مسئله طبقه‌بندی سوال را می‌توان با استفاده از یک روش یادگیری کاملاً دقیق حل کرد و مزایای ویژگی‌های مبتنی بر تحلیل معنایی را نشان داد. شیونگ<sup>۱</sup> و همکاران [۷]، شبکه‌های حافظه پویا جدیدی پیشنهاد می‌کند که از یک رمزگذار دو سطحی با یک جمله‌خوان و لایه ترکیبی ورودی استفاده می‌شود تا جریان اطلاعات بین جملات را امکان‌پذیر کند که با نام شبکه‌های حافظه پویا پلاس معرفی شده‌است. برای بخش حافظه، از لایه واحد بازگشتی گیتی<sup>۲</sup> استفاده می‌شود. با این حساب، شبکه جدید حافظه پویا پلاس برخلاف قبل، نیازی به برچسب‌گذاری حقایقی که برای پاسخ به یک سوال خاص مرتبط است ندارد.

در [۸]، نمونه‌ای از یک ربات اندرویدی خودکار ارائه شده است که قادر به برقراری گفتگوی طبیعی و روزانه با انسان است. سیستم گفتگوی پیشنهادی از سیستم‌های گفتگوی وظیفه‌محور متمایز می‌شود، زیرا از هدف یا اطلاعات مشخصی بهره‌برداری نمی‌کند و در نتیجه، در شرایطی که از سوی انسان درخواست مشخصی نباشد، همچنان قادر به تولید جملات مناسب است. این نیازمند تغییر سیاست‌های مدیریت گفتگو نسبت به سیستم‌های گفتگوی کنونی است. هدف از این تحقیق، پیدا کردن سازنده‌ای برای سیستم گفتگو است که گفتگوی روزانه را با استفاده از ربات گفتگوی خودکار و طبیعی تحقق بخشد. این ربات توانست گفتگوی روزانه را برای مدت ۱۰ دقیقه در محیطی که ربات و مخاطب برای اولین بار به آزمایش گذاشته شده‌اند، انجام دهد. علاوه بر این، یک آزمون تورینگ چندجانبه نشان داد که نیمی از شرکت‌کنندگان احساس می‌کردند که ربات اندرویدی به طور جزئی از راه دور کنترل می‌شود، یعنی رفتار آن با انسان قابل مقایسه است. این یافته‌ها نشان می‌دهند که رویکرد پیشنهادی در این تحقیق برای تحقق گفتگوی روزانه موثر است. تشخیص خودکار گفتار<sup>۳</sup> در سال‌های اخیر پیشرفت‌های چشمگیری داشته است. کونیو<sup>۴</sup> و همکاران [۹] توانسته‌اند رویکردی جدید برای مدلی که بازنمایی گفتار بین زبانی را از طریق پیش‌آموزش بر روی شکل موج خام داده‌های گفتار در چندین زبان، یاد می‌گیرد معرفی کنند. این مقاله نشان داده می‌شود که پیش‌آموزش چند زبانه به‌طور قابل توجهی بهتر از پیش‌آموزش تک‌زبانه است. آزمایش‌های آنها بهبودهای قابل توجهی را در کاهش میزان خطای واج و میزان خطای کلمه در مقایسه با روش‌های پیشرفته نشان می‌دهد. علاوه بر این، این مطالعه بازنمایی گفتار گسسته پنهان

5 Automatic Speech Recognition (ASR)

6 metaverse

7 Brain-Computer Interface (BCI)

8 Brain-To-Speech (BTS)

9 imagined speech

10 Chang

11 policy

1 Xiong

2 Gated Recurrent Unit (GRU)

3 Automatic Speech Recognition (ASR)

4 Conneau



شکل ۱. ربات خیمه‌شب‌باز [۱۲]

Fig. 1. puppeteer robot

عصبی انسان بهره می‌برد و ارتباط طبیعی میان انسان و ربات را فراهم می‌کند. این مطالعه از این روش خلاقانه برای کنترل ربات و کاربردهای صنعتی از راه دور بحث می‌کند و مسیری راحت برای برقراری ارتباط آسان بین افراد غیرمتخصص و ربات ارائه می‌دهد. این پژوهش به طراحی و ساخت یک مکانیزم گفت‌وگو با یادگیری عمیق برای ربات خیمه‌شب‌باز مرجع [۱۲] اشاره دارد که در شکل ۱ نشان داده شده است.

در این پژوهش یک مکانیزم گفت‌وگو با رویکردی نوین برای ربات خیمه شب باز ارائه شده است که وجه تمایز مکانیزم گفت‌وگو آن با موارد مشابه قبلی، قابلیت ادراک گفتار فارسی از طریق استدلال عمیق در شبکه عصبی طراحی شده برای آن است. از دیگر وجوه تمایز این سامانه، امکان کنترل برخط ربات تعاملی توسط کاربر انسانی در موارد خاص و یادگیری شیوه تفکر کاربر انسانی در موارد عام می‌باشد.

از کاربردهای راهبردی که این مکانیزم برای ربات ایجاد می‌کند، می‌توان به موارد زیر اشاره کرد:

- ۱) جایگزینی جهت مشاوره روانشناسی، خانوادگی، آموزشی و غیره
- ۲) کاربردهای تبلیغات و بازاریابی و سرگرمی
- ۳) تدریس دروس به کودکان و نوجوانان
- ۴) آموزش مفاهیم به بیماران کم‌توان ذهنی

در این پژوهش، روشی برای تعامل بین انسان و ربات با استفاده از مکانیزم گفت‌وگو پیاده‌سازی شده است که بر مبنای ساختار یک عروسک پارچه‌ای انجام گرفته است. در این روش، ارتباط صوتی بین انسان و ربات از طریق ماژول بلوتوث برقرار می‌شود. مطابق شکل ۲، ابتدا، داده‌های صوتی حاصل از گفتار انسان توسط میکروفون ربات دریافت و از طریق ماژول بلوتوث به نرم‌افزار توسعه داده شده ربات منتقل می‌شوند. سپس، مدل پیش‌آموزش دیده‌ی گفت‌وگویی که در این پژوهش توسعه داده شده است، به شکل یک نرم‌افزار گرافیکی بر روی کامپیوتر اجرا می‌شود و به تحلیل محتوای داده‌های صوتی ورودی و تبدیل به متن می‌پردازد. پس از تحلیل متن، پاسخ مناسب از مدل گرفته می‌شود که به صورت صوتی به ربات ارسال می‌شود. این پاسخ با استفاده از بلندگوی ربات پخش می‌شود و انسان می‌تواند پاسخ را به صورت صوتی از ربات دریافت کند. این روش تعاملی و دوطرفه، امکان یک تبادل اطلاعات مؤثر و طبیعی بین انسان و ربات را فراهم می‌کند و از جدیدترین پیشرفت‌ها در حوزه گفت‌وگوی ماشینی بهره می‌برد. به منظور ارتباط بیشتر بین اپراتور ربات و کاربر، از ابزار جانبی دیگری مانند سروو موتور (برای حرکت سر و دست‌ها)، دوربین شلنگی (برای مشاهده

کاربر توسط اپراتور) و پردازنده esp32 جهت کنترل سروو موتورها استفاده شده است.

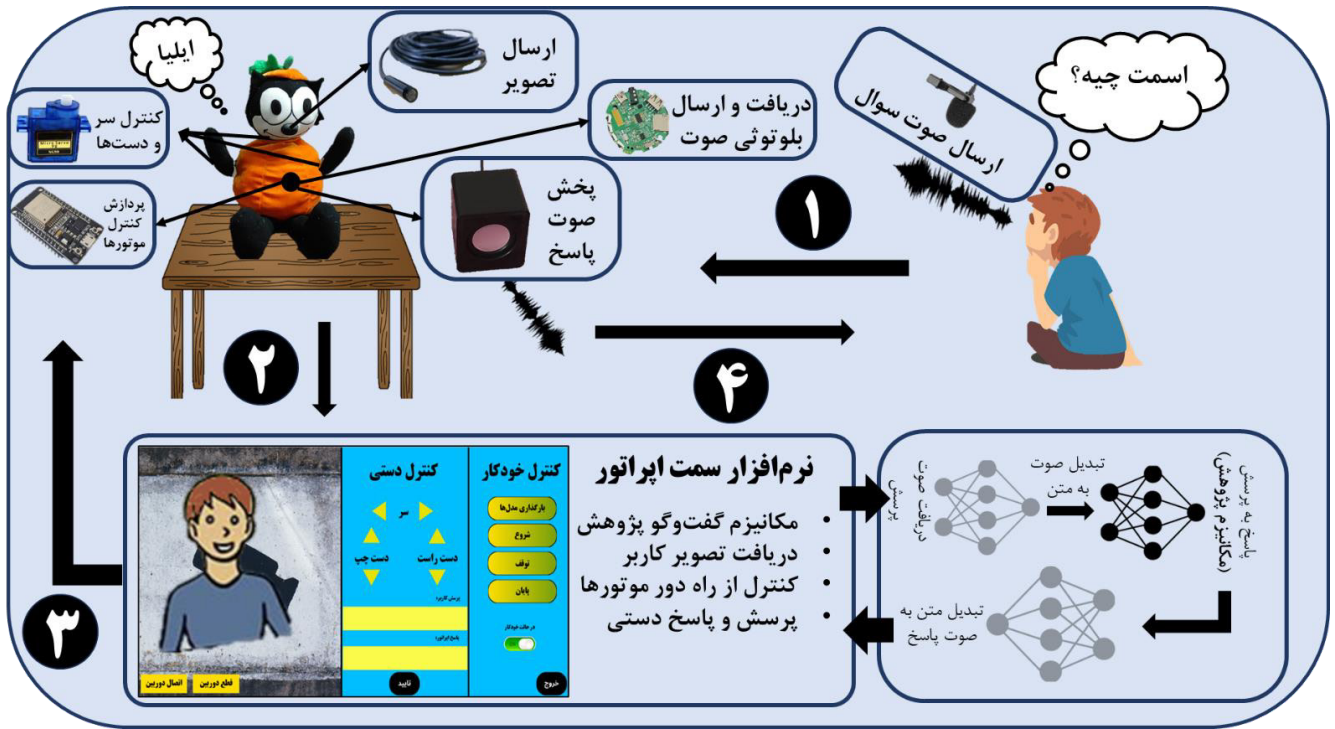
رویکرد این مقاله به این صورت است که ابتدا روش‌های مختلفی پیرامون مکانیزم گفت‌وگو بیان می‌شود و در آخر، قابل اجرایی‌ترین روش جهت قرارگیری در ربات خیمه‌شب‌باز با توجه به مزایا و معایبی که هر روش داشته است، انتخاب می‌شود. سپس، با تغییر پارامترهای مختلف مدل شبکه یادگیری عمیق (مثل تغییر در تعداد لایه‌ها، نرون‌های عصبی، توابع فعال‌ساز و غیره)، بهترین مدل ساخته می‌شود و نتایج آن ارائه می‌گردد.

