

Sosyal Bilimlerde Yapay Zeka Kullanımı: Bir Kılavuz

Emre TOROS

2025-09-09



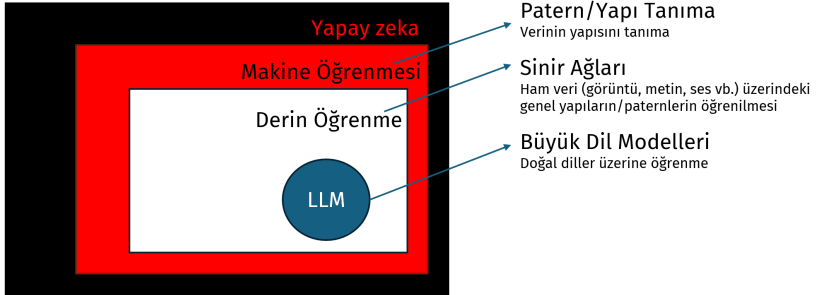
Büyük Dil Modelleri (LLM) nedir ve nasıl çalışırlar?

- Büyük dil modellerinin nasıl çalıştığı, veri bilimciler veya yapay zekâ alanında çalışanlar dışında çok az kişi tarafından biliniyor.

Büyük Dil Modelleri (LLM) nedir ve nasıl çalışırlar?

- ▶ Büyük dil modellerinin nasıl çalıştığı, veri bilimciler veya yapay zekâ alanında çalışanlar dışında çok az kişi tarafından biliniyor.
- ▶ Açıklayıcı çalışmalar ya fazla teknik olduğu için ön bilgi gerektiriyor ya da o kadar yüzeysel ki okuduktan sonra yeni bir şey öğrenmiş olunmuyor.

Yapay Zekâ Katmanları



Burada dış katmanı (AI/YZ) geçip doğrudan makine öğrenmesine odaklanıyoruz.

Makine Öğrenmesi

- Burada amaç verilerdeki paternleri/desenleri keşfetmektir. Daha spesifik olarak, bir girdiyi bir çıktıyla ilişkilendiren bir patern keşfedilir.

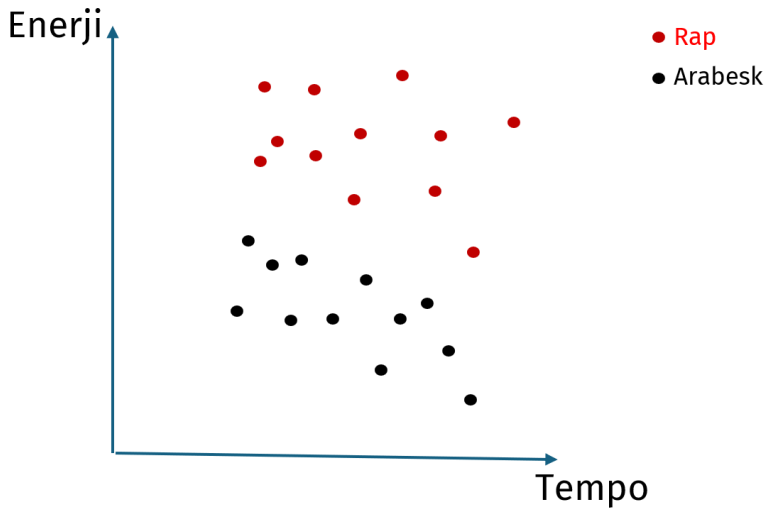
Makine Öğrenmesi

- ▶ Burada amaç verilerdeki paternleri/desenleri keşfetmektir. Daha spesifik olarak, bir girdiyi bir çıktıyla ilişkilendiren bir patern keşfedilir.
- ▶ Acaba dinlediğimiz şarkı hangi tür? Arabesk mi, rap mi?

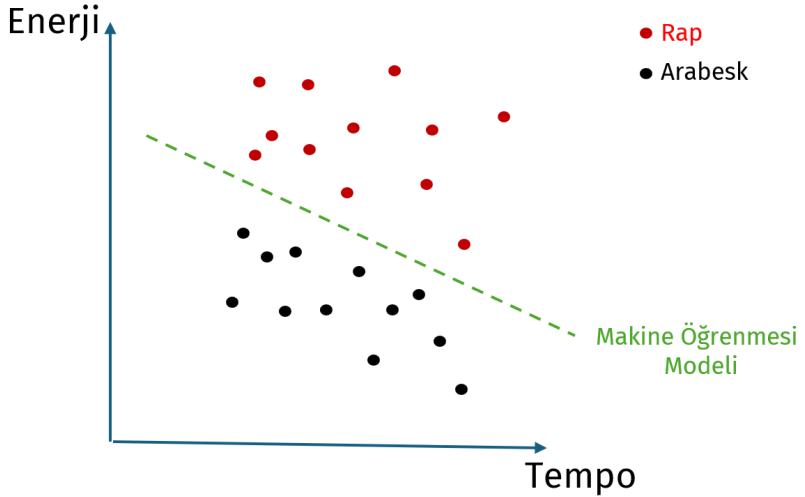
Makine Öğrenmesi

- ▶ Burada amaç verilerdeki paternleri/desenleri keşfetmektir. Daha spesifik olarak, bir girdiyi bir çıktıyla ilişkilendiren bir patern keşfedilir.
- ▶ Acaba dinlediğimiz şarkı hangi tür? Arabesk mi, rap mi?
 - ▶ Diyelim ki iki favori müzik türü olan arabesk ve rap'i ayırt etmek istiyoruz. Rap, canlı ritimleri ve dans edilebilir temposuyla bilinirken arabesk ortadoğu müzikal geleneklerine dayanan, duygusal vokallerin ve daha yavaş tempolu parçaların bir araya geldiği bir tür olarak tanımlanabilir.

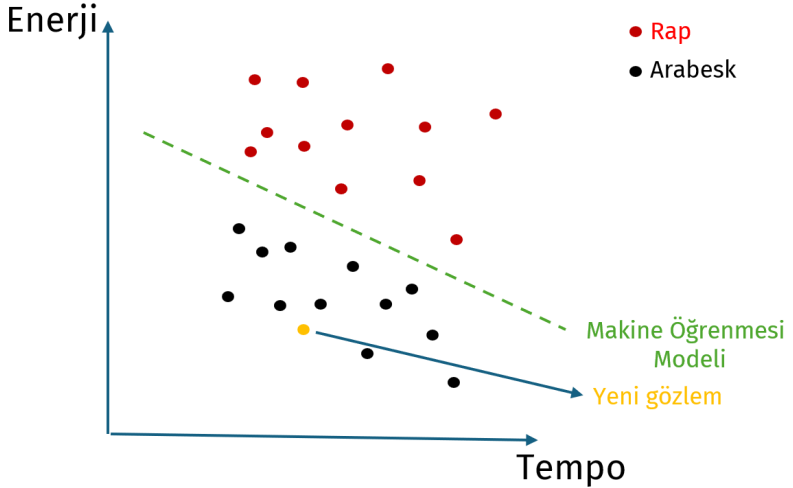
Makine Öğrenmesi



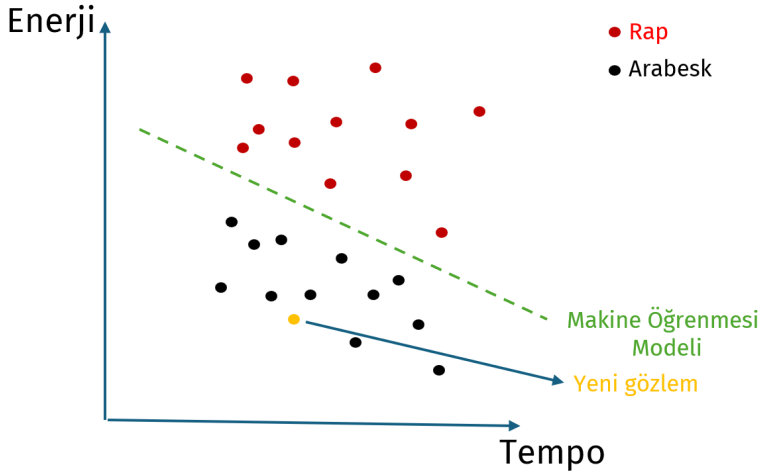
Makine Öğrenmesi



Makine Öğrenmesi



Makine Öğrenmesi



Yeni Gözlem:
Tempo: 40
Enerji: 20

Makina
Öğrenmesi

Sınıf	Olasılık
Rap	%25
Arabesk	%75

Makine Öğrenmesi

- ▶ Makine Öğrenmesi terimleriyle, bu bir sınıflandırma problemidir.

Makine Öğrenmesi

- ▶ Makine Öğrenmesi terimleriyle, bu bir sınıflandırma problemidir.
- ▶ Çünkü çıktı değişkeni (müzik türü) yalnızca sabit bir sınıf/etiket kümesinden birini alabilir — burada rap veya arabesk.

Makine Öğrenmesi

- ▶ Makine Öğrenmesi terimleriyle, bu bir sınıflandırma problemidir.
- ▶ Çünkü çıktı değişkeni (müzik türü) yalnızca sabit bir sınıf/etiket kümesinden birini alabilir — burada rap veya arabesk.
- ▶ Bunun aksine, regresyon probleminde çıktı sürekli bir değer olur (örneğin sıcaklık ya da mesafe gibi).

Makine Öğrenmesi

- ▶ Makine Öğrenmesi terimleriyle, bu bir sınıflandırma problemidir.
- ▶ Çünkü çıktı değişkeni (müzik türü) yalnızca sabit bir sınıf/etiket kümesinden birini alabilir — burada rap veya arabesk.
- ▶ Bunun aksine, regresyon probleminde çıktı sürekli bir değer olur (örneğin sıcaklık ya da mesafe gibi).
- ▶ Artık etiketli veri setimizi kullanarak bir Makine Öğrenmesi modeli (ya da “sınıflandırıcı”) eğitebiliriz. Yani türünü bildiğimiz şarkılardan oluşan bir seti kullanırız. Görsel olarak, modelin eğitimi aslında bu iki sınıfı en iyi ayıran çizgiyi bulmaktır.

Makine Öğrenmesi

- Peki bu ne işe yarar? Artık bu çizgiyi bildiğimiz için, elimizdeki yeni bir şarkının hangi tarafta kaldığına bakarak onun rap mi yoksa arabesk mi olduğunu tahmin edebiliriz. Bunun için sadece tempo ve enerji bilgisine ihtiyaç duyarız; bu da her şarkının türünü insanın tek tek etiketlemesine kıyasla çok daha kolay ve ölçeklenebilir bir yöntemdir.

Makine Öğrenmesi

- ▶ Peki bu ne işe yarar? Artık bu çizgiyi bildiğimiz için, elimizdeki yeni bir şarkının hangi tarafta kaldığına bakarak onun rap mi yoksa arabesk mi olduğunu tahmin edebiliriz. Bunun için sadece tempo ve enerji bilgisine ihtiyaç duyarız; bu da her şarkının türünü insanın tek tek etiketlemesine kıyasla çok daha kolay ve ölçeklenebilir bir yöntemdir.
- ▶ Ayrıca tahminin doğruluğuna olan güvenimizi de bu çizgiye olan uzaklığa göre ölçebiliriz. Şarkı çizgiden ne kadar uzaksa, tahminimizin o kadar doğru olduğundan emin olabiliriz. Örneğin, düşük enerjili ve düşük tempolu yeni bir şarkı için %98 olasılıkla bunun bir arabesk şarkısı olduğunu, %2 ihtimalle de aslında rap olabileceğini söyleyebiliriz.

Makine Öğrenmesi

- ▶ Ama elbette, gerçeklik bundan çok daha karmaşıktır.

Makine Öğrenmesi

- ▶ Ama elbette, gerçeklik bundan çok daha karmaşıktır.
- ▶ Sınıfları ayıran en iyi sınır her zaman doğrusal olmayabilir. Başka bir deyişle, girdiler ile çıktı arasındaki ilişki daha karmaşık olabilir.

Makine Öğrenmesi

- ▶ Ama elbette, gerçeklik bundan çok daha karmaşıktır.
- ▶ Sınıfları ayıran en iyi sınır her zaman doğrusal olmayabilir. Başka bir deyişle, girdiler ile çıktı arasındaki ilişki daha karmaşık olabilir.
- ▶ Gerçeklik başka bir açıdan da genellikle daha karmaşıktır. Örneğimizde yalnızca iki girdi (enerji ve tempo) vardı; ancak çoğu durumda onlarca, yüzlerce hatta binlerce girdi değişkeni bulunur.

Makine Öğrenmesi

- ▶ Ayrıca çoğu zaman yalnızca iki değil, daha fazla sınıf vardır. Ve tüm bu sınıflar, inanılmaz derecede karmaşık, doğrusal olmayan ilişkiler aracılığıyla bu girdilerin tümüne bağlı olabilir.

Makine Öğrenmesi

- ▶ Ayrıca çoğu zaman yalnızca iki değil, daha fazla sınıf vardır. Ve tüm bu sınıflar, inanılmaz derecede karmaşık, doğrusal olmayan ilişkiler aracılığıyla bu girdilerin tümüne bağlı olabilir.
- ▶ Bizim örneğimizde bile, gerçekte iki müzik türünden daha fazlası olduğunu biliyoruz ve tempo ile enerji dışında çok daha fazla ölçüte ihtiyacımız var. Bunlar arasındaki ilişki de muhtemelen bu kadar basit değildir.

Makine Öğrenmesi

- ▶ Ayrıca çoğu zaman yalnızca iki değil, daha fazla sınıf vardır. Ve tüm bu sınıflar, inanılmaz derecede karmaşık, doğrusal olmayan ilişkiler aracılığıyla bu girdilerin tümüne bağlı olabilir.
- ▶ Bizim örneğimizde bile, gerçekte iki müzik türünden daha fazlası olduğunu biliyoruz ve tempo ile enerji dışında çok daha fazla ölçüte ihtiyacımız var. Bunlar arasındaki ilişki de muhtemelen bu kadar basit değildir.
- ▶ Dolayısıyla girdi ile çıktı arasındaki ilişki ne kadar karmaşıkça, bu ilişkiyi öğrenebilmek için o kadar karmaşık ve güçlü bir Makine Öğrenmesi modeline ihtiyaç duyarız. Genellikle, karmaşıklık girdilerin ve sınıfların sayısı arttıkça artar. Buna ek olarak, daha fazla veriye de ihtiyacımız olur.

Girdimiz metin olursa ne olur?

- Girdi-çıktı ilişkilerinin bir başka son derece karmaşık türünü ele alalım: bir cümle ile onun duygusu (sentiment) arasındaki ilişki. Burada duygu ile kastedilen, cümlenin aktardığı duygudur; örneğin olumlu ya da olumsuz.

Gençlerbirliği müthiş bir spor kulübüdür

.4	.4	.1	.5	.3
.2	.7	.05	.2	.2
-.5	.8	-.1	.1	.3
...	...	...	...	...

Makina
Öğrenmesi

Sınıf	Olasılık
Pozitif	%85
Negatif	%15

Girdimiz metin olursa ne olur?

- Girdi olarak elimizde bir kelime dizisi, yani bir cümle var; çıktı değişkeni ise cümlenin duygusu. Daha önce olduğu gibi bu da bir sınıflandırma problemidir, ancak bu sefer olası etiketler iki tanedir: olumlu veya olumsuz.

Girdimiz metin olursa ne olur?

- ▶ Girdi olarak elimizde bir kelime dizisi, yani bir cümle var; çıktı değişkeni ise cümlenin duygusu. Daha önce olduğu gibi bu da bir sınıflandırma problemidir, ancak bu sefer olası etiketler iki tanedir: olumlu veya olumsuz.
- ▶ İnsan olarak bu ilişkiyi doğal biçimde anlayabiliyoruz, peki bir Makine Öğrenmesi modeline de aynı şeyi öğretebilir miyiz?

Girdimiz metin olursa ne olur?

- ▶ Girdi olarak elimizde bir kelime dizisi, yani bir cümle var; çıktı değişkeni ise cümlenin duygusu. Daha önce olduğu gibi bu da bir sınıflandırma problemidir, ancak bu sefer olası etiketler iki tanedir: olumlu veya olumsuz.
- ▶ İnsan olarak bu ilişkiyi doğal biçimde anlayabiliyoruz, peki bir Makine Öğrenmesi modeline de aynı şeyi öğretebilir miyiz?
- ▶ Buna yanıt vermeden önce, kelimelerin Makine Öğrenmesi modeli için nasıl sayısal girdilere dönüştürüleceği sorunu ortaya çıkıyor. Burada bilmemiz gereken şey, her kelimenin bir gömülü kelime değeri (word embedding) ile temsil edilebilmesidir.

Girdimiz metin olursa ne olur?

- Kısaca, gömülü kelime değeri, bir kelimenin anlamsal ve sözdizimsel karşılığını —çoğu zaman bağlam içinde— sayısal biçimde ifade eder. Bu değerler, modelin eğitimi sırasında ya da ayrı bir eğitim süreci ile elde edilebilir. Genellikle, bir gömülü kelime değeri vektörü onlarca, yüzlerce hatta binlerce değişkenden oluşabilir.

Girdimiz metin olursa ne olur?

- ▶ Kısaca, gömülü kelime değeri, bir kelimenin anlamsal ve sözdizimsel karşılığını —çoğu zaman bağlam içinde— sayısal biçimde ifade eder. Bu değerler, modelin eğitimi sırasında ya da ayrı bir eğitim süreci ile elde edilebilir. Genellikle, bir gömülü kelime değeri vektörü onlarca, yüzlerce hatta binlerce değişkenden oluşabilir.
- ▶ Özetle, bir cümleyi alıp onu sayısal girdiler dizisine, yani anlamsal ve sözdizimsel anlam taşıyan kelime değerlerine dönüştürebiliriz. Bu diziyi de Makine Öğrenmesi modeline besleyebiliriz.

