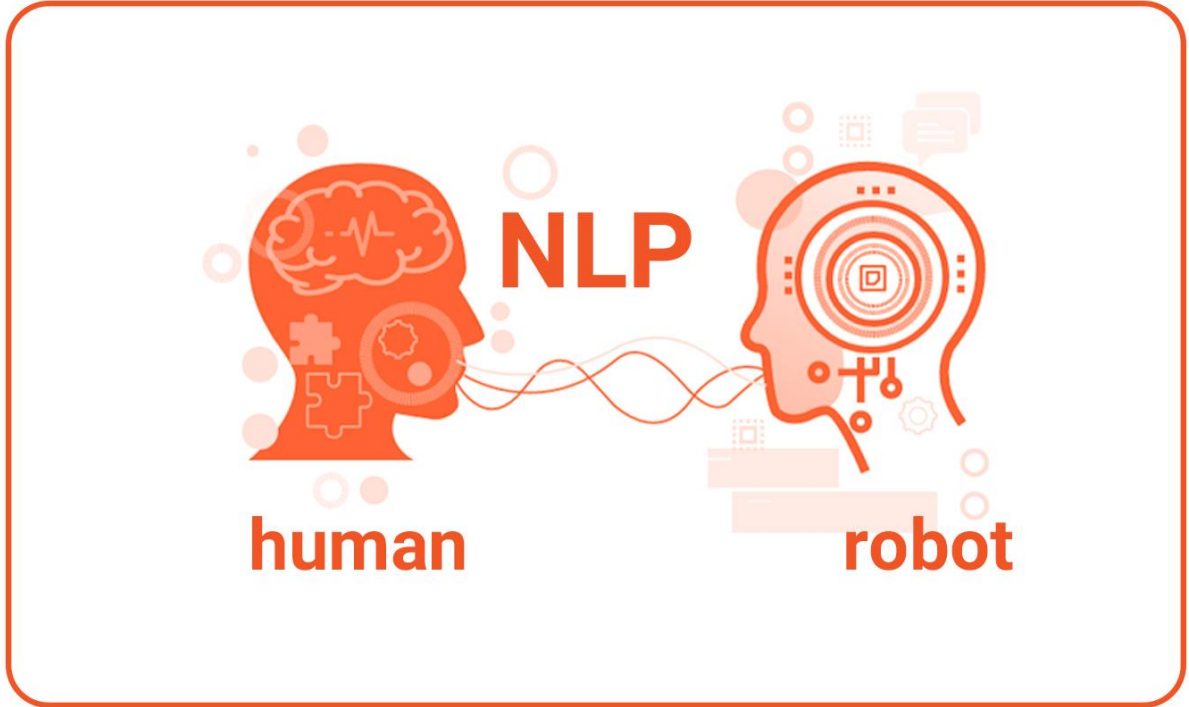




مجموعه شرکت های مهندسی دانش بنیان رها

## پردازش زبان طبیعی (NLP)؛ جذاب ترین مبحث در هوش مصنوعی

شرکت رهاکو



rahaco.net/mag

## فهرست

- 3 ..... پردازش زبان طبیعی یعنی چه؟
- 4 ..... NLP چه وظایفی دارد؟
- 4 ..... موارد استفاده از پردازش زبان طبیعی NLP
- 5 ..... ابزارها و رویکردهای پردازش زبان طبیعی NLP
- 6 ..... NLP آماری، یادگیری ماشینی و یادگیری عمیق
- 6 ..... نحوه کاربرد پردازش زبان های طبیعی NLP چگونه است؟
- 6 ..... مزایای پردازش زبان طبیعی
- 7 ..... نتیجه گیری



پردازش زبان طبیعی (NLP) زیر شاخه‌ای از زبان‌شناسی، علوم رایانه و هوش مصنوعی است که به تعامل بین رایانه‌ها و زبان انسان برای پردازش و تجزیه و تحلیل داده‌ها اشاره دارد. هدف رایانه‌ای است که قادر به "درک" خواهد بود. انسان‌ها به ندرت دقیق یا با صراحت صحبت می‌کنند. برای درک زبان انسانی، نه تنها کلمات، بلکه مفاهیم و نحوه پیوند آن‌ها با یکدیگر نیز بسیار اهمیت دارد. علیرغم این که زبان یکی از ساده‌ترین مهارت‌ها برای یادگیری ذهن انسان است. اما ابهام زبان همان چیزی است که پردازش زبان طبیعی را برای کامپیوترها بسیار دشوار می‌کند.

این فناوری اطلاعات را به طور دقیق استخراج کند سپس دسته بندی و سازماندهی آن‌ها را به طور کامل انجام دهد. چالش‌های پردازش زبان طبیعی بیشتر شامل تشخیص گفتار، درک و تولید زبان طبیعی است. NLP در تلاش است تا بتواند متن یا داده‌های صوتی را درک کند و به آن‌ها پاسخ دهد.

## پردازش زبان طبیعی یعنی چه؟

پردازش زبان طبیعی یا NLP یکی از زیر شاخه‌های هوش مصنوعی می‌باشد. این فناوری توانایی درک متن و کلمات را به همان روشی که در انسان وجود دارد، به رایانه‌ها یاد می‌دهد. در حقیقت، رشته تحصیلی که بر تعاملات بین زبان انسان و رایانه متمرکز است، پردازش زبان طبیعی یا به اختصار NLP نامیده می‌شود. این رشته در دسته علوم کامپیوتر، هوش مصنوعی، و زبان‌شناسی محاسباتی قرار دارد.

NLP زبان‌شناسی محاسباتی، مدل‌سازی مبتنی بر قواعد زبان انسانی را با مدل‌های آماری، یادگیری ماشینی و یادگیری عمیق ترکیب می‌کند. این فناوری‌ها در کنار هم، رایانه‌ها را قادر می‌سازند تا زبان انسان را به صورت متن یا داده‌های صوتی پردازش کنند و معنای کامل آن را کاملاً با هدف و احساسات گوینده یا نویسنده «درک» کنند.

پردازش زبان طبیعی NLP این قابلیت را دارد که یک متن را از یک زبان به زبان دیگر ترجمه کند و همچنین می‌تواند دستورات گفتاری را کدگشایی کند و پردازش متن‌های زیاد را با سرعت بالا انجام دهد. کاربران با استفاده از پردازش زبان طبیعی NLP به سیستم‌های GPS، دستیارهای دیجیتال، نرم‌افزار تبدیل گفتار به متن، چت ربات‌های خدمات مشتری و سایر امکانات دسترسی خواهند داشت. بنابراین، NLP نقش رو به رشدی در سازمان ایفا می‌کند که به ساده‌سازی عملیات تجاری، افزایش بهره‌وری کارکنان و ساده‌سازی فرآیندهای کسب‌وکار کمک می‌کند.



## NLP چه وظایفی دارد؟

زبان انسان مملو از ابهامات می باشد و نوشتن نرم افزاری که دقیقا معنای متن یا داده های صوتی را تشخیص دهد بسیار دشوار است. ارایه ها، استعاره ها، کنایه ها، آواها، اصطلاحات، قواعد کاربردی در ساختار جملات فقط کمی از پیچیدگی های زبان انسان را نشان می دهد. به همین دلیل یادگیری آن برای ماشین ها تا سال ها طول می کشد! اما برنامه نویسان در تلاشند تا برنامه های کاربردی مبتنی بر زبان طبیعی را ارائه می دهند.

**تشخیص گفتار به متن:** ابزار تشخیص گفتار به متن داده های صوتی را به داده های متنی تبدیل می کند. سیستم تشخیص گفتار برای هر برنامه ای که دستورات صوتی را دنبال می کند یا به سوالات گفتاری پاسخ می دهد، لازم است.

**تجزیه و تحلیل احساسات:** تجزیه و تحلیل احساسات سعی می کند نگرش های ذهنی، احساسات، کلیدی، سردرگمی و سو ظن را در متن تشخیص دهد.

## موارد استفاده از پردازش زبان طبیعی NLP

پردازش زبان طبیعی، نیروی محرکه هوش ماشینی است که در دنیای واقعی مدرن کاربرد بسیاری دارد. در این قسمت به چند نمونه اشاره می کنیم:

**تشخیص اسپم:** ممکن است تشخیص اسپم به عنوان یک راهکار NLP به نظر نرسد، اما فناوری تشخیص اسپم متن را برای اسکن ایمیل ها بررسی می کند. به همین ترتیب، زبانی که اغلب نشان دهنده اسپم یا فیشینگ است پیدا خواهد شد. این عناصر تهدید کننده شامل استفاده بیش از حد از اصطلاحات مالی، جملات تهدید آمیز، نام برندها با املای اشتباه و غیره هستند.

**ترجمه ماشینی:** Google Translate نمونه ای از فناوری NLP می باشد که در همه ی مکان ها در دسترس است. ترجمه ماشینی واقعا مفید و کاربردی است و در واقع چیزی بیش از جایگزینی ساده کلمات یک زبان با زبان دیگر می باشد. ترجمه موثر ترجمه ای است که معنی و لحن زبان ورودی را به دقت دریافت کند و آن را به متنی با همان معنا و تاثیر در زبان خروجی ترجمه کند.

ابزارهای ترجمه ماشینی از نظر دقت پیشرفت خوبی داشته اند. یک راه عالی برای آزمایش ابزارهای ترجمه ماشینی، ترجمه متن به یک زبان و سپس بازگشت همان متن به زبان اصلی است.



**عوامل مجازی و ربات های گفت و گو:** برنامه های مجازی مانند: سیری اپل و الکسای آمازون از سیستم ترجمه برای تشخیص الگوهای دستورات صوتی و تولید زبان طبیعی استفاده می کنند. سیری Siri یک دستیار هوشمند است که از هوش مصنوعی استفاده می کند. این نرم افزار کاربردی در iOS وجود دارد.

بهترین این برنامه ها سرخ های متنی در مورد درخواست های انسانی را تشخیص دهند و از آن ها برای ارائه پاسخ ها یا گزینه های بهتر در طول زمان استفاده کنند. پیشرفت بعدی برای این برنامه ها پاسخگویی به سؤالات است، توانایی پاسخگویی به سؤالات ما با جواب های مرتبط و مفید است.

**تجزیه و تحلیل احساسات در رسانه های اجتماعی:** NLP به یک ابزار تجاری و ضروری برای کشف بینش داده های پنهان، از کانال های رسانه های اجتماعی تبدیل شده است. تجزیه و تحلیل احساسات می تواند زبان مورد استفاده در پست های رسانه های اجتماعی، پاسخ ها، بررسی ها است و موارد دیگر را برای استخراج نگرش ها و احساسات در پاسخ به محصولات، تبلیغات و رویدادها تجزیه و تحلیل کند.

**خلاصه سازی متن:** خلاصه سازی متن از تکنیک های NLP برای هضم حجم عظیمی از متن دیجیتالی و ایجاد خلاصه ها و همچنین خلاصه هایی برای نمایه ها، پایگاه های اطلاعاتی پژوهشی یا خوانندگان پرمشغله ای که وقت خواندن متن کامل را ندارند، استفاده می کند.

## ابزارها و رویکردهای پردازش زبان طبیعی NLP

### Python and the Natural Language Toolkit (NLTK)

زبان برنامه نویسی پایتون طیف وسیعی از ابزارها و کتابخانه ها را برای انجام وظایف خاص NLP فراهم می کند. بسیاری از این موارد در Natural Language Toolkit یا NLTK، مجموعه ای open source از کتابخانه ها، برنامه ها و منابع آموزشی برای ساخت برنامه های NLP یافت می شوند.

NLTK شامل کتابخانه هایی برای بسیاری از وظایف NLP ذکر شده در قسمت بالا، به علاوه کتابخانه هایی برای وظایف فرعی، مانند تجزیه جملات، تقسیم بندی کلمات، ریشه گذاری و واژه سازی و نشانه سازی (برای شکستن عبارات، جملات، پاراگراف ها) است.

همچنین شامل کتابخانه هایی برای پیاده سازی، قابلیت هایی مانند: استدلال معنایی، توانایی رسیدن به نتایج منطقی بر اساس حقایق استخراج شده از متن است.



## NLP آماری، یادگیری ماشینی و یادگیری عمیق

NLP آماری را وارد می کند که بتواند الگوریتم های کامپیوتری را با مدل های یادگیری ماشینی و یادگیری عمیق ترکیب کند. تا به طور خودکار عناصر متن و داده های صوتی را استخراج، طبقه بندی و برچسب گذاری کند. و سپس احتمال آماری را به هر معنای احتمالی آن عناصر اختصاص دهد.

امروزه مدل های یادگیری عمیق و تکنیک های یادگیری مبتنی بر شبکه های عصبی کانولوشن (CNN) و شبکه های عصبی مکرر (RNN) سیستم های NLP را قادر می سازند که در حین کار، «یاد بگیرند» و معنای دقیق تری را از حجم عظیمی از متن خام، بدون ساختار، بدون برچسب و مجموعه داده های صوتی استخراج کنند.

### نحوه کارکرد پردازش زبان های طبیعی NLP چگونه است؟

NLP کامپیوترها را قادر می سازد تا زبان طبیعی را مانند انسان درک کنند. چه زبان گفتاری باشد چه زبان نوشتاری، پردازش زبان طبیعی از هوش مصنوعی برای دریافت ورودی های دنیای واقعی، پردازش آن و درک آن به گونه ای استفاده می کند که کامپیوتر بتواند آن را بفهمد.

همانطور که انسان ها حسگرهای مختلفی دارند مانند گوش برای شنیدن و چشم برای دیدن کامپیوترها هم برنامه های برای خواندن و میکروفون های برای جمع آوری صدا دارند. همانطور که انسان ها برای پردازش این ورودی مغز دارند، کامپیوترها نیز برنامه ای برای پردازش ورودی های مربوطه خود دارند. در مرحله ای از پردازش، ورودی به کدی تبدیل می شود که کامپیوتر می تواند آن را درک کند.

### مزایای پردازش زبان طبیعی

مزیت اصلی NLP این است که نحوه ارتباط انسان و کامپیوتر با یکدیگر را بهبود می بخشد. مستقیم ترین راه برای دستکاری کامپیوتر از طریق کد زبان کامپیوتر است. با قادر ساختن رایانه ها به درک زبان انسان، تعامل با رایانه برای انسان بسیار شهودی تر می شود.

مزایای دیگر عبارتند از:

- بهبود دقت و کارایی اسناد؛
- توانایی ایجاد خودکار خلاصه ای قابل خواندن از یک متن اصلی بزرگ تر و پیچیده تر.
- برای دستیاران شخصی مانند الکسا مفید است، زیرا آن را قادر به درک کلمات گفتاری می کند.
- سازمان را قادر می سازد تا از چت بات ها برای پشتیبانی مشتری استفاده کند.



- آسان تر برای انجام تجزیه و تحلیل احساسات، بینش های پیشرفته ای را از تجزیه و تحلیل ها ارائه می دهد که قبلا به دلیل حجم داده ها غیرقابل دسترسی بودند.

## نتیجه گیری

پردازش زبان طبیعی (NLP) توانایی یک برنامه کامپیوتری برای درک زبان انسان به همان صورتی که گفته و نوشته می شود، به عنوان زبان طبیعی شناخته شده است. پردازش زبان طبیعی یکی از اجزای هوش مصنوعی (AI) می باشد.

NLP بیش از 50 سال است که وجود دارد و ریشه در زمینه ی زبان شناسی دارد. NLP برای تجزیه و تحلیل متن استفاده می شود و به ماشین ها اجازه می دهد تا نحوه صحبت انسان ها را درک کنند. این تعامل انسان و رایانه، برنامه های کاربردی دنیای واقعی مانند خلاصه سازی خودکار متن، تجزیه و تحلیل احساسات، استخراج موضوع، ریشه یابی و غیره را ممکن می سازد. NLP معمولا برای ترجمه ماشینی و پاسخگویی خودکار به سوالات استفاده می شود.