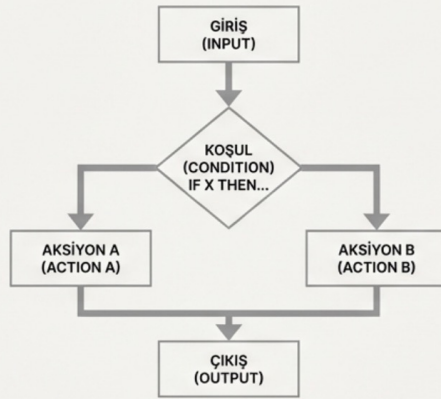


MAKİNE ÖĞRENİMİ: TEORİK TEMEL VE SPOTIFY

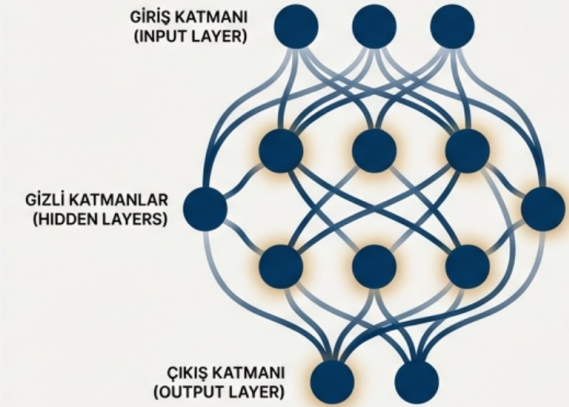
1030510924-Hayrunnisa YILDIZ

MAKİNE ÖĞRENİMİNİN TEORİK TEMELLERİ VE MODERN MİMARİLERİ



Klasik Programlama

Klasik programlama, geliştiricinin tüm kuralları, koşulları ve mantığı açıkça tanımladığı; bilgisayarın bu önceden belirlenmiş talimatları aynen uyguladığı bir yazılım geliştirme yaklaşımıdır.



Makine Öğrenimi

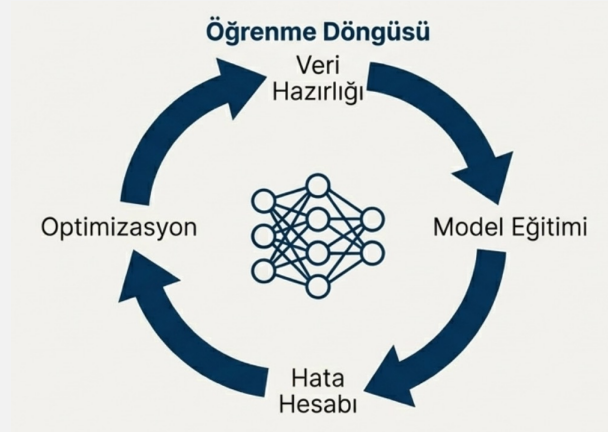
Makine öğrenimi, bilgisayara kuralları tek tek vermek yerine çok sayıda örnek veri üzerinden istatistiksel modeller öğrenmesini sağlayarak belirsizlik içeren durumlarda bile tahmin ve karar üretebilen bir yaklaşımı ifade eder.

BİR MAKİNE NASIL ÖĞRENİR?

- 1- Makine öğrenimi, bir algoritmanın çok sayıda örneği görerek bir kalıp çıkarmasıdır. Makine verilerdeki tekrar eden ilişkileri bulur. Algoritma başlangıçta hiçbir şey bilmez, sadece veriye bakarak tahmin üretmeyi öğrenir.
- 2- Makine veriyi aldığı anda önce rastgele bir tahmin yapar. Daha sonra gerçek değer ile tahmin arasındaki fark bir **hata (loss)** fonksiyonu ile ölçülür. Bu hata, modelin ne kadar kötü tahmin yaptığını gösterir. Amaç, bu hatayı her iterasyonda biraz daha azaltmaktır.
- 3- Modelin içindeki sayısal parametreler hatayı düşürmek için küçük adımlarla güncellenir. Eğer model yanlış bir tahmin yaptıysa sistem parametreleri değiştirerek daha iyi tahmin yapmayı öğrenir.

BİR MAKİNE NASIL ÖĞRENİR?

- 4- Bu tahmin–hata–güncelleme süreci **binlerce kez** tekrar edilir. Bu tekrarlar sonunda model, verinin içindeki ilişkileri öğrenmiş olur. Model tekrar eden ilişkileri ne kadar iyi kavarsa, o kadar doğru tahmin yapar.
- Model, verilerden örüntü çıkararak kendi iç parametrelerini optimize eder ve zamanla daha doğru tahminler üretir. Yani “öğrenme” aslında hatayı azaltmak için yapılan sürekli matematiksel ayarlamadır.



