



Bölüm 10: Algoritmik Etik

Algoritmik Oyun Kuramı



Giriş

- Algoritmik kararlar → kredi onayı, işe alım, sosyal medya önerileri
- Neden önemlidir?
 - Adalet ve etik ihlalleri → toplumsal güven kaybı
- Örnekler:
 - Amazon işe alım algoritması
 - Kredi skarlama sistemleri



Algoritmik Adalet Kavramları

- Demografik Parite: Gruplar eşit oranda olumlu karar almalı
- Eşit Doğruluk (Equalized Odds): Yanlış pozitif / negatif oranları dengeli
- Predictive Parity: Tahmin doğruluğu gruplar arasında eşit
- Örnek: Kredi skorlama → farklı etnik grupların adil değerlendirilmesi



Önyargı ve Ayrımcılık

- Veri kaynaklı önyargı: Tarihsel eşitsizlikler yansır
- Algoritma tasarım hatası: Model adaletsizliği
- Örnekler:
 - Yüz tanıma sistemlerinde etnik ayrımcılık
 - İşe alım algoritmalarında cinsiyet önyargısı



Etik İlkeler ve Standartlar

- Şeffaflık: Karar süreçleri anlaşılır olmalı
- Hesap verebilirlik: Hatalı karar sorumlusu net olmalı
- Adalet ve eşitlik: Sosyal etkiler göz önünde olmalı
- Örnekler: GDPR, AI Ethics Guidelines



Stratejik Tasarım ve Teşvikler

- Kullanıcılara doğru teşvikler verilmesi
- Kötü niyetli manipülasyonu önleme
- Örnek: Sosyal medya → tıklama tuzağı yerine doğru öneri



Gerçek Hayat Uygulamaları

- Kredi skortlama ve bankacılık
- İşe alım ve insan kaynakları
- Sağlık sektörü → teşhis ve tedavi önerileri
- Etkinlik: Etik bir algoritmanın tasarlanması ve test edilmesi



Tartışma Soruları

- Algoritmalar her zaman tarafsız olabilir mi?
- Etik ve verimlilik arasındaki denge nasıl sağlanır?
- Algoritmik adaletin sağlanmasında en büyük zorluk nedir?



Algoritmalar her zaman tarafsız olabilir mi?

- Teorik olarak:
- Mümkün, ama gerçek veriler ve tasarım kısıtları nedeniyle zordur
- Neden zor:
 - Tarihsel eşitsizlikler ve önyargılar veriye yansır
 - Algoritma tasarımı hataları → model adaletsizliği
 - Stratejik manipülasyon ve bilinçsiz öncelikler
- Örnekler:
 - Yüz tanıma sistemlerinde etnik ayrımcılık
 - İşe alım algoritmalarında cinsiyet önyargısı



Etik ve verimlilik arasındaki denge nasıl sağlanır?

- Etik → adil karar, şeffaflık, hesap verebilirlik
- Verimlilik → hızlı ve doğru sonuç, maliyet etkinliği
- Denge yöntemleri:
 - Teşvik tasarımı → kullanıcıyı etik davranmaya yönlendirme
 - Adil algoritmik optimizasyon → bazı doğruluk kayıplarına razı olma
 - Regülasyon ve standartlar → etik sınırları belirleme
- Örnek: Kredi skorelama algoritmasında düşük gelirli kullanıcıları adil değerlendirmek



Algoritmik adaletin sağlanmasında en büyük zorluk?

- Veri kaynaklı önyargı: Tarihsel eşitsizlikler ve eksik veri
- Algoritma tasarım hatası: Model karmaşıklığı, yanlış optimizasyon
- Stratejik davranış: Kullanıcı veya sistem manipülasyonu
- Sosyal ve kültürel farklılıklar: Aynı algoritma farklı gruplarda farklı etkiler yaratabilir
- Örnekler:
 - Amazon işe alım algoritması → cinsiyet önyargısı
 - Yüz tanıma sistemleri → etnik farklılıklar



SON