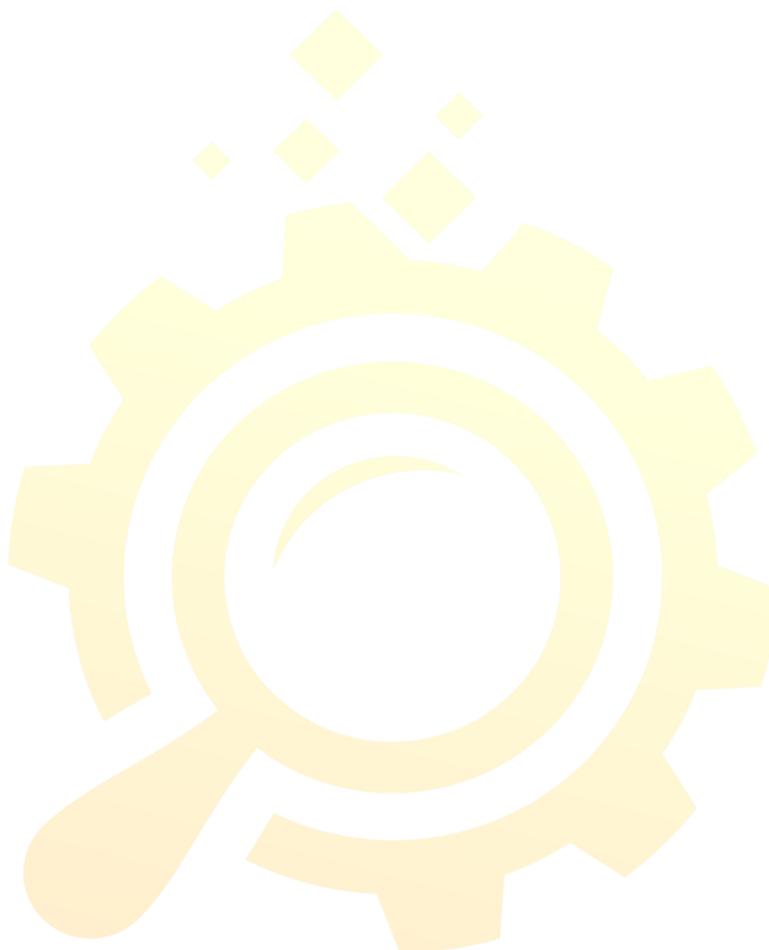


بیان نیاز تفصیلی

عنوان	عنوان
<p>فناوری اطلاعات و ارتباطات، رسانه و سرگرمی، آموزش، تبلیغات، دستیارهای هوشمند، متاورس و تولید محتوای صوتی.</p>	<p>۱. حوزه‌ی کاربری</p>
<p>«نیاز به سامانه‌ی هوش مصنوعی بومی برای تولید گفتار فارسی طبیعی، احساسی، بلادرنگ و قابل تغییر به صدای دلخواه»</p>	<p>۲. عنوان نیاز / فرصت</p>
<p>صدای تولیدشده توسط TTS های فعلی فارسی ربّاتی، خشک و غیرطبیعی است. فقدان پشتیبانی از احساسات، لحن و بیان انسانی در خروجی صوت. عدم وجود قابلیت تقلید صدا (Voice Cloning) با حفظ ویژگی‌های گفتاری مبدا. زمان پاسخ‌گویی بالا و عدم امکان استفاده‌ی بلادرنگ در چت، بازی یا دوبله زنده. دلایل: کمبود دیتاست گفتاری احساسی فارسی، نبود مدل‌های عمیق بومی‌سازی‌شده، ضعف پردازش هم‌زمان صدا، و کمبود ابزارهای کنترل دقیق بر لحن و شخصیت صوتی.</p>	<p>۳. مشکلات و نواقص موجود و دلیل بروز مشکل</p>
<ul style="list-style-type: none"> • موتور تبدیل متن به گفتار (TTS) فارسی با خروجی طبیعی، روان و دارای احساسات. • قابلیت تقلید صدای انسان با ورودی چندثانیه‌ای از صدای مبدا (Voice Cloning). • عملکرد بلادرنگ (Real-time) برای کاربردهای زنده (چت، بازی، دوبله هم‌زمان). • پشتیبانی از کنترل لحن، احساس، سرعت و بلندی صدا. • سازگاری با لهجه‌ها و گویش‌های مختلف فارسی. • ماژول تحلیل احساسات متن برای تنظیم خودکار لحن گفتار. • API و SDK برای یکپارچه‌سازی با سایر پلتفرم‌ها و نرم‌افزارهای تولید محتوا. 	<p>۴. الزامات کلیدی و حیاتی مربوط به نیاز</p>
<ul style="list-style-type: none"> • نبود دیتاست احساسی بزرگ و باکیفیت از گفتار فارسی. • تنوع زیاد در لهجه و گویش‌ها. • نیاز به پردازش سریع و سبک برای کاربرد بلادرنگ. • نیاز به فاین‌تیون مدل‌های خارجی (در صورت مجاز بودن) با داده‌ی فارسی. • ملاحظات حریم خصوصی و جلوگیری از جعل صدا (Voice Deepfake Protection). 	<p>۵. محدودیت‌ها و قیود</p>