## ۲- روش‌های مکانیزم گفت‌وگو

روش‌های بررسی شده جهت گفت‌وگو هوشمند و تعامل فارسی ربات با انسان، شامل سه دسته می‌شود:

- ۱) مدل‌هایی که پیاده‌سازی و ارزیابی شده‌اند و به عنوان مدل نهایی و به عنوان مکانیزم مناسب در ربات خیمه‌شب‌باز قرار گرفته‌اند.
  - ۲) مدل‌هایی که پیاده‌سازی و ارزیابی شده‌اند اما به دلیل معایب آنها در ربات خیمه‌شب‌باز قرار ندارند.
  - ۳) مدل‌هایی که بر مبنای توضیح و شرحی از آنها می‌باشند و نتایج آن مورد بررسی واقع نشده است.
- مطابق با بخش‌های بیان شده در مکانیزم‌های گفت‌وگویی بررسی شده در این مقاله، روش‌های زیر مورد بررسی قرار گرفته‌اند:
- ۱) ماشین حالت متناهی





شکل ۲. ربات طراحی و ساخته شده جهت اجرای مکانیزم پرسش و پاسخ پژوهش

Fig. 2. A robot designed and built to execute the question-and-answer mechanism of research

در نظر گرفته می شود. به این ترتیب، بر اساس مکانیزم گفت و گویی که از قبل پیش بینی شده است، می توان برای پاسخ به سوالات محدودی، جواب هایی تنظیم کرد.

مفهوم این بخش در کدنویسی، یک نوع کنترل فازی به حساب می آید. به این صورت که باید برای هر وضعیت، دستورات شرطی نوشته شود. این دستورات نه تنها باید وضعیت بعدی را مشخص کند، بلکه باید بتواند شرایط مختلفی مثل تکرار وضعیت ها، بروز خطاها و بازگشت به وضعیت قبلی را تحت کنترل قرار دهد تا سیستم دچار اشتباه نشود. در شکل ۳، نمونه ساده ای از نمودار ماشین حالت محدود نشان داده شده است.

آنچه که از نمودار شکل ۳ مشخص است، مرحله شروع همواره با نقطه توپر شروع شده و در پایان یک توالی، نقطه پایان دارای دو دایره هم مرکز می باشد. هر دایره منفرد، نشانگر وضعیت گفت و گو به ازای پاسخی که دریافت کرده است و هر پیکان، پاسخی است که برای وضعیت قبلی در نظر گرفته شده است. در این حالت باید مشخص شود که گفت و گو به چه صورت خواهد بود. اگر فردی که پاسخی را بیان می کند، مجبور به گفتن کلمات خاصی

(۲) همسایگی نیوتن

(۳) همسایگی متن سریع<sup>۱</sup>

(۴) سوال و پاسخ تصادفی

(۵) تشخیص سوال از متن

(۶) پیاده سازی احساسات (تکامل یافته روش چهارم)

این بخش سعی دارد با بیان مزایا و معایبی هر یک از روش های بیان شده، قابل اجرا ترین روش را برای مکانیزم گفت و گوی هوشمند ربات خیمه شب باز پیشنهاد دهد.

## ۲-۱- روش ماشین های حالت محدود<sup>۲</sup>

در این روش، برای تشخیص پاسخ مناسب، باید کلماتی که قرار است به عنوان ورودی و خروجی (پاسخ) در سیستم داده می شود، به صورت مشخص بیان شود. در هر بخش، ماشین انتظار دارد که کلمه ورودی که برای آن تعریف شده است به آن داده شود. پس از آن، شاخه بعدی توسط ماشین

1 Fasttext

2 Finite State Machine (FSM)

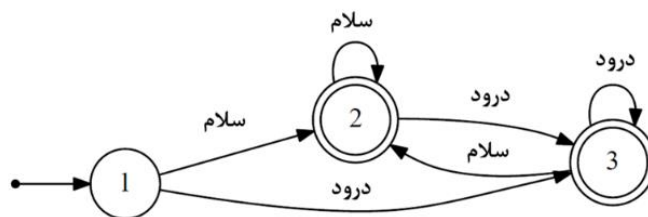
(۷) هر دو مرحله ۲ و ۳ در این نمودار، پایان این حلقه می‌باشند و اتفاق دیگری رخ نمی‌دهد. می‌توان برای پایان نهایی، یک پاسخ یکتا تعریف کرد و از شروع مجدد حلقه جلوگیری کرد. این روش، بستگی به ساختاری دارد که برای نمودار تعریف شده‌است.

در این نمونه ساده از ماشین حالت محدود سعی شد پیچیدگی این روش برای اجرای گفت‌وگوهای حجیم و نزدیک‌تر به انسان نشان داده شود. بنابراین، اگر گفت‌وگویی که ربات انجام می‌دهد، از کلمات محدودی استفاده می‌شود و مکانیزم‌های گفت‌وگوی آن کوتاه و دارای پیچیدگی‌های بازگشت‌پذیر بین عبارات بیان‌شده نیست، می‌توان از این روش استفاده کرد اما همچنان در پایین‌ترین سطح هوشمندی قرار دارد و پیاده‌سازی آن نیاز به شناخت کلیه حالت‌های یک گفت‌وگو است که تقریباً غیرممکن به نظر می‌رسد.

این روش، مزایایی نیز دارد. یکی از چالش‌های گفت‌وگو، احساساتی است که در بیان کلمات منتقل می‌شود. یک کلمه ثابت، در هر وضعیت، با لحن مختلف، پیام‌های متفاوتی را بیان می‌کند. برای مثال، اگر کلمه «سلام»، سه مرتبه به شخصی گفته شود، پیام دریافتی آن می‌تواند برخلاف مفهوم خود کلمه باشد و شخص احساس کند که مورد تمسخر واقع شده‌است. در این لحظه، واکنشی که در پاسخ به «سلام» سوم دیده می‌شود، می‌تواند هر چیزی باشد. روش ماشین حالت محدود می‌تواند این چالش را در محدوده کوچکی، حل کند و سرعت بیان این مطلب را بسیار سریع‌تر از روش‌های پردازش زبان طبیعی و مدل‌های حجیم آن انجام دهد.

### ۲-۲- روش همسایگی نیوتن

یکی از روش‌ها برای اینکه یک مکانیزم گفت‌وگو شکل بگیرد، استفاده از بردار اعداد به جای کلمات واقع در گفت‌وگو است. هر کلمه متناظر یک بردار قرار می‌گیرد و یک جمله شامل بردارهایی از کلمات می‌شود. حال می‌توان از این بردار کلمات مستقیم و با استفاده از روش‌های یادگیری ماشین به یادگیری کلمات و تشخیص آنها پرداخت. همچنین این بردارها می‌توانند در تکنیک‌های یادگیری عمیق نیز مورد استفاده قرار گیرند. در همسایگی نیوتن، کلمات و حروف دو جمله تفکیک می‌شوند و با یکدیگر مقایسه می‌گردند و طبق رابطه (۱)، میزان فاصله آنها از یکدیگر محاسبه می‌شود. در نهایت یک بازه برای کلمه و حروف در کنترل نزدیکی دو جمله مطرح می‌شود.



شکل ۳. نمودار روش FSM در گفت‌وگو

Fig. 3. Diagram of FSM method in conversation

باشد، می‌توان از این روش برای پیاده‌سازی یک گفت‌وگو ساده استفاده کرد. یکی از نکات مهم در این نمودارها، وضعیت‌هایی است که یک کلمه خارج از لیست کلمات گفت‌وگو باشد. برای رفع این مشکل، باید دستورات به‌گونه‌ای تنظیم شود که اگر کلمه‌ای غیر از کلمه مورد انتظار نمودار مطرح شد، درخواست کلمه دیگر یا خطا را نشان دهد. این موضوع در شکل ۳ رعایت نشده‌است. برای مثال، اگر در وضعیت ۱، کلمه «چطوری» بیان شود، هیچ راهی برای حرکت در نمودار باقی نمی‌ماند و گفت‌وگو دچار ایستایی می‌شود.

برای بیان عملکرد نمونه‌ای که در شکل ۳ نشان داده شده‌است، می‌توان به موارد زیر اشاره کرد:

- (۱) در مرحله ۱، کلمه «سلام» گفته می‌شود.
- (۲) با دریافت کلمه «سلام»، مرحله دوم آغاز می‌شود. در این مرحله، می‌توان پاسخی برای کلمه «سلام» گفته‌شده در نظر گرفت.
- (۳) در این مرحله، دو انتخاب باعث ورود به مرحله جدید می‌شود.
- (۴) اگر مجدد کلمه «سلام» گفته شود، سیستم در مرحله ۲، باقی می‌ماند و می‌تواند مجدد پاسخی که در این مرحله دریافت می‌کرد را مجدد دریافت کند.
- (۵) اگر به جای کلمه سلام، کلمه «درود» بیان شود، مکانیزم گفت‌وگو وارد مرحله ۳ می‌شود و می‌تواند پاسخ مربوط به آن را از این مرحله دریافت کند.
- (۶) شرایط مرحله ۳ مانند مرحله ۲ است با دو تفاوت.
  - a. اگر از مرحله ۱ مستقیماً کلمه «درود» بیان شود، سیستم مستقیماً وارد مرحله ۳ می‌شود و پاسخ‌گویی از آن وضعیت ادامه می‌یابد.
  - b. این مرحله قابلیت بازگشت به مرحله ۲ را با بیان کلمه «سلام» دارد.

