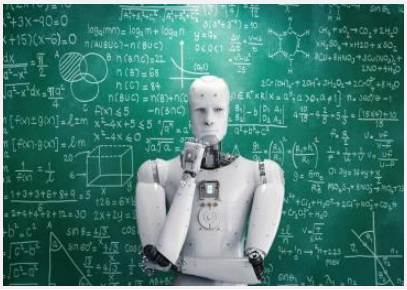

پنجشنبه 02 اسفند 1403 - 03:50

هوش مصنوعی صدای افراد را شبیه سازی می کند

یک مدل هوش مصنوعی ابداع شده که با کلیپ چندثانیه ای از صدای فرد آن را شبیه سازی می کند.

یک مدل هوش مصنوعی ابداع شده که با کلیپ چندثانیه ای از صدای فرد آن را شبیه سازی می کند.

به گزارش خبرگزاری مهر به نقل از رجیستر، یک استارت آپ آمریکایی به نام Zyphra از یک مدل هوش مصنوعی متن به گفتار(TTS) رونمایی کرده که می تواند با دریافت نمونه صوتی ۵ ثانیه ای از فرد، صدای او را شبیه سازی کند.

دنی مارتینلی و کریتیک پوتالات این استارت آپ را در ۲۰۲۱ میلادی با هدف ساخت یک سیستم عامل چند حالتی به نام MaiaOS راه اندازی کردند. این نتیجه این تلاش ها به شکل عرضه خانواده مدل های زبانی کوچک Zamba و اکنون عرضه مدل های متن به گفتار Zonos نمایش داده شده است.

هر یک از این مدل ها ۱.۶ میلیارد پارامتر دارند و براساس ۲۰۰ هزار ساعت داده گفتاری شامل حرف زدن با لحن صدای خنثی مانند خوانش کتاب صوتی و همچنین گفتار با لحن احساسی آموزش دیدند. بخش اعظم داده های آموزشی آن به زبان انگلیسی بوده اما مقدار زیادی داده به زبان چینی، ژاپنی، فرانسوی، اسپانیایی و آلمانی نیز بین این موارد وجود داشته است. به گفته شرکت اطلاعات مذکور از وب جمع آوری شده اند و از دلال های داده خریداری نشده اند.

هر دو مدل عملکردی مشابه دیگر مدل های هوش مصنوعی تبدیل متن به گفتار هستند.