Olasılıksal Grafik Modeller

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

Bayes ağları | Markov ağları | Koşullu bağımsızlık | Çıkarım | Parametre öğrenme | 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_12_Olasiliksal_Grafik_Modeller.html

GİRİŞ

Ders Hangi Soruyu Sorar?

> **Birbirine bağlı, belirsiz çok sayıda değişken nasıl modellenir ve sorgulanır?**

Olasılıksal grafik modeller, çok sayıda belirsiz değişkenin ortak dağılımını bir **çizge** ile gösterir: düğümler değişkenler, kenarlar doğrudan bağımlılıklardır. Çizgenin söylediği koşullu bağımsızlıklar dev bir dağılımı küçük parçalara böler; böylece hem modellemek hem de sorgulamak mümkün olur.

Bir bakışta

YAKLAŞIM olasılık ve çizge bir arada	**GİRDİ** değişkenler, bağımlılık yapısı, veri	**ÇIKTI** koşullu olasılık, en olası açıklama
BAŞARI ÖLÇÜTÜ olabilirlik, çıkarım doğruluğu	**MATEMATİK YÜKÜ** yüksek: olasılık	**VERİ İHTİYACI** orta

Kısa Tarihçe

1906
Markov, bir sonraki durumun yalnız şimdiki duruma bağlı olduğu zincirleri inceler.

1966
Baum ve Petrie saklı Markov modelinin (HMM) istatistiksel temelini yayımlar.

1977
Dempster, Laird ve Rubin, eksik veriyle öğrenmek için EM algoritmasını yayımlar.

1988
Pearl "Probabilistic Reasoning in Intelligent Systems" kitabıyla Bayes ağlarını YZ'ye taşır.

1989
Rabiner'in HMM eğitici makalesi konuşma tanımada standart başvuru olur.

2001
Lafferty, McCallum ve Pereira dizi etiketleme için koşullu rastgele alanları (CRF) önerir.

DERS

Temel Kavramlar

Ortak dağılım
Bütün değişkenlerin birlikte alabileceği değerlerin olasılığı.

Koşullu bağımsızlık
Bir değişken bilinince ötekiler birbirini etkilemez.

Bayes ağı
Yönlü, döngüsüz çizge; her düğümde koşullu olasılık tablosu.

Markov ağı
Yönsüz çizge; ilişkiler potansiyel fonksiyonlarıyla.

Saklı Markov modeli
Gözlenmeyen durum zinciri ve her durumdan bir gözlem.

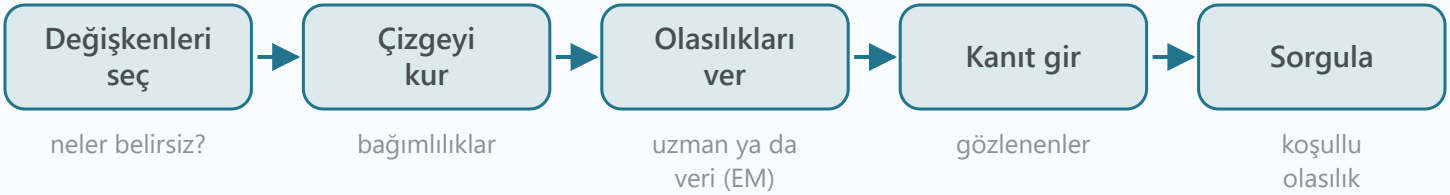
Kesin çıkarım
Değişken eleme, kavşak ağacı (junction tree).

Yaklaşık çıkarım
Örnekleme (MCMC), döngülü inanç yayılımı, değişimsel.

Öğrenme
Parametreler (EM) ve yapı: çizgenin kendisini veriden bulmak.

DERS

Tipik İş Akışı



Örnek: Bir tanı ağında sigara, kanser ve öksürük düğümleri vardır. Hekim tablolarını literatürden doldurur. Hasta öksürüyor ve sigara içiyorsa bu iki kanıt girilir; ağ kanser olasılığını yeniden hesaplar ve hangi testin belirsizliği en çok azaltacağını gösterir.

DERS

Derste Neler İşlenir?

1. Olasılık tekrarı; ortak dağılım ve çarpanlara ayırma
2. Bayes ağları ve d-ayrılma
3. Markov ağları ve çarpan çizgeleri
4. Kesin çıkarım: değişken eleme
5. İnanç yayılımı ve kavşak ağacı
6. Örneklemeyle çıkarım: Gibbs, MCMC
7. Değişimsel çıkarım
8. Saklı Markov modelleri: ileri-geri, Viterbi
9. Koşullu rastgele alanlar ve dizi etiketleme
10. Parametre öğrenme: en çok olabilirlik, Bayesçi
11. Eksik veriyle öğrenme: EM algoritması
12. Yapı öğrenme ve nedensel çıkarıma giriş

Kazandırdığı beceriler

pgmpy ile Bayes ağı kurmak

hmmlearn ile dizi modellemek

PyMC ya da Stan ile Bayesçi model

koşullu bağımsızlığı çizgeden okumak

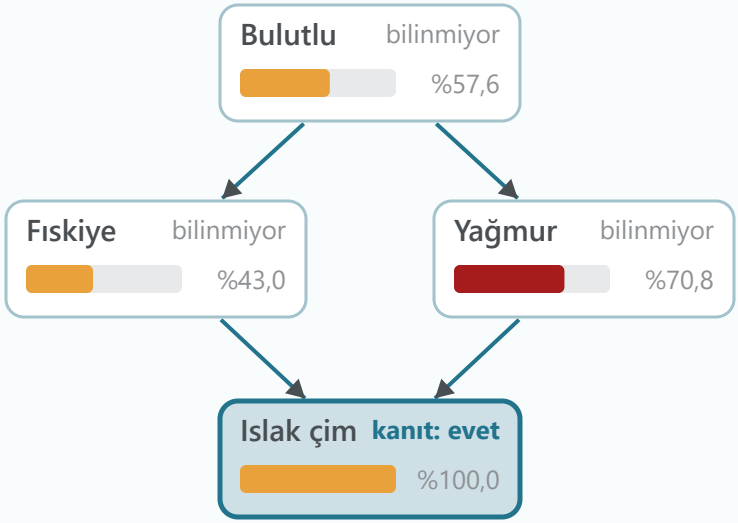
çıkarım sonucunu yorumlamak

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

DENE

Çim Neden Islak?

Sabah çim ıslak: yağmur mu yağdı, fıskiye mi çalıştı? Dört değişkenli Bayes ağında bir düğüme tıklayarak kanıt verin (evet / hayır / bilinmiyor); diğer düğümlerin olasılıkları tam olasılık hesabıyla güncellenir.



Kanıtları sil | Çim ıslak | Çim ıslak + fıskiye açık

Kanıt: Islak çim = evet
P(Yağmur | kanıt) = **%70,8**
P(Fıskiye | kanıt) = %43,0
16 durumdan uyumlu 8 tanesi toplandı · P(kanıt) = 0,647

Okların yönü nedenselliği gösterir ama çıkarım iki yöne de akar: ıslak çimden geriye, yağmura ve buluta doğru. Önce çimi ıslak yapın, sonra fıskiyeyi açın.

Dört değişkenin 16 birleşik durumu tablo olarak tutulmadı; ağ yalnız **yerel koşullu olasılıkları** saklar. Kanıt gelince numaralandırma bu küçük tablolardan istenen sonsal olasılığı hesaplar.

KONUM

Ön Koşullar

Bu dersten önce bilinmesi yararlı olan dersler ve neden gerektikleri.

Ders	Neden gerekir
Olasılık ve İstatistik	koşullu bağımsızlık, ortak dağılım
Algoritmalar	çizge dolaşma, dinamik programlama (Viterbi, ileri-geri)
Veri Yapıları	çizge gösterimi, ağaç ve öncelik kuyruğu

KONUM

Bu Ders Sana Göre mi?