مشاهده است. در این دسته‌بندی، شهرهای ایرانی، زبان‌ها، احساسات، میوه‌ها و حیوانات در یک محدوده نزدیک به هم قرار گرفته‌اند و توسط مدل تشخیص داده شده‌اند.

یکی از روش‌های مرسوم در تبدیل کردن کلمات به بردار، روش کدگذاری یک‌بعدی<sup>۲</sup> است. در این روش، به تعداد کلمات مدنظر، یک بردار با آرایه‌های صفر در نظر گرفته می‌شود. در مواقعی که تعداد کلمات زیادی مورد بررسی قرار می‌گیرد، بردارهای روش کدگذاری یک‌بعدی بسیار بزرگ می‌شوند و سرعت فرآیند یادگیری را کند می‌کنند. از طرفی، بردارهای این روش به علت منحصربه‌فرد بودن، عمودبرهم هستند بنابراین، ارزش کلمات نسبت به هم به یک مقدار سنجیده می‌شود (به عنوان مثال کلمه «ملکه» و «سگ» ارزش یکسانی با کلمه «ملکه» و «پادشاه» دارد).

این مدل می‌تواند در انتخاب کلمات مناسب برای مکانیزم گفت‌وگو مناسب باشد. به این ترتیب که اگر کلمه‌ای در متن پاسخ بیان شد، با نزدیک‌ترین کلمات به آن جواب خود را پیدا کند. برای مثال اگر کلمه «سلام» به این مدل داده شود، در خروجی به تعداد قابل تنظیم، همسایگی دریافت می‌شود که شامل کلماتی مانند «باسلام»، «درد» و غیره است و اگر دامنه همسایگی افزایش یابد، کلماتی مانند «سلام دوستان» یا «سلام مرسی» را هم تشخیص دهد. در شکل ۵ نمونه‌ای از نتایج این مدل نشان داده شده‌است. پس از تست و اجرا، مشکلاتی که بر سر راه این مدل وجود داشت، استفاده از آن را برای این مورد، لغو کرد. این مشکلات شامل موارد زیر می‌شدند:

- ۱) مشکل در دریافت همسایگی درست و موردنظر از خروجی مدل
- ۲) احتمال کم دریافت پاسخ‌های مناسب از کلمه پرسشی با دریافت همسایگی
- ۳) مدل حجیم حدود ۶ گیگابایت و مشکل در لود کردن مدل
- ۴) درگیر کردن ۱۰ گیگابایت از رم کامپیوتر
- ۵) کاهش حجم مدل جهت لود سریع‌تر اما با دقت کم‌تر
- ۶) عدم کاربری مناسب از آنچه انتظار می‌رفت.

آنچه که می‌توان از این روش انتظار داشت، استفاده از بردارهای از پیش آموزش دیده است که می‌تواند برای هر کلمه، یک بردار با ۳۰۰ ویژگی مختلف که به صورت یادگیری نظارت‌نشده انجام گرفته‌است تحویل دهد. با استفاده از این بردارها می‌توان در یادگیری‌های مختلف که نیازمند جاسازی از پیش انجام شده هستند استفاده کرد. زمانی که مباحثی مثل دسته‌بندی

$$d = \frac{n_{wq} - n_{sw}}{\max(n_{wq} - n_{sw})} \times 100 \quad (1)$$

مشکلات این روش شامل موارد زیر می‌شود:

- ۱) عدم تشخیص شباهت کلمات مترادف
- ۲) تشخیص ضعیف کلمات داری پسوند (مثل بهار و بهاری)
- ۳) وابستگی به جایگاه کلمه در جمله

استفاده از همسایگی نیوتن دقت چندانی ندارد اما روش استخراج کلمات از جمله و شمارش آن که در ابتدای این روش استفاده می‌شود، می‌تواند برای روش‌های یادگیری عمیق مفید باشد.

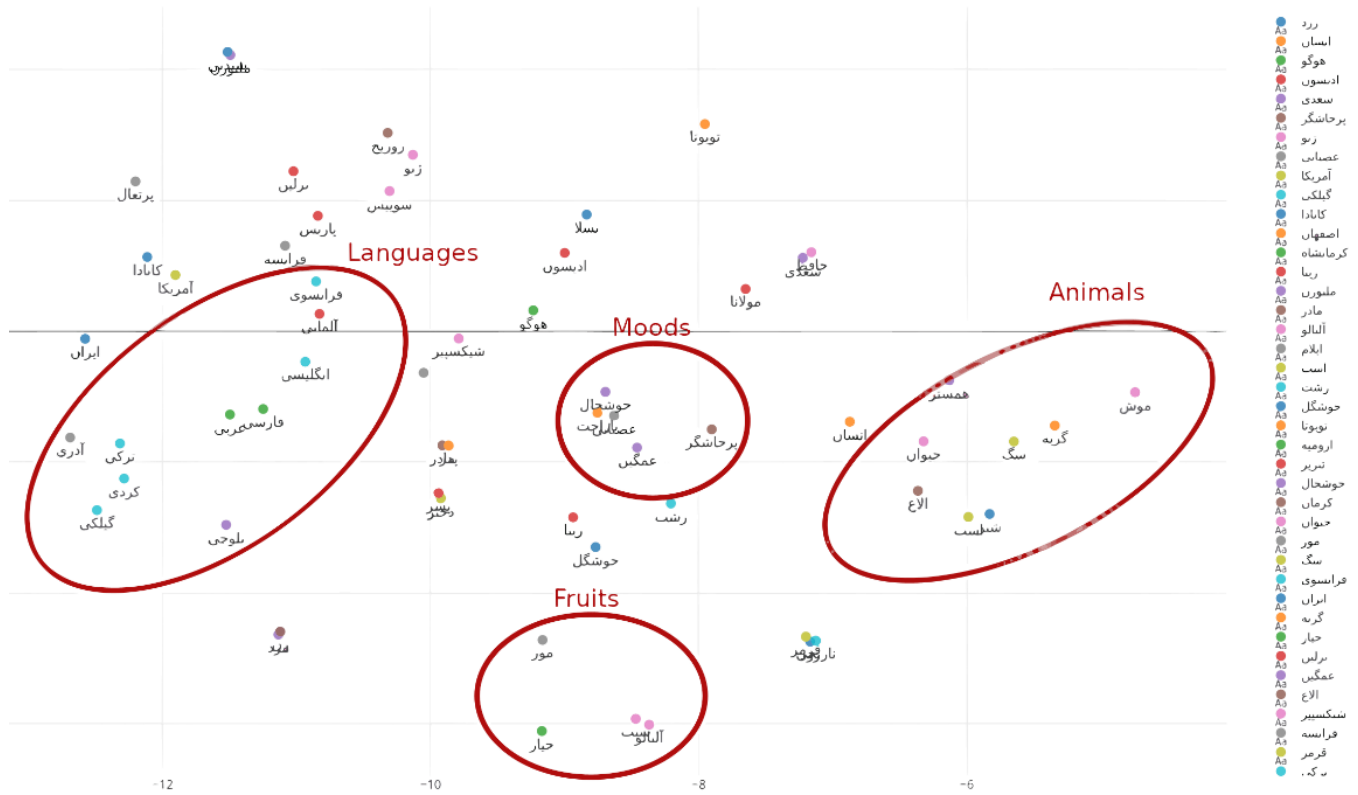
### ۲-۳ روش همسایگی متن سریع

روش همسایگی متن سریع، یک کتابخانه منبع‌باز رایگان و سبک است که به کاربران اجازه می‌دهد تا بازنمایی متن و طبقه‌بندی متن را یاد بگیرند. این کتابخانه روی سخت‌افزارهای عمومی کار می‌کند. از ویژگی‌هایی که مدل‌های همسایگی متن سریع دارند این است که می‌توان اندازه مدل‌ها را کاهش داد تا حتی در دستگاه‌های تلفن همراه هم قابل اجرا باشند. یکی دیگر از قابلیت‌های جذاب این کتابخانه، آموزش مدل اختصاصی با توابع از پیش نوشته آن می‌باشد [۱۳].

این کتابخانه مدلی برای برای زبان فارسی نیز ارائه داده‌است و از آن در این پژوهش، برای دسترسی به همسایگی کلمات استفاده شد. یکی از مشکلاتی که در اجرای این مدل به‌وجود می‌آید، حجم بالای آن است. اگرچه مدل‌های همسایگی متن سریع قابلیت کاهش ابعاد و حجم را دارند اما با کاهش حجم مدل، دقت آن نیز کاهش می‌یابد.

همسایگی متن سریع نسبت به روش‌های مشابه کشف همسایگی‌های کلمات نظیر Word2vec و GloVe انعطاف‌پذیرتر است زیرا هر کلمه ترکیبی از حروف یا کلمات پشت‌سرهم است در حالی که در دو مدل دیگر، کلمات کوچک‌ترین واحدها هستند. در واقع، در همسایگی متن سریع با یک مقداردهی اولیه و تشکیل شبکه عصبی بازگشتی<sup>۱</sup>، تجزیه و تحلیل متن با دقت بهتری انجام می‌شود، به‌خصوص در مواردی که اندازه مجموعه داده کوچک باشد [۱۴].

حامد همتی، یک نمونه از همسایگی و طبقه‌بندی کلمات را با مدل همسایگی متن سریع انجام داده‌است که در شکل ۴، نتیجه موردنظر قابل



شکل ۴. همسایگی برخی کلمات از مدل متن سریع حامد همتی [۱۵]

Fig. 4. The neighborhood of some words from Hamed Hemti's Fasttext model

در سوال عنوان جدا و پاسخ‌های منحصر به فرد خودش را داشته باشد تا در تشخیص بهترین خروجی مشاهده شود.