MAKİNE ÖĞRENİMİNİN 3 ANA YAKLAŞIMI

- **I) Denetimli Öğrenme (Supervised Learning)**
- Model eğitim sırasında hem girdi hem çıktı görür ve bunlar arasındaki ilişkiyi öğrenir.
- **Algoritma Örnekleri:**
- Lineer Regresyon (Linear Regression): Sürekli değer tahmini
- Random Forest: Karar ağacı tabanlı sınıflandırma ve regresyon
- XGBoost / Gradient Boosting: Ensemble tabanlı güçlü tahmin modelleri

MAKİNE ÖĞRENİMİNİN 3 ANA YAKLAŞIMI

- **2) Denetimsiz Öğrenme (Unsupervised Learning)**
 - Modelde çıktı etiketi yoktur, veri içinde gizli yapıları veya örüntüleri keşfeder.
 - **Algoritma Örnekleri:**
 - K-Means: Kümeleme (Clustering)
 - PCA (Principal Component Analysis): Boyut indirgeme
- **3) Pekiştirmeli Öğrenme (Reinforcement Learning)**
 - Model, bir ajan gibi ortamda aksiyon alır ve ödül veya ceza ile geri bildirim alır; hedef, toplam ödülü maksimize etmektir.

SPOTIFY MÜZİK ZEVKİNİZİN HARİTASINI NASIL ÇIKARIYOR?

- Spotify'nın öneri sistemi, kullanıcı deneyimini optimize etmek için arka planda birden çok makine öğrenimi modelini eş zamanlı olarak çalıştırır. Bu sistem üç temel veri kaynağını analiz eder:
- **Şarkıların özellikleri:** Tempo, enerji, ton, ritim gibi ses sinyali tabanlı parametreler çıkarılır ve her parça bir vektör temsili ile modellenir.
- **Kullanıcı davranışları:** Atlama geçmişi, dinleme süreleri, tekrarlanan parçalar gibi davranışlar izlenir ve feature engineering ile modele beslenir.
- **İşbirlikçi filtreleme (Collaborative Filtering):** Milyonlarca kullanıcı arasında benzer dinleme alışkanlıkları tespit edilir, benzer kullanıcıların tercihleri temel alınarak yeni öneriler üretilir.



HAFTALIK KEŞİF

- Her hafta Spotify'ın sana özel hazırladığı müzik listesi, aslında tek bir algoritmanın ürünü değildir. Arkada paralel çalışan üç farklı makine öğrenimi modeli vardır ve bu modellerden gelen sinyaller birleştirilerek nihai öneri listesi oluşturulur.
- **I. Collaborative Filtering (CF) – Kullanıcı Tabanlı Benzerlik Modellemesi**
- Bu model, seni milyonlarca kullanıcı arasındaki davranış örüntülerine göre konumlandırır. Temel fikir “benzer kullanıcılar benzer şeyleri sever” prensibidir.
- Dinleme geçmişin, atlama davranışların ve playlist etkileşimlerin vektörleştirilir ve yüksek boyutlu bir kullanıcı–şarkı matrisinde seninle **istatistiksel olarak en benzer kullanıcı kümeleri** bulunur. Eğer 5 milyon kişilik bir grup seninle çok benzer dinleme örüntülerine sahipse, onların yüksek puan verdiği veya sıkça dinlediği şarkılar senin için öneri adaylarına eklenir.