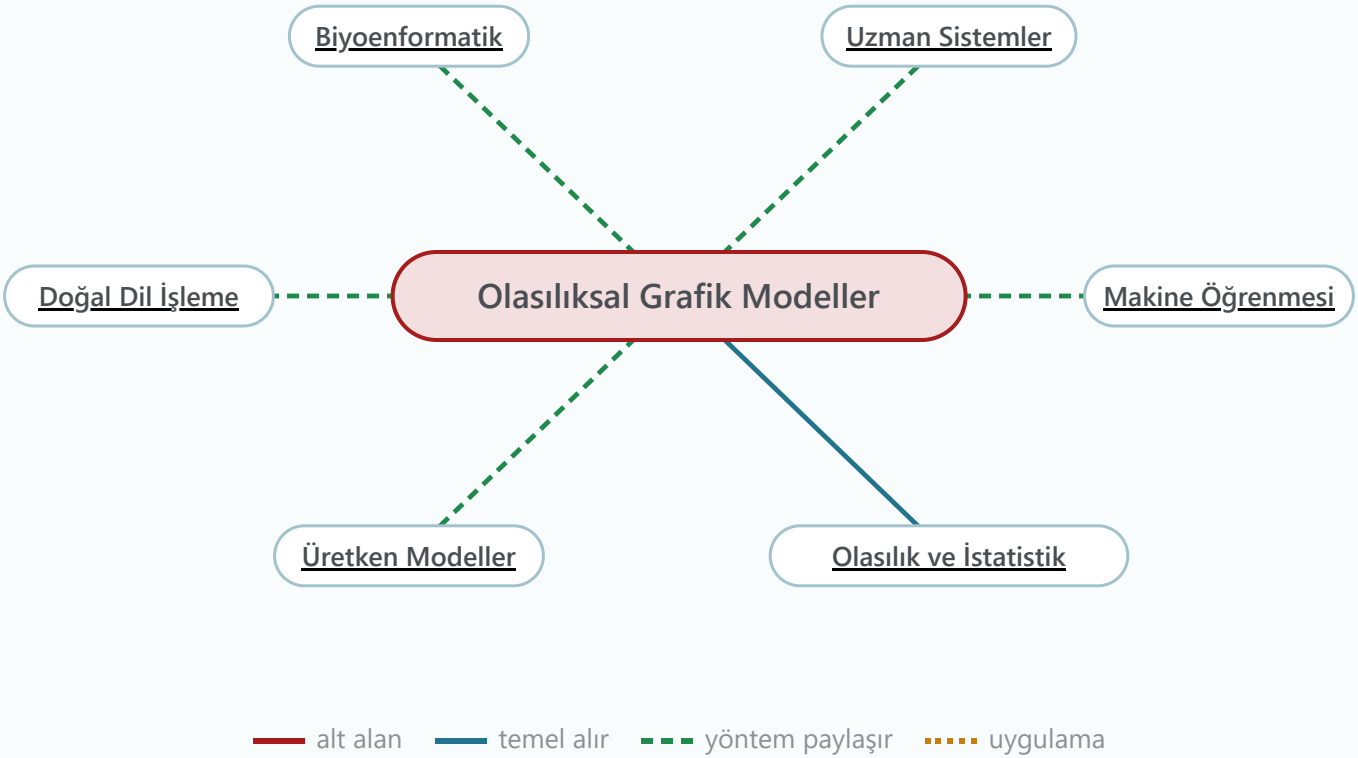
Kendine sor; "hayır" dediklerin için sağdaki sütun nereden başlayacağını gösterir.