به منظور آموزش مدل یادگیری عمیق در این روش، ابتدا مجموعه داده داده‌ها تهیه می‌شود. بخشی از مجموعه داده آماده شده برای آموزش در شکل ۶ نشان داده شده است.

همانطوری که در شکل ۶ مشخص است، هر موضوعی که پیرامون آن جوابی در نظر گرفته شده است در یک برچسب به خصوصی ارائه شده است. مدل با برچسب‌ها و سوالاتی که در هر برچسب مطرح شده است آموزش می‌بیند. برای اینکه مدل دقت بالایی داشته باشد، باید تعداد سوالات هر برچسب بیشتر از یک یا دو عدد انتخاب شود. برای این کار نیاز است سوالاتی که مطرح می‌شود، دارای یک مفهوم باشد و در آن کلمات مشابهی که قرار است در فاز پیش‌بینی مطرح شود، توسط سازنده مجموعه داده در نظر گرفته شود.

خلاصه مدل استفاده شده در این روش در شکل ۷ نشان داده شده است.

متون مطرح باشد، می‌توان با استفاده از این روش، روند یادگیری را بهبود و با سرعت بیشتری انجام داد.

## ۲-۴ روش سوال و پاسخ تصادفی

یکی از به‌روزترین روش‌های چت‌بات، استفاده از روش تشخیص مفهوم سوال و انتخاب پاسخی مناسب برای آن است. در این روش، ابتدا باید سوال تشخیص داده شود. بنابراین، حروف و کلماتی که در سوال گفته شده‌اند، به صورت یک لیست در می‌آید. سپس یک سری عملیات‌های پیش‌پردازش بر روی آنها انجام می‌شود. این عملیات‌ها می‌توانند شامل مواردی مثل حذف کلمات غیر ضروری مثل «در»، «از»، «با»، «می» و «...» باشند یا دریافت ریشه برخی از کلمات مانند «رفتن به جای رفتن».

پس از تشخیص سوال و عنوانی که برای آن سوال در داده‌های آموزش در نظر گرفته شده است، یک پاسخ از سری پاسخ‌های از پیش تعیین شده را به صورت تصادفی انتخاب کرده و در خروجی ارائه می‌دهد. بنابراین، باید



```
ft.get_nearest_neighbors('200', 'سلام')
```

```
[(0.7062562704086304, 'سلامو'),
 (0.7010449767112732, 'باسلام'),
 (0.6782733201980591, 'درود'),
 (0.6445516347885132, 'سلام.ضمن'),
 (0.6404579281806946, 'عزیز.بنده'),
 (0.6391009092330933, 'عزیز سلام'),
 (0.6373170018196106, 'سلاممم'),
 (0.6371632814407349, 'سلام.آقا'),
 (0.637017011642456, 'عزیز.با'),
 (0.6366755962371826, 'شمامن'),
 (0.632809579372406, 'سلام.آقا'),
 (0.6300079226493835, 'سسلام'),
 (0.6289799809455872, 'سلام5'),
 (0.6267051100730896, 'سلام.امروز'),
 (0.6260673403739929, 'سلام.با'),
 (0.6258944869041443, 'سلام.تشکر'),
 (0.622391939163208, 'سلاممرسی'),
 (0.620347797870636, 'سلام.اولا'),
 (0.6199771761894226, 'سلام.ی'),
 (0.6198668479919434, 'همگی.من'),
 (0.6179640293121338, 'سلام.اول'),
 (0.6178856492042542, 'سلام.پس'),
 (0.6136710047721863, 'سلاممم'),
 (0.6132506132125854, 'سلام.خدمت'),
 (0.6124227046966553, 'سلامسلام')]
```

شکل ۵. نتایج همسایگی کلمه «سلام» در همسایگی متن سریع متناسب با دقت همسایگی آن

Fig. 5. The results of the neighborhood of the word "Hello" in FastText according to the accuracy of its neighborhood

<pre>{   "tag": "هتل",   "input": [     "کجا میخوابی",     "جای خوابت کجاست",     "اصلاً تو میخوابی",     "تو کی خسته میشی",     "من رو زمین میخوابم"   ],   "responses": [     "پیش تو روی تخت",     "هرجا فقط پیش تو",     "من خواب ندارم",     "سر جام میخوابم",     "هرجا فقط پیش تو"   ] },</pre>	<pre>{   "tag": "آب و هوا",   "input": [     "هوا چطوره",     "هوا خوبه",     "هوا عالییه",     "امروز هوا چطوره",     "فردا هوا چطوره",     "دیروز هوا چطور بود"   ],   "responses": [     "خوبه",     "بد نیست",     "دوستش ندارم این هوا رو",     "مردانه سرد است u200Cاناجوان",     "مگه من هواشناسم"   ] },</pre>
<pre>{   "tag": "مکان",   "input": [     "کجایی",     "اهل کجایی",     "کجا زندگی میکنی",     "اهل چه جایی هستی",     "الان کجایی"   ],   "responses": [     "ایرانی هستم",     "تو چیکار داری",     "ایران اهل اصفهان",     "سبز و سفید و قرمز",     "جایی که عشق باشه",     "هرجا که خنده هست",     "کره زمین"   ] },</pre>	<pre>{   "tag": "بای",   "input": [     "خداحافظ",     "خدانگهدار",     "به امید دیدار",     "بازم ببینمت",     "بای بای",     "مچکرم ازت",     "برو دیگه",     "روز خوش"   ],   "responses": [     "بای بای",     "خداحافظ",     "خدانگهدار",     "خداحافظ باشه",     "خواهش میکنم خداحافظ",     "بازم بیا سراغ ما فعلاً"   ] },</pre>

شکل ۶. نمونه مجموعه داده روش تشخیص سوال و پاسخ تصادفی

Fig. 6. Sample dataset of random question and answer detection method

```

Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
dense (Dense)                (None, 4096)                4808704
dropout (Dropout)            (None, 4096)                0
dense_1 (Dense)              (None, 64)                  262208
-----
Total params: 5,070,912
Trainable params: 5,070,912
Non-trainable params: 0
    
```

شکل ۷. خلاصه مدل پرسش و پاسخ تصادفی

Fig. 7. Summary of random question and answer model

#### ۲-۵- روش تشخیص سوال از متن

یک راه برای تشخیص سوال و پاسخ به آن، استفاده از متن کمکی است. به این صورت که در مجموعه داده‌ها، باید یک جمله در نظر گرفته شود و متناسب با آن سوال مطرح شود و پاسخ آن از متن استخراج شود. با این روش، در یادگیری، با استخراج ویژگی‌های بین متن، سوال و آنچه که در پاسخ به آن داده شده‌است، مدل یاد می‌گیرد که در صورت داشتن یک متن ورودی و پرسش سوال از آن، چه جوابی به عنوان خروجی بدهد. سجاد ایوبی برای داده‌های فارسی، چنین مدلی را استخراج کرده‌است و در قالب مدلی با نام PersianQA ارائه شده‌است [۱۶].

داده‌های PerisanQA مجموعه داده‌های درک خواندن است که در ویکی‌پدیا فارسی جمع‌آوری شده‌است. این داده‌ها دارای ۹۰ ورودی است که هر ورودی می‌تواند دارای پاسخی غیرممکن یا یک تا چند پاسخ متفاوت باشد که پاسخ‌دهنده به سوال، آن را پیشنهاد کرده‌است. در شکل ۸ نمونه‌ای از این داده‌ها نشان داده شده‌است.

حجم این مجموعه داده شامل حدود ۹۰۰۰ عنوان مختلف است که در هر کدام بین ۵ تا ۱۰ سوال مطرح شده‌است. در جدول ۱، مشخصات دیگری مثل تعداد عنوان‌های داده‌های تست، متوسط طول سوالات، پاسخ‌ها و متن

این مدل دارای ۶۴ کلاس مختلف است و می‌تواند به ۶۴ سوال مختلف پاسخ دهد. معماری مدل شامل دو لایه تراکمی<sup>۱</sup> و یک لایه حذف تصادفی<sup>۲</sup> است که با نرخ حذف ۰/۴ برای نورون‌های لایه تراکمی اول تنظیم شده است. با توجه به آزمایش‌های که بر روی برچسب‌هایی با یک سوال انجام شد، مشاهده شد که مدل به اصطلاح بیش‌برازش<sup>۳</sup> کرده‌است. در این حالت، مدل بر روی سوالات همگرا نمی‌شود بلکه برای تشخیص، نیاز صد درصدی به مشابهت سوال پیش‌بینی با سوال داده‌های آموزش را می‌طلبد. با این حساب، حتی اگر یک حرف مثل «؟» در پایان سوالی قرار داده شود، مدل نمی‌تواند به درستی تشخیص دهد.

در واقع، افزایش تعداد سوالات هر برچسب باعث می‌شود که مدل بتواند بر روی کلمات کلیدی هر جمله یاد بگیرد و بر اساس آن اگر در جمله جدیدی آن کلمات دیده شود، تصمیم درستی در تشخیص برچسب آن صورت دهد. جهت تشخیص تک سوالی، یک راه جابه‌جایی کلمات در یک جمله و ارائه آن به عنوان جمله‌های مشابه در یک برچسب بود اما نتایج نشان داد که جابه‌جا کردن کلمات هیچ تاثیری در آموزش بهتر ندارد.

