Üretken Modeller

Yapay Zekâ ve Veri Dersleri Rehberi · Kısım II: Yapay Zekânın Dalları

Veri dağılımı | Gizil uzay | VAE ve GAN | Difüzyon | Özbağlanımlı modeller | Komşu dersler

Sercan KÜLCÜ · Giresun Üniversitesi · Bilgisayar Mühendisliği

© Sercan Külcü. Tüm hakları saklıdır. · Sunumun etkileşimli HTML sürümü: https://sercankulcu.github.io/files/ai_courses/sunumlar/Bolum_13_Uretken_Modeller.html

GİRİŞ

Ders Hangi Soruyu Sorar?

> **Bir model gerçek veriye benzeyen yeni örnekleri nasıl üretir?**

Üretken modeller, verinin olasılık dağılımını öğrenip o dağılımdan **yeni örnekler üreten** modelleri inceler. Ayırt edici bir model "bu görüntü kedi mi?" diye sorarken üretken model "kedi görüntüleri nasıl dağılır, yenisini nasıl çizerim?" diye sorar.

Bir bakışta

YAKLAŞIM veri dağılımını öğrenme	**GİRDİ** çok büyük, çoğu etiketsiz veri	**ÇIKTI** yeni örnek: görüntü, metin, ses
BAŞARI ÖLÇÜTÜ gerçekçilik ve çeşitlilik (FID vb.)	**MATEMATİK YÜKÜ** yüksek	**VERİ İHTİYACI** çok fazla

Kısa Tarihçe

2013
Kingma ve Welling değişimsel otokodlayıcıyı (VAE) önerir.

2014
Goodfellow ve arkadaşları üretici ile ayırıcının yarıştığı çekişmeli ağları (GAN) tanıtır.

2015
Sohl-Dickstein ve arkadaşları, gürültüyü adım adım geri alan difüzyon modeli fikrini yayımlar.

2018
OpenAI'ın ilk GPT modeli: büyük metin üzerinde önceden eğitilmiş özbağlanımlı dönüştürücü.

2020
Ho, Jain ve Abbeel'in DDPM çalışması difüzyon modellerini görüntü üretiminde öne çıkarır.

2022
Stable Diffusion açık ağırlıklarla yayımlanır; ChatGPT üretken modelleri geniş kitlelere taşır.

DERS

Temel Kavramlar

Olabilirlik
Modelin gerçek veriye verdiği olasılık; eğitim onu büyütür.

Gizil uzay
Örnekleri özetleyen düşük boyutlu gizli değişkenler.

Otokodlayıcı ve VAE
Kodla, çöz; VAE gizil uzayı düzenli bir dağılıma zorlar.

GAN
Üretici sahte örnek üretir, ayırıcı onu gerçekten ayırmaya çalışır.

Difüzyon
Veriye gürültü ekler, ağ gürültüyü adım adım silmeyi öğrenir.

Özbağlanımlı model
Her belirteci öncekilere bakarak tahmin eder: GPT.

Koşullu üretim
Metin, sınıf ya da çizimle yönlendirilen üretim.

Değerlendirme
FID, şaşkınlık (perplexity) ve insan değerlendirmesi.

Tipik İş Akışı



Örnek: Bir difüzyon modeli milyonlarca görüntüye adım adım gürültü ekleyip ağı bu gürültüyü tahmin etmeye eğitir. Üretim sırasında saf gürültüden başlar, metin isteminin yönlendirmesiyle gürültüyü onlarca adımda siler ve hiç görülmemiş bir görüntüye varır.

DERS

Derste Neler İşlenir?

1. Üretken modelleme; olabilirlik ve KL ıraksaması
2. Özbağlanımlı modeller: PixelCNN, dönüştürücü
3. Değişimsel otokodlayıcı (VAE) ve ELBO
4. Normalleştiren akışlar
5. Çekişmeli ağlar (GAN) ve eğitim kararlılığı
6. Enerji tabanlı modeller
7. Skor eşleme ve difüzyon modelleri
8. Hızlı örnekleme ve gizil difüzyon
9. Koşullu üretim ve sınıflandırıcısız yönlendirme
10. Büyük dil modelleri: ön eğitim ve ince ayar
11. Çok kipli modeller: metinden görüntü
12. Değerlendirme: FID, şaşkınlık, insan yargısı

Kazandırdığı beceriler

PyTorch ile VAE ve GAN eğitmek

Diffusers ile difüzyon modeli kullanmak

Transformers ile dil modeline ince ayar

FID ile örnek kalitesini ölçmek

eğitim deneylerini izleyip karşılaştırmak

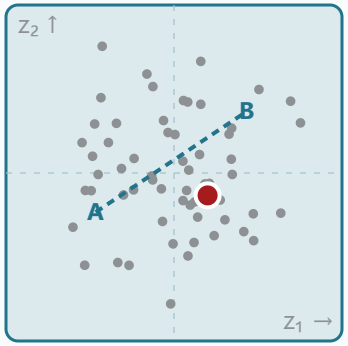
Tipik bir dönemin akışıdır; sıra ve kapsam üniversiteden üniversiteye değişir.

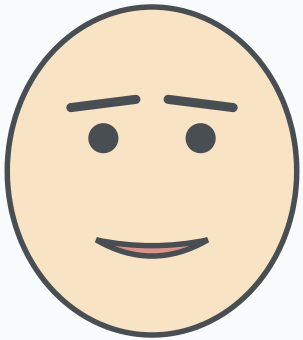
DENE

Gizil Uzaydan Yeni Yüzler

Bir oyun stüdyosu 60 çizgi yüzden yeni karakterler üretmek istiyor. Model her yüzü iki sayıya (z_1, z_2) sıkıştırmayı öğrendi; kaydırıcılarla ya da soldaki düzleme tıklayarak yeni yüz üretin.

Gizil uzay (tıklayın)





z_1 = **0,6**

z_2 = **-0,4**

Rastgele yüz üret | A → B ara değerle

İki sayı → 6 yüz ölçüsü. Gülümseme **0,60**, göz boyu **0,34**, ağız açıklığı **0,21**.
En yakın eğitim yüzüne uzaklık 0,05: eğitim yüzüne çok yakın.
İki boyut, verideki çeşitliliğin %98'ini taşıyor.

A'dan B'ye ara değerleme:



Model yüzleri ezberlemedi; verinin **gizil yapısını** öğrendi. Düzlemdeki her nokta yeni ama inandırıcı bir yüz verir; VAE ve difüzyon aynı fikri milyonlarca boyuta taşır.

KONUM

Ön Koşullar

Bu dersten önce bilinmesi yararlı olan dersler ve neden gerektikleri.

Ders	Neden gerekir
Derin Öğrenme	GAN, VAE, difüzyon, büyük dil modelleri
Bilgi Kuramı	KL ıraksaması, olabilirlik
Analiz (Kalkülüs)	integral, türev ve olasılık yoğunluklarıyla rahat çalışmak
Programlama	Python ve bir derin öğrenme kütüphanesi (PyTorch)

KONUM

Bu Ders Sana Göre mi?

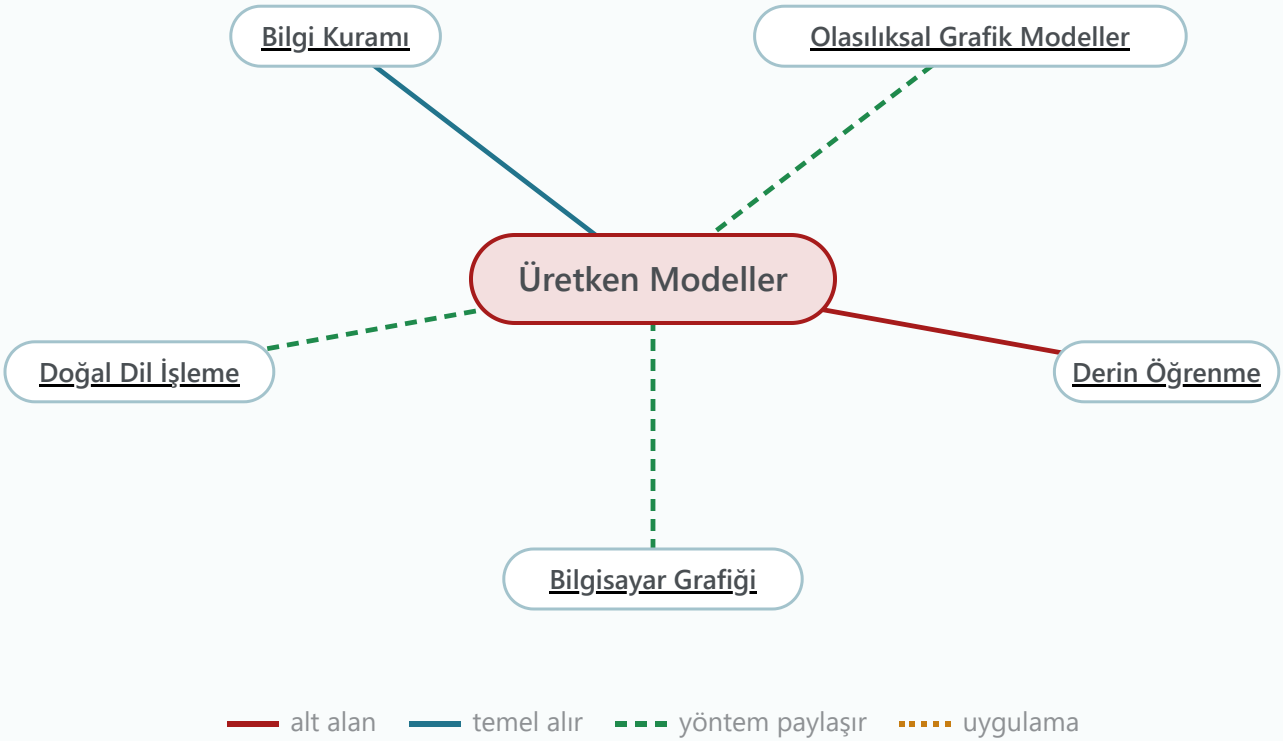
Kendine sor; "hayır" dediklerin için sağdaki sütun nereden başlayacağını gösterir.

