Bilgi Kuramı

Yapay Zekâ ve Veri Dersleri Rehberi · Kısım V: Matematik Temeli

Entropi | Sıkıştırma | Kanal kapasitesi | Hata düzelten kodlar | MÖ'deki ölçüler | 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_34_Bilgi_Kurami.html

GİRİŞ

Ders Hangi Soruyu Sorar?

> **Bilgi nasıl ölçülür, en az kayıpla nasıl sıkıştırılır ve iletilir?**

Bilgi kuramı, bir mesajın taşıdığı bilgiyi olasılıkla ölçer: az beklenen olay daha çok bilgi taşır ve ortalama bilgi **entropi** ile ifade edilir. Ders, verinin en çok ne kadar sıkıştırılabileceğini ve gürültülü bir kanaldan ne hızla güvenle iletilebileceğini gösteren sınırları kurar.

Bir bakışta

YAKLAŞIM olasılıkla bilgi ölçümü	**GİRDİ** kaynak ve kanal modeli	**ÇIKTI** sınır, kod, sıkıştırma oranı
BAŞARI ÖLÇÜTÜ bit, kanal kapasitesi	**MATEMATİK YÜKÜ** yüksek: olasılık	**VERİ İHTİYACI** gerekmez

GİRİŞ

Kısa Tarihçe

1928
Hartley bilgiyi olası mesaj sayısının logaritmasıyla ölçmeyi önerir.

1948
Shannon, "A Mathematical Theory of Communication" ile entropiyi ve kanal kapasitesini tanımlar.

1951
Kullback ve Leibler iki olasılık dağılımı arasındaki ıraksama ölçüsünü yayımlar.

1952
Huffman, ortalama uzunluğu en kısa önek kodunu kuran algoritmayı yayımlar.

1977
Lempel ve Ziv sözlük tabanlı LZ77 sıkıştırmasını yayımlar; ZIP ve PNG'nin atası.

1993
Turbo kodlar tanıtılır: pratik bir kod Shannon sınırına yarım desibel kadar yaklaşır.

DERS

Temel Kavramlar

Entropi
Bir kaynağın sembol başına ortalama bilgisi, bit cinsinden.

Koşullu entropi
Bir değişken bilindiğinde ötekinde kalan belirsizlik.

Ortak bilgi
İki değişkenin birbiri hakkında ne kadar bilgi taşıdığı.

KL ıraksaması
Bir dağılımı ötekiyle kodlamanın fazladan maliyeti.

Çapraz entropi
Sınıflandırmada en yaygın kayıp fonksiyonu.

Kaynak kodlama
Kayıpsız sıkıştırma: Huffman, aritmetik kodlama, LZ.

Kanal kapasitesi
Gürültülü kanaldan güvenle geçebilecek en yüksek hız.

Hata düzelten kodlar
Artıklık ekleyerek hataları bulmak ve düzeltmek.

Tipik İş Akışı



Örnek: Telefonla gönderilen bir fotoğraf önce JPEG ile sıkıştırılır; ardından kablosuz ağın kanal kodu buna denetimli artıklık ekler. Havada bozulan bitler alıcıda bu artıklıkla düzeltilir ve fotoğraf açılır. Sıkıştırma ve koruma birbirinin tersi gibi görünse de ikisi de aynı kuramdan çıkar.

DERS

Derste Neler İşlenir?

1. Entropi, ortak ve koşullu entropi
2. Ortak bilgi ve KL ıraksaması
3. Jensen ve veri işleme eşitsizlikleri
4. Tipik kümeler ve asimptotik eşbölüşüm (AEP)
5. Kaynak kodlama teoremi ve Kraft eşitsizliği
6. Huffman kodları ve aritmetik kodlama
7. Evrensel sıkıştırma: Lempel-Ziv
8. Kanal kapasitesi ve kanal kodlama teoremi
9. Diferansiyel entropi ve sürekli değişkenler
10. Gauss kanalı ve Shannon-Hartley formülü
11. Hız-bozulma kuramı ve kayıplı sıkıştırma
12. Hata düzelten kodlar: Hamming, LDPC

