

درخواست پیشنهاد (RFP) برای پروژه‌های تولید گفتار

(Speech Synthesis) در حوزه هوش مصنوعی

کد مسئله : SI-16

۱. مقدمه

این پروژه به منظور توسعه سیستمی طراحی شده است که قادر به تشخیص و تبدیل متن به گفتار به صورت دقیق، کارآمد و واقع‌گرایانه باشد. هدف از این سیستم بهبود تعاملات کاربری، تسهیل دسترسی به اطلاعات، و ارتقای عملکرد در محیط‌های متنوع است.

۲. اهداف پروژه

این پروژه توسعه یک سیستم تولید گفتار است که بتواند با دقت بالا و کیفیت مناسب، متن‌ها را به گفتار تبدیل کند. اهداف فرعی عبارتند از:

a. **تولید گفتار طبیعی:** توسعه سیستمی که توانایی تولید گفتاری طبیعی و واقع‌گرایانه از متن‌های ورودی را داشته باشد.

b. **پشتیبانی از چندین زبان:** سیستم باید قادر به تولید گفتار در زبان‌های مختلف (به ویژه زبان‌های مورد استفاده در بازار هدف) باشد.

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

d. **درک و تبدیل متن به گفتار:** سیستم باید قادر به درک ساختار زبان و تبدیل صحیح متن به گفتار باشد.

e. **پردازش گفتار در زمان واقعی:** سیستم باید قابلیت یکپارچگی با سایر سیستم‌ها و ابزارهای داخلی شرکت را داشته باشد.

f. **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 تولید گفتار، شامل انتخاب مدل‌های یادگیری، پردازش زبان طبیعی و تعیین منابع داده.
- c. **توسعه و پیاده‌سازی:** توسعه الگوریتم‌ها و مدل‌های تولید گفتار با استفاده از تکنیک‌های پیشرفته هوش مصنوعی.
- d. **آزمون و ارزیابی:** ارزیابی عملکرد سیستم از نظر دقت، کیفیت گفتار تولیدی 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 Tacotron ، WaveNet، و سایر مدل‌های تبدیل متن به گفتار برای تولید گفتار طبیعی.
- b. **پایگاه داده:** استفاده از پایگاه‌های داده صوتی مقیاس‌پذیر برای آموزش و بهبود مدل‌ها.
- 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 کاربران.

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

f. یکپارچگی API : ارائه API های استاندارد برای یکپارچگی با سیستم‌های داخلی و برنامه‌های جانبی.

۵. الزامات غیر فنی

a. تجربه و توانمندی‌ها : ارائه‌دهندگان باید تجربه کافی در توسعه سیستم‌های تولید گفتار و پردازش زبان طبیعی 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 نیازها 4 هفته

b. طراحی معماری سیستم 6 هفته

c. توسعه و پیاده‌سازی 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

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.