

VERİ MADENCİLİĞİ

SINIFLANDIRMA (CLASSIFICATION)

Sınıflandırma

Ele alınan veri madenciliği algoritmalarından ilki olan sınıflandırma, temel olarak veritabanındaki gizli örüntülerin ortaya çıkarılması yani verinin ortak özelliklerine göre ayrıştırılması için kullanılır. Sınıflandırma bir öğrenme algoritmasına dayanır, öncelikle veritabanının bir kısmı örnek veri kümesi olarak belirlenerek eğitim amacıyla kullanılır ve sınıflandırma kuralları oluşturulur, daha sonra bu kurallar yardımıyla yeni bir durum ortaya çıktığında nasıl karar verileceği belirlenir. Böylece hangi sınıfa ait olduğu bilinmeyen bir kayıt için bir sınıf belirlenebilir.

Sınıflandırma

Amaç: Yeni bir kayıt geldiğinde, bu kaydı geliştirilen modeli kullanılarak mümkün olduğunca doğru bir sınıfa atamak.

- Verinin dağılımına göre bir model bulunur
- Bulunan model, başarımı belirlendikten sonra niteliğin gelecekteki değerini
- Ya da bilinmeyen değerini tahmin etmek için kullanılır
 - model başarımı: doğru sınıflandırılmış sınıfa kümesi örneklerinin oranı
- Veri madenciliği uygulamasında:
 - ayrık nitelik değerlerini tahmin etmek: sınıflandırma
 - sürekli nitelik değerlerini tahmin etmek: öngörü

Özet

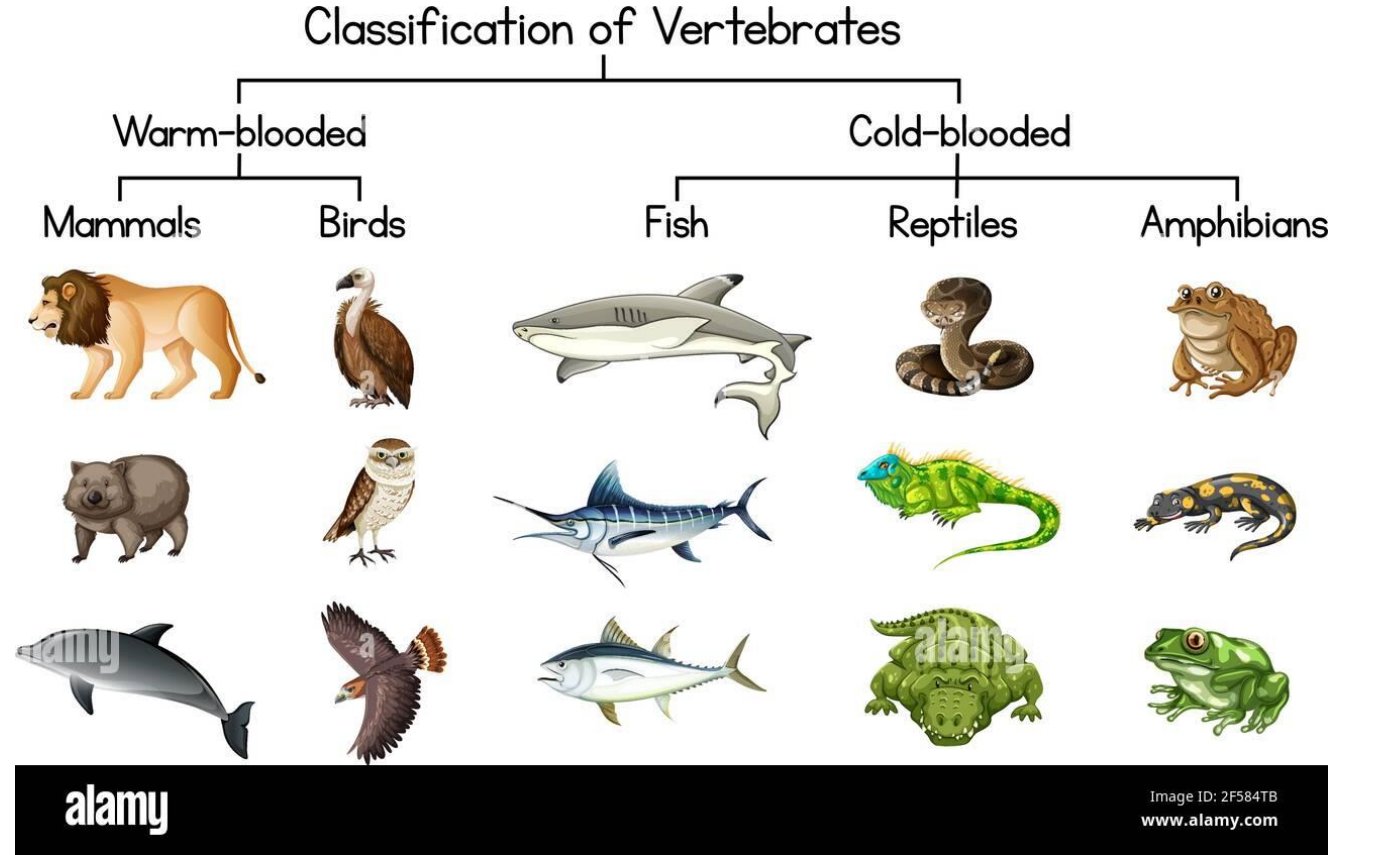
Sınıflandırma;

