

درخواست پیشنهاد (RFP) برای پروژه‌های تشخیص گوینده

(Speaker Recognition) در حوزه هوش مصنوعی

کد مسئله : SI-17

۱. مقدمه

پروژه تشخیص گوینده به منظور بهبود امنیت و تجربه کاربران در ارتباط با سیستم‌ها و خدمات ما طراحی می‌شود. این سیستم‌ها به تشخیص و شناسایی دقیق گویندگان کمک می‌کنند و باعث می‌شوند که تعامل با کاربران به صورت شخصی‌سازی شده و با دقت بیشتری صورت گیرد.

۲. اهداف پروژه

هدف اصلی این پروژه توسعه یک سیستم تشخیص گوینده است که بتواند به صورت دقیق و با کارایی بالا، گویندگان مختلف را از طریق تحلیل صدای آن‌ها شناسایی کند. اهداف فرعی عبارتند از:

a. **تشخیص گوینده خودکار:** توسعه سیستمی که بتواند گویندگان را بر اساس تحلیل ویژگی‌های صوتی شناسایی کند.

b. **درک زبان طبیعی:** سیستم باید قادر به درک به زبان طبیعی (به صورت نوشتاری و گفتاری) باشد.

c. **پشتیبانی از چندین زبان:** سیستم باید توانایی تشخیص گویندگان در زبان‌های مختلف را داشته باشد.

d. **یکپارچگی با سیستم‌های موجود:** سیستم باید قابلیت یکپارچگی با سایر سیستم‌ها و ابزارهای داخلی شرکت را داشته باشد.

e. **قابلیت یادگیری و بهبود:** سیستم باید بتواند از طریق بازخورد کاربران و داده‌های جدید به صورت مستمر بهبود یابد.

۳. دامنه کار

پروژه شامل مراحل زیر خواهد بود:

- a. **تحلیل نیازها:** تحلیل نیازهای شرکت و کاربران نهایی، شناسایی موارد استفاده و تعیین مشخصات سیستم.
- b. **طراحی معماری سیستم:** طراحی معماری سیستم تشخیص گوینده، شامل انتخاب مدل‌های یادگیری، پردازش صوت و تعیین منابع داده.
- c. **توسعه و پیاده‌سازی:** توسعه الگوریتم‌ها و مدل‌های تشخیص گوینده با استفاده از تکنیک‌های پیشرفته هوش مصنوعی، پیاده‌سازی ماژول‌های پردازش صوتی، شامل استخراج ویژگی‌های صوتی مانند ویژگی‌های (MFCC) و (prosody) مدل‌سازی احساسات.
- d. **آزمون و ارزیابی:** ارزیابی عملکرد سیستم از نظر دقت، سرعت تشخیص و رضایت کاربران.
- e. **آموزش و مستندسازی:** آموزش تیم داخلی برای استفاده و مدیریت سیستم و ارائه مستندات فنی و کاربری.
- f. **پشتیبانی و بهبود مستمر:** ارائه خدمات پشتیبانی و بهبود مستمر پس از تحویل سیستم.

۴. الزامات فنی

- a. **الگوریتم‌ها و مدل‌ها:** استفاده از مدل‌های یادگیری عمیق مانند CNN، RNN، و مدل‌های (Transformer) برای تشخیص گوینده.
- b. **پایگاه داده:** استفاده از پایگاه‌های داده مقیاس‌پذیر و مقاوم برای ذخیره‌سازی و بازیابی اطلاعات صوتی.
- c. **پلتفرم‌های توسعه:** سیستم باید به گونه‌ای طراحی شود که قابل اجرا بر روی سرورها و محیط‌های ابری مختلف باشد.
- d. **استانداردهای امنیتی:** تضمین امنیت داده‌ها و حفاظت از اطلاعات حساس صوتی.

e. قابلیت مقیاس پذیری : سیستم باید توانایی مدیریت حجم بالای درخواستها و صدهای مختلف را داشته باشد.

f. یکپارچگی API : ارائه API های استاندارد برای یکپارچگی با سیستم های داخلی و برنامه های جانبی.

۵. الزامات غیر فنی

a. تجربه و توانمندی ها : ارائه دهندگان باید تجربه کافی در توسعه سیستم های تشخیص گوینده و پردازش صوت داشته باشند.

b. کیفیت و پشتیبانی: ارائه دهندگان باید متعهد به ارائه محصولات با کیفیت بالا و خدمات پشتیبانی پس از تحویل باشند.

c. مدیریت پروژه : استفاده از روش های مدیریت پروژه استاندارد برای تضمین تحویل به موقع و در چارچوب بودجه.

۶. زمانبندی و تحویل ها

پیشنهاد می شود که زمانبندی پروژه به صورت زیر تنظیم شود:

a. تحلیل نیازها 4 هفته

b. طراحی معماری سیستم 6 هفته

c. توسعه و پیاده سازی 12 هفته

d. آزمون و ارزیابی 4 هفته

e. آموزش و مستندسازی 2 هفته

f. پشتیبانی و بهبود مستمر 6 ماه پس از تحویل

۷. نحوه ارائه پیشنهاد

ارائه دهندگان باید پیشنهادات خود را حداکثر تا [تاریخ] ارسال نمایند. پیشنهادها باید شامل موارد زیر باشند:

- a. توضیحات شرکت و تجربه‌های مرتبط : اطلاعاتی درباره شرکت و نمونه‌های پروژه‌های مشابه انجام شده.
- b. پیشنهاد فنی :توضیحات کامل درباره روش‌ها، مدل‌ها و تکنولوژی‌هایی که قرار است به کار گرفته شوند.
- c. پیشنهاد مالی :جزئیات هزینه‌ها و بودجه مورد نیاز برای اجرای پروژه.
- d. برنامه زمانی :جدول زمانی برای اجرای هر یک از مراحل پروژه.
- e. پشتیبانی و نگهداری :توضیحات درباره خدمات پشتیبانی و نگهداری پس از تحویل سیستم.