



هوش مصنوعی Eleven v3 معرفی شد / طبیعی‌ترین مدل تبدیل متن به گفتار

هوش مصنوعی شرکت «الون لبز» که متن را به گفتار تبدیل می‌کند، امکان کنترل بیان را برای تولید صدا به ارمغان می‌آورد و به جای خوانش‌های ساده، امکان اجرای واقعی را فراهم می‌کند.

هوش مصنوعی شرکت «الون لبز» که متن را به گفتار تبدیل می‌کند، امکان کنترل بیان را برای تولید صدا به ارمغان می‌آورد و به جای خوانش‌های ساده، امکان اجرای واقعی را فراهم می‌کند. به گزارش ایسنا، هوش مصنوعی «الون وی۳» (Eleven v3) شرکت «الون لبز» (ElevenLabs) که براساس یک ساختار جدید ساخته شده است، واقع‌گرایی و کنترل بی‌سابقه‌ای را برای تولید گفتار به ارمغان می‌آورد. این نرم‌افزار می‌تواند لحن را در اواسط جمله تغییر دهد، به طور یکپارچه بین شخصیت‌ها حرکت کند و به نشانه‌هایی مانند نجوا، خنده و آه کشیدن پاسخ دهد. همه این موارد بدون ایجاد اختلال در جریان گفتار صورت می‌گیرند.

به نقل از دیجیتال ترمنال، این نسخه از هوش مصنوعی الون لبز یک نسخه «آلفا» است. این نسخه نسبت به مدل‌های پیشین به مهندسی سریع‌تری نیاز دارد اما خروجی آن یک تغییر اساسی را در بیان، ظرافت و واقع‌گرایی انسانی ارائه می‌دهد. این نسخه جدید، پشتیبانی از زبان‌های گوناگون را از ۳۳ به بیش از ۷۰ زبان افزایش داده است. با تنظیم دقیق‌تر، قابلیت اطمینان و کنترل پذیری نیز افزایش خواهد یافت.

قابلیت‌های مهم الون وی۳ به شرح زیر هستند.
۱. پوشش دهی بیش از ۷۰ زبان؛ زبان‌های پوشش داده شده در الون وی۳ از ۳۳ زبان به بیش از ۷۰ زبان گسترش یافته‌اند و پوشش دهی جمعیت جهان از ۶۰ درصد به ۹۰ درصد افزایش داشته است.
۲. حالت گفتگو؛ الون وی۳ وقفه‌های طبیعی، تغییر لحن و جریان احساسی را بین چندین گوینده مدیریت می‌کند.
۳. برجسب‌های صوتی؛ الون وی۳ ارائه خود را با برجسب‌هایی مانند زمزمه، عصبانیت، خنده یا صدای جیرجیر در هدایت می‌کند و امکان کنترل عملکرد را با جزئیات دقیق فراهم می‌آورد.
۴. پشتیبانی از پخش زنده؛ به زودی ویژگی پشتیبانی از پخش زنده به صورت در لحظه برای مراکز تماس و اپراتورهای مکالمه ارائه می‌شود.

۵. API عمومی برای الون وی۳ (آلفا) به زودی منتشر می‌شود. کاربران برای دسترسی زودهنگام می‌توانند با بخش فروش شرکت تماس بگیرند.
الون وی۳ برای سازندگان، توسعه‌دهندگان و شرکت‌هایی طراحی شده است که محتوای شامل داستان‌ها، کتاب‌های صوتی، گفت‌وگوی شخصیت‌ها و تعاملات رسانه‌ای را تولید می‌کنند. این مدل به آزمایش و ارائه بازخورد آگاهانه از متن ارجح می‌نهد. برای موارد استفاده در لحظه و با تأخیر کم مانند هوش مصنوعی محاوره‌ای توصیه شده که کاربران فعلاً با مدل‌های «توربو» (Turbo) و «فلش» (Flash) الون وی۲.۵ کار کنند. نسخه در لحظه الون وی۳ در حال توسعه است.
الون وی۳ کنترل بیان را برای تولید صدا به ارمغان می‌آورد و به جای خوانش‌های ساده، اجراهای واقعی را ممکن می‌سازد. این نسخه می‌تواند احساسات را تغییر دهد، نحوه بیان را تنظیم کند و یک حرکت روان را بین شخصیت‌های یک نسل داشته باشد.

برای اولین بار، گفتار هوش مصنوعی می‌تواند ریتم و ظرافت احساسی مکالمه انسانی را در بیش از ۷۰ زبان دنبال کند. «ماتی استانیسزفسکی» (Mati Staniszewski) از بنیان‌گذاران الون لبز و مدیرعامل این شرکت گفت: الون وی۳ رساترین مدل تبدیل متن به گفتار تا به امروز است که کنترل کامل را بر احساسات، نحوه بیان و نشانه‌های غیر کلامی ارائه می‌دهد. کاربران با استفاده از برجسب‌های صوتی می‌توانند آن را به زمزمه، خنده، تغییر لهجه یا حتی آواز خواندن ترغیب کنند. آنها می‌توانند سرعت، احساسات و سبک را برای مطابقت با هر نوع نوشتار کنترل کنند. با مأموریت جهانی خود، ما خوشحالیم که این مدل را با پشتیبانی از بیش از ۷۰ زبان گسترش می‌دهیم. این نسخه نتیجه چشم‌انداز و رهبری همکار من «پیوتر» (Piotr) و گروه پژوهشی فوق‌العاده‌ای است که او تشکیل داده بود.

ایجاد یک محصول خوب سخت است اما ایجاد یک الگوی کاملاً جدید تقریباً غیرممکن است. من و همه ما در الون لبز از دیدن جادویی که این تیم به زندگی می‌بخشد، احساس خوش‌شانسی می‌کنیم و با این نسخه هیجان زده‌ایم که یک بار دیگر مرزها را جابه‌جا کنیم.

الون وی۳ (آلفا) هم اکنون در آدرس <https://elevenlabs.io> موجود است.