Girdimiz metin olursa ne olur?

- ▶ Fakat şimdi bir seri zorluklarla karşılaşıyoruz. Uzun bir cümle (ya da paragraf, hatta tüm bir belge) söz konusu olduğunda, gömülü kelime değeri vektör boyutları nedeniyle çok hızlı biçimde yüz binlerce girdiye ulaşabiliriz.

Girdimiz metin olursa ne olur?

- ▶ Fakat şimdi bir seri zorluklarla karşılaşıyoruz. Uzun bir cümle (ya da paragraf, hatta tüm bir belge) söz konusu olduğunda, gömülü kelime değeri vektör boyutları nedeniyle çok hızlı biçimde yüz binlerce girdiye ulaşabiliriz.
- ▶ İkinci problem ise, dil ile duygu arasındaki ilişkinin çok ama çok karmaşık olmasıdır. “Ne harika bir gol yedik” gibi bir cümlenin tüm olası anlamlarını düşünün (ironi ya da alay da dahil).

Girdimiz metin olursa ne olur?

- ▶ Fakat şimdi bir seri zorluklarla karşılaşıyoruz. Uzun bir cümle (ya da paragraf, hatta tüm bir belge) söz konusu olduğunda, gömülü kelime değeri vektör boyutları nedeniyle çok hızlı biçimde yüz binlerce girdiye ulaşabiliriz.
- ▶ İkinci problem ise, dil ile duygu arasındaki ilişkinin çok ama çok karmaşık olmasıdır. “Ne harika bir gol yedik” gibi bir cümlenin tüm olası anlamlarını düşünün (ironi ya da alay da dahil).
- ▶ Dolayısıyla ihtiyacımız olan şey son derece güçlü bir Makine Öğrenmesi modeli ve çok fazla veridir. İşte bu noktada Derin Öğrenme devreye girer.

Derin Öğrenme ve sinir ağları

- Girdi ile çıktı arasındaki ilişkinin çok karmaşık olması durumunda — ve girdi ya da çıktı değişkenlerinin sayısı fazla olduğunda daha esnek ve güçlü modellere ihtiyaç duyarız.

Derin Öğrenme ve sinir ağları

- ▶ Girdi ile çıktı arasındaki ilişkinin çok karmaşık olması durumunda — ve girdi ya da çıktı değişkenlerinin sayısı fazla olduğunda daha esnek ve güçlü modellere ihtiyaç duyarız.
- ▶ İşte sinir ağları bu noktada devreye girer. Sinir ağları, karmaşık ilişkilerin modellenmesine olanak tanıyan çok güçlü Makine Öğrenimi modelleridir.

Derin Öğrenme ve sinir ağları

- ▶ Girdi ile çıktı arasındaki ilişkinin çok karmaşık olması durumunda — ve girdi ya da çıktı değişkenlerinin sayısı fazla olduğunda daha esnek ve güçlü modellere ihtiyaç duyarız.
- ▶ İşte sinir ağları bu noktada devreye girer. Sinir ağları, karmaşık ilişkilerin modellenmesine olanak tanıyan çok güçlü Makine Öğrenimi modelleridir.
- ▶ Aslında sinir ağlarının temel mimarileri nispeten basittir. Girdi sinyalinin sonuç değişkenini tahmin etmek üzere geçtiği, birbirine bağlı “nöron” katmanlarından oluşan bir dizilim içerirler. Bunları, aralarına doğrusal olmayanlıklar eklenmiş çok katmanlı doğrusal regresyonlar gibi düşünebilirsiniz; bu da sinir ağının son derece doğrusal olmayan ilişkileri modellemesini sağlar.

Derin Öğrenme ve sinir ağları

- Sinir ağları genellikle birçok katmandan oluşur (bu nedenle Derin Öğrenme adını alır), bu da onları son derece büyük hale getirebilir. Örneğin ChatGPT, yaklaşık 176 milyar nörondan oluşan bir sinir ağına dayanır; bu sayı, insan beynindeki yaklaşık 100 milyar nörondan fazladır.

Derin Öğrenme ve sinir ağları

- ▶ Sinir ağları genellikle birçok katmandan oluşur (bu nedenle Derin Öğrenme adını alır), bu da onları son derece büyük hale getirebilir. Örneğin ChatGPT, yaklaşık 176 milyar nörondan oluşan bir sinir ağına dayanır; bu sayı, insan beynindeki yaklaşık 100 milyar nörondan fazladır.
- ▶ Bu noktada Büyük Dil Modellerini (LLM)'leri anlamak için gerekli tüm alt yapıya sahibiz.

Derin Öğrenme ve sinir ağları

- ▶ Sinir ağları genellikle birçok katmandan oluşur (bu nedenle Derin Öğrenme adını alır), bu da onları son derece büyük hale getirebilir. Örneğin ChatGPT, yaklaşık 176 milyar nörondan oluşan bir sinir ağına dayanır; bu sayı, insan beynindeki yaklaşık 100 milyar nörondan fazladır.
- ▶ Bu noktada Büyük Dil Modellerini (LLM)’leri anlamak için gerekli tüm alt yapıya sahibiz.
- ▶ “Büyük” kelimesinin ne anlama geldiğini zaten biliyoruz; bu durumda, sinir ağındaki nöron sayısını — yani parametreleri — ifade ediyor. Bir modelin “Büyük Dil Modeli” sayılması için net bir sayı yok, ancak 1 milyar nöronun üzerindeki her şeyi büyük olarak düşünebiliriz.

Dil modellemesi

- Belirli bir kelime dizisinde — yani bir cümlede ya da paragrafta — sıradaki kelime nedir? Başka bir deyişle, herhangi bir anda bir sonraki kelimeyi nasıl tahmin ederiz. Bu soruyu bir makina öğrenimi problemi olarak ele alalım.

Dil modellemesi

- ▶ Belirli bir kelime dizisinde — yani bir cümlede ya da paragrafta — sıradaki kelime nedir? Başka bir deyişle, herhangi bir anda bir sonraki kelimeyi nasıl tahmin ederiz. Bu soruyu bir makina öğrenimi problemi olarak ele alalım.
- ▶ Aslında bu görev, daha önce gördüğümüz duygu sınıflandırma görevine oldukça benziyor. O örnekte olduğu gibi, sinir ağına verilen girdi bir kelime dizisidir; ancak bu kez çıktı sadece bir sonraki kelimedir.

Dil modellemesi

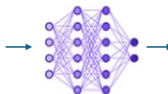
- ▶ Belirli bir kelime dizisinde — yani bir cümlede ya da paragrafta — sıradaki kelime nedir? Başka bir deyişle, herhangi bir anda bir sonraki kelimeyi nasıl tahmin ederiz. Bu soruyu bir makina öğrenimi problemi olarak ele alalım.
- ▶ Aslında bu görev, daha önce gördüğümüz duygu sınıflandırma görevine oldukça benziyor. O örnekte olduğu gibi, sinir ağına verilen girdi bir kelime dizisidir; ancak bu kez çıktı sadece bir sonraki kelimedir.
- ▶ Yine de bu bir sınıflandırma görevidir. Tek fark, yalnızca iki ya da birkaç sınıf yerine, şimdi kelime sayısı kadar sınıfımız var — örneğin yaklaşık 50.000 kelime. İşte dil modellemesi tam olarak budur: Bir sonraki kelimeyi tahmin etmeyi öğrenmek.

Dil modellemesi

Dil Modelleme

Bir sonraki kelimeyi tahmin et! (Bu bir sınıflandırma problemi!)

Kedilerin uyumayı sevdiği yerlerden biri



Girdi/Input

Sinir Ağları (LLM)

Kelime	Olasılık
Balkon	0.02
Kutu	0.7
....	
Zebra	0.0001

Çıktı/Output

Dil modellemesi

- Görevi belli olduğuna göre ve şimdi sinir ağını eğitmek için veriye ihtiyacımız var. Aslında “bir sonraki kelimeyi tahmin etme” görevimiz için çok fazla veri oluşturmak zor değil. İnternette, kitaplarda, akademik makalelerde ve daha fazlasında bol miktarda metin var. Bunların hepsinden kolayca devasa bir veri kümesi oluşturabiliriz.

Dil modellemesi

- ▶ Görevi belli olduğuna göre ve şimdi sinir ağını eğitmek için veriye ihtiyacımız var. Aslında “bir sonraki kelimeyi tahmin etme” görevimiz için çok fazla veri oluşturmak zor değil. İnternette, kitaplarda, akademik makalelerde ve daha fazlasında bol miktarda metin var. Bunların hepsinden kolayca devasa bir veri kümesi oluşturabiliriz.
- ▶ Veriyi etiketlememize bile gerek yok, çünkü bir sonraki kelime zaten etiketin kendisidir — bu nedenle buna kendiliğinden denetimli öğrenme (self-supervised learning) denir. Önemli olan, bunu hem kısa hem de uzun dizilimler için yapmamızdır (bazıları binlerce kelimeye kadar çıkabilir), böylece her bağlamda bir sonraki kelimenin ne olması gerektiğini öğreniriz.

Dil modellemesi

- Özetlemek gerekirse, burada yaptığımız şey, bir sinir ağını (LLM) verilen bir kelime diziliminde bir sonraki kelimeyi tahmin etmesi için eğitmek — bu dizilim ister uzun ister kısa olsun, ister Almanca ister İngilizce ya da başka bir dilde olsun, ister bir tweet, ister matematiksel bir formül, bir şiir ya da bir kod parçası olsun. Bunların hepsi eğitim verisinde karşılaştığımız dizilimlerdir.

Dil modellemesi

- ▶ Özetlemek gerekirse, burada yaptığımız şey, bir sinir ağını (LLM) verilen bir kelime diziliminde bir sonraki kelimeyi tahmin etmesi için eğitmek — bu dizilim ister uzun ister kısa olsun, ister Almanca ister İngilizce ya da başka bir dilde olsun, ister bir tweet, ister matematiksel bir formül, bir şiir ya da bir kod parçası olsun. Bunların hepsi eğitim verisinde karşılaştığımız dizilimlerdir.
- ▶ Yeterince büyük bir sinir ağına ve yeterli miktarda veriye sahipsek, LLM bir sonraki kelimeyi tahmin etme konusunda gerçekten iyi hale gelir. Mükemmel olur mu?

- ▶ Artık bir kelimeyi tahmin edebildiğimize göre, genişletilmiş dizilimi tekrar LLM'e (Büyük Dil Modeli) verebilir ve bir sonraki kelimeyi tahmin edebiliriz — ve bunu tekrar tekrar yapabiliriz. Başka bir deyişle, eğitilmiş LLM'imizi kullanarak artık yalnızca tek bir kelime değil, metin üretebiliriz. İşte bu nedenle LLM'ler, “Üretici Yapay Zekâ” (Generative AI) olarak adlandırılan şeyin bir örneğidir.

Doğal Dil Yaratımı

- ▶ Artık bir kelimeyi tahmin edebildiğimize göre, genişletilmiş dizilimi tekrar LLM'e (Büyük Dil Modeli) verebilir ve bir sonraki kelimeyi tahmin edebiliriz — ve bunu tekrar tekrar yapabiliriz. Başka bir deyişle, eğitilmiş LLM'imizi kullanarak artık yalnızca tek bir kelime değil, metin üretebiliriz. İşte bu nedenle LLM'ler, “Üretici Yapay Zekâ” (Generative AI) olarak adlandırılan şeyin bir örneğidir.
- ▶ LLM'e adeta konuşmayı öğrettik — her seferinde bir kelimeyle.

GPT ne demek?

- ▶ “G” “Generative” yani üretici — bu, modelin dil üretimi ön görevine göre eğitildiği anlamına geliyor

GPT ne demek?

- ▶ “G” “Generative” yani üretici — bu, modelin dil üretimi ön görevine göre eğitildiği anlamına geliyor
- ▶ “T” harfı “Transformer” anlamına gelir: burada kullanılan sinir ağı mimarisinin türüdür.

GPT ne demek?

- ▶ “G” “Generative” yani üretici — bu, modelin dil üretimi ön görevine göre eğitildiği anlamına geliyor
- ▶ “T” harfı “Transformer” anlamına gelir: burada kullanılan sinir ağı mimarisinin türüdür.
- ▶ “P” harfı de “Pre-training” yani ön eğitim anlamına gelir. Neden artık sadece “eğitim” değil de “ön eğitim”? Çünkü ChatGPT gibi Büyük Dil Modelleri aslında birkaç aşamada eğitilir.

Büyük Dil Modellerinin Eğitim Aşamaları: Ön eğitim

- İlk aşama, az önce detaylıca ele aldığımız ön eğitimidir. Bu aşama, bir sonraki kelimeyi tahmin etmeyi öğrenmek için devasa miktarda veriye ihtiyaç duyar. Bu süreçte model yalnızca dilin dilbilgisi ve sözdizimini öğrenmekle kalmaz, aynı zamanda dünya hakkında büyük miktarda bilgi edinir.

Büyük Dil Modellerinin Eğitim Aşamaları: Ön eğitim

- İlk aşama, az önce detaylıca ele aldığımız ön eğitimidir. Bu aşama, bir sonraki kelimeyi tahmin etmeyi öğrenmek için devasa miktarda veriye ihtiyaç duyar. Bu süreçte model yalnızca dilin dilbilgisi ve sözdizimini öğrenmekle kalmaz, aynı zamanda dünya hakkında büyük miktarda bilgi edinir.
- Sorunlar: LLM'nin gerçekte ne öğrenir?. Yani, model esasen bir konu hakkında uzun uzun konuşmayı öğrenmiştir. Bunu son derece iyi yapıyor olabilir, ancak yapmadığı şey, genellikle bir yapay zekâya vermek istediğimiz türden girdilere — örneğin bir soru ya da talimata — iyi yanıt vermektir.

Büyük Dil Modellerinin Eğitim Aşamaları: Ön eğitim

- Bu aşamada model bir asistan olmayı öğrenmemiştir ve dolayısıyla bir asistan gibi davranmaz. Örneğin, sadece ön eğitilmiş bir LLM'ye “Adın ne?” diye sorarsanız, size “Soyadın ne?” diye yanıt verebilir — çünkü ön eğitim sırasında gördüğü veri türü bu şekildedir, örneğin boş formlar gibi.

Büyük Dil Modellerinin Eğitim Aşamaları: Ön eğitim

- ▶ Bu aşamada model bir asistan olmayı öğrenmemiştir ve dolayısıyla bir asistan gibi davranmaz. Örneğin, sadece ön eğitilmiş bir LLM'ye “Adın ne?” diye sorarsanız, size “Soyadın ne?” diye yanıt verebilir — çünkü ön eğitim sırasında gördüğü veri türü bu şekildedir, örneğin boş formlar gibi.
- ▶ Model yalnızca girdi dizisini tamamlamaya çalışır. Talimatları takip etme konusunda iyi değildir, çünkü bu tür dil yapıları — yani talimat ve ardından gelen yanıt — eğitim verisinde çok yaygın değildir. B.

Büyük Dil Modellerinin Eğitim Aşamaları: Ön eğitim

- ▶ Bu aşamada model bir asistan olmayı öğrenmemiştir ve dolayısıyla bir asistan gibi davranmaz. Örneğin, sadece ön eğitilmiş bir LLM'ye “Adın ne?” diye sorarsanız, size “Soyadın ne?” diye yanıt verebilir — çünkü ön eğitim sırasında gördüğü veri türü bu şekildedir, örneğin boş formlar gibi.
- ▶ Model yalnızca girdi dizisini tamamlamaya çalışır. Talimatları takip etme konusunda iyi değildir, çünkü bu tür dil yapıları — yani talimat ve ardından gelen yanıt — eğitim verisinde çok yaygın değildir.
- ▶ Bu aşamada, LLM'nin insan niyetleriyle uyumlu olmadığını söyleriz. “Uyum” (alignment), LLM'ler için önemli bir konudur ve bunu büyük ölçüde nasıl düzeltebileceğimizi öğreneceğiz. Çünkü görünen o ki, ön eğitilmiş LLM'ler aslında oldukça yönlendirilebilir.

Büyük Dil Modellerinin Eğitim Aşamaları: İnce ayar

- İşte tam bu noktada talimatlarla ince ayar devreye girer. Önceden eğitilmiş LLM'yi (Büyük Dil Modeli) mevcut yetenekleriyle alırız ve temelde daha önce yaptığımız şeyi tekrarlarız — yani her seferinde bir kelime tahmin etmeyi öğrenmesini sağlarız — ancak bu kez yalnızca yüksek kaliteli talimat ve yanıt çiftlerini içeren bir eğitim verisi kullanırız.

Büyük Dil Modellerinin Eğitim Aşamaları: İnce ayar

- ▶ İşte tam bu noktada talimatlarla ince ayar devreye girer. Önceden eğitilmiş LLM'yi (Büyük Dil Modeli) mevcut yetenekleriyle alırız ve temelde daha önce yaptığımız şeyi tekrarlarız — yani her seferinde bir kelime tahmin etmeyi öğrenmesini sağlarız — ancak bu kez yalnızca yüksek kaliteli talimat ve yanıt çiftlerini içeren bir eğitim verisi kullanırız.
- ▶ Bu şekilde, model yalnızca metin tamamlayıcısı olmayı “unutur” ve kullanıcı talimatlarını takip eden, yardımcı bir asistan olmayı öğrenir. Bu talimat veri kümesinin boyutu genellikle ön eğitim veri kümesine göre çok daha küçüktür.