<p>استفاده از سرویس‌های خارجی مانند ElevenLabs یا Play.ht (که از فارسی پشتیبانی کامل ندارند) یا استفاده از موتورهای TTS ساده و غیراحساسی فارسی.</p>	<p>۶. راه حل فعلی</p>
<p>پروژه‌های محدود TTS فارسی در دانشگاه‌های شریف، تهران و شرکت‌های فناپ و آبروان. برخی پژوهش‌ها در حوزه‌ی Voice Cloning انگلیسی (مانند VITS) اما بومی‌سازی فارسی هنوز در سطح تحقیقاتی است. نمونه خارجی: Descript، Play.ht، Meta Voicebox، OpenAI Voice Engine، ElevenLabs، Overdub. در ایران: چند نمونه‌ی ساده‌ی TTS فارسی وجود دارد (مثل گوگل ترنسلیت فارسی یا برخی سرویس‌های محدود بومی) اما کیفیت احساسی و بلادرنگ ندارند. بنابراین نیاز بالا برای توسعه‌ی مدل بومی با درک احساسات، لهجه‌ها و تطبیق‌پذیری بالا در زبان فارسی کاملاً وجود دارد.</p>	<p>۷. برنامه‌ها، پروژه‌ها و اقدامات مرتبط</p>
<p>پلتفرم هوش مصنوعی بومی برای تولید گفتار فارسی طبیعی و بلادرنگ با ویژگی‌های زیر:</p> <p>تبدیل متن به صوت انسانی و احساسی با درک لحن متن.</p> <p>قابلیت تغییر صدا به صدای هر فرد با ورودی نمونه صوت کوتاه.</p> <p>پاسخ‌گویی در لحظه برای استفاده در مکالمات زنده، بازی و دوبله.</p> <p>پشتیبانی از چنداحساس (شاد، غمگین، هیجان‌زده، رسمی و...)</p> <p>امکان تنظیم دستی پارامترهای گفتار (سرعت، لحن، بلندی، استرس).</p> <p>ماژول امنیتی برای جلوگیری از سوءاستفاده صوتی.</p>	<p>۸. محصول / راه حل پیشنهادی</p>

کاربر	مرکز عملیات رشد
تاریخ:	تاریخ:

<input checked="" type="checkbox"/> سرمایه‌گذاری برای توسعه محصول <input checked="" type="checkbox"/> بازاریابی و فروش محصول <input type="checkbox"/> تسهیل فروش محصول (از طریق وضع مقررات و ...) <input type="checkbox"/> قرارداد خرید تضمینی محصول <input type="checkbox"/> قرار دادن در لیست تأمین‌کنندگان (List Vendor) و تعامل مانند سایر تأمین‌کنندگان <input checked="" type="checkbox"/> سایر: توزیع در سطح کشور	۹. نحوه حمایت بهره‌بردار از حل مسئله
---	--



مرکز عملیات رشد	کاربر
تاریخ:	تاریخ: