



ERZURUM
TEKNİK ÜNİVERSİTESİ
2010

DOĞAL DİL İŞLEMEYE GİRİŞ ÖDEVİ

AD VE SOYADI: MECİD BEDRAN ELBAŞA

ÖĞRENCİ NO: 190303077

Doğal Dil İşleme Teknolojileri

❖ Doğal Dil İşlemenin Temel Prensipleri:

Dil binlerce yıldır insan iletişiminin temeli olmuştur ve dille ilgili tüm gelişmeler, insan toplumu tarihinde önemli kilometre taşlarıdır.

Bundan 5000 yıl önce yazının bulunması bugün yaşadığımız dünyanın temellerinin atılmasında önemli rol oynamıştır.

Bu nedenle bilgisayarlara ve makinelere insan benzeri dilleri anlama ve üretme yeteneğinin kazandırılması şüphesiz geleceğimiz için büyük bir işarettir. Bunu başardığımızı iddia edemeyiz ama doğal dil işlemedeki (NLP) hızlı gelişmeler bizi bu noktaya çok yaklaştırdı.

Doğal Dil İşleme (NLP) Nedir?

Çoğumuz Amazon'un Alexa'sı Google Asistanı veya Apple'ın Siri'si gibi akıllı asistanlar kullanmışızdır ve bu asistanlarla bir programlama dilinde değil, doğal dilimizde, hepimizin iletişim kurduğu dilde etkileşime giriyoruz.

Doğal dil eski çağlardan beri insanlar arasındaki temel iletişim aracı olmuştur ancak bilgisayarlar verileri yalnızca ikili (0 ve 1) olarak işleyebilir ve dil verilerini ikili olarak temsil edebilmek de makinelerin dili anlamasını nasıl sağlayabiliriz?

İşte bu noktada Doğal Dil İşleme (NLP) devreye giriyor.

NLP :

*İnsan dilini analiz etme, modelleme ve anlamaya yönelik yöntemlerle ilgilenen bir bilgisayar bilimi alanıdır.

*Doğal dil işleme (NLP), dil bilimini makine öğrenimi, derin öğrenme ve yapay sinir ağları dahil olmak üzere bilgisayar biliminin birçok alanıyla birleştiren bir disiplindir.

NLP, makinelerin insanın yazılı ve sözlü dilini anlamasını ve yeniden üretmesini amaçlamaktadır. Ayrıca insan-insan etkileşimi gibi insan-makine iletişimini güçlendirmeye çalışıyoruz.

NLP, bilişsel bilimin en önemli ve zorlu alanlarından biri olarak kabul edilir. Akıllı cihazların ve sistemlerin performansını artırmak önemlidir.

* Doğal Dil İşleme (NLP), hesaplamalı dilbilim algoritmalarını makine öğrenimi ve derin öğrenmeyle ilgili istatistiksel algoritmalarla bütünleştirmeyi amaçlayan eski bir bilimdir.

Bu entegrasyon, makinelerin karmaşık dili ve anlamlarını anlamasını sağlamayı amaçlıyor.

NLP, arama motorları, yazı sistemleri, cep telefonu klavyelerindeki hataların düzeltilmesi, tıbbi araştırmalar, finans ve işletme alanları gibi çeşitli uygulamalarda kullanılmaktadır.

NLP'de "doğal dil", çeşitli yerel lehçeler ve konuşma dili varyasyonları da dahil olmak üzere önceden planlanmadan veya kurallar belirlenmeden gelişen İngilizce, Türkçe ve Fransızca gibi insan dillerini ifade eder.

Öte yandan Python, Java ve C# gibi programlama dilleri gibi yapay diller de insanlar tarafından önceden tanımlanmış kurallar ve terminolojilerle oluşturulur.

Bu diller tipik olarak insanlar arasında kullanılmaz, ancak insanlar ve bilgisayarlar arasında bir iletişim aracı olarak hizmet eder, açıklık avantajı ve dilsel belirsizlik eksikliği veya komutları için alternatif anlam olasılığı sunar.

Ancak bu, bu makalenin ilerleyen kısımlarında ele alacağımız çeşitli zorlukları da beraberinde getirir.

Doğal Dil İşleme Türleri (NLP):

Yüzlerce kullanım örneği ve uygulaması olan, günlük hayatımızı kolaylaştıran onlarca Doğal Dil İşleme (NLP) türü ve tekniği vardır.

Aşağıda iki ana Doğal Dil İşleme (NLP) türüne genel bir bakış yer almaktadır:

1) Doğal Dil Anlama (NLU):

Doğal Dil Anlama, Doğal Dil İşleme'nin (Hesaplamalı Dilbilim) insan dilinin bilgisayar tarafından anlaşılmasıyla ilgilenen bir alt alanıdır.

Niyetler, adlandırılmış varlıklar veya insan dilinin diğer yönleri şeklini alabilen uygun anlamı çıkarmak için insan dilinin işlenmesini içerir.

2) Doğal Dil Üretimi (NLG):

Doğal Dil Oluşturma, Yapay Zeka ve Doğal Dil İşlemenin (NLP) bir alt alanıdır ve yapılandırılmış verilerden insan benzeri çıktıya benzeyen metin veya konuşmanın otomatik olarak üretilmesine odaklanır.

NLG sistemleri verileri analiz eder, bağlamını anlar ve uygun tutarlı anlatılar veya metinler üretir ve bu genellikle raporlar, makaleler, özetler veya özel mesajlar biçiminde olur ve NLG, gazetecilik, iş zekası, sohbet robotları, içerik oluşturma, süreç basitleştirme ve veriye dayalı içgörüler için etkili iletişimin sağlanması dahil olmak üzere çeşitli sektörlerde uygulamalar bulur.

Temel kavramlar ve işlemler:

1) Konuşma tanıma (Speech recognition):

Konuşma tanıma, bir programın insan konuşmasını tanımasını, anlamasını ve metne çevirmesini sağlayan bir teknolojidir. Konuşma tanıma süreci genellikle makine öğrenimi ve doğal dil işleme (NLP) kullanılarak gerçekleştirilir.

Genellikle konuşma tanıma programları iki parametreye göre değerlendirilir:

Konuşma Tanıma Hızı: Programın konuşmacıya ayak uydurması için geçen süre analiz edilerek değerlendirilir.

Konuşma Tanıma Doğruluğu: Konuşulan kelimelerin dijital verilere dönüştürülmesi sırasında hataların yüzdesi belirtilerek belirlenir.

Bu teknoloji, ses verilerinin metinsel verilere dönüştürülmesi amacıyla günümüzde yaygın olarak kullanılmakta olup, yazmaya gerek kalmadan sesli komutlar verilmesi veya sözlü olarak soru sorulması amaçlanmaktadır.

Bu teknoloji, birden fazla lehçe, yanlış telaffuz, benzer kelimeler, dil hataları ve çok daha fazlası gibi çeşitli zorluklarla karşı karşıyadır. Ancak bu zorluklara rağmen Alexa ve Siri gibi uygulamalarında önemli ilerlemelere tanık oluyoruz. Konuşma tanıma teknolojisi, diğer birçok kullanım durumunun yanı sıra sanal asistan sistemlerinde, müşteri hizmetlerinde ve şikayetlerinin ele alınmasında kullanılabilir.

2) Dil modelleme (Language modeling):

Bu, önceki kelimelerin geçmişine dayanarak bir cümledeki bir sonraki kelimenin ne olacağını tahmin etme görevidir.

Bu görevin amacı belirli bir dilde bir dizi kelimenin görünme olasılığını belirlemektir.

Dil modelleme, konuşma tanıma, optik karakter tanıma, el yazısı tanıma, makine çevirisi ve yazım denetimi gibi çeşitli sorunlara çözüm üretmek için kullanışlıdır.

3) Metin sınıflandırması (Text Classification):

Bu, metni içeriğine göre bilinen bir kategoriler kümesine atama görevidir.

Metin sınıflandırma, Doğal Dil İşleme'de en yaygın görevdir ve spam e-postaların tespitinden duygu analizine kadar çeşitli araçlarda kullanılır.

1. Konu Sınıflandırması(Topic Classification):

Bu teknoloji, insanların okuması zor olan büyük miktarda metinle uğraşırken, karmaşıklığı nedeniyle veya önemli ölçüde zaman ve emekten tasarruf sağlamak için metni belirli konulara göre kategorize ederken faydalıdır.

2. Niyet Tespiti(Intent Detection):

Bu teknolojiye yapay zeka algoritmaları metinlerin veya konuşmaların ardındaki hedefleri ve niyetleri tanıyor. Bu teknoloji, birçok şirketin müşteri hizmetleri veya satış gibi departmanları iyileştirmesine ve yönetmesine yardımcı olur.

3. Yazarlık Atfı(Authorship Attribution):

Bu teknoloji, yazarları hakkında soru işaretleri uyandıran yaratıcı çalışmalarla ilgilendiği için son derece ilgi çekicidir. Algoritmayı söz konusu yazarların çeşitli eserleriyle besleyip ardından ihtilafli eseri sunarak eseri gerçek yazarına atfedebilir.

4) Konu modelleme (Topic modeling):

Bu, geniş bir belge kümesi içindeki yapısal konuların tespit edilmesi görevidir. Konu modelleme, metin madenciliği için yaygın bir araçtır ve edebiyattan biyoenformatiğe kadar çok çeşitli alanlarda kullanılmaktadır.

5) Metin Çıkarma (Text Extraction):

Doğal Dil İşleme'nin (NLP) bir diğer önemli tekniği Metin Çıkarma'dır. Bu teknik, kitap veya içerik gibi büyük ve kapsamlı metinlerde bilgi veya belirli bölümleri ararken çok fazla zaman ve çaba tasarrufu sağlamaya yardımcı olur.

Başlıca metin çıkarma teknolojisinin kapsamına giren birçok önemli algoritma ve teknik vardır:

1) Anahtar Kelime Çıkarma(Keyword Extraction):

Bu teknik, bir metinden veya bir dizi metinden önemli anahtar kelimeleri ve ifadeleri otomatik olarak çıkarabilir. Arama Motorlarında yaygın olarak kullanılır.

2) Adlandırılmış Varlık Tanıma (Named Entity Recognition):

Bu son derece önemli teknik, makinelerin, ünlü bir kişinin adı, bir film veya dizinin adı, hatta bir yerin adı gibi belirli anlamlara sahip kelimeleri çıkararak doğal dile daha iyi yanıt vermesini sağlar.