- 1 Dense
- 2 Dropout
- 3 overfit

```
{
  "title": "شرکت فولاد مبارکه اصفهان",
  "paragraphs": [
    {
      "qas": [
        {
          "answers": [
            {
              "answer_start": 114,
              "answer_end": 131,
              "text": "در شرق شهر مبارکه"
            }
          ],
          "question": "شرکت فولاد مبارکه در کجا واقع شده است",
          "is_impossible": false,
          "id": 1
        }
      ],
      {
          "answers": [
            {
              "answer_start": 263,
              "answer_end": 264,
              "text": "۶"
            }
          ],
          "question": "فولاد مبارکه چند بار برنده جایزه شرکت دانشی را کسب کرده است؟",
          "is_impossible": false,
          "id": 2
        }
      ]
    }
  ]
}
```

شکل ۸. نمونه مجموعه داده PersianQA [۱۶]

Fig. 8. Sample PersianQA dataset

جدول ۱. مشخصات داده‌های PersianQA [۱۶]

Table 1. PersianQA data specification

تقسیم‌بندی	تعداد نمونه‌ها	تعداد بی‌پاسخ‌ها	متوسط طول سوال	متوسط طول پاراگراف	متوسط طول پاسخ
داده‌های آموزش	۹۰۰۰	۲۷۰۰	۸/۳۹	۲۲۴/۵۸	۹/۶۱
داده‌های تست	۹۳۸	۲۸۰	۸/۰۲	۲۲۰/۱۸	۵/۹۹

```

10 response = 'کاش ای کاش که دنیای عطر میفهمید. \
11             ' آب مهریه زهرا است بیا تا بیرویم. \
12             ' کربلا منتظر ماست بیا تا بیرویم '
13 while True:
14     question = input('You: ')
15     answerM = find(response, question)['answer']
16     print('ROB: ', answerM)

```

You: مهریه حضرت زهرا چیه؟  
 ROB: آب  
 You: کجا منتظر ما می باشد؟  
 ROB: کربلا  
 You: چه چیزی را ای کاش دنیای عطر میفهمید؟  
 ROB: آب مهریه زهرا

شکل ۹. تست مدل persianQA

Fig. 9. persianQA model test

به آن بدهد. اگرچه بیان احساسات به صورت دقیق باید از طریق پردازش صدا و تصویر در کنار جمله بیان شده قضاوت شود، اما در این پژوهش صرفاً به بررسی احساسات ناشی از کلمات در جمله پرداخته شده است.

تشخیص احساسات در این روش، توسط یک مدل با سری زمانی نشان داده شده است. برای این کار، کلمات موجود در یک جمله به صورت عدد گذاری می شوند و هر کدام از این کلمات موجود در جمله در زمان اولیه، کلمه اول را دریافت کرده و وارد یک لایه جاسازی<sup>۱</sup> می کند. استفاده از این لایه به این دلیل است که بردارهای یک بعدی تشکیل شده از کلمات فارسی بزرگ نشوند و ضرب های ماتریسی شبکه قابل آموزش توسط رم کامپیوتر باشد. خروجی از این لایه، یک بردار عددی است. این بردار با کد کلمه لینک می شود و وارد یک لایه حافظه طولانی کوتاه مدت می شود. حافظه طولانی کوتاه مدت دارای یک حالت پنهان و روی است که در ابتدای کار می تواند یک بردار تصادفی یا صفر باشد و با ورودی بردار عددی کلمه، یک حالت پنهان جدید تولید می کند. کلمه بعدی به همین شکل با کد مخصوص به خود وارد لایه جاسازی می شود و بردار عددی لینک شده وارد لایه حافظه

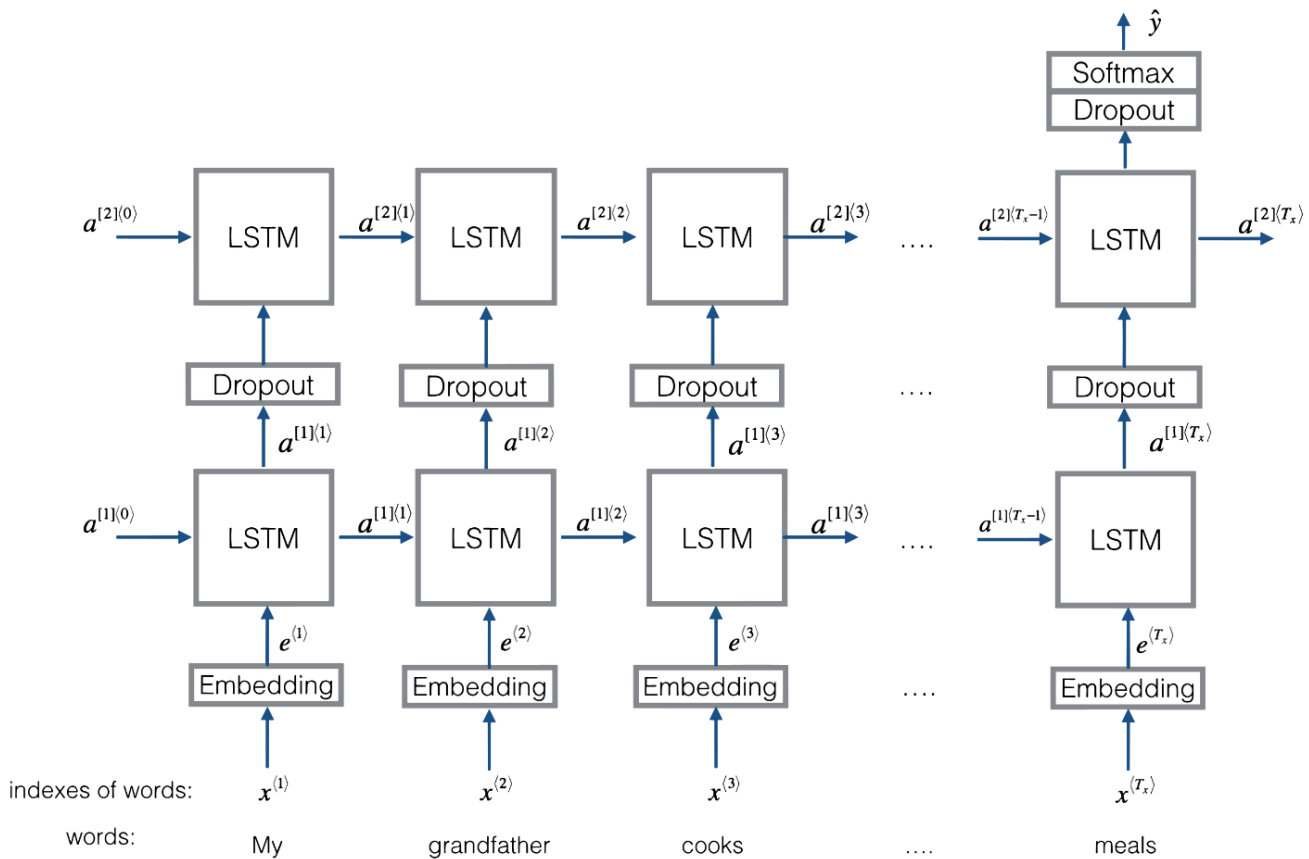
مربوط به آن و همچنین تعداد سوالات بدون جواب آورده شده است. یک عیب از این مجموعه داده آن است که تنها توانایی دریافت متنی با حداکثر ۵۰۰ کلمه را دارد و با افزایش بیشتر کلمات، مدل قادر به اجرای صحیح نمی باشد [۱۶]. راه استفاده از این مدل در مکانیزم پرسش و پاسخ ربات به این شرح است که تعدادی متن با طول های کوتاه آماده می شود و در حافظه ربات قرار می گیرد. هنگامی که سوالی توسط کاربر پرسیده می شود، ابتدا بررسی می شود که پاسخ مناسبی برای آن وجود داشته باشد. در صورتی که پاسخ مناسبی یافت نشود، سوال وارد مدل از پیش آموزش دیده Per-sianQA می شود و پاسخی برای آن استخراج می گردد و پخش می شود. در شکل ۹، یک نمونه از متن، سوال ورودی و پاسخی که مدل داده است، دیده می شود.

## ۲-۶- روش پیاده سازی احساسات (ارتقایافته سوال و پاسخ تصادفی)

برخی از کلمات دارای بار احساسی هستند و هنگامی که در یک جمله قرار می گیرند بر روی مفهوم جمله تاثیر می گذارند. در این موارد نیاز است تا هوش ربات بتواند این احساسات را از کلمات شناسایی و پاسخ مناسبی نسبت

1 Embedding





شکل ۱۰. معماری شبکه تشخیص احساسات متن [۱۷]

Fig. 10. Network architecture for text sentiment recognition

- عدد ۲ برای احساس ترس
- عدد ۳ برای احساس تنفر
- عدد ۴ برای احساس خشم
- عدد ۵ برای احساس تعجب

#### ۲-۷- مقایسه مکانیزم‌های گفت‌وگو

هر یک از روش‌های بیان شده برای ساخت مکانیزم گفت‌وگو در این بخش، دارای مزایا و معایبی است که جهت انتخاب مکانیزم بهینه برای روبو خیمه‌شب‌باز مورد نیاز است. در جدول ۲، مزایا و معایب روش‌های بیان شده نشان داده شده‌است. بر اساس این موارد، روش تشخیص سوال و پاسخ تصادفی برای مکانیزم گفت‌وگو روبو خیمه‌شب‌باز در نظر گرفته شده‌است.

طولانی کوتاه‌مدت می‌گردد با این تفاوت که حالت پنهان ورودی این لایه، از کلمه قبلی استخراج شده‌است. به منظور عدم حفظ داده‌ها توسط شبکه، یک لایه حذف تصادفی<sup>۱</sup> جهت خاموش کردن درصدی از داده‌های وارد شده به آن پس از هر حافظه طولانی کوتاه‌مدت قرار داده می‌شود خروجی این لایه، به یک لایه حافظه طولانی کوتاه‌مدت دیگر داده می‌شود. در شکل ۱۰، معماری شبکه مطرح شده نشان داده شده‌است.

