

Doğal dil işleme

Vikipedi, özgür ansiklopedi

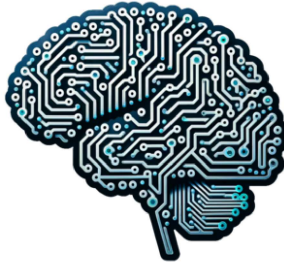


Bu maddede kaynak listesi bulunmasına karşın **metin içi kaynakların yetersizliği nedeniyle bazı bilgilerin hangi kaynaktan alındığı belirsizdir**. Lütfen kaynakları uygun biçimde metin içine yerleştirerek maddenin geliştirilmesine (https://tr.wikipedia.org/w/index.php?title=Do%C4%9Fal_dil_i%C5%9Fleme&action=edit) yardımcı olun. (Ekim 2021) *(Bu şablonun nasıl ve ne zaman kaldırılması gerektiğini öğrenin)*

Başlığın diğer anlamları için Nörolinguistik Programlama sayfasına bakınız.

Makale serisidir

Yapay zekâ



Gelişim süreci

[Genişlet]

Yaklaşımlar

[Genişlet]

Felsefe

[Genişlet]

Tarihçe

[Genişlet]

Teknoloji

[Genişlet]

Sözlük

[Genişlet]

G · T · D (https://tr.wikipedia.org/w/index.php?title=%C5%9Eablon:Yapay_zek%C3%A2&action=edit)

Doğal Dil İşleme, yaygın olarak **NLP** (*Natural Language Processing*) olarak bilinen yapay zekâ ve dilbilim alt kategorisidir. Türkçe, İngilizce, Almanca, Fransızca gibi doğal dillerin işlenmesi ve kullanılması amacı ile araştırma yapan bilim dalıdır.

Uzman Sistemler ve Doğal Dil İşleme

NLP yani **Doğal Dil İşleme**, doğal dillerin kurallı yapısının çözümlenerek anlaşılması veya yeniden üretilmesi amacını taşır. Bu çözümlemenin insana getireceği kolaylıklar, *yazılı dokümanların otomatik çevrilmesi, soru-cevap makineleri, otomatik konuşma ve komut anlama, konuşma sentezi, konuşma üretme, otomatik metin özetleme, bilgi sağlama* gibi birçok başlıkla özetlenebilir. Bilgisayar teknolojisinin yaygın kullanımı, bu başlıklardan üretilen uzman yazılımların gündelik hayatımızın her alanına girmesini sağlamıştır. Örneğin, tüm kelime işlem yazılımları birer imlâ düzeltme aracı taşır. Bu araçlar aslında yazılan metni çözümleyerek dil kurallarını denetleyen **doğal dil işleme** yazılımlarıdır.

Batı dillerinde SAPI (Microsoft şirketinin konuşma sentezleyici üretmek amacı ile satışa sunduğu geliştirici program) tabanlı Konuşma sentezleyici bileşenleri, yazılımcıların multimedia (çoklu ortam) sunuları hazırlamaları için hizmete sunulmuştur.

Konuşma ve komut anlama yazılımları ise gelecekte insan ve bilgisayar arasındaki klavye, fare gibi veri girişi aygıtlarını ortadan kaldıracak yazılımlardır. Bu gelişmeler makine-insan iletişimde yeni ve devrimci değişimlere yol açacak ve bilgisayarların daha çok insan tarafından kabul görmesine yol açacaktır.

Yapay Zekâ ve Doğal Dil İşleme

Gelecekte, konuşma sentezleyiciler ve konuşma anlama alanındaki gelişmeler ve makine-insan iletişiminin gelişmesi, insanın makineden beklentilerini yükseltecektir. İnsanlar makinelere kendisini anlamalarını isteyecek, karmaşık kullanımı olan makineler pazar bulamayacaktır. Giderek gelişen ve insanı anlayan makinelerin daha zeki olması insanın yaşam kalitesini yükselteceğinden, vazgeçilmez olması kaçınılmazdır. Zeki makine kavramı, yapay zekâ çalışmalarının hızlanmasına yol açmıştır. Geleceğin en önemli sektörlerinden biri olan yapay zekâ ile insanın iletişim kuracağı tek araç dildir.



Bir web sayfasında müşteri hizmetleri sağlayan otomatik sanal asistan, doğal dil işlemenin bir örneğidir.

Dil, insanlığının uygarlaşmasını sağlamakla kalmamış, onun zekâsının doğada daha önce görülmemiş şekilde parlamasını sağlamıştır. Kültür dediğimiz insanlık birikimi, **dil** kullanan ve iletişim kuran insanın sosyalleşme sürecinin ürünüdür.

Dilin Matematik Modeli

Dilin işlenmek üzere çözümlenebilmesi için, matematik modelinin oluşturulması gerekmektedir.

Genişletilmiş Geçiş Ağları

ATN Genişletilmiş Geçiş Ağları (*Augmented Transition Network*), Woods tarafından 1970 ve 1973 yılları arasında geliştirilmiş bir yaklaşımdır.

Genişletilmiş geçiş ağları (GGA) üç bileşenden oluşur:

- En az *başlangıç* ve *son* (/s) durumları olan sonlu sayıdaki durumlar kümesi,
- Belli bir metindeki mümkün olan harflerden oluşan alfabe (e),
- Sonlu sayıdaki bir durumdan diğer bir duruma geçişi sağlayacak geçişler kümesi.

Genişletilmiş geçiş ağlarında, bir durumdan diğer bir duruma geçmek için gerekli harf okunur ve bu harf geçilecek olan duruma geçmek için gereken harfle karşılaştırılır; uygun ise diğer duruma geçilir. Geçiş ağlarında doğru bir yol, bir başlangıç durumundan başlayıp, son duruma ulaşan geçişler sağlandığında tamamlanır. Harflerin birbirine eklenmesiyle oluşan metin, ağın kabul etmesi için verilen metin ise, bu metin ağ tarafından kabul edilmiş demektir.

Yanda: "Bal" metnini kabul eden Genişletilmiş Geçiş Ağı.

Fonetik ve fonoloji

Fonetik, konuşulurken, dil, gırtlak, ses telleri, damak, dişler ve dudaklar ile çıkarılan sesleri ve bu seslerin dil ile olan ilişkilerini tanımlamak için kullanılan bir terimdir. Doğal dillerde anlam ayırıcı olarak kullanılan en küçük ses **fondur** (*phon*) dur. Fonetik terimi bu kökten gelmektedir.

Fon kavramı evrensel değildir ve her dilde farklı seslere karşılık gelir. Farklı dillerdeki **fonların** tek ortak özelliği ayırıcı temel sesler olmalarıdır. Sesle ifade edilen dili, yani konuşmayı kaydetmek için yazı icad edilmişti. Konuşmayı yazı ile ifade etmek için ses birim veya **fonları** harflerle eşleştirmek gerekliydi. Bazı dillerde, örneğin Türkçe, Fince ve Japoncada, sesbirimler doğrudan harflere karşılık gelmektedir. Bu tip dillere fonetik diller denir. İngilizce, Almanca, Fransızca gibi dillerde ise **Fonlar** harflere karşılık gelmezler. Bu yaklaşımın yerine uluslararası olarak geçerliliği olan fonetik bir alfabe ses birimleri ifade etmek için kullanılır. Ses birimlerin simgesel olarak ifade edilmesi sonucu oluşan simgeler fonem (phoneme) olarak adlandırılır. Bir başka deyişle aslında fonemlerin seslendirilmesiyle ses birimler (phon) oluşur.

Dildeki ses birimler belirlenirken iki yaklaşım kullanılır. Bunlar,

- Parçalı sesbirimler (segmental) ve,
- Parçalarüstü ses birimler (supra-segmental, prosodic) dir.

Dilin morfolojisi

Dil bilime terim olarak 1859 yılında August Schleicher tarafından kazandırılan morfoloji, dilde biçimi oluşturan öğelerin türlerini tanımlamak ve özetle dil bilgisi kuralları denen biçimsel öğelerin sınıflandırmasını yapmaktır.

Morfolojik çözümlemede analitik yaklaşımlar

Doğal dil işleme çalışmalarında anlam bütünsel çözümleme yapabilmek için, bazı yaklaşımlar belirmiştir. Bu yaklaşımlar aşağıdaki süreçlerden oluşur.

Sözdizimsel (sentaktik) analiz

Sözdizimsel analiz, sözdizimini (syntax) veya cümleyi oluşturan morfolojik öğelerin hiyerarşik kurallara uyumunu karşılaştırarak ölçümlemektir. Böylece söz dizimin anlamlı olup olmadığının ölçülebilmesi için düzenleyici bir süreç gerçekleşmiş olur.

Türkçede cümleler en genel şekliyle özne, nesne ve yüklem bileşenlerinden oluşur. Cümleye eklenmek istenen anlamlar arttıkça cümleler, özne, yer tamlayıcısı, zarf tamlayıcısı, nesne ve yüklem gibi bileşenleri içerir. Ayrıca cümlelerin anlamını kuvvetlendiren cümle dış bileşenler de (bağlaç, edat, vb) cümlede bulunabilir. Bunlara örnek olarak "ile, için, ama, çünkü" kelimeleri verilebilir. Türkçede özne ile yüklem cümlelerin temel bileşenleridir ve genelde tüm cümlelerde yer alırlar. Yer tamlayıcısı, zarf tamlayıcısı, nesne gibi bileşenler bazı cümlelerde yer almayabilirler veya bazı cümlelerde sadece biri, bazılarında sadece ikisi bulunabilir. Bu bileşenlerin cümle içindeki sıralanışları da değişebilir.

Bilgisayarla doğal dilin modellenmesinde anlamsal analizden önce kelimelerden oluşturulan yapının cümle olup olmadığının test edilmesi faydalıdır. Bu işlem sentaktik eşleştirme işleminde anlamsız eşleşmelerin önlenmesine faydalı olur.

Yandaki Şekil : *Sözdizimsel Analiz.*

Simgeler: **Ö:** özne, **D:** dolaylı tümleş, **Z:** zarf tümleci, **N:** nesne, **Y:** yüklem, **İG:** isim grubu, **SG:** sıfat grubu, **İN:** isim nesnesi, **SN:** sıfat nesnesi, **DZ:** diğer zarflar, **S:** sıfat, **İ:** isim, **ZB:** zaman belirteçleri, **T:** tamlayan, **TN:** tamlanan, **ZM:** zamir, **NE:** nesne eki, **TE:** tamlayan eki, **TNE:** tamlanan eki, **KE:** kip eki, **ZE:** zaman eki, **DE:** dolaylı tümleş eki, **EF:** ek fiil

Anlambilimsel (semantik) analiz

Anlambilimsel analiz, sözdizimini oluşturan morfolojik öğelerin ayrılması, yani sözdizimsel analiz ile anlam taşıyan kelimelerin sınıflandırılması işleminden sonra gelen anlamlandırma veya anlama sürecidir. Bu süreçte anlam taşıyan kelimelerin, ekler ve cümle hiyerarşisi içindeki konumlarının saptanması sayesinde birbirleri ile ilişkileri kurulabilir. Bu ilişkiler **anlam çıkarma**, **fikir yürütme** gibi ileri seviye bilişsel fonksiyonların oluşturulmasında ham bilgi olarak kullanılacaktır.

Yapay konuşma

Morfolojik çözümleme aşamalarından sonra **sözdizimsel kurgu** veya **yapay konuşma** süreci ile yapay zekâ ya veya uzman sistemlere iletişim becerisi kazandırılacaktır. Sözdizimsel çözümlemenin tersi süreçlerden oluşan birleştirme sürecinde, önceki süreçlerde ele geçen bilgi yine morfolojik kurallar dahilinde birleştirilir.