I. Yeni bir verinin yer alacağı grubun (sınıfın) belirlenmesi,

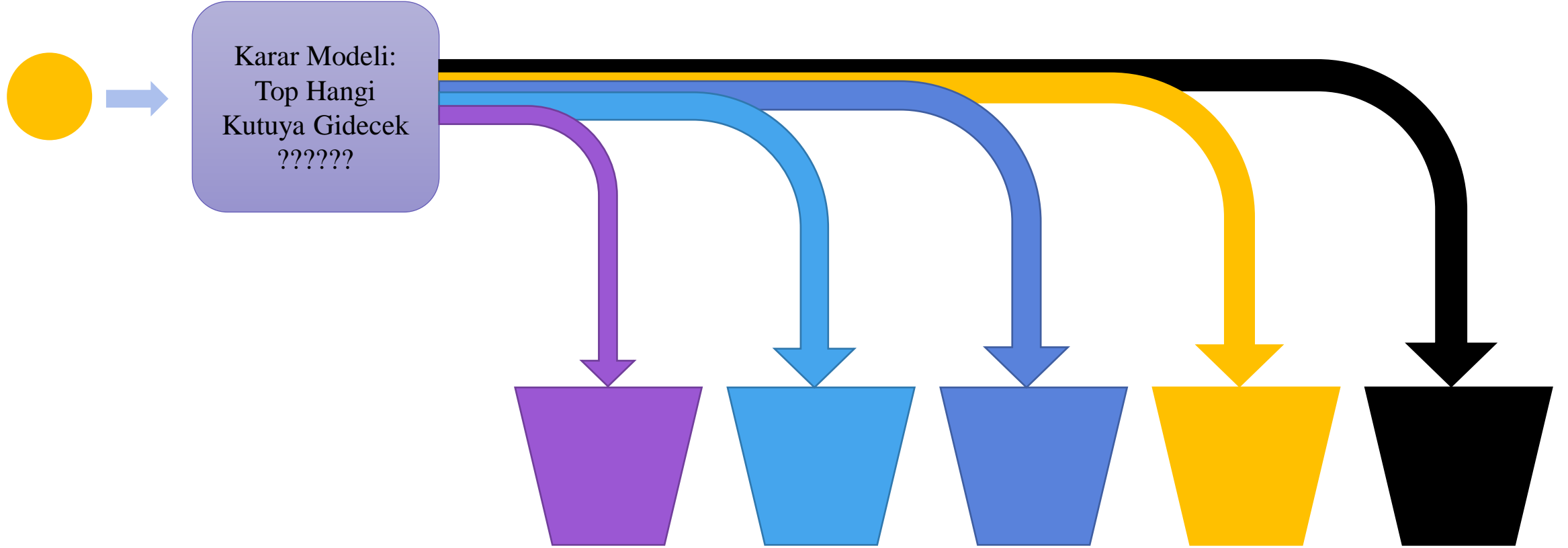
II. Önceden gruplanmış (sınıflandırılmış) verilerin doğru sınıflandırılıp sınıflandırılmadığının tespiti ve yanlış sınıflandırma varsa gözlemin (nesnenin) doğru gruba atanması

Temel olarak yaptığı şey yeni bir nesnenin niteliklerini inceleme ve bu nesneyi önceden tanımlanmış bir sınıfa atamaktadır. Burada önemli olan ,bir sınıfın özelliklerinin önceden net bir şekilde belirtilmiş olması gerektiğidir.

Dr.Günay TEMÜR



Örnek



Sınıflandırma Uygulamaları

- Kredi başvurusu değerlendirme
- Kredi kartı harcamasının sahtekarlık olup olmadığına karar verme
- Hastalık teşhisi
- Ses tanıma
- Karakter tanıma
- Gazete haberