

پردازش زبان طبیعی یکی از زیر شاخه‌های با اهمیت در حوزه گسترده هوش مصنوعی و دانش زبان‌شناسی است. تلاش عمده در این زمینه، ماشینی کردن فرایند درک و برداشت مفاهیم بیان شده توسط یک زبان طبیعی انسانی است. به تعریف دقیق‌تر پردازش زبانهای طبیعی عبارت است از استفاده از رایانه به منظور پردازش زبان گفتاری و نوشتاری. پردازش زبان ها و مکالمات طبیعی یکی از اموری است که با ورود فناوری رایانه‌ای به زندگی بشر مورد توجه بسیاری از دانشمندان و پژوهشگران حوزه هوش مصنوعی قرار گرفته است. حتی اندیشه‌ای که تورینگ از ماشین هوشمند خود و تعریفی که او از هوش مصنوعی داشت، در مرحله اول مربوط به پردازش زبان طبیعی می‌شد. پردازش زبان طبیعی رهیافت بسیار جذابی برای ارتباط بین انسان و ماشین به شمار می‌آید و در صورت عملی شدنش به طور کامل می‌تواند تحولات شگفت‌انگیزی را در پی داشته باشد. مسئله پردازش زبان طبیعی یک مسئله **AI-Complete** به شمار می‌آید، چرا که محقق شدن آن به طور کامل مستلزم سطح بالایی از درک جهان خارج و حالات انسان برای ماشین است.

هدف اصلی در پردازش زبان طبیعی، ایجاد تئوری‌های محاسباتی از زبان، با استفاده از الگوریتم‌ها و ساختارهای داده‌ای موجود در علوم رایانه‌ای است. بدیهی است که در راستای تحقق این هدف، نیاز به دانشی وسیع از زبان بوده و علاوه بر پژوهشگران دانش رایانه‌ای، نیاز به دانش زبان‌شناسان نیز در این حوزه خواهد بود. کاربردهای پردازش زبان طبیعی به دو دسته کلی قابل تقسیم است:

- کاربردهای نوشتاری
- کاربردهای گفتاری

از کاربردهای نوشتاری آن می‌توان به استخراج اطلاعاتی خاص از یک متن، ترجمه یک متن به زبانی دیگر و یا یافتن مستندات خاص در یک پایگاه داده نوشتاری (مثلا یافتن کتاب‌های مرتبط به هم در یک کتابخانه) اشاره کرد. از نمونه‌های کاربردی گفتاری پردازش زبان می‌توان به سیستم‌های پرسش و پاسخ انسان با رایانه، سرویس‌های خودکار ارتباط با مشتری از طریق تلفن و یا سیستم‌های کنترلی توسط صدا اشاره کرد. در سال‌های اخیر این حوزه تحقیقاتی توجه پژوهشگران زیادی را به خود جلب کرده است و پژوهش‌های قابل ملاحظه‌ای در این زمینه صورت گرفته است. از لحاظ رده بندی، علم پردازش زبان طبیعی از شاخه‌های هوش مصنوعی به حساب می‌آید. اطلاعات و دانشی که در پردازش زبان طبیعی از آنها استفاده می‌شود به شش رده مختلف تقسیم بندی می‌شوند:

- آوا شناسی که به تشخیص آواها و صداها و بازشناسی گفتار می‌پردازد.
- ریخت شناسی که به ساختارهای کلمات و ریشه‌یابی واژگان می‌پردازد.
- نحو که به ارتباط کلمات به همدیگر و مباحث دستوری آنها در گروه‌ها و جملات می‌پردازد.
- معناشناسی که در آن به ارتباطات معنایی کلمات پرداخته می‌شود.
- عمل‌گرایی که کاربردهای زبان برای رساندن یک مطلب به مخاطب یا مخاطبان، در حالت عملی و یا در نوشتار و گفتار طبیعی می‌پردازد.
- مباحثه که به ارتباطات کلی یک زبان فرای یک یا چند جمله خاص می‌پردازد.

بر همین مبنا الگوریتم‌های بسیاری برای رسیدن به برنامه‌هایی هوشمندتر توسط دانشمندان و پژوهشگران علوم رایانه، زبانشناسی و ریاضیدانان، طراحی و پیشنهاد شده است. به عنوان مثال الگوریتم‌های الگوی مارکوف و الگوی مخفی مارکوف و نیز تلاش‌های چندین ساله نوام چامسکی در این راه، نمونه خوبی برای این امور است. روز به روز بر پیشرفت‌های دانشمندان در این امر افزوده شده و دانشمندان در سراسر دنیا سعی بر بهبود روش‌ها و پیاده سازی این روش‌ها در زبان‌های بومی خودشان هستند.