Kazandırdığı beceriler

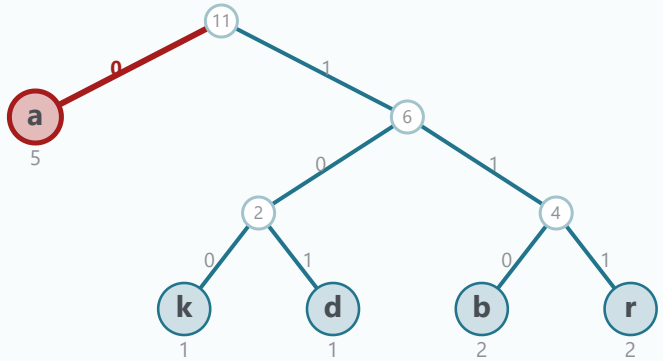
entropi ve ortak bilgi hesaplamak · Huffman kodlayıcısı yazmak · zlib ile sıkıştırma ölçmek · kanal kapasitesini hesaplamak · eşitsizliklerle kanıt yazmak

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

DENE

Mesajı Daha Az Bitle Gönder

Pille çalışan bir sensör mesajlarını radyoyla yolluyor; her bit enerji harcar. Bir metin yazın ya da hazır metin seçin: Huffman ağacı sık harfe kısa kod verir. Ağaçta bir harfe tıklayıp kodunu görün.



Gönderilecek metin

abrakadabra

abrakadabra | tekerleme | çok tekrarlı | hepsi eşit

Harf “a”: 5 kez geçiyor, kodu **0** (1 bit).

5 farklı harf, 11 harf
Entropi H = **2,04** bit/harf
Huffman ortalaması L = **2,09** bit/harf
8 bitlik kodla 88 bit → Huffman **23 bit** (%74 kazanç)

Gönderilen bitler (harf harf):

0110**111****0**100**0**10**1****0**1101110

Bilgi kuramı sınırı koyar: kayıpsız hiçbir kod ortalamada **entropiden** kısa olamaz. Huffman bu sınıra bir bitten yakın kalır. Harfler eşit sıklıktaysa kazanç yalnız alfabenin küçüklüğünden gelir.

KONUM

Ön Koşullar

Bu dersten önce bilinmesi yararlı olan dersler ve neden gerektikleri.

Ders	Neden gerekir
Ayrık Matematik	logaritma, kümeler, sayma ve ikili gösterim
Algoritmalar	açgözlü algoritmalar ve öncelik kuyruğu: Huffman kodunun kuruluşu

KONUM

Bu Ders Sana Göre mi?

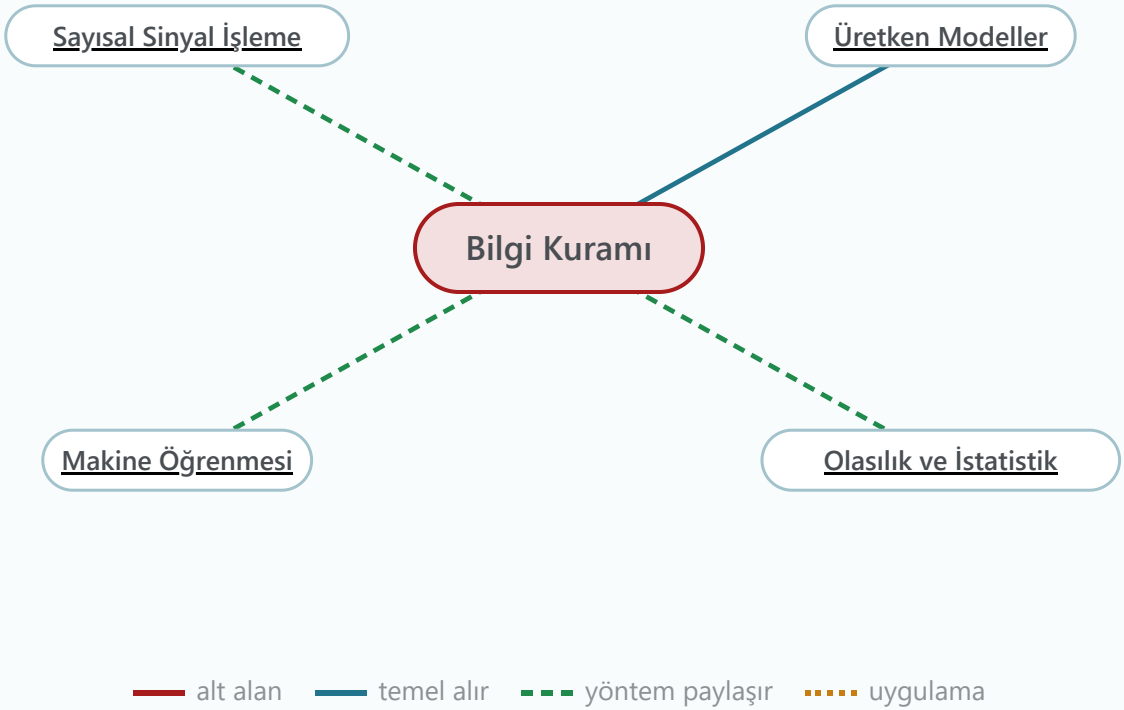
Kendine sor; "hayır" dediklerin için sağdaki sütun nereden başlayacağını gösterir.

Kendine sor	Hayırsa önce
Logaritma kurallarını kullanıp 2 tabanında hesap yapabiliyor musun?	Lise matematiğindeki logaritma konusu
Koşullu olasılığı ve beklenen değeri açıklayabiliyor musun?	Olasılık ve İstatistik
Bir sayıyı ikili tabanda yazıp kaç bit gerektiğini bulabiliyor musun?	Ayrık matematikteki sayı tabanları
Öncelik kuyruğu kullanan açgözlü bir algoritma yazabiliyor musun?	Algoritmalar dersindeki açgözlü yöntemler ve yığınlar

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

KONUM

Dersin Komşuları



ÖRTÜŞME

Örtüştüğü Dersler

İlişki	Ders	Ortak konu
temel alır	Üretken Modeller bu dersi temel alır	KL ıraksaması, olabilirlik
yöntem paylaşır	Makine Öğrenmesi ile yöntem paylaşır	çapraz entropi kaybı, karar ağacı bölme ölçütü
yöntem paylaşır	Sayısal Sinyal İşleme ile yöntem paylaşır	sıkıştırma, örnekleme, kanal
yöntem paylaşır	Olasılık ve İstatistik ile yöntem paylaşır	entropi, ortak bilgi

AYRIŞMA

Ayrıştığı Noktalar (1/2)

Ölçüt	Bilgi Kuramı	Olasılık ve İstatistik	Sayısal Sinyal İşleme
Temel soru	Bilgi nasıl ölçülür, en az kayıpla nasıl sıkıştırılır ve iletilir?	Belirsizlik nasıl ölçülür ve sınırlı veriden nasıl güvenilir sonuç çıkarılır?	Örneklenmiş sinyaller nasıl çözümlenir, süzülür ve dönüştürülür?
Yaklaşım	olasılıkla bilgi ölçümü	olasılık modeli ve çıkarım	matematiksel dönüşüm ve süzgeç
Girdi	kaynak ve kanal modeli	örneklem verisi	örneklenmiş sinyal
Çıktı	sınır, kod, sıkıştırma oranı	kestirim, güven aralığı, test sonucu	süzülmüş sinyal, spektrum
Başarı ölçütü	bit, kanal kapasitesi	yansızlık, güç, anlamlılık	sinyal-gürültü oranı, frekans yanıtı
Matematik yükü	yüksek: olasılık	kuramın kendisi	yüksek: karmaşık sayı, Fourier
Veri ihtiyacı	gerekmez	örneklem	öğrenme yok; sinyal yeter

- **Olasılık ve İstatistik:** Olasılık ve istatistik dağılımları ve belirsizliği tanımlar; bu ders aynı dağılımlardan bilgi miktarını ve iletim sınırlarını çıkarır.
- **Sayısal Sinyal İşleme:** Sinyal işleme sinyali süzer ve dönüştürür; bu ders o sinyalin taşıyabileceği bilginin ve iletim hızının sınırını sorar.

AYRIŞMA

Ayrıştığı Noktalar (2/2)

Ölçüt	Bilgi Kuramı	Makine Öğrenmesi	Üretken Modeller
Temel soru	Bilgi nasıl ölçülür, en az kayıpla nasıl sıkıştırılır ve iletilir?	Bir program örneklerden nasıl genelleme yapar?	Bir model gerçek veriye benzeyen yeni örnekleri nasıl üretir?
Yaklaşım	olasılıkla bilgi ölçümü	istatistiksel öğrenme	veri dağılımını öğrenme
Girdi	kaynak ve kanal modeli	etiketli ya da etiketsiz örnekler	çok büyük, çoğu etiketsiz veri
Çıktı	sınır, kod, sıkıştırma oranı	tahmin yapan model	yeni örnek: görüntü, metin, ses
Başarı ölçütü	bit, kanal kapasitesi	görülmemiş veride hata	gerçekçilik ve çeşitlilik (FID vb.)
Matematik yükü	yüksek: olasılık	yüksek: doğrusal cebir, olasılık, optimizasyon	yüksek
Veri ihtiyacı	gerekmez	orta ve çok	çok fazla

- **Makine Öğrenmesi:** MÖ çapraz entropiyi kayıp fonksiyonu olarak kullanır; bu ders o ölçünün nereden geldiğini ve ne anlama geldiğini açıklar.
- **Üretken Modeller:** Üretken modeller KL ıraksamasını eğitimde küçültür; bu ders ıraksamanın kendisini ve kod uzunluğuyla bağını inceler.