جهت تست و بررسی این روش، تعدادی سوال با پاسخ ایجاد شد و متناسب با نوع سوال، بین اعداد ۰ تا ۵ احساسات تقسیم‌بندی شده‌است که هر عدد شرح احساس زیر را بر عهده دارد:

- عدد ۰ برای احساس خوشحالی
- عدد ۱ برای احساس ناراحتی

1 Dropout

جدول ۲. مقایسه روش‌های مکانیزم گفت‌وگو ربات خیمه‌شب‌باز

Table 2. Comparing puppeteer robot dialogue mechanism methods

روش	مزایا	معایب
ماشین‌های حالت محدود	<ul style="list-style-type: none"> <li>سادگی در اجرا</li> <li>دارای سلسله متوالی ساده</li> <li>سرعت در پاسخ به پرسش</li> </ul>	<ul style="list-style-type: none"> <li>ساخت شبکه سخت</li> <li>گفت‌وگو ضعیف</li> <li>تنوع پایین در اجرا</li> </ul>
همسایگی نیوتن	<ul style="list-style-type: none"> <li>تشخیص سریع</li> <li>دقت در تشخیص با وجود کلمات ناهمگون</li> </ul>	<ul style="list-style-type: none"> <li>تشخیص ضعیف در کلمات مشابه</li> <li>تشخیص ضعیف در معنای واقعی جمله</li> </ul>
متن سریع	<ul style="list-style-type: none"> <li>تشخیص کلمات مشابه در جمله</li> <li>تشخیص نزدیکی کلمات به یکدیگر</li> </ul>	<ul style="list-style-type: none"> <li>نامناسب در گفت‌وگو مستقیم</li> <li>عدم ایجاد پاسخ در گفت‌وگو</li> </ul>
تشخیص سوال و پاسخ تصادفی	<ul style="list-style-type: none"> <li>تشخیص سوال با وجود کلمات نامشابه</li> <li>ارائه پاسخ‌های مختلف برای یک سوال</li> </ul>	<ul style="list-style-type: none"> <li>نیاز به تنوع‌سازی سوالات برای تشخیص</li> <li>پاسخ‌های تصادفی بدون احساسات</li> </ul>
تشخیص سوال از متن	<ul style="list-style-type: none"> <li>تشخیص سوال از متن آماده</li> <li>ارائه پاسخ با دقت خوب از متن</li> </ul>	<ul style="list-style-type: none"> <li>تشخیص سوال از متن حداکثر ۵۰۰ کلمه</li> <li>نیازمندی به متن جهت پرسش و پاسخ</li> </ul>
تشخیص احساسات در پاسخ	<ul style="list-style-type: none"> <li>تشخیص توالی کلمات جمله</li> <li>ارائه پاسخ مناسب با احساسات پرسش</li> </ul>	<ul style="list-style-type: none"> <li>نیاز به مجموعه داده با تحلیل احساسات</li> <li>عدم توجه به جملات دیگر گفت‌وگو</li> </ul>

### ۳- نتایج و بحث

زمان و تسریع در همگرایی الگوریتم هستند و به بهینه‌سازی عملکرد کمک می‌کنند. تابع هزینه در این مدل به صورت آنتروپی متقاطع است، که برای مسائل دسته‌بندی مناسب است. هدف اصلی در اینجا کاهش هزینه است تا پیش‌بینی‌های مدل به‌طور مکمل با برچسب‌های واقعی تطابق پیدا کنند. به منظور بررسی پارامترهای قابل تغییر در معماری شبکه و رسیدن به دقت بهینه، ۱۰ آزمایش با تعداد لایه‌ها و نرون‌های مختلف، در ۱۰۰ دوره آموزشی با ۶۴ دسته سوال انجام شد. در شکل ۱۱، نمودار خطا و در شکل ۱۱ نمودار دقت بر حسب افزایش تعداد لایه‌ها نشان داده شده است. مدت زمان رسیدن به نتایج هر دوره آموزشی بین ۸۰ تا ۱۲۰۰ میلی‌ثانیه متغیر بوده است.

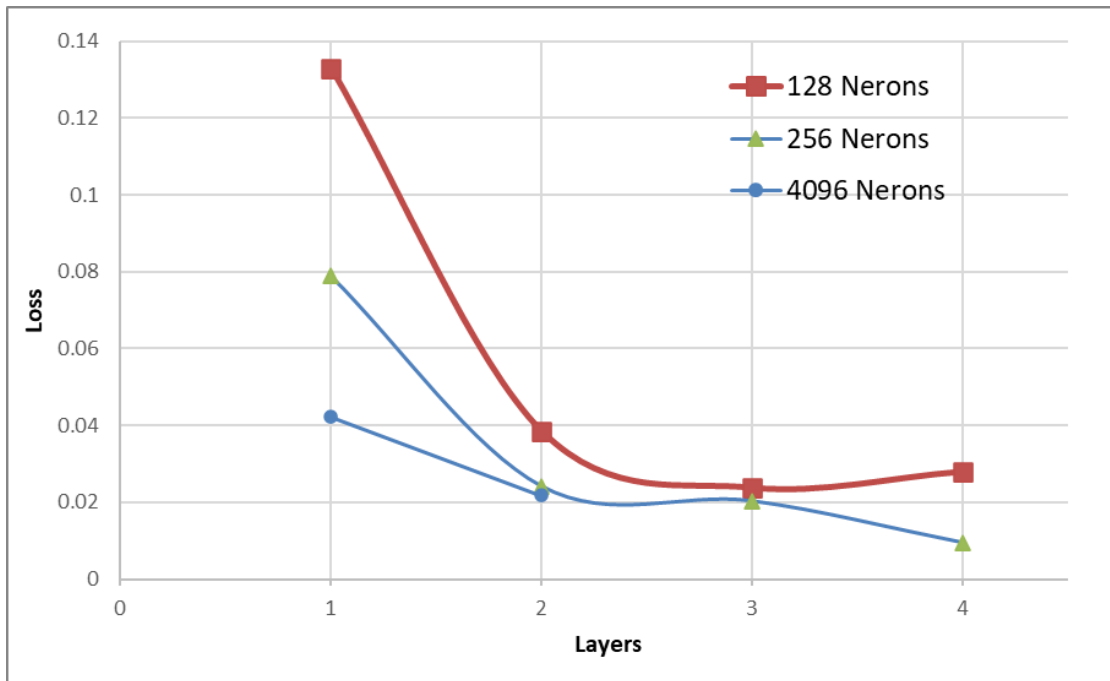
همانطور که در شکل ۱۱ مشخص است، افزایش تعداد لایه‌ها به صورت کلی باعث کاهش خطای یادگیری می‌شود. همچنین، با افزایش نرون‌های آموزشی، نرخ کاهش خطا با شدت بیشتری انجام می‌شود. از شکل ۱۱ این نتیجه بیان می‌شود که افزایش تعداد لایه‌ها در نرون‌های پایین، لزوماً باعث کاهش خطا نمی‌شود.

در شکل ۱۲، دقت شبکه یادگیری نشان داده شده است. مطابق این شکل، افزایش تعداد لایه‌ها در تعداد نرون‌های بالای ۱۲۸، تاثیری در دقت

با توجه به روش‌های بررسی‌شده در این مقاله، در نهایت، روش سوال و پاسخ تصادفی به عنوان روش نهایی در پرسش و پاسخ ربات خیمه‌شب‌باز استفاده شده است. با فرارگیری این مکانیزم در ربات، می‌توان امکان گفت‌وگوی هوشمند در آن را مهیا کرد. در معماری شبکه عصبی از لایه‌های کاملاً متصل<sup>۱</sup> یا فشرده<sup>۲</sup>، یک لایه حذف تصادفی جهت جلوگیری از بیش‌برازش داده‌ها استفاده شده است.

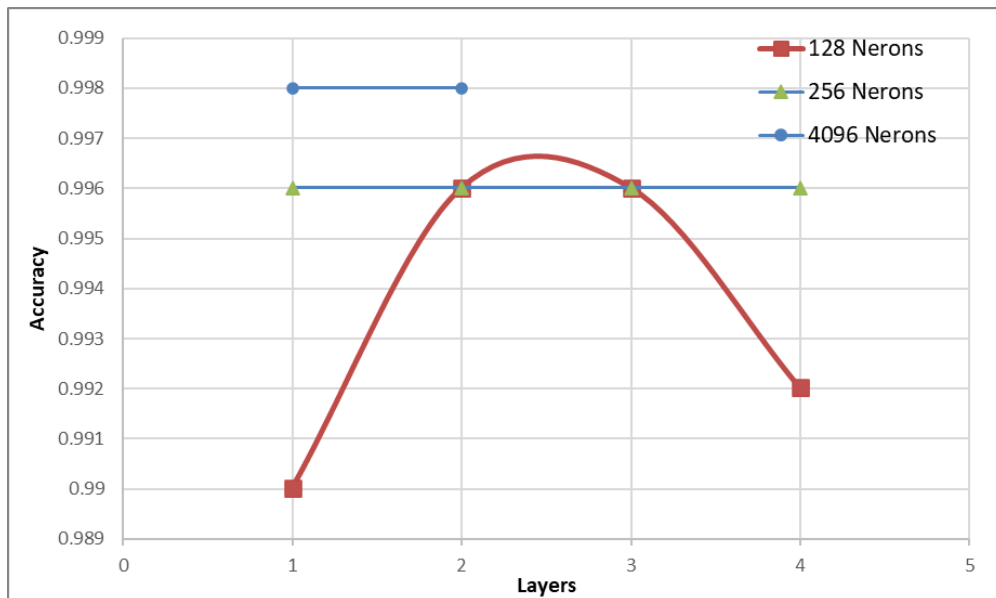
در این مدل، از الگوریتم نزول گرادینت تصادفی<sup>۳</sup> به عنوان بهینه‌ساز استفاده شده است. هدف این الگوریتم کمینه‌کردن یک تابع هزینه به‌وسیله‌ی گرفتن گام‌های کوچک در جهت منفی گرادینت تابع است. تابع هزینه، میزان عدم تطابق بین برچسب‌های پیش‌بینی‌شده و برچسب‌های واقعی را اندازه‌گیری می‌کند. سه فرآیند کلیدی برای این بهینه‌ساز وجود دارد: نرخ یادگیری ۰/۰۱، واپاشی نرخ یادگیری ۰/۰۰۰۰۰۱ و تکانه ۰/۹، که هر کدام به ترتیب مسئول تعیین اندازه گام، کاهش تدریجی نرخ یادگیری با گذر