Kendine sor	Hayırsa önce
Bayes kuralıyla bir koşullu olasılığı hesaplayabiliyor musun?	Olasılık ve İstatistik
Bir çizgeyi komşuluk listesiyle gösterip dolaşabiliyor musun?	Veri Yapıları dersindeki çizge gösterimi
Bir problemi dinamik programlamayla alt problemlere bölüp çözebiliyor musun?	Algoritmalar dersindeki dinamik programlama
En çok olabilirlik kestirimini bir örnekle açıklayabiliyor musun?	Makine Öğrenmesi

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

KONUM

Dersin Komşuları



Örtüştüğü Dersler

İlişki	Ders	Ortak konu
temel alır	Bu ders Olasılık ve İstatistik dersini temel alır	koşullu bağımsızlık, ortak dağılım
yöntem paylaşır	Makine Öğrenmesi ile yöntem paylaşır	olasılıksal sınıflandırıcılar, EM algoritması
yöntem paylaşır	Doğal Dil İşleme ve Metin Madenciliği ile yöntem paylaşır	saklı Markov modeli, koşullu rastgele alan
yöntem paylaşır	Uzman Sistemler ve Bilgi Temsili ile yöntem paylaşır	belirsizlik altında çıkarım
yöntem paylaşır	Üretken Modeller ile yöntem paylaşır	gizil değişkenli modeller, değişimsel çıkarım
yöntem paylaşır	Biyoenformatik ile yöntem paylaşır	dizi hizalamada saklı Markov modeli

Ayrıştığı Noktalar (1/2)

Ölçüt	Olasılıksal Grafik Modeller	Olasılık ve İstatistik	Makine Öğrenmesi
Temel soru	Birbirine bağlı, belirsiz çok sayıda değişken nasıl modellenir ve sorgulanır?	Belirsizlik nasıl ölçülür ve sınırlı veriden nasıl güvenilir sonuç çıkarılır?	Bir program örneklerden nasıl genelleme yapar?
Yaklaşım	olasılık ve çizge bir arada	olasılık modeli ve çıkarım	istatistiksel öğrenme
Girdi	değişkenler, bağımlılık yapısı, veri	örneklem verisi	etiketli ya da etiketsiz örnekler
Çıktı	koşullu olasılık, en olası açıklama	kestirim, güven aralığı, test sonucu	tahmin yapan model
Başarı ölçütü	olabilirlik, çıkarım doğruluğu	yansızlık, güç, anlamlılık	görülmemiş veride hata
Matematik yükü	yüksek: olasılık	kuramın kendisi	yüksek: doğrusal cebir, olasılık, optimizasyon
Veri ihtiyacı	orta	örneklem	orta ve çok

- **Olasılık ve İstatistik:** Olasılık ve istatistik birkaç değişkenin dağılımını inceler; bu ders yüzlerce değişkeni çizgeyle parçalayarak modellemeyi öğretir.
- **Makine Öğrenmesi:** MÖ çoğu zaman tek bir hedefi tahmin eder; OGM bütün değişkenlerin ortak dağılımını kurar ve herhangi birini sorgulatır.

Ayrıştığı Noktalar (2/2)

Ölçüt	Olasılıksal Grafik Modeller	Üretken Modeller	Uzman Sistemler
Temel soru	Birbirine bağlı, belirsiz çok sayıda değişken nasıl modellenir ve sorgulanır?	Bir model gerçek veriye benzeyen yeni örnekleri nasıl üretir?	Bir uzmanın bilgisi nasıl yazıya dökülür ve çıkarım için kullanılır?
Yaklaşım	olasılık ve çizge bir arada	veri dağılımını öğrenme	sembolik mantık ve kurallar
Girdi	değişkenler, bağımlılık yapısı, veri	çok büyük, çoğu etiketsiz veri	olgular ve uzman kuralları
Çıktı	koşullu olasılık, en olası açıklama	yeni örnek: görüntü, metin, ses	sonuç ve gerekçesi
Başarı ölçütü	olabilirlik, çıkarım doğruluğu	gerçekçilik ve çeşitlilik (FID vb.)	uzmanla uyum, açıklanabilirlik
Matematik yükü	yüksek: olasılık	yüksek	düşük: önerme ve yüklem mantığı
Veri ihtiyacı	orta	çok fazla	az; bilgi tabanı

- **Üretken Modeller:** Üretken modeller örnek üretmek için büyük sinir ağları kullanır; bu ders açık, yorumlanabilir yapı ve kesin çıkarım üzerinde durur.
- **Uzman Sistemler:** Uzman sistemler bilgiyi keskin kurallarla yazar; OGM aynı bilgiyi olasılıklarla yazar ve kanıtları tutarlı biçimde birleştirir.