6) Makine Çevirisi (Machine Translation):

Doğal dil işlemenin tarihi hakkındaki tartışmamızdan bildiğimiz gibi makine çevirisi, NLP'nin çözüm bulmaya çalıştığı ilk sorunlardan biriydi. "Google Translate" gibi araçlarda görüldüğü gibi yüksek düzeyde profesyonellik ve doğruluk elde edilmesine rağmen, insan çevirisinin doğruluğuna ulaşmak için hala yapılması gereken çok iş var.

7) Duygu Analizi (Sentiment Analysis):

Duygu analizi, mesajın duygusal tonunun olumlu, olumsuz veya nötr olup olmadığını belirlemek için dijital metni analiz etmeyi içerir.

Günümüzde şirketler, e-postalar, müşteri desteği sohbet kayıtları, sosyal medya yorumları ve incelemeler dahil olmak üzere çok büyük miktarda metinsel veriye sahiptir.

Duygu analizi araçları, yazarın belirli bir konudaki duruşunu otomatik olarak belirlemek için bu metin verilerini tarayabilir.

Şirketler müşteri hizmetlerini geliştirmek ve marka itibarını artırmak için duygu analizi içgörülerini kullanıyor.

8) Otomatik Metin Özetleme (Automatic Text Summarization)

Özetleme süreçlerinin otomatikleştirilmesi, doğal dil işleme uzmanlarının onlarca yıldır öncelikleri arasında yer alıyor.

Bu teknoloji, insanlara bu zorlu görevi yerine getirirken önemli miktarda zaman ve çaba tasarrufu sağlayabilir.

Artık araştırmacılar, çeşitli disiplinlerdeki düzinelerce bilimsel çalışmayı, bunları işlemek ve ele almak için önemli bir çaba harcamadan özetleyebiliyor.

Bu, araştırmacıların deneyler yürütmesinde ve araştırmalarını sürdürmelerinde önemli miktarda zaman tasarrufu sağlamasına yardımcı oldu.

9) Konuşma bölümlerinin sınıflandırılması (Part of speech tagging):

Makinelerin dil veya metinlerle uğraşırken konuşmanın bölümlerini tanımlaması veya sınıflandırması gerekir.

Bu, cümlelerin fiiller, sıfatlar, isimler, zarflar ve diğerleri gibi bileşenlerine bölündüğü Konuşma Kısmı Etiketleme tekniği aracılığıyla gerçekleştirilir.

Bu, makinenin, insanlar tarafından kendisine gönderilen metni veya konuşmayı doğru şekilde anlamasına yardımcı olur.

10) Kelime Anlamı Belirsizliğinin Giderilmesi (Word Sense Disambiguation):

Bazı kelimelerin birden fazla anlamı ve hatta çeşitli gramer biçimleri (isimler, fiiller vb.) vardır ve bu da makineler için kafa karışıklığına neden olabilir. Bu sorunla baş etmek, belirli algoritmaların kullanılmasını gerektirir.

Kelime Anlamı Belirsizliği Giderme (WSD) teknolojisi cümlelerin bağlamını ele alır. Bu sayede, karmaşık istatistiksel modeller kullanarak, olasılığı daha yüksek olan kelimeleri tercih ederek bir kelimenin kastedilen anlamını belirleyebilir.

❖Endüstri Uygulamaları:

Finans sektöründe NLP kullanımı

Finans sektöründe Nöro-Dilsel Programlama (NLP) kurumların riskleri yönetme dolandırıcılığı tespit etme ve yatırım stratejilerini geliştirme biçimlerinde bir dönüşüm yarattı.

Kredi riski değerlendirme modelleri borçluların kredi değerliliğini doğru bir şekilde değerlendirmek için geçmiş verilerden yararlanır ve bu borç verenlerin bilinçli kararlar almasına olanak tanıyarak temerrüt risklerini ve potansiyel mali kayıplarını azaltır.

Dolandırıcılık tespiti Nöro-Dilsel Programlamanın (NLP) finans alanındaki bir diğer önemli uygulamasıdır. Makine öğrenimi algoritmaları işlem modellerini ve kullanıcı davranışını analiz ederek şüpheli etkinlikleri ve potansiyel dolandırıcılık işlemlerini hızlı bir şekilde tespit edebilir.

Bu finansal kuruluşların hizmetlerine olan güveni korurken müşterilerini ve varlıklarını korumalarına yardımcı olur.

Nöro-dilsel programlama (NLP) doğal dili anlamak ve analiz etmek için yapay zekayı kullanan yeni bir teknolojidir. Bunun finans sektörü üzerinde çeşitli uygulama alanlarında önemli bir etkisi olacaktır.

NLP'nin finans sektöründeki olası kullanımları şunları içerir:

1) Mali duyarlılık ve haber ve sosyal medya analizi:

NLP'yi kullanarak finansal ve ekonomik raporları izleyin ve analiz edin.

Yatırımcı görüşlerini, fikirlerini ve pazar eğilimlerini anlamak için sosyal medyayı izleyin.

Hisse senedi piyasasında alım satım, belirli şirketler hakkındaki bir dizi bilgiye bağlıdır ve bu bilgi, hisselerin satın alınacağını, tutulacağını veya satılacağını belirleyen bir dizi eylemin oluşturulmasına yardımcı olur.

Bu analiz şirketlerin üç aylık mali raporlarına veya analistlerin raporlarında şirketler hakkında söylediklerine dayandırılabilir gibi sosyal medyadan da gelebilir.

Sosyal medya analizi, sosyal medya gönderilerinin izlenmesine ve potansiyel ticaret fırsatlarının belirtilmesine yardımcı olur.

Örneğin bir CEO istifa ederse bu duygular genellikle olumsuz olur ve bu durum şirketin hisse senedi fiyatını olumsuz etkileyebilir ve öte yandan CEO'nun performansı iyi değilse ve piyasalar onun istifasını memnuniyetle karşılıyorsa bu durum hisse senedi fiyatlarında artışa neden olabilir hisse senedi fiyatı.

Ticaret için bu bilgileri sağlayan şirketlere örnek olarak DataMinr ve Bloomberg verilebilir.

2) Mali raporların analizi:

Mali raporları incelemek ve önemli bilgileri çıkarmak için NLP'yi kullanın ve iş ve piyasa performansını tahmin etmek için mali tablo analizi.

3) Müşteri hizmetleri ve dil etkileşimi:

Müşteri sorularını ve desteğini karşılamak için otomatik sohbet sistemleri geliştirin.

Basit bankacılık görevlerini gerçekleştirmek ve müşteri sorularını yanıtlamak için ses robotlarını kullanıyoruz.

4) Dolandırıcılık tespiti ve risk yönetimi:

Dolandırıcılığı tanımlamak amacıyla mali belgelerdeki olağandışı kalıpları tespit etmek için NLP'yi kullanın.

Potansiyel riskleri anlamak ve verilere dayalı kararlar almak için belgeleri analiz edin.

5) Pazar ve yatırım beklentileri:

Piyasa ve yatırım eğilimleri hakkında doğru tahminlerde bulunmak için haber ve raporları sürekli analiz ediyoruz.

6) Hisse senedi piyasasında duyarlılık analizi:

Hisse senedi piyasası üzerindeki etkisini anlamak için çevrimiçi finans forumlarında yatırımcı duyarlılığını araştırın.

Hisse senedi piyasasında alım satım, belirli şirketler hakkındaki bir dizi bilgiye bağlıdır ve bu bilgi, hisselerin satın alınacağını, tutulacağını veya satılacağını belirleyen bir dizi eylemin oluşturulmasına yardımcı olur.

Bu analiz şirketlerin üç aylık mali raporlarına veya analistlerin raporlarında şirketler hakkında söylediklerine dayandırılabilir de gelebilir.

7) Yatırım tahsisi ve portföy yönetimi:

Kişiselleştirilmiş yatırım tekliflerinin sağlam bir analizini sunuyor ve bunları finansal yönetim stratejilerine entegre ediyoruz.

Bunlar olası kullanımlardan sadece birkaçıdır ve model doğruluğu ve veri koruması gibi olası konuların dikkate alınması önemlidir.

8) Risk değerlendirmesi:

Kredi riski, başarılı bir kredi geri ödemesi şansını ölçmenin bir yoludur ve genellikle bireyin geçmiş harcamalarından ve kredi geri ödeme geçmişinden hesaplanır, ancak bu bilgi birçok senaryoda sınırlıdır.

Doğal dil işleme teknikleri, kredi riskini değerlendirmek için kullanılacak birçok veri noktası ekleyebilir.

Örneğin ticari kredilerde, girişimcilik yeteneği ve tutumu doğal dil işleme kullanılarak ölçülebilir ve bu yaklaşım Capital Float ve Microbnk tarafından kullanılmaktadır.

Benzer şekilde, borçlu tarafından sağlanan verilerdeki tutarsızlıklar da daha fazla inceleme için gündeme getirilebilir ve borç verenlerin ve borçluların kredi başvurusu sırasındaki duyguları gibi diğer daha incelikli hususlar da dahil edilebilir.

Çoğu zaman bireysel kredi sözleşmelerinde, kredi belgelerinden çeşitli bilgilerin elde edilmesi ve bunların daha sonra kredi riski modellerine beslenmesi gerekir.

Elde edilen bilgiler kredi riskinin belirlenmesine yardımcı olur ve bu belgelerden yanlış verilerin çıkarılması hatalı değerlendirmelere yol açabilir.

Sonuç olarak:

Finansal hizmetler, kamu şirketlerinin izlenmesinden yatırım bankacılığı işlem akışına kadar geniş bir yelpazeyi kapsayan çok çeşitli bir alandır.

Küresel olarak finansal hizmetler sektörünün gelecekte büyümesi bekleniyor.

NLP'nin diğer teknolojilerle etkin entegrasyonu finans sektöründe kullanımının verimliliğini artırabilir.

Sağlık sektöründe hastane kayıtları ve rapor analizi

Gelişmiş ekonomilerin çoğunda sağlık hizmetleri, Gayri Safi Yurtiçi Hasıla'nın (GSYH) önemli bir bölümünü oluşturur ve genellikle %10'u aşar. Büyük boyutu göz önüne alındığında, bu süreçleri ve sistemleri otomatikleştirmenin ve iyileştirmenin çok büyük faydaları vardır ve işte bu noktada Doğal Dil İşleme (NLP) devreye giriyor. Bir endüstri olarak sağlık hizmetleri, terapötik, önleyici, palyatif ve rehabilite edici bakıma yönelik hem ürünleri (ilaç ve ekipman gibi) hem de hizmetleri (danışma veya teşhis testleri) kapsar.

Sağlık hizmetleri büyük miktarlarda yapılandırılmamış metin verileriyle ilgilenir ve bu tür bağlamlarda sağlık sonuçlarını iyileştirmek için Doğal Dil İşleme kullanılabilir. Doğal Dil İşleme'nin yardımcı olabileceği geniş alanlar arasında tıbbi kayıtların, faturaların analizi ve ilaç güvenliğinin sağlanması yer alır. İlerleyen bölümlerde bu uygulamalardan bazılarını kısaca ele alacağız.

1) Sağlık Kayıtları ve Tıbbi Belgeler:

Çoğunlukla sağlık ve tıbbi verilerin önemli bir kısmı yapılandırılmamış metin formatlarında toplanır ve saklanır. Buna tıbbi notlar, reçeteler, sesli metinlerin yanı sıra patoloji ve radyoloji raporları da dahildir.

Bu, verileri ham haliyle arama, organize etme, inceleme ve anlama konusunda zorluklara yol açar. Standartlaştırılmış depolamanın olmayışı karmaşıklığı artırır. Doğal Dil İşleme (NLP), sağlık profesyonellerinin bu verileri daha etkili bir şekilde aramasına ve analiz etmesine yardımcı olabilir. Hatta soruları yanıtlamak için otomatik sistemler oluşturmak, ilgili hasta bilgilerini aramak için harcanan zamanı azaltmak gibi bazı iş akışı görevlerini bile otomatikleştirebilir.

2) Hastalara ve Faturalara Öncelik Vermek:

Hastanın durumunu ve çeşitli sağlık prosedürlerine ve muayenelerine öncelik verme ihtiyacını anlamak için doktorların notlarında Doğal Dil İşleme (NLP) teknikleri kullanılabilir.

Bu, gecikmeleri, idari hataları azaltabilir ve süreçleri otomatikleştirebilir. Benzer şekilde, tıbbi kodları tanımlamak için yapılandırılmamış notların analiz edilmesi ve bilgilerin çıkarılması, faturalandırma sürecini kolaylaştırabilir.

3) Farmakovijilans (Pharmacovigilance):

Farmakovijilans, ilaçların güvenliğini sağlamak için gerekli tüm faaliyetleri kapsar. Bu, ilaçlara karşı olumsuz reaksiyonların toplanmasını, tespit edilmesini ve izlenmesini içerir. Tıbbi prosedürler veya ilaçlarla istenmeyen veya zararlı etkiler ortaya çıkabilir ve ilacın amaçlandığı gibi çalıştığından emin olmak için bu etkilerin izlenmesi ve önlenmesi önemlidir.

Sosyal medya kullanımının artmasıyla birlikte sosyal medya mesajlarında bu yan etkilerden daha fazla bahsediliyor.

Bu yan etkilerin izlenmesi ve tanımlanması çözümün bir parçasıdır. Bu teknolojilerin bazılarını sosyal medyanın genel analizine odaklanan önceki bir makalede ele almıştık.

Sosyal medyanın yanı sıra tıbbi kayıtlara uygulanan Doğal Dil İşleme (NLP) teknikleri de farmakovijilansı kolaylaştırmaktadır.

4) Klinik Karar Destek Sistemleri:

Klinik Karar Destek Sistemleri (CDSS), sağlık profesyonellerinin sağlıkla ilgili kararlar almasına yardımcı olur. Buna muayeneler, teşhisler, tedavi ve izleme dahildir. Elektronik sağlık kayıtları, sütun planlı laboratuvar sonuçları ve operasyonel notlar dahil olmak üzere çeşitli metinsel veriler bu sistemler için girdi olarak kullanılabilir. Karar destek sistemlerini geliştirmek için tüm bu yönlerde Doğal Dil İşleme (NLP) kullanılmaktadır.

5) Sağlık Asistanları:

Sağlık asistanları ve konuşabilen robotlar, uzman sistemlerin ve doğal dil işlemenin çeşitli yönlerinden yararlanarak hasta ve bakıcı deneyimlerini geliştirebilir.

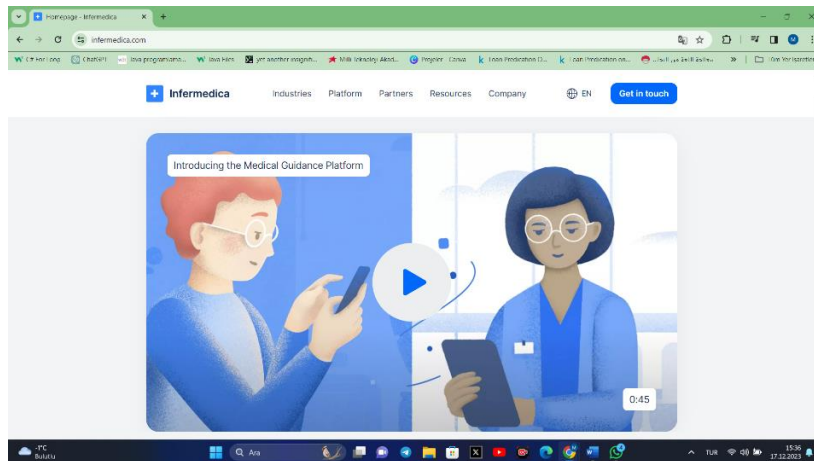
Örneğin Woebot gibi hizmetler, akıl hastalıkları ve depresyondan muzdarip hastaların moralinin korunmasına yardımcı olabilir ve Woebot, olumlu düşünceleri teşvik eden ilgili soruları sorarak doğal dil işlemeyi bilişsel terapiyle birleştirir.



Benzer şekilde sağlık asistanları, potansiyel tıbbi sorunları teşhis etmek için hastaların semptomlarını değerlendirebilir ve otomatik sohbet programları, teşhisin ciddiyetine ve aciliyetine bağlı olarak ilgili doktorlarla randevu ayarlayabiliyor ve böyle bir sistemin bir örneği (Buoy).

Bu sistemler aynı zamanda mevcut teşhis çerçeveleri kullanılarak kullanıcıların özel ihtiyaçlarına göre de oluşturulabilmektedir.

Infermedica, sohbet arayüzünün kullanıcı semptomlarını ortaya çıkarabildiği ve olasılıklarıyla birlikte olası hastalıkların bir listesini sunabildiği böyle bir çerçevenin bir örneğidir.



6) Ayrıca aşağıdakiler de dahil olmak üzere başka birçok kullanım alanı vardır:

- 1) Hasta muayenesi ve teşhis.
- 2) Tedavi yanıtını izle.
- 3) Hastane performans analizi.
- 4) İlaç etkileşimlerini ve yan etkilerini izleyin.
- 5) Epidemiyolojik araştırma.

Sonuç olarak:

Doğal dil işlemenin faydalı olabileceği bir dizi sağlık hizmeti uygulamasını tartıştık.

E-ticarete müşteri hizmetleri ve öneri sistemleri

Günümüz dünyasında e-ticaret alışverişi eş anlamlı hale geldi. Perakende mağazalarının sunduklarıyla karşılaştırıldığında zengin müşteri deneyimi, e-ticaretin büyümesini artırdı.

Bu hızlı büyümede makine öğrenmesi ve doğal dil işlemedeki son gelişmeler önemli rol oynadı.

Bu makalede, ürün kataloğu oluşturma, inceleme toplama ve öneri de dahil olmak üzere bu alandaki bazı önemli doğal dil işleme görevlerini tartışacağız.

1) E-ticaret Kataloğu:

Her büyük e-ticaret kuruluşunun kolayca erişilebilen bir ürün kataloğuna ihtiyacı vardır ve ürün kataloğu, işletmenin ilgilendiği veya kullanıcıların satın alabileceği ürünlerin yer aldığı bir veritabanıdır ve her ürün için ürün açıklama özelliklerini ve görsellerini içerir.

İlgili bilgileri içeren daha iyi ürün açıklamaları, müşterilerin katalog aracılığıyla doğru ürünü seçmesine yardımcı olur ve bu bilgiler aynı zamanda ürün arama ve önerilerde de yardımcı olabilir.

Mor rengini sevdiğinizi otomatik olarak bilen bir öneri motoru hayal edin! ve bu yalnızca motor, son satın alma veya arama etkinliklerinizin çoğunun mor renkli öğelerle ilgili olduğunu algılasa mümkündür.

Bunu başarmanın ilk adımı "mor"un ürünlerle bir renk özelliği olarak ilişkilendirildiğini belirlemektir ve bu bilginin otomatik olarak çıkarılmasına nitelik çıkarma denir. Ürün açıklamalarından özelliklerin çıkarılması, ilgili tüm ürün bilgilerinin indekslenmesini ve her ürün için doğru şekilde görüntülenmesini sağlayarak ürünün keşfedilebilirliğini artırabilir.

2) İnceleme Analizi:

Bir e-ticaret platformunun en göze çarpan kısmı, her ürün için kullanıcı inceleme bölümüdür ve incelemeler, kalite, kullanılabilirlik, diğer ürünlerle karşılaştırmalar, teslimat notları gibi yalnızca ürün özelliklerinden elde edilemeyecek şekilde ürüne farklı bir bakış açısı sağlar.

İncelemelerin tümü yararlı olmayabilir veya güvenilir kullanıcılardan gelmeyebilir. Ayrıca, belirli bir ürüne ilişkin birden fazla incelemenin manuel olarak işlenmesi zorlayıcı olabilir.

Doğal dil işleme (NLP) teknikleri, duygu analizi, incelemeleri özetleme ve bir incelemenin kullanılabilirliğini belirleme gibi görevleri gerçekleştirerek tüm incelemelerin kapsamlı bir görünümünü sağlar.

3) Ürün Arama:

E-ticaret arama sistemleri Google, Bing ve Yahoo gibi genel arama motorlarından farklıdır. Bir e-ticaret arama motoru, mevcut ürünlerle ve bunlarla ilişkili çeşitli bilgi türleriyle yakından bağlantılıdır.

Örneğin, normal bir arama motorunda, e-ticaret için organize satış ve inceleme verileri yerine, büyük ölçüde serbest metin verileriyle (haber makaleleri veya bloglar gibi) ilgileniriz. "Düğün için mor kareli gömlek" diye arama yapabiliriz ve e-ticaret arama motorunun onu getirmesi gerekir.