Kendine sor	Hayırsa önce
PyTorch ile bir sinir ağı kurup eğitebiliyor musun?	Derin Öğrenme
Olasılık yoğunluğunu, beklenen değeri ve Bayes kuralını açıklayabiliyor musun?	Olasılık ve İstatistik
Geri yayılımın zincir kuralıyla gradyanı nasıl hesapladığını açıklayabiliyor musun?	Yapay Sinir Ağları
Bir integrali ve logaritmanın türevini hesaplayabiliyor musun?	Analiz dersindeki integral ve türev

Çoğuna "evet" diyebiliyorsan derse hazırsın; bir ya da iki "hayır" dönem içinde kapanabilir.

KONUM

Dersin Komşuları



Örtüştüğü Dersler

İlişki	Ders	Ortak konu
alt alan	Bu ders, Derin Öğrenme dersinin alt alanıdır	GAN, VAE, difüzyon, büyük dil modelleri
temel alır	Bu ders Bilgi Kuramı dersini temel alır	KL ıraksaması, olabilirlik
yöntem paylaşır	Doğal Dil İşleme ve Metin Madenciliği ile yöntem paylaşır	metin üretimi
yöntem paylaşır	Olasılıksal Grafik Modeller ile yöntem paylaşır	gizil değişkenli modeller, değişimsel çıkarım
yöntem paylaşır	Bilgisayar Grafiği ile yöntem paylaşır	görüntü ve sahne sentezi

Ayrıştığı Noktalar (1/2)

Ölçüt	Üretken Modeller	Derin Öğrenme	Olasılıksal Grafik Modeller
Temel soru	Bir model gerçek veriye benzeyen yeni örnekleri nasıl üretir?	Çok katmanlı ağlar ham veriden kullanışlı temsilleri nasıl kendisi öğrenir?	Birbirine bağlı, belirsiz çok sayıda değişken nasıl modellenir ve sorgulanır?
Yaklaşım	veri dağılımını öğrenme	çok katmanlı temsil öğrenme	olasılık ve çizge bir arada
Girdi	çok büyük, çoğu etiketsiz veri	ham veri: görüntü, metin, ses	değişkenler, bağımlılık yapısı, veri
Çıktı	yeni örnek: görüntü, metin, ses	öğrenilmiş temsil ve tahmin	koşullu olasılık, en olası açıklama
Başarı ölçütü	gerçekçilik ve çeşitlilik (FID vb.)	doğruluk, kayıp, ölçeklenme	olabilirlik, çıkarım doğruluğu
Matematik yükü	yüksek	yüksek: türev, matris, olasılık	yüksek: olasılık
Veri ihtiyacı	çok fazla	çok; GPU gerekir	orta

- **Derin Öğrenme:** Derin öğrenme genel araç takımıdır ve çoğunlukla tahmin yapar; bu ders o araçlarla veri dağılımından yeni örnek üretir.
- **Olasılıksal Grafik Modeller:** Grafik modeller değişkenler arası yapıyı açıkça kurar; bu dersin derin üretken modelleri yapıyı çoğu zaman örtük öğrenir.