Büyük Dil Modellerinin Eğitim Aşamaları: İnce ayar

- ▶ İşte tam bu noktada talimatlarla ince ayar devreye girer. Önceden eğitilmiş LLM'yi (Büyük Dil Modeli) mevcut yetenekleriyle alırız ve temelde daha önce yaptığımız şeyi tekrarlarız — yani her seferinde bir kelime tahmin etmeyi öğrenmesini sağlarız — ancak bu kez yalnızca yüksek kaliteli talimat ve yanıt çiftlerini içeren bir eğitim verisi kullanırız.
- ▶ Bu şekilde, model yalnızca metin tamamlayıcısı olmayı “unutur” ve kullanıcı talimatlarını takip eden, yardımcı bir asistan olmayı öğrenir. Bu talimat veri kümesinin boyutu genellikle ön eğitim veri kümesine göre çok daha küçüktür.
- ▶ Bunun nedeni, yüksek kaliteli talimat-yanıt çiftlerinin genellikle insanlar tarafından oluşturulması ve dolayısıyla çok daha pahalı olmasıdır. Bu, ucuz ve kendiliğinden etiketlenmiş verilerle yapılan ön eğitimden oldukça farklıdır.

Büyük Dil Modellerinin Eğitim Aşamaları: İnce ayar

- Bu nedenle bu aşamaya “denetimli talimatla ince ayar, supervised instruction fine tuning” adı verilir.

Büyük Dil Modellerinin Eğitim Aşamaları: İnce ayar

- ▶ Bu nedenle bu aşamaya “denetimli talimatla ince ayar, supervised instruction fine tuning” adı verilir.
- ▶ Bazı LLM’ler — örneğin ChatGPT — üçüncü bir aşamadan da geçer: İnsan Geri Bildirimiyle Pekiştirme (Reinforcement Learning from Human Feedback, RLHF). Amacı talimatla ince ayarla aynıdır ancak RLHF, modelin insan değerleri ve tercihleriyle uyumlu yanıtlar üretmesini sağlar. Erken dönem araştırmalar, bu aşamanın insan seviyesini yakalamak ya da aşmak için kritik olduğunu göstermektedir.

Başarılı Kullanım alanları

- Özetleme: Bunun nedenini anlamak için, eğitim verisini düşünmemiz gerekiyor. LLM'nin öğrenim malzemesindeki içerik sıklıkla özetlemeler yapar — internette, akademik makalelerde, kitaplarda ve daha fazlasında. Sonuç olarak, bu verilerle eğitilmiş bir LLM (Büyük Dil Modeli) bunu nasıl yapacağını da öğrenir. Ana noktalara odaklanmayı ve bunları kısa bir metne sıkıştırmayı öğrenir.

Başarılı Kullanım alanları

- ▶ Özetleme: Bunun nedenini anlamak için, eğitim verisini düşünmemiz gerekiyor. LLM'nin öğrenim malzemesindeki içerik sıklıkla özetlemeler yapar — internette, akademik makalelerde, kitaplarda ve daha fazlasında. Sonuç olarak, bu verilerle eğitilmiş bir LLM (Büyük Dil Modeli) bunu nasıl yapacağını da öğrenir. Ana noktalara odaklanmayı ve bunları kısa bir metne sıkıştırmayı öğrenir.
- ▶ Soru cevaplama: Bir asistan gibi davranma ve uygun şekilde yanıt verme yeteneği, talimatla ince ayar ve RLHF (İnsan Geri Bildirimiyle Pekiştirme) sayesinde kazanılır. Ancak soruları yanıtlamak için gereken bilginin büyük kısmı zaten ön eğitim sırasında edinilmiştir.

Başarılı Kullanım alanları

- ▶ Özetleme: Bunun nedenini anlamak için, eğitim verisini düşünmemiz gerekiyor. LLM'nin öğrenim malzemesindeki içerik sıklıkla özetlemeler yapar — internette, akademik makalelerde, kitaplarda ve daha fazlasında. Sonuç olarak, bu verilerle eğitilmiş bir LLM (Büyük Dil Modeli) bunu nasıl yapacağını da öğrenir. Ana noktalara odaklanmayı ve bunları kısa bir metne sıkıştırmayı öğrenir.
- ▶ Soru cevaplama: Bir asistan gibi davranma ve uygun şekilde yanıt verme yeteneği, talimatla ince ayar ve RLHF (İnsan Geri Bildirimiyle Pekiştirme) sayesinde kazanılır. Ancak soruları yanıtlamak için gereken bilginin büyük kısmı zaten ön eğitim sırasında edinilmiştir.
- ▶ Ya LLM yanıtı bilmiyorsa? Ne yazık ki, böyle bir durumda cevabı uydurabilir. Bunun nedenini anlamak için yine eğitim verisini ve eğitim hedefini düşünmemiz gerekir.

Halüsinasyonlar

- ▶ LLM'lerin gerçekte olmayan bilgileri uydurması olgusu

Halüsinasyonlar

- ▶ LLM'lerin gerçekte olmayan bilgileri uydurması olgusu
- ▶ Neden? Çünkü LLM yalnızca metin üretmeyi öğrenir, gerçeğe dayalı metin üretmeyi değil.

Halüsinasyonlar

- ▶ LLM'lerin gerçekte olmayan bilgileri uydurması olgusu
- ▶ Neden? Çünkü LLM yalnızca metin üretmeyi öğrenir, gerçeğe dayalı metin üretmeyi değil.
- ▶ Eğitim sürecinde modelin, verinin doğruluğu ya da güvenilirliği hakkında hiçbir göstergesi yoktur. Çoğunluk bilgi Internet'teki çöplükten çıkıyor olabilir.

İyi kullanım örnekleri

- ▶ Sorulan soruya bağlam sağlama: LLM sorduğunuz soru ile ilgili eğitilmemiş olabilir; akademik araştırmada sıklıkla karşılaştığımız sorun budur. LLM'ler bu yüzden gerçek olmayan referans verir.