4) Ürün Önerileri:

Öneri motoru olmadan herhangi bir e-ticaret platformu eksik kalır. Müşteriler, platformun seçimlerini akıllıca anlamasını ve daha sonra satın alınacak ürünleri önermesini seviyor. Aslında müşterinin alışverişiyle ilgili düşüncelerini düzenlemesine yardımcı oluyor ve daha iyi bir alışveriş deneyimine katkıda bulunuyor.

İndirimli ürün, aynı marka veya tercih edilen özelliklere sahip ürün önerileri, gerçekten müşterinin sitede ilgisini çekerek daha fazla zaman geçirmesini sağlayabilir. İşlem tabanlı öneri motorlarının yanı sıra, ürün içeriğinin ve yorumların metinsel yapısına dayalı olarak geliştirilen zengin bir algoritma dizisi de bulunmaktadır ve bu öneri sistemlerini oluşturmak için NLP kullanılır.

Bu genel bakışla, doğal dil işlemenin e-ticaretteki rolünü daha ayrıntılı olarak keşfetmeye hazırız.

E-ticaret için arama motorlarının oluşturulmasında nasıl kullanıldığını başlatalım.

A. E-ticaret Araması:

Müşteriler istedikleri ürünleri bulmak ve hızlı bir şekilde satın almak için bir e-ticaret sitesini ziyaret ederler ve İdeal olarak, arama özelliği hızlı, doğru olmalı ve müşterinin ihtiyaçlarına mükemmel şekilde uygun sonuçlar getirmelidir.

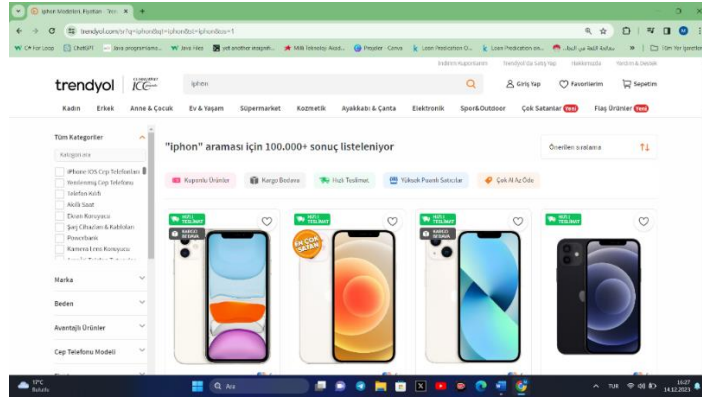
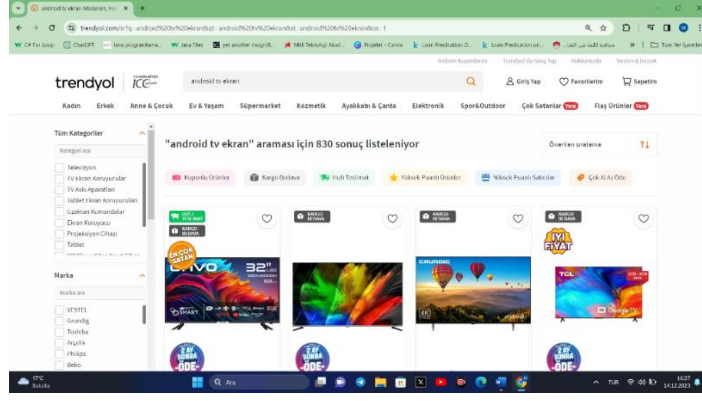
İyi bir arama mekanizması, dönüşüm oranını olumlu yönde etkileyerek e-ticaret perakendecisinin gelirini doğrudan etkiler.

Arama, bu bilginin renk, desen veya kategori gibi çeşitli yönleri üzerinde çalışır. E-ticarette bu tür aramaya genel olarak "çok yönlü arama" denir.

B. Yan arama:

Yan arama, müşterinin filtreleri kullanarak basitleştirilmiş bir şekilde gezinmesine olanak tanıyan özel bir arama türüdür.

Örneğin bir TV almayı planlıyorsak marka, fiyat, TV boyutu vb....



Her iki görselin en sol kısmı, müşterinin aramasını satın alma ihtiyaçlarına uygun şekilde yönlendirmesine olanak tanıyan bir dizi filtreyi (alternatif olarak "görüntüler") gösterir.

Yukarıdaki şekilde TV modellerine yönelik bir arama görüyoruz, dolayısıyla filtreler çözünürlük ve ekran boyutu gibi hususları gösteriyor.

Bu özel filtrelerin yanı sıra marka, fiyat aralığı, gönderim şekli gibi birçok ürün aramasında geçerli olan bazı genel özellikler de bulunmaktadır.

Bu filtreler ürün algısının net boyutlarıdır ve bu yönlendirmeli arama, kullanıcının aradığını bulmak için tonlarca sonucu incelemek yerine, alışverişleri üzerinde daha fazla kontrole sahip olması için arama sonuçlarını kendisinin sıralamasına olanak tanır.

Filtreler, yan aramayı tanımlayan anahtardır ve ancak bunlar her zaman tüm ürünler için hazır olmayabilir ve bunun bazı nedenleri şunlardır:

Sınırlı Ürün Bilgisi: Satıcılar, özellikle çok sayıda satıcının hızlı bir şekilde katılımını sağladıklarında, bir ürünü listelerken gerekli tüm ayrıntıları sağlayamayabilir.