UYGULAMA

Uygulama Alanları

Tıbbi tanı
Belirti ve testlerden hastalık olasılığı.

Konuşma tanıma
HMM ile sesten sözcük dizisi.

Dizi etiketleme
Sözcük türü ve varlık adı bulma (CRF).

Biyoenformatik
Gen bulma ve dizi hizalama.

Arıza teşhisi
Gözlenen belirtilerden en olası arıza.

Görüntü bölütleme
Komşu piksellerin etiket uyumu (MRF).

UYGULAMA

Riskler ve Etik

Tahminle doldurulan tablolar

Veri yoksa olasılık tabloları uzman tahminiyle dolar; yüzde 83 gibi kesin görünen sonuç bu tahminlerin belirsizliğini gizler.

Yanlış bağımsızlık varsayımı

Gerçekte bağımlı iki kanıtı bağımsız saymak aynı bilgiyi iki kez sayar; model aşırı emin bir sonuca varır.

Mahkemede olasılık hatası

Kanıtın olasılığını suçluluğun olasılığı sanmak (savcı yanılgısı) ve taban oranı unutmak masum birini suçlu gösterebilir.

Yakınsamamış örnekleme

MCMC zinciri erken durdurulursa sonsal dağılım yanlış çıkar; yakınsama tanıları raporlanmadan sonuç güvenilmez.

Ayrıntılı tartışma: Açıklanabilir Yapay Zekâ ve Yapay Zekâ Etiği.

UYGULAMA

Araçlar ve Örnek Veriler

Araçlar ve kütüphaneler

pgmpy | PyMC | Stan | hmmlearn | Pyro | GeNIe ve SMILE | bnlearn (R)

Klasik ağlar ve veri setleri

Asia (göğüs kliniği) ağı | ALARM ağı | Yağmur-fıskiye ağı | Insurance ağı | CoNLL-2003 | TIMIT

UYGULAMA

Sık Karıştırılanlar

Yaygın düşünce	Doğrusu
Kenar neden-sonuç ilişkisi demektir.	Bayes ağında kenar yalnız olasılıksal bağımlılıktır; nedensellik ayrıca varsayım ya da deney ister.
Bağlantı yoksa değişkenler bağımsızdır.	Bağımsızlık koşulludur; iki düğüm ortak bir çocukları gözlenince bağımlı hâle gelebilir.
Çıkarım her zaman kolayca hesaplanır.	Genel çizgelerde kesin çıkarım NP-zordur; büyük modellerde yaklaşık yöntemler gerekir.
Derin öğrenme bu alanı gereksiz kıldı.	Az veri, açıklanabilirlik ve belirsizlik ölçümü gereken yerlerde OGM hâlâ güçlüdür; ikisi birleştirilir de.

SONRASI

Bu Dersten Sonra

Bu dersi temel alan ya da genişleten dersler

Bu ders çoğunlukla uç noktadır; öteki dersleri temel alır.

İlgili meslekler

- Makine öğrenmesi araştırmacısı
- Veri bilimci
- Biyoistatistikçi
- Konuşma ve dil teknolojileri mühendisi

Olasılıksal grafik modeller olasılığı yapıyla buluşturur. Bu dersten sonra gizil değişkenli üretken modellere, dizi modelleriyle doğal dil işlemeye ve biyoenformatiğe ya da nedensel çıkarıma geçmek doğal adımlardır.

SONRASI

Kaynaklar

- Koller ve Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, 2009.
- Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, 1988.
- Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.
- Darwiche, *Modeling and Reasoning with Bayesian Networks*, Cambridge University Press, 2009.
- Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proceedings of the IEEE* 77(2), 1989.

Son

Bölüm 12: Olasılıksal Grafik Modeller

← Önceki bölüm · Alanın haritası · Sonraki bölüm →