UYGULAMA

Uygulama Alanları

Dosya sıkıştırma
ZIP, gzip ve PNG: kayıpsız sıkıştırma.

Görüntü ve ses
JPEG ve MP3: kayıplı sıkıştırma.

Mobil iletişim
4G'de turbo, 5G'de LDPC ve polar kodlar.

Veri depolama
QR kod, CD ve SSD'de hata düzelten kodlar.

Makine öğrenmesi
Çapraz entropi kaybı, karar ağacında bilgi kazancı.

Öznitelik seçimi
Ortak bilgiyle en bilgilendirici öznitelikler.

UYGULAMA

Riskler ve Etik

Az veriyle ortak bilgi

Az örnekte histogramla kestirilen ortak bilgi yukarı yanlıdır; ilgisiz bir öznitelik bilgilendirici görünür.

Yanlış KL yönü

KL(p||q) ile KL(q||p) farklı modeller üretir: biri bütün kipleri örtmeye, öteki tek bir kipe kilitlenmeye iter.

Sıkıştırmadan sızan sır

Gizli veri ile saldırganın metni birlikte sıkıştırılınca boyut sırrı ele verir; TLS'e yönelik CRIME saldırısı buydu.

Bağımsızlık varsayımı

Sembolleri bağımsız saymak entropiyi olduğundan büyük gösterir; sıkıştırma ve rastgelelik değerlendirmesi yanılır.

Ayrıntılı tartışma: Açıklanabilir Yapay Zekâ ve Yapay Zekâ Etiği.

UYGULAMA

Araçlar ve Örnek Veriler

Araçlar ve kütüphaneler

Python · NumPy · SciPy · zlib · scikit-learn · MATLAB

Klasik örnek problemler

Hileli zar ve yazı-tura · İkili simetrik kanal · İngilizce metnin entropisi · Huffman ağacı kurma · Yirmi soru oyunu · Canterbury derlemi

Bilgi kuramı çoğunlukla ispat ve hesapla çalışılır; kod, entropiyi ölçmek ve sıkıştırma oranını gerçek dosyalarda denemek için yazılır.

UYGULAMA

Sık Karıştırılanlar

Yaygın düşünce	Doğrusu
Bilgi, mesajın anlamıyla ölçülür.	Shannon bilgisi anlamla ilgilenmez; mesajın ne kadar beklenmedik olduğuyla, yani olasılıkla ölçülür.
İyi bir program her dosyayı küçültebilir.	Hiçbir kayıpsız yöntem her girdiyi küçültemez; veri entropisinin altına sıkıştırılamaz.
Gürültülü kanalda hatasız iletim imkânsızdır.	Kapasitenin altındaki hızlarda hata olasılığı istenildiği kadar küçültülebilir (Shannon).
KL ıraksaması bir uzaklıktır.	Simetrik değildir ve üçgen eşitsizliğini sağlamaz: p'den q'ya ve q'dan p'ye değerler farklıdır.

SONRASI

Bu Dersten Sonra

Bu dersi temel alan ya da genişleten dersler

Üretken Modeller

İlgili meslekler

Haberleşme mühendisi
Sıkıştırma ve kodek geliştiricisi
Makine öğrenmesi araştırmacısı
Kriptografi uzmanı
Veri depolama mühendisi

Bilgi kuramı, matematik temelinin MÖ'ye açılan kapısıdır: çapraz entropi, KL ıraksaması ve ortak bilgi üretken modellerde, karar ağaçlarında ve temsil öğrenmede yeniden karşınıza çıkar. Haberleşme yönünde ise sayısal sinyal işleme ve kodlama kuramıyla devam edilir.

SONRASI

Kaynaklar

- Cover ve Thomas, *Elements of Information Theory*, 2. baskı, Wiley, 2006.
- MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.
- Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal* 27, 1948.
- Shannon ve Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, 1949.

Son

Bölüm 34: Bilgi Kuramı

← Önceki bölüm · Alanın haritası · Sonraki bölüm →