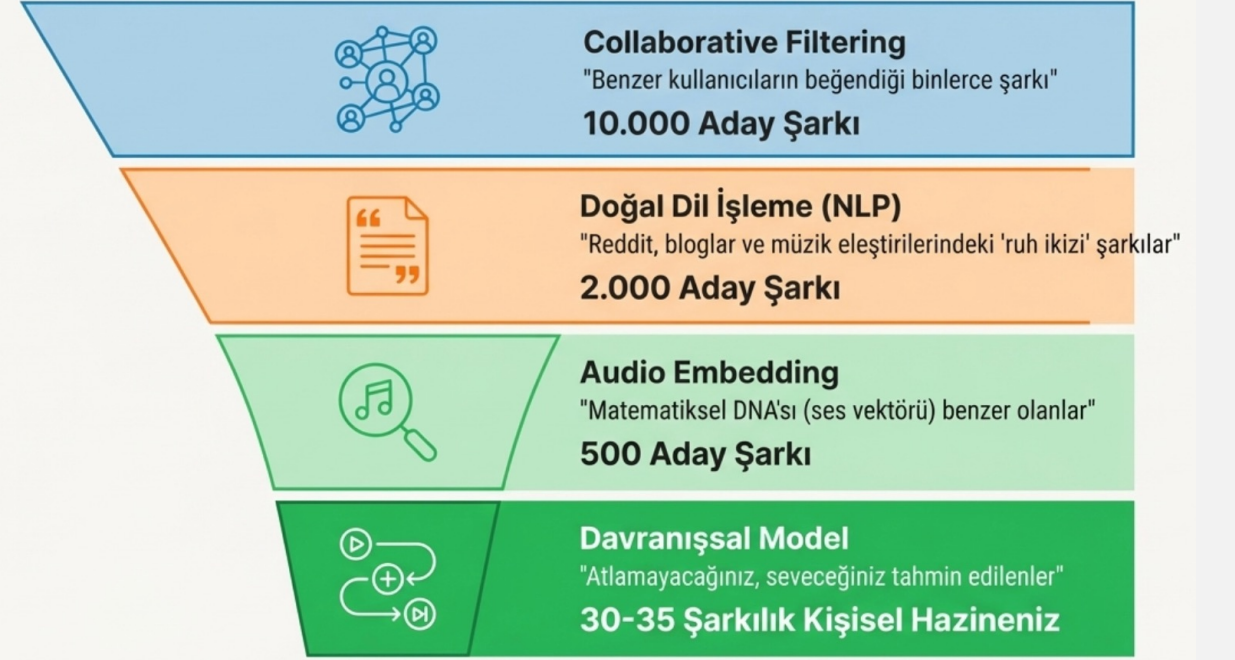
- **2. NLP Tabanlı “Music2Vec” – Şarkılar Arası Anlamsal İlişki**

- Spotify, yalnızca sayısal dinleme verilerini değil, internetteki şarkı yorumlarını, incelemeleri ve metadatanın dilsel içeriğini de işler. Burada Word2Vec veya Transformer tabanlı embedding tekniklerinin müzik nesnelerine uygulanmış bir versiyonu kullanılır. Şarkılar, kelimeler gibi düşünülür ve “semantic embedding” uzayında konumlandırılır:
- Bir şarkı için yapılan yorumlarda sürekli “dreamy” veya “nostalgic” gibi benzer kelimeler geçiyorsa, model bu kelimelerden yola çıkarak şarkının anlamını temsil eden bir vektör oluşturur.
- Aynı kelimeler başka bir şarkının bağlamında da sık geçiyorsa, bu iki şarkı embedding uzayında birbiriyle yakınlaşır.
Bu model, müziklerin duygusal tonu, atmosferi ve kültürel bağlamı gibi soyut kavramları yakalar.

- **3. Audio Embedding – Ses Sinyallerinden 128 Boyutlu Vektör**

- Bu aşamada Spotify herhangi bir metin ya da kullanıcı davranışı olmaksızın, doğrudan ham ses dalga formundan özellik çıkarır. CNN tabanlı bir ses modeli, her şarkıyı frekans spektrumu, ritim, timbre ve enerji gibi yüzlerce sinyali işleyerek 128 boyutlu bir vektöre dönüştürür.
Ardından bu vektörler arasındaki kosinüs benzerliği veya öklidyen mesafe hesaplanır.
Böylece sistem, “Bu iki şarkı frekans yapıları ve ritmik patternleri bakımından birbirine çok yakın” diyebilir. Bu model tamamen içerik tabanlıdır.

- Discover Weekly final olarak şöyle çalışır:
- Benzer kullanıcılar → 10 bin aday
- NLP benzerliği → 2 bin aday
- Ses benzerliği → 500 aday
- Kullanıcı davranış modeli → En iyi 30-35 şarkı
- Spotify her hafta bunu yeniden hesaplar.



- KAYNAKÇA
- <https://research.atspotify.com/2025/9/generalized-user-representations-for-large-scale-recommendations>
- <https://ijisrt.com/user-based-spotify-recommendation-system-using-machine-learning-algorithms>
- <https://arxiv.org/abs/2008.11432>
- <https://www.sciencedirect.com/science/article/pii/S2772941925000924>

BENİ DİNLEDİĞİNİZ İÇİN TEŞEKKÜR EDERİM.

HAYRUNNİSA YILDIZ