****

**دانشگاه شهید بهشتی**

**دانشکده مهندسی و علوم کامپیوتر**

**اطلاعیه دفاع**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **نام استاد راهنما: دکتر یاسر شکفته** | | | **نام دانشجو: سوگل علیپور اسگندانی** | |
| **رشته: مهندسی کامپیوتر** | **گرایش: هوش مصنوعی** | | | **مقطع: کارشناسی ارشد** |
| **تاریخ: 29/6/1401** | | | **نوع دفاع:**   * **دفاع پروپوزال □** * **دفاع پایان نامه** * **دفاع رساله دکترا □** | |
| **ساعت: 16 الی 18** | | |
| **مکان: کلاس 117** | | |
| **عنوان : ارائه سیستم پیشنهاد صداپیشه انگلیسی به فارسی با استفاده از ویژگی های صوتی** | | | | |
| **داوران داخلی: دکتر آرمین سلیمی بدر** | | **داوران خارجی: دکتر حسین زینلی** | | |
| **چکیده:**  هم زمان با گسترش صنعت رسانه ای و افزایش تولید محصولات مالتی مدیا (مانند فیلم، سریال، انیمیشن و بازی کامپیوتری)، تقاضای آن نیز، به سرعت در حال رشد است. تفاوت زبانی، یکی از چالش های پیش رو در عرضه محصولات مالتی مدیا از یک کشور مبدا در یک کشور مقصد است، لذا بومی سازی این محصولات می تواند در این زمینه کمک شایانی به عمل آورد. یکی از روش های متداول بومی سازی، دوبلاژ است و قطعا یکی از مهمترین چالش ها در دوبلاژ نیز، انتخاب گوینده است. این انتخاب معمولا توسط مدیر دوبلاژ انجام می شود ولی با افزایش نیاز به بومی سازی محصولات، وجود یک سیستم خودکار انتخاب صداپیشه، می تواند در تسریع فرآیند دوبلاژ نقش مهمی را ایفا کند. یک سیستم خودکار انتخاب صداپیشه، با استفاده از شاخصه های فردی گویندگان، بهترین صداپیشه در زبان مقصد را برای یک گوینده از زبان مبدا انتخاب می کند.  در این پژوهش، با توجه به فقدان دادگان مناسب برای دوبلاژ از زبان انگلیسی به زبان فارسی، دادگان E2PCast با مجموعا 5 ساعت داده گفتاری و 60 گوینده (30 زن و 30 مرد) در زبان های انگلیسی و فارسی، جمع آوری گردیده است. از آنجایی که نگاشت فضای زبانی-هویتی گویندگان در فضای مبدا و مقصد پیچیدگی بالایی دارد، سیستم انتخاب خودکار صداپیشه با استفاده از نگاشت غیرخطی مبتنی بر رمزگذار-رمزگشا طراحی شده است.  سیستم پایه انتخاب خودکار صداپیشه در حالتی که مبتنی بر شبکه سیامی با ورودی x-vector است، بهترین عملکرد را از خود نشان داده است. دقت آزمون 4-best character این سیستم برابر با 25/99 درصد و دقت آزمون شبکه سیامی در آن، 32/96 درصد است. بهترین عملکرد، توسط سیستم پیشنهادی با استفاده از ویژگی های آوایی و با ورودی x-vector به دست آمد. در این سیستم، دقت آزمون شبکه سیامی برابر 98/98 درصد و دقت آزمون 4-best character آن، برابر 100 درصد است.  گفتار صداپیشگان از نظر بیانی و احساسی با گفتار مردم عادی متفاوت است و این شاخصه ها می تواند در بردار های ویژگی صوتی نمایان شود. لذا بر این مبنا، سیستم تشخیص صداپیشه در زبان فارسی با استفاده از ویژگی های صوتی را طراحی نمودیم. دقت آزمون این سیستم، در آزمایش 5-fold برای ورودی های غیر نویزی با بازنمایی x-vector و i-vector و طبقه بند SVM با کرنل RBF برابر با 100 درصد است. | | | | |