Eksik Satıcı Bilgileri: Bazı durumlarda satıcılar belirli filtreleme faktörlerini sağlayacak eksiksiz bilgilere sahip olmayabilir.

Yeni E-ticaret Platformları: Çok sayıda satıcının hızlı bir şekilde katılımını teşvik eden yeni e-ticaret platformları, tüm filtreleme faktörlerini elde etmede zorluklarla karşılaşabilir.

Bu zorluklara rağmen, etkili yönlü arama, kullanıcıların arama deneyimlerini geliştirmeleri ve ürünleri daha verimli bir şekilde bulmaları için çeşitli filtreleme faktörlerine sahip olmaya dayanır.

5) E-ticaret kataloğu oluşturma

Bu yazının başlarında da gördüğümüz gibi katalog oluşturmak e-ticaretin temel sorunlarından biridir. Birkaç alt probleme ayrılabilir:

A. Özellik çıkarma :

Özellikler, ürünü tanımlayan özelliklerdir ve örneğin yukarıdaki görselde ilgili özellikler olarak markayı, çözünürlüğü, TV'nin boyutunu vb.

Bu özelliklerin doğru bir şekilde görüntülenmesi, e-ticaret web sitesinde ürüne ilişkin bir genel bakış sunacak ve böylece müşterinin karar vermesine olanak tanıyacaktır.

B. Ürün sınıflandırması ve taksonomi oluşturma:

Ürün sınıflandırması, ürünleri gruplara ayırma işlemidir.

Bu gruplar benzerliğe dayalı olarak tanımlanabilir; örneğin aynı markanın ürünleri veya aynı türdeki ürünler birlikte gruplandırılabilir.

Genel olarak e-ticarete elektronik, kişisel bakım ürünleri ve gıdalar gibi geniş, önceden tanımlanmış ürün kategorileri bulunur.

Yeni bir ürün geldiğinde kataloğa yerleştirilmeden önce sıralanması gerekir.

C. Ürün zenginleştirme (Product enrichment):

Daha iyi öneriler ve araştırmalar için ürün hakkında daha zengin bilgilerin toplanması önemlidir.

Bu bilgilerin olası kaynaklarından bazıları kısa ve uzun başlıklar, ürün resimleri ve ürün açıklamalarıdır. Ancak bu bilgiler çoğunlukla ya yalıttır ya da eksiktir.

Örneğin yanıltıcı bir başlık, bir e-ticaret platformunda yan arama yapılmasını engelleyebilir.

Ürün başlığını optimize etmek yalnızca aramadaki tıklama oranını artırmakla kalmayacak, aynı zamanda ürünü satın alma açısından dönüşüm oranını da artıracaktır.

D. Ürün tekilleştirme ve eşleştirme:

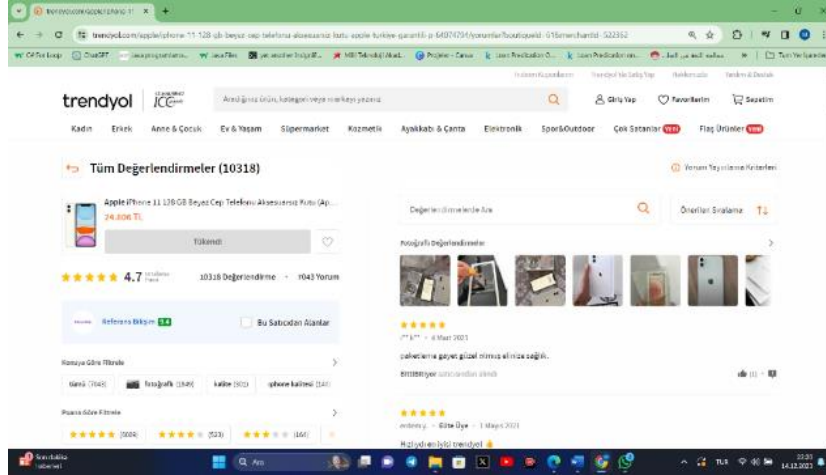
Ürünler genellikle üçüncü taraf satıcılar tarafından platforma eklenir ve farklı satıcılar aynı ürüne farklı isimlerle atıfta bulunabilir.

Nadiren aynı terminolojiyi takip ederler; bu da aynı ürünün birden fazla başlık ve ürün görseliyle listelenmesine yol açabilir.

Ürün sınıflandırma ve özellik çıkarmanın yanı sıra ürün tekilleştirme de e-ticaretin önemli bir yönüdür.

Dikkat edilmesi gereken bazı görevler:

1. Özellik Eşleştirme
2. Adres Eşleştirme
3. Görüntü Eşleştirme
4. İnceleme Analizi
5. Duygu Analizi



Sonuç olarak:

E-ticaret sektörünün büyük başarısının ardındaki temel itici güç, büyük miktarda veri toplama ve veriye dayalı karar verme olmuştur. NLP teknolojileri, e-ticaret ve perakende sektörlerinde kullanıcı deneyiminin iyileştirilmesinde ve gelirlerin artırılmasında önemli bir rol oynamıştır.

Bu makalede, e-ticarete sinirsel dil işlemenin çeşitli yönlerini ele aldık.