

درخواست پیشنهاد (RFP) برای پروژه‌های تبدیل گفتار به متن

(Speech-to-Text) در حوزه هوش مصنوعی

کد مسئله : SI- 2

۱. مقدمه

هدف از این پروژه، تسهیل دسترسی کاربران به اطلاعات از طریق تبدیل گفتار به متن و بهبود تعاملات درونی و بیرونی شرکت است.

۲. اهداف پروژه

هدف اصلی این پروژه توسعه یک سیستم تبدیل گفتار به متن با دقت و کارایی بالا است که توانایی تبدیل گفتار به متن در شرایط مختلف محیطی و با زبان‌ها و لهجه‌های مختلف را داشته باشد. اهداف فرعی عبارتند از:

a. تبدیل گفتار به متن با دقت بالا: سیستم باید قادر باشد گفتار را به متن با دقت بالا و حداقل خطا تبدیل کند.

b. پشتیبانی از چندین زبان و لهجه: سیستم باید قابلیت شناسایی و تبدیل گفتار به متن برای چندین زبان و لهجه مختلف را داشته باشد.

c. قابلیت کار در شرایط محیطی متنوع: سیستم باید توانایی کار در محیط‌های مختلف با سطوح نویز متفاوت را داشته باشد.

d. سرعت پردازش بالا: سیستم باید به صورت همزمان و با حداقل تأخیر گفتار را به متن تبدیل کند.

e. یکپارچگی با سیستم‌های موجود: سیستم باید قابلیت یکپارچگی با سیستم‌های نرم‌افزاری موجود در شرکت را داشته باشد.

f. قابلیت یادگیری و بهبود مستمر: سیستم باید بتواند با استفاده از داده‌های جدید و بازخورد کاربران، به مرور زمان بهبود یابد.

۳. دامنه کار

پروژه شامل مراحل زیر خواهد بود:

- a. **تحلیل و بررسی نیازها:** تحلیل نیازهای تجاری و فنی شرکت، شناسایی موارد استفاده مختلف و تعیین مشخصات سیستم.
- b. **طراحی معماری سیستم:** طراحی معماری سیستم تبدیل گفتار به متن، شامل انتخاب مدل‌های یادگیری، تکنیک‌های پردازش گفتار و ابزارهای مورد نیاز.
- c. **توسعه و پیاده‌سازی:** توسعه مدل‌های یادگیری برای تبدیل گفتار به متن، شامل آموزش مدل‌ها با استفاده از مجموعه داده‌های صوتی گسترده و متنوع.
- d. **آزمون و ارزیابی:** ارزیابی سیستم از نظر دقت تبدیل، مقاومت در برابر نویز، و سرعت پردازش.
- e. **آموزش و مستندسازی:** آموزش تیم داخلی برای استفاده و مدیریت سیستم و ارائه مستندات فنی و کاربری.
- f. **پشتیبانی و نگهداری:** ارائه خدمات پشتیبانی و نگهداری پس از تحویل سیستم برای به‌روزرسانی و بهبودهای آینده.

۴. الزامات فنی

- a. **الگوریتم‌ها و مدل‌ها:** استفاده از مدل‌های پیشرفته یادگیری عمیق مانند RNN، LSTM، CNN و Transformer برای تبدیل گفتار به متن.
- b. **پایگاه داده‌های آموزشی:** استفاده از مجموعه داده‌های صوتی با کیفیت بالا برای آموزش و بهبود مدل‌ها.
- c. **پلتفرم‌های توسعه:** سیستم باید به گونه‌ای طراحی شود که بتواند بر روی سرورهای محلی و همچنین در محیط‌های ابری (Cloud) اجرا شود.
- d. **استانداردهای امنیتی:** رعایت استانداردهای امنیتی برای حفاظت از داده‌های صوتی و اطلاعات کاربران.
- e. **مقیاس‌پذیری و کارایی:** سیستم باید توانایی مدیریت حجم بالای داده‌ها و درخواست‌ها را داشته باشد.
- f. **یکپارچه‌سازی API:** ارائه API‌های مناسب برای یکپارچه‌سازی با سایر سیستم‌های نرم‌افزاری موجود.

۵. الزامات غیر فنی

- a. **تجربه و توانمندی‌ها:** ارائه‌دهندگان باید تجربه کافی در توسعه سیستم‌های تبدیل گفتار به متن و کار با داده‌های صوتی بزرگ را داشته باشند.
- b. **کیفیت و پشتیبانی:** ارائه‌دهندگان باید متعهد به ارائه محصولات با کیفیت بالا و خدمات پشتیبانی پس از تحویل باشند.
- c. **مدیریت پروژه:** استفاده از روش‌های مدیریت پروژه استاندارد برای تضمین تحویل به‌موقع و در چارچوب بودجه.

۶. زمانبندی و تحویل‌ها

پیشنهاد می‌شود که زمانبندی پروژه به صورت زیر تنظیم شود:

- a. تحلیل و بررسی نیازها: ۴ هفته
- b. طراحی معماری سیستم: ۶ هفته
- c. توسعه و پیاده‌سازی: ۱۴ هفته
- d. آزمون و ارزیابی: ۴ هفته
- e. آموزش و مستندسازی: ۲ هفته
- f. پشتیبانی و نگهداری: ۶ ماه پس از تحویل

۷. نحوه ارائه پیشنهاد

ارائه‌دهندگان 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.

