

مروری بر تلاش شرکت ها برای آموزش زبان چینی به کامپیوترها - دیجیاتو

مریم موسوی | شنبه، ۰۷ مرداد ۱۳۹۶

۲۰۱۷ را از بسیاری جهات سال می توان سال رونق گرفتن کامپیوترهای سخنگو در چین دانست.

با در نظر داشتن پیشرفت های صورت گرفته در تشخیص کلام و پردازش زبان طبیعی، شرکت هایی نظیر علی بابا (غول تجارت الکترونیک در سرزمین اژدهای سرخ) و همچنین بایدو دست به توسعه تکنولوژی هایی زده اند تا ارتباط کلامی را به کامپیوترها و سیستم های چینی بیاورند.

به پاس همین تلاش ها نیز محصولات مبتنی بر این فناوری که توسط دو شرکت بایدو و علی بابا توسعه داده شده اند به تدریج در حال وارد شدن به بازار چین هستند.

اسپیکر Tmall Genie که دستیار صوتی علی بابا به نام **AliGenie** را در خود دارد یکی از این گجت هاست که شباهت زیادی هم به اکوی آمازون دارد. این اسپیکر می تواند در وبسایت های خرید و فروش آنلاین برای کاربرانش سفارش گذاری کند، وضعیت آب و هوا را بررسی نماید و موسیقی دلخواه دارندگان را پخش کند. البته کنترل دیگر لوازم و وسایل هوشمند درون منزل از طریق فرامین صوتی را هم باید به فهرست قابلیت های این دستگاه اضافه نمود.

پلتفرم محاوره ای بایدو به نام **DuerOS** نیز به عنوان قابلیت تازه برای محصولات نظیر ربات های دستیار خانگی، ستاپ باکس تلویزیون و یکی از موبایل های شرکت اچ تی سی توسعه یافته. این محصول قابلیت هایی مشابه به **AliGenie** و دیگر دستیارهای صوتی دارد و علاوه بر این برای انجام مکالمات تصادفی برخی توانمندی های ابتدایی برایش در نظر گرفته شده. شرکت سازنده مدعی شده سفارشات زیادی را برای دریافت کیت توسعه دهندگان **DuerOS** دریافت نموده.

کان جینگ مدیرکل واحد تجارت **Duer** در بایدو پیش بینی نموده که امسال شرکت های بیشتری به این حوزه ورود نمایند که البته این مساله تا حدودی به موفقیت گجت های مشابهی چون اکو در آمریکا باز می گردد که سرمایه گذاران چینی را راغب به انجام این کار کرده است.

شرکت تحقیقاتی IDC پیش بینی کرده که تا سال ۲۰۲۰ میلادی، ۵۱ درصد از صنعت رانندگی هوشمند و ۶۸ درصد از موبایل ها و پوشیدنی ها در چین همراه با پشتیبانی نوعی سیستم مکالمه محور و مبتنی بر هوش مصنوعی توسعه پیدا کنند.

به گفته کان جینگ:

در واقع درست همانطور که نمایشگرهای تاچ تعامل با تلفن های همراه را راحت تر کردند، اینترفیس های محاوره ای نیز شکلی طبیعی تر به تعاملات کاربران می بخشند و افراد بیشتری را به حضور در دنیای هوشمند ترغیب می نمایند.

اما رایانش مبتنی بر صدا فواید زیادی در بازار چین دارد. امروزه تایپ کردن حروف روی کیبوردهای مرسوم QWERTY از طریق سیستمی به نام pinyin انجام می گیرد که تکیه بر تلفظ حروف دارد اما از آنجا که زبان چینی چهار لحن و آهنگ مختلف دارد که هر کدام نیز معنایی متفاوت از دیگری دارد، به همین خاطر فرایند انتخاب کاراکتر صحیح برای کاربر از منویی کشویی که بعد از تایپ تلفظ آن حرف باز می گردد کاری اندک دشوار و البته زمان بر است.



در زبان چینی یک هجای معمولی مانند Yi ممکن است به ۶۰ یا تعداد بیشتری کاراکتر پرکاربرد مربوط باشد. البته در برخی روش های ورودی دیتا، محتمل ترین کاراکتر براساس متن به کاربر پیشنهاد داده می شود اما این پیشنهادات همیشه هم درست نیستند. از این رو جای تعجب ندارد اگر بگوییم کاربران تکنولوژی های موبایلی نظیر اپ محبوب وی چت معمولاً تمایل دارند که پیام های صوتی برای یکدیگر بگذارند تا اینکه متن را تایپ کنند.

در چین امروز، تکنولوژی دستیار صوتی ابتدا فرامین کاربر را به متن تبدیل کرده و بعد پاسخ مناسب را براساس معنای متن ارائه می دهد. این فرایند در فرامین مبتنی بر وظیفه (نظیر بررسی وضعیت آب و هوا یا جستجوی معادل انگلیسی کلمات چینی) بسیار سودمند واقع می شود اما با

تکیه بر آن نمی توان مکالمه ای دو سویه در رابطه با موضوعات متعدد داشت.

لذا حل مشکلاتی که در مسیر رایانش محاوره ای وجود دارد نیازمند برطرف کردن برخی پیچیدگی های چالشی زبان چینی است. در این زبان برای نمونه برخی کاراکترها در صورتی که با ترتیب خاصی کنار هم چیده شده باشند بسته به اینکه قبل یا بعد از آنها چه حرفی قرار داشته باشد می توانند معانی مختلفی به خود بگیرند. علاوه بر این، افعال زمان گذشته، حال یا آینده ندارند که این موضوع تشخیص تقدم و تاخر زمانی برای کامپیوترها را دشوار می کند.

اما چالش های پردازش زبان چینی به این موارد ختم نمی شوند و دانشمندان با مشکلات دیگری هم در این مسیر دست و پنجه نرم می کنند؛ در این زبان گویش ها متعددند و برخی از آنها برای دیگر چینی زبان ها غیرقابل درک. نکته دیگر آنکه برخی عبارات در شرایط مختلف معانی گوناگونی به خود می گیرند.

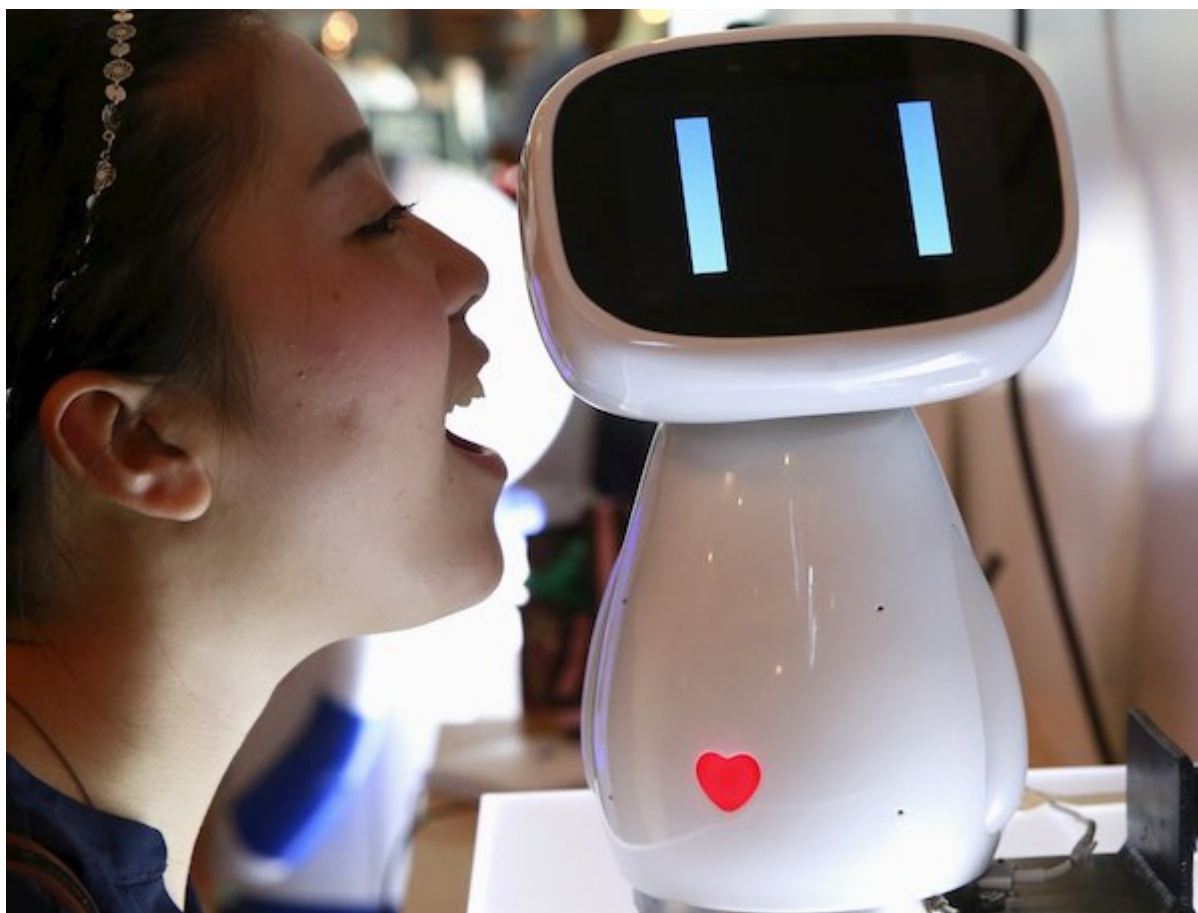


Zhiyong Wu از استارادیاران دانشگاه Tsinghua که در زمینه درک زبان طبیعی تحقیق میکند در این باره گفت:

برای آنکه کامپیوترها واقعا قصد و منظور متکلم های انسانی را درک کنند و ارتباطی مناسب با آن برقراری نمایند لازم است که به درک مناسبی از جزئیات مختلف این زبان نظیر فراز و فرود صدا و تکیه روی کلمات دست پیدا کنند.

به گفته پروفیسور Jia Jia که روی موضوع رایانش موثر اجتماعی تحقیق می کند و در دانشگاه Tsinghua درس می دهد ماشین ها همچنین باید درک خوبی از احساسات پیدا کنند چراکه تصمیم گیری انسان صرفا بر اساس منطق نیست.

بایدو برای آنکه این سیستم ها را هوشمندتر نماید نوعی مد جدید موسوم به Trainer یا مربی را به پلتفرم خود اضافه کرده که به توسعه دهندگان نرم افزار امکان می دهد به صورت آنی و از طریق ربات در گسترش دیتای زبانی آن مشارکت نمایند. این ربات بازخورد توسعه دهندگان (نظیر توضیحات مربوط به پرسش های مطرح شده در رابطه با عبارتی که سیستم بار نخست متوجه نشده) را دریافت کرده، از آن یاد می گیرد و بعد سیستم را اصلاح می کند.



اما یکی از مزایایی که محققان چینی همزمان با تلاش برای رفع این مشکل از آن بهره می گیرند حجم بالای دیتاست. شبکه های عصبی که درک زبان های طبیعی را در کامپیوترهای امروزی بر عهده دارند نیازمند حجم بالایی دیتا برای یادگیری هستند. بنابراین هرچه میزان دیتای تحت اختیار یک شرکت بیشتر باشد، شبکه توسعه یافته توسط آن نیز هوشمندتر است و این مساله به خصوص در مورد شرکت های علی بابا و بایدو صادق است که طیف وسیعی از کاربران را دارند. تا پایان سال ۲۰۱۶ میلادی، بایدو ۶۶۵ میلیون کاربر فعال ماهانه در بستر موبایل داشته و علی بابا هم اعلام کرده که تا پایان سه ماهه نخست امسال ۵۰۷ میلیون کاربر ماهانه موبایلی داشته است.

اما به گفته گنگ ونگ از دانشمندان هوش مصنوعی علی بابا ضروریست محققان شبکه هایی را طراحی کنند که برای رسیدن به سطح مطلوب از مهارت های زبانی به حجم بالایی از دیتا نیاز ندارند.

در دنیای واقعی افراد به طرق مختلف معنا و مفهوم مدنظرشان را به مخاطب انتقال می دهند و به همین خاطر آموزش تمامی این شیوه های ارتباطی به کامپیوترها کاری غیرممکن است.

وی که پیشتر به عنوان یک پژوهشگر دانشگاهی فعالیت کرده به همراه تیم همکارانش به روشی دست پیدا کرده بود که با آن می شد موضوعات مختلف را با حداقل دیتای موجود به کامپیوترها آموزش داد و برای این منظور دیتای به دست آمده از دیگر موضوعات مرتبط را به کار برد.

برای مثال اگر قرار باشد درک متون در زمینه های ورزشی و پزشکی را به یک کامپیوتر آموزش دهیم سیستم از دیتایی کمک می گیرد که در این دو حوزه موجود است. البته مزایای این روش به مراتب کمتر از کاربرد دیتای اورگانیک یا اصیل است با این همه آنطور که ونگ می گوید: در نبود این دیتا می توان از این روش برای آموزش شبکه های عصبی پیرامون یک موضوع خاص بهره گرفت.

در نهایت هم طبق گفته Chenfeng Song بنیانگذار استارتآپی به نام Ainemo که ربات دستیار Little Fish را ساخته آنچه باعث موفقیت یک دستیار صوتی در چین می شود محتوا و خدماتش است. سانگ در نظر دارد که به تدریج برنامه های آموزشی و سلامتی را به هم در داخل دستیار خانگی اش به کار ببرد. در مورد ربات این شرکت باید بگوییم که از پلتفرم محاوره ای DuerOS بهره می گیرد.

سانگ در ادامه توضیحاتش می گوید: صدا راهی برای ارائه محتوا به مردمی (خصوصاً کودکان و افراد سالخورده) است که نمی توانند از طریق سیستم های رومیزی یا اسمارت فون ها به خوبی به اینترنت دسترسی پیدا کنند.

[دیجیاتو](#)