Ayrıştığı Noktalar (2/2)

Ölçüt	Üretken Modeller	Doğal Dil İşleme	Bilgisayar Grafiği
Temel soru	Bir model gerçek veriye benzeyen yeni örnekleri nasıl üretir?	Bilgisayar insan dilini nasıl çözümler, anlar ve üretir?	Bir sahnenin tanımından gerçekçi bir görüntü nasıl üretilir?
Yaklaşım	veri dağılımını öğrenme	dilbilim, istatistik ve derin öğrenme	geometri ve ışık benzetimi
Girdi	çok büyük, çoğu etiketsiz veri	metin	sahne: model, ışık, kamera
Çıktı	yeni örnek: görüntü, metin, ses	etiket, çeviri, özet, yanıt	görüntü ya da animasyon
Başarı ölçütü	gerçekçilik ve çeşitlilik (FID vb.)	doğruluk, BLEU, insan değerlendirmesi	gerçekçilik, kare hızı
Matematik yükü	yüksek	orta	yüksek: doğrusal cebir, geometri
Veri ihtiyacı	çok fazla	çok metin (derlem)	gerekmez

- **Doğal Dil İşleme:** NLP dili çözümleme ve anlamayı da kapsar; bu ders metni, görüntüyü ve sesi üretmenin ortak ilkelerine bakar.
- **Bilgisayar Grafiği:** Bilgisayar grafiği görüntüyü sahnenin fiziksel tanımından hesaplar; üretken modeller görüntüyü veriden öğrenerek üretir.

UYGULAMA

Uygulama Alanları

Metinden görüntü
İstemden resim ve illüstrasyon üretmek.

Sohbet ve yazım
Büyük dil modelleriyle metin üretimi.

Ses sentezi
Doğal sesli konuşma ve müzik üretimi.

Veri artırma
Az veri olan sınıflar için yapay örnekler.

İlaç tasarımı
Yeni molekül adayları önermek.

Görüntü düzenleme
Eksik bölgeyi doldurma, çözünürlük artırma.

UYGULAMA

Riskler ve Etik

Derin sahte (deepfake)
Bir kişinin yüzü ve sesi izinsiz taklit edilir; dolandırıcılıkta, şantajda ve seçim dezenformasyonunda kullanılır.

İzinsiz eğitim verisi
İnternetten toplanan LAION-5B gibi veri setleri eserleri sahiplerinin izni olmadan içerir; telif davaları sürüyor.

Model çöküşü
Üretilmiş veriyle yeniden eğitilen modeller nesiller boyunca dağılımın kuyruklarını kaybeder; çıktılar tekdüzeleşir.

Ölçüt yanılsaması
FID gibi ölçütler insan yargısıyla her zaman örtüşmez; ölçüte göre ayarlanan model gözle daha kötü görünebilir.

Ayrıntılı tartışma: Açıklanabilir Yapay Zekâ ve Yapay Zekâ Etiği.

UYGULAMA

Araçlar ve Örnek Veriler

Araçlar ve kütüphaneler

PyTorch | Hugging Face Diffusers | Hugging Face Transformers | JAX | Stable Diffusion | Weights & Biases

Tanınmış veri setleri

MNIST | CelebA | CIFAR-10 | LSUN | LAION-5B | FFHQ

UYGULAMA

Sık Karıştırılanlar

Yaygın düşünce	Doğrusu
Üretken model eğitim verisini kopyalar.	Amaç dağılımı öğrenmektir; yine de ezberleme olabilir ve ayrıca denetlenmelidir.
Üretken YZ yalnız sohbet robotlarıdır.	Görüntü, ses, video, molekül ve kod üreten modeller de bu dersin konusudur.
Gerçekçi görünen çıktı doğru çıktıdır.	Model akla yatkın ama yanlış içerik üretebilir; doğruluk ayrıca denetlenir.
GAN eski, difüzyon her yerde daha iyi.	Her ailenin hız, kalite ve çeşitlilik dengesi farklıdır; GAN tek adımda üretir.

SONRASI

Bu Dersten Sonra

Bu dersi temel alan ya da genişleten dersler

Bu ders çoğunlukla uç noktadır; öteki dersleri temel alır.

İlgili meslekler

Üretken YZ mühendisi | Araştırmacı | Makine öğrenmesi mühendisi | Yaratıcı teknoloji uzmanı

Üretken modellerden sonra metin üretimi için doğal dil işlemeye, görüntü ve sahne sentezi için bilgisayar grafiğine ya da modellerin güvenli ve sorumlu kullanımı için açıklanabilir YZ ve etik dersine geçilebilir. Bilgi kuramı ve olasılıksal grafik modeller ise kuramsal temeli derinleştirir.

SONRASI

Kaynaklar

- Foster, *Generative Deep Learning*, 2. baskı, O'Reilly, 2023.
- Tomczak, *Deep Generative Modeling*, Springer, 2022.
- Goodfellow ve ark., "Generative Adversarial Nets", *Advances in Neural Information Processing Systems* 27, 2014.
- Kingma ve Welling, "Auto-Encoding Variational Bayes", *ICLR*, 2014.
- Ho, Jain ve Abbeel, "Denoising Diffusion Probabilistic Models", *Advances in Neural Information Processing Systems* 33, 2020.

Son

Bölüm 13: Üretken Modeller

← Önceki bölüm · Alanın haritası · Sonraki bölüm →