Ayrıca bakınız

- Doğal dil üretme

Kaynakça

- Vasif Nabiyeu - Yapay Zekâ: Problemler, Yöntemler, Algoritmalar, 764 say., Seçkin, Ankara, 2005
- Devrim Çamoğlu - **D.U.Y.G.U.** Projesi araştırma tezleri.
- Ünal Çakıroğlu - (KTU) Şekiller, Sözdizimsel Analiz ve matematik model bölümü

Dış bağlantılar

- ITÜ Doğal Dil İşleme Takımı (<http://ddi.itu.edu.tr/>) 23 Ekim 2020 tarihinde [Wayback Machine](https://web.archive.org/web/20201023081720/http://ddi.itu.edu.tr/) sitesinde [arşivlendi](https://web.archive.org/web/20201023081720/http://ddi.itu.edu.tr/) (<https://web.archive.org/web/20201023081720/http://ddi.itu.edu.tr/>).
- Stanford Üniversitesi Doğal Dil İşleme Öbeği (<http://nlp.stanford.edu/>) 29 Kasım 2005 tarihinde [Wayback Machine](https://web.archive.org/web/20051129025046/http://nlp.stanford.edu/) sitesinde [arşivlendi](https://web.archive.org/web/20051129025046/http://nlp.stanford.edu/) (<https://web.archive.org/web/20051129025046/http://nlp.stanford.edu/>).
- Survey of the State of the Art in Human Language Technology (<https://web.archive.org/web/20051124031210/http://cslu.cse.ogi.edu/HLTsurvey/>)
- Natural Language Processing Group at the Johns-Hopkins University (<https://web.archive.org/web/20051124031103/http://nlp.cs.jhu.edu/nlp/>)
- DNLP - Dalhousie Natural Language Processing Group (<https://web.archive.org/web/20031207013200/http://www.cs.dal.ca/~nlp/>)
- 2004 International Workshop on Natural Language Understanding and Cognitive Science (<https://web.archive.org/web/20040402005315/http://www.iceis.org/workshops/nlu/nlu2004-cfp.html>)
- CLAC: Computational Linguistics At Concordia (<http://www.cs.concordia.ca/CLAC/>) 4 Aralık 2005 tarihinde [Wayback Machine](https://web.archive.org/web/200511204085823/http://www.cs.concordia.ca/CLAC/) sitesinde [arşivlendi](https://web.archive.org/web/200511204085823/http://www.cs.concordia.ca/CLAC/) (<https://web.archive.org/web/200511204085823/http://www.cs.concordia.ca/CLAC/>).
- TCC: Cognitive and Communication Technologies (TCC) at ITC-Irst (<https://web.archive.org/web/20040613021113/http://tcc.itc.it/>)
- YTÜ Doğal Dil İşleme Araştırma Grubu (<https://web.archive.org/web/20190831034622/http://www.kemik.yildiz.edu.tr/>)
- Fatih Ü. Doğal Dil İşleme Grubu (<https://web.archive.org/web/20110302062810/http://nlp.ceng.fatih.edu.tr/tr/>)
- Cognitive Science Society of Trakya (CSST) (<http://tbbt.trakya.edu.tr/>) 10 Mart 2012 tarihinde [Wayback Machine](https://web.archive.org/web/20120310070732/http://tbbt.trakya.edu.tr/) sitesinde [arşivlendi](https://web.archive.org/web/20120310070732/http://tbbt.trakya.edu.tr/) (<https://web.archive.org/web/20120310070732/http://tbbt.trakya.edu.tr/>).

G · T · D (https://tr.wikipedia.org/w/index.php?title=%C5%9Eablon:Bilgisayar_bilimi&action=edit)	Bilgisayar biliminin alt dalları	[Genişlet]
Otorite kontrolü	LCCN: sh88002425 (https://id.loc.gov/authorities/subjects/sh88002425) · NDL: 00562347 (https://id.ndl.go.jp/auth/ndlna/00562347) · NKC: ph427562 (https://aleph.nkp.cz/F/?func=find-c&local_base=aut&ccl_term=ica=ph427562&CON_LNG=ENG) · NLI: 987007536703305171 (http://olduli.nli.org.il/F/?func=find-b&local_base=N LX10&find_code=UID&request=987007536703305171)	

"https://tr.wikipedia.org/w/index.php?title=Doğal_dil_işleme&oldid=31169798" sayfasından alınmıştır

-