İyi kullanım örnekleri

- ▶ Sorulan soruya bağlam sağlama: LLM sorduğunuz soru ile ilgili eğitilmemiş olabilir; akademik araştırmada sıklıkla karşılaştığımız sorun budur. LLM'ler bu yüzden gerçek olmayan referans verir.
- ▶ LLM'lerin sorulara yanlış cevap vermesinin önüne bağlam sağlayarak geçildiğinde çok şaşırtıcı sonuçlar elde edilebiliyor: Örneğin, GPT-4'ün toplamda bir trilyondan fazla parametreye sahip olduğu söyleniyor ve bu model baro sınavını ya da AP Biyoloji testini en iyi yüzde 10'luk dilimde geçebiliyor.

İyi kullanım örnekleri

- ▶ Sorulan soruya bağlam sağlama: LLM sorduğunuz soru ile ilgili eğitilmemiş olabilir; akademik araştırmada sıklıkla karşılaştığımız sorun budur. LLM'ler bu yüzden gerçek olmayan referans verir.
- ▶ LLM'lerin sorulara yanlış cevap vermesinin önüne bağlam sağlayarak geçildiğinde çok şaşırtıcı sonuçlar elde edilebiliyor: Örneğin, GPT-4'ün toplamda bir trilyondan fazla parametreye sahip olduğu söyleniyor ve bu model baro sınavını ya da AP Biyoloji testini en iyi yüzde 10'luk dilimde geçebiliyor.
- ▶ Sıfır Örnekle (Zero-Shot) Görev Çözme: LLM'ler eğitim sırasında hiç karşılaşmadıkları görevleri tamamen yeni bir şekilde yerine getirebilirler. “Damlaya damlaya göl olur”

İyi kullanım örnekleri

- Birkaç Örnekle (Few-Shot) Görev Çözme: Biri size yeni bir görev yapmanızı söylese, muhtemelen bu görevin nasıl yapıldığını gösteren birkaç örnek ya da uygulama istersiniz. LLM'ler de böylesi girdilerden fayda sağlar

İyi kullanım örnekleri

- ▶ Birkaç Örnekle (Few-Shot) Görev Çözme: Biri size yeni bir görev yapmanızı söylese, muhtemelen bu görevin nasıl yapıldığını gösteren birkaç örnek ya da uygulama istersiniz. LLM'ler de böylesi girdilerden fayda sağlar
- ▶ Daha karmaşık görevlerde, sıfır örnekle (zero-shot) yönlendirme genellikle çok ayrıntılı talimatlar gerektirir — ve buna rağmen performans çoğu zaman mükemmel olmaktan uzaktır.

İyi kullanım örnekleri

- ▶ Birkaç Örnekle (Few-Shot) Görev Çözme: Biri size yeni bir görev yapmanızı söylese, muhtemelen bu görevin nasıl yapıldığını gösteren birkaç örnek ya da uygulama istersiniz. LLM'ler de böylesi girdilerden fayda sağlar
- ▶ Daha karmaşık görevlerde, sıfır örnekle (zero-shot) yönlendirme genellikle çok ayrıntılı talimatlar gerektirir — ve buna rağmen performans çoğu zaman mükemmel olmaktan uzaktır.
- ▶ Para birimi örneği: burada örneğin çözümünü özellikle vermedik. Unutmayın ki bir LLM özünde hâlâ bir metin tamamlayıcısıdır, bu yüzden tutarlı bir yapı korumalıdır. Modelin yalnızca istediğiniz biçimde yanıt vermesini neredeyse zorlamalısınız — tıpkı yukarıdaki örnekte yaptığımız gibi.

İyi kullanım örnekleri

- Zincirleme Düşünme: LLM'lerin Adım Adım Akıl Yürütme Yeteneği:
LLM'lerin bir diğer ilginç yeteneği de insan zekâsını andıran bir özellik taşır. Bu özellik, görev daha karmaşıksa ve çözüm için birden fazla akıl yürütme adımı gerekiyorsa oldukça işe yarar.

,



İyi kullanım örnekleri

- ▶ Zincirleme Düşünme: LLM'lerin Adım Adım Akıl Yürütme Yeteneği:
LLM'lerin bir diğer ilginç yeteneği de insan zekâsını andıran bir özellik taşır. Bu özellik, görev daha karmaşıksa ve çözüm için birden fazla akıl yürütme adımı gerekiyorsa oldukça işe yarar.
- ▶ Sorumuz şu olsun: “Gençlerbirliği en son Türkiye kupasını aldığı yıl ligde kim şampiyon oldu?” Soru adım adım çözmeyi gerektiriyor. Bir LLM'ye sadece “adım adım düşün” demenin bile birçok görevde performansı önemli ölçüde artırabildiğini göstermiştir.

İyi kullanım örnekleri

- ▶ Zincirleme Düşünme: LLM’lerin Adım Adım Akıl Yürütme Yeteneği: LLM’lerin bir diğer ilginç yeteneği de insan zekâsını andıran bir özellik taşır. Bu özellik, görev daha karmaşıksa ve çözüm için birden fazla akıl yürütme adımı gerekiyorsa oldukça işe yarar.
- ▶ Sorumuz şu olsun: “Gençlerbirliği en son Türkiye kupasını aldığı yıl ligde kim şampiyon oldu?” Soru adım adım çözmeyi gerektiriyor. Bir LLM’ye sadece “adım adım düşün” demenin bile birçok görevde performansı önemli ölçüde artırabildiğini göstermiştir.
- ▶ Peki bu neden işe yarıyor? Çünkü bu tür alışılmadık bileşik bilgiler genellikle LLM’in iç belleğinde doğrudan yer almaz. Ancak parçalandığında bilgiler büyük ihtimalle LLM içerisinde bulunur.*

*Bu sunumun içeriği Andreas Stöffelbauer’in "How Large Language Models work" isimli çalışmasından adapte edilerek hazırlanmıştır.