****

**دانشگاه شهید بهشتی**

**دانشکده مهندسی و علوم کامپیوتر**

**اطلاعیه دفاع**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **نام استاد راهنما:دکتر یاسر شکفته** | | | **نام دانشجو:علی یزدانی** | |
| **گرایش:هوش مصنوعی، رباتیک و رایانش شناختی** | **رشته: مهندسی کامپیوتر** | | | **مقطع:کارشناسی ارشد** |
| **تاریخ:30/07/1401** | | | **نوع دفاع:**   * **دفاع پروپوزال□** * **دفاع پایان نامه ◼** * **دفاع رساله دکترا □** | |
| **ساعت: 8:30** | | |
| **مکان: دانشکده علوم و مهندسی کامپیوتر** | | |
| **عنوان: شناسایی احساس در گفتار فارسی با استفاده از ویژگی­های صوتی و زبانی** | | | | |
| **داوران داخلی: دکتر آرمین سلیمی بدر** | | **داوران خارجی:دکتر بابک ناصر شریف** | | |
| **چکیده:**  شناسایی احساسات در گفتار (SER) یکی از مهمترین روش­های ادراکی انسان در درک موقعیت و نحوه تعامل با دیگران است، از این رو در سال­های اخیر سعی شده است تا با افزودن قابلیت شناسایی احساسات به سیستم­های ارتباطی انسان-ماشین، امکان افزایش بازدهی این سیستم­ها در برقراری ارتباط مؤثر با انسان­ها را فراهم نمود. تنها از طریق این ظرفیت، می­توان یک مکالمه معنادار و اصولی بر پایه اعتماد و درک متقابل انسان– ماشین انجام داد. شناسایی احساسات در گفتار به سه دلیل اصلی یک امر چالش برانگیز است: 1)احساس انسانی انتزاعی است و به این معنی است که تشخیص دادن آن دشوار است؛ 2)به طور کلی احساسات انسانی تنها در برخی لحظات خاص در طول یک کلام طولانی قابل تشخیص است؛ 3)داده‌های گفتاری با برچسب احساسی معمولاً محدود است.  با وجود اینکه در سال­های اخیر پیشرفت قابل ملاحظه­ای در شناسایی احساس با استفاده از شبکه­های عصبی عمیق صورت گرفته، اما در زبان فارسی، هنوز هم جای پیشرفت وجود دارد. یکی از راهکارهای بهبود عملکرد سیستم­های شناسایی احساس در گفتار، ترکیب اطلاعات صوتی و متنی است. تمرکز ما در این پایان نامه، بر استفاده از ویژگی­های متنی است، به طوری که بتوان آن­ها را به صورت خودکار از دادگان صوتی استخراج نمود و تلاش شد تا این اطلاعات بتوانند کارایی­ای در حد اطلاعات دقیق متنی فایل صوتی باشد. آزمایشات مختلفی با استفاده از رویکردهای عمیق بر روی مجموعه داده گفتار احساسی فارسی شریف (ShEMO) انجام شده است. نتایج نشان می­دهد استفاده از ویژگی­های متنی به همراه ویژگی­های صوتی، می­تواند عملکرد مدل را نسبت به استفاده تنها از ویژگی­های صوتی یا متنی، به طور قابل توجهی بهبود ببخشد. همچنین با بکارگیری مناسب یک سیستم تشخیص گفتار خودکار و بررسی اثر مدل­های زبانی به دست آمده از پیکره­های متنی فارسی قابل دسترس، برخی اشکالات در مجموعه داده مورد نظر نمایان شد که در این پایان نامه به اصلاح آن­ها نیز پرداخته شده است. در نهایت با ترکیب ویژگی­های صوتی و اطلاعات مربوط به سیستم تشخیص گفتار مانند اطلاعات خروجی دیکدر، اطلاعات مرز کلمات و بکارگیری ویژگی­های پروزودیک کلمات، 72.14% دقت بدون وزن بدست آمد که در مقایسه با نتایج گزارش شده در مقاله پایه مربوط به این مجموعه داده، 13.94% بهبود حاصل شده است.  **واژگان کلیدی:** شناسایی احساسات در گفتار، پردازش گفتار، بازشناسی گفتار، جاسازی کلمات، ویژگی­های پروزودیک، یادگیری عمیق | | | | |