- 1 Fully Connected
- 2 Dense
- 3 Stochastic Gradient Descent (SGD)



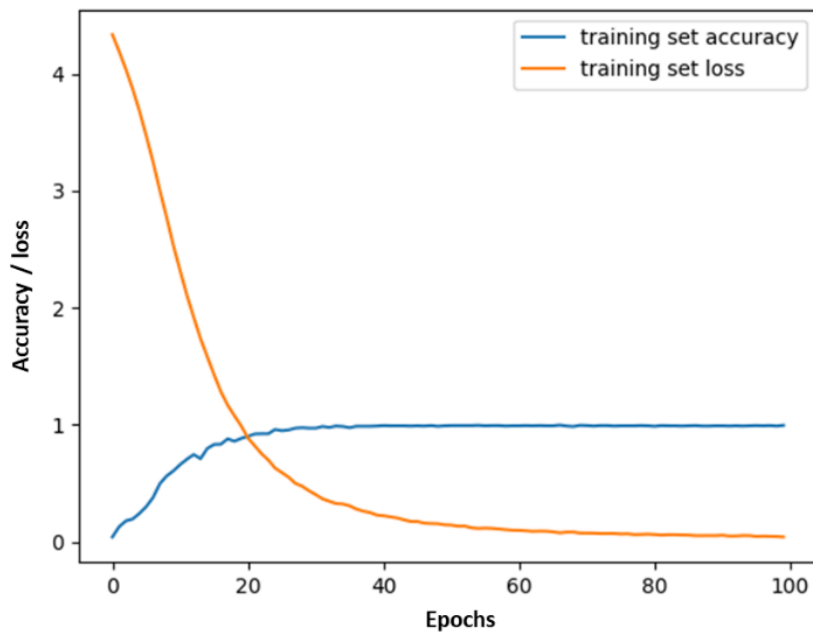
شکل ۱۱. نمودار خطا برحسب تعداد لایه در سه مقدار ۱۲۸، ۲۵۶ و ۴۰۹۶ نرون شبکه فشرده

Fig. 11. Error diagram according to the number of layers in three values of 128, 256 and 4096 Dense network neurons



شکل ۱۲. نمودار دقت برحسب تعداد لایه در سه مقدار ۱۲۸، ۲۵۶ و ۴۰۹۶ نرون شبکه فشرده

Fig. 12. Accuracy graph according to the number of layers in three values of 128, 256 and 4096 Dense network neurons



شکل ۱۳. دقت و خطا در ۱۰۰ دوره آموزشی با ۴۰۹۶ نرون آموزشی از شبکه فشرده

Fig. 13. Accuracy and error in 100 training courses with 4096 training neurons from Dense network

سطح گفتار فارسی می‌گردد.

با توجه به آزمایش‌های انجام‌شده، مکانیزم گفتگوی پرسش و پاسخ برای ۶۴ کلاس مختلف در حالت صوتی محاسبه و تنظیم شده است. در این مکانیزم، تلاش شده تا تعداد پرسش‌های بدون پاسخ به حداقل برسد. این استراتژی از آن جهت که حتی در صورت عدم قطعیت در درستی پاسخ، ربات همچنان پاسخی ارائه می‌دهد، طراحی شده است. دلیل این رویکرد این است که یکی از فاکتورهای مهم در افزایش جذابیت ربات‌های گفتگویی، قابلیت پاسخگویی به سوالات، حتی با احتمال بروز خطاست. این امر می‌تواند تجربه کاربری را بهبود بخشد و تعامل بین کاربر و ربات را افزایش دهد.

با بررسی دقیق داده‌های ارائه شده در جدول ۳، یک سری از نتایج قابل توجه در رابطه با عملکرد مکانیزم گفت‌وگو پرسش و پاسخ تصادفی به چشم می‌خورد. از لحاظ صحت، مدل مذکور با نرخ ۰/۸۵٪ عملکردی برتر از سایر مدل‌های ارائه‌شده داشته‌است. این امر، تصدیق‌کننده‌ی توانایی مدل در تولید پاسخ‌های صحیح با کاهش حداقلی خطا می‌باشد. از جنبه بازیابی، مدل روبو خیمه‌شب‌باز با نرخ ۰/۶۵٪، مجدداً بالاترین عملکرد را در میان تمام مدل‌های موجود ارائه داده است. بازیابی با این میزان، شاهد بر توانایی مدل در شناسایی و پاسخگویی به دسته‌ای گسترده از پرسش‌ها است، حتی اگر

آموزشی ندارد. این می‌تواند به خاطر تعداد کم کلاس‌های پرسش و پاسخ باشد. همچنین، با افزایش تعداد نرون‌های آموزشی، دقت افزایش می‌یابد. در معماری شبکه با تعداد نرون پایین، افزایش لایه‌ها لزوماً باعث افزایش دقت نمی‌شود. مطابق با نتایج به دست‌آمده از مدل شکل ۷، معماری شبکه پرسش و پاسخ تصادفی، با تعداد نرون ۴۰۹۶ و یک لایه آموزشی فشرده، به خطای ۰/۴۲ و دقت ۰/۹۹۸ رسیده‌است. در شکل ۱۳، نمودار خطا و دقت آموزش در ۱۰۰ دور آموزشی نشان داده شده‌است.

برای پیشبرد تحلیل موثرتر و دقیق‌تر این مدل، از معیارهای ارزیابی موثر و معتبری نظیر صحت<sup>۱</sup>، بازیابی<sup>۲</sup> و معیار اندازه‌گیری F1<sup>۳</sup> استفاده می‌شود. در جدول ۳، ارزیابی مکانیزم گفت‌وگو پرسش و پاسخ تصادفی با سایر مدل‌های موجود در این ادبیات مقایسه شده‌است. بررسی و تحلیل‌ها نشان می‌دهد که اهمیت برجسته مکانیزم پرسش و پاسخ تصادفی در روبو خیمه‌شب‌باز، بر خلاف دیگر مدل‌های موجود که بر روی زبان انگلیسی تمرکز دارند، در پشتیبانی از زبان فارسی است. این توجه اختصاصی به زبان فارسی موجب ایجاد قابلیت پیشرفته‌ای برای پردازش و درک معناشناختی در

- 
- 1 Precision
  - 2 Recall
  - 3 F1-measurement

جدول ۳. مقایسه معیارهای ارزیابی مکانیزم گفت‌وگو پرسش و پاسخ تصادفی

Table 3. Comparing evaluation criteria of random question and answer dialogue mechanism

معیار اندازه‌گیری F1	بازیابی	صحت	مدل‌ها / معیار ارزیابی
۵۵٪/۹	٪۸۰/۴۱	۵۱٪/۵۰	مرجع [۱۸]
٪۷۰/۷۵	٪۷۲/۵۰	٪۶۲/۶۲	مرجع [۱۹] بدون روش T5
٪۷۷/۸۷	٪۷۷/۷۵	٪۷۷/۸۷	مرجع [۱۹] با روش T5
٪۷۴	٪۷۱	٪۷۸	مرجع [۲۰] روش چند جمله ای ساده بیز
٪۷۴	٪۷۵	٪۷۳	مرجع [۲۰] روش دسته‌بندی ساده بیز
٪۸۲	٪۸۲	٪۸۳	مرجع [۲۰] روش رگرسیون لجستیک
٪۸۰	٪۸۰	٪۸۰	مرجع [۲۰] روش ماشین بردار پشتیبانی خطی
٪۸۳	٪۸۳	٪۸۳	مرجع [۲۰] روش مدل هیبریدی
٪۸۷/۷۳	٪۹۰/۶۵	٪۸۵	مدل روبرو خیمه‌شب‌باز

شده‌است.

با در نظر گرفتن دوره‌های آموزشی ۵۰ و اندازه دسته ۳۲ مدل به صورت شکل ۱۶ آموزش دیده‌است.

بر طبق شکل ۱۷، خط نارنجی مربوط به تابع هزینه<sup>۱</sup> و خط آبی مربوط به دقت داده‌ها در هر دوره آموزشی است. ماکزیمم دقت در این آموزش مقدار ۱/۰۰ و مینیمم خطا ۰/۰۰۴۶ می‌باشد. سرعت آموزش با مشخصات ذکر شده نیز ۷۷ ثانیه بوده‌است.

در این مدل، یک جمله به عنوان ورودی وارد می‌شود و خروجی عددی بین ۰ تا ۵ می‌دهد که بیانگر احساسی است که از این جمله استخراج شده‌است. در شکل ۱۷ یک نمونه چالشی برای مدل در نظر گرفته شده‌است. همانطور که در شکل ۱۷ مشخص است، خروجی عدد ۴ بیان شده که مربوط به احساس خشم است. البته انتظار نمی‌رود که این مدل آموزشی عملکرد عالی داشته باشد زیرا تنها برای ۱۰ دسته از سوالات آموزش دیده‌است. بنابراین، اگر داده‌های آموزشی افزایش یابد، تشخیص احساسات دقیق‌تر خواهند شد. حال می‌توان با در نظر گرفتن پاسخ سوال از روش‌های دیگر و دسته‌بندی آنها بر اساس احساسات، متناسب با درکی که از سوال داده می‌شود، پاسخ مناسب بیان شود.

برخی از این پاسخ‌ها شامل خطا باشند.

در نهایت، از نظر معیار معیار اندازه‌گیری F1، مدل روبرو خیمه‌شب‌باز ارزش ۸۷/۷۳٪، به دست آمده‌است. این معیار، نمایانگر میانگین هارمونیک صحت و بازیابی است که نشان‌دهنده تعادل میان این دو معیار مهم می‌باشد. بر اساس این بررسی، می‌توان ادعا کرد که مدل روبرو خیمه‌شب‌باز به خوبی عملکرد خود را نشان داده است، با داشتن دقت بالا و همچنین توانایی در شناسایی و پاسخگویی به تعداد قابل ملاحظه‌ای از پرسش‌ها. همچنین، معیار معیار اندازه‌گیری F1 این نکته را تأیید می‌کند که مدل موفق بوده است و توانسته‌است تعادلی مناسب بین صحت و بازیابی برقرار کند که این خود یکی از معیارهای مهم در ارزیابی عملکرد مدل‌ها می‌باشد.

روش پیاده‌سازی احساسات نیز یک مدل ارتقایافته از روش پرسش و پاسخ تصادفی است اما به دلیل مجموعه داده‌سازی سنگین آن با توجه به تعداد کلاس‌های روش پرسش و پاسخ تصادفی، تنها بر روی ۱۰ دسته از آن مجموعه داده آماده شده که در شکل ۱۴، نمونه‌ای از این داده‌های احساسات آورده شده‌است.

جهت استفاده از بردار جاسازی از داده‌های متن سریع که به صورت بدون ناظر برای چند میلیون کلمه فارسی آموزش داده شده‌است استفاده. معماری شبکه مدل بر اساس آنچه در توضیحات اشاره شد در شکل ۱۵ نشان داده

1 Loss function



```
"tag": "حالیبرس",  
"input": [  
  ["چطوری", 0],  
  ["چه طوری", 0],  
  ["حالت چه طوره", 0],  
  ["میگم چه طوری", 4],  
  ["خوبی", 0],  
  ["حالت خوبه", 0],  
  ["سلامتی", 0],  
  ["چه خبرا", 0],  
  ["سلام چه خبر", 0],  
  ["سلام چطوری", 0],  
  ["از این ورا", 3],  
  ["به به خوبی", 5],  
  ["حال شما", 0],  
  ["احوال چطوره", 0],  
  ["خوبی یا نه", 3],  
  ["نگفتی چطوری", 3],  
  ["میگم بگو چطوری", 4],  
  ["احوال را می پرسم احمق", 4],  
  ["این احوال بدت چطوره", 4],  
  ["هوی خوبی یا نه", 4],  
  ["بابو جواب بده چطور هستی نفهم", 4],  
  ["مریض خنگ چطوری", 4],  
  ["چطوری عزیزم", 0]
```

شکل ۱۴. مجموعه داده تست احساسات متن

Fig. 14. Text Sentiment Test Dataset

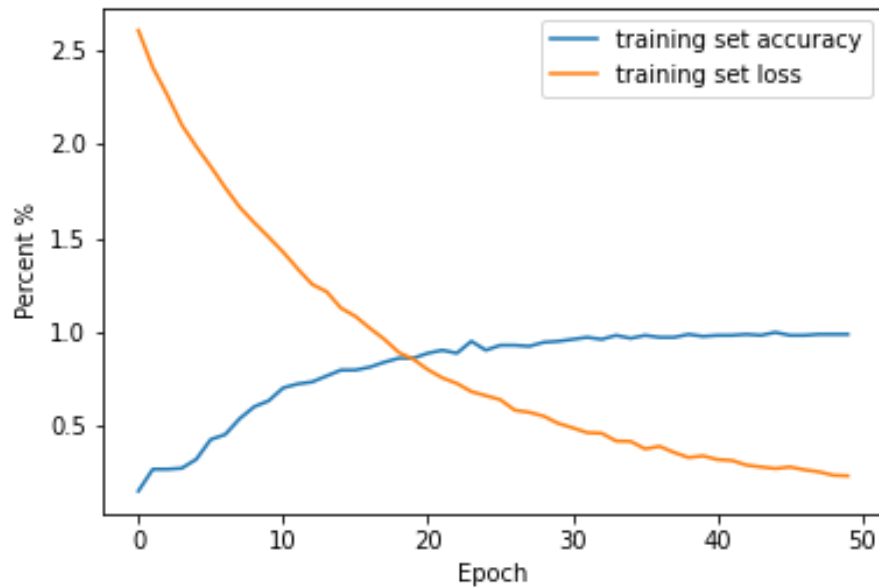
Model: "model"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 8)]	0
embedding (Embedding)	(None, 8, 300)	90000000
lstm (LSTM)	(None, 8, 128)	219648
dropout (Dropout)	(None, 8, 128)	0
lstm_1 (LSTM)	(None, 128)	131584
dropout_1 (Dropout)	(None, 128)	0
dense (Dense)	(None, 6)	774

=====  
 Total params: 90,352,006  
 Trainable params: 352,006  
 Non-trainable params: 90,000,000  
 =====

شکل ۱۵. معماری شبکه تحلیل احساسات متن

Fig. 15. Network architecture of text sentiment analysis



شکل ۱۶. نمودار آموزش مدل احساسات متنی

Fig. 16. Diagram of textual emotion model training

```
x_test = np.array(['دیوانه خر سلام گلم عزیزم بیخشد'])
X_test_indices = sentences_to_indices(x_test, word_to_index, maxLen)
print(np.argmax(model.predict(X_test_indices)))
```

```
w0: 0
w0: 1
w0: 2
w0: 3
w0: 4
w0: 5
1/1 [=====] - 0s 47ms/step
4
```

شکل ۱۷. تست خروجی مدل احساسات متنی

Fig. 17. Test the output of the text sentiment model

#### ۴- نتیجه گیری

در این مقاله، مکانیزم روش‌های مکانیزم گفت‌وگو برای یک ربات خیمه‌شب‌باز طراحی و پیاده‌سازی شد. از بین این مکانیزم‌ها، روش تشخیص سوال با پاسخ تصادفی و در نظر گرفتن احساسات بررسی گردید. کلیه پرسش و پاسخ‌ها به زبان فارسی اجرا شده‌اند و کاربر می‌تواند به وسیله آن با ربات خیمه‌شب‌باز ارتباط برقرار کند. پیاده‌سازی روش تشخیص سوال با پاسخ تصادفی با ۶۴ کلاس مختلف از پرسش و پاسخ‌های فارسی و ایجاد یک شبکه عصبی عمیق بر روی آن، به دقت ۰/۹۹۸ و خطای ۰/۰۴۲ رسیده‌است. بنابراین، با تشخیص سوال کاربر، کلاس مدنظر توسط مدل یادگیری عمیق انتخاب می‌شود و پاسخ تصادفی از مجموعه پاسخ‌های در نظر گرفته شده برای آن کلاس به عنوان خروجی نشان داده می‌شود.

#### منابع

- networks for visual and textual question answering, in: International conference on machine learning, PMLR, 2016, pp. 2397-2406.
- [8] T. Minato, K. Sakai, T. Uchida, H. Ishiguro, A study of interactive robot architecture through the practical implementation of conversational android, *Frontiers in Robotics and AI*, 9, (2022)
- [9] A. Conneau, A. Baevski, R. Collobert, A. Mohamed, M. Auli, Unsupervised cross-lingual representation learning for speech recognition, *arXiv preprint arXiv:2006.13979*, (2020)
- [10] S.-H. Lee, Y.-E. Lee, S.-W. Lee, Toward imagined speech based smart communication system: potential applications on metaverse conditions, in: 2022 10th International Winter Conference on Brain-Computer Interface (BCI), IEEE, 2022, pp. 1-4.
- [11] P. Chang, S. Liu, K.D. Campbell, Robot Sound Interpretation: Learning Visual-Audio Representations for Voice-Controlled Robots, *CoRR*, (2021)
- [12] S. Hamed, puppeteer robot: Designing a construction of a new interactive game based on imitative learning of a humanoid robot from a human, in: The first national conference of computer games; Opportunities and challenges, 1394 (in Persian).
- [13] fastText, fastText: Library for efficient text classification and representation learning, in, Facebook.
- [14] H. Hemati, fastText Word Embedding: The Persian Approach to Text Embedding, (2018)
- [15] H. Hemati, fastText-Persian, in, Github, 2019.
- [16] S. Ayobi, PersianQA, in, Github, 2021.
- [17] A. Akhavan, Text Classification Emojify, in, Github, 1399.
- [18] J.C. Vasquez-Correa, J.C. Guerrero-Sierra, J.L. Pemberty-Tamayo, J.E. Jaramillo, A.F. Tejada-Castro, One system to rule them all: A universal intent recognition system for customer service chatbots, *arXiv preprint arXiv:2112.08261*, (2021)
- [1] [1] A. Chowanda, A.D. Chowanda, Recurrent neural network to deep learn conversation in Indonesian, *Procedia computer science*, 116 (2017) 579-586.
- [2] K. Karpagam, K. Madusudanan, A. Saradha, Deep learning approaches for answer selection in question answering system for conversation agents, *ICTACT Journal on Soft Computing*, 10(2) (2020) 2040-2044.
- [3] R. Yan, Y. Song, H. Wu, Learning to respond with deep neural networks for retrieval-based human-computer conversation system, in: *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, 2016, pp. 55-64.
- [4] A. Engineered, Ameca The future face of robotics, in, *Engineered Arts*, (2022).
- [5] Y. Sharma, S. Gupta, Deep learning approaches for question answering system, *Procedia computer science*, 132 (2018) 785-794.
- [6] M. Tan, C. Dos Santos, B. Xiang, B. Zhou, Improved representation learning for question answer matching, in: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2016, pp. 464-473.
- [7] C. Xiong, S. Merity, R. Socher, Dynamic memory

- Humanized Computing, 14(4) (2023) 3129-3144.
- [20] E.H. Almansor, F.K. Hussain, O.K. Hussain, Supervised ensemble sentiment-based framework to measure chatbot quality of services, Computing, 103 (2021) 491-507.
- [19] J.J. Bird, A. Ekárt, D.R. Faria, Chatbot Interaction with Artificial Intelligence: human data augmentation with T5 and language transformer ensemble for text classification, Journal of Ambient Intelligence and

چگونه به این مقاله ارجاع دهیم

M. Amirkhani, H. Shahbazi, Designing and building a dialogue mechanism suitable for RoboPuppet with using deep inference learning, Amirkabir J. Mech Eng., 55(8) (2023) 919-942.

DOI: 10.22060/mej.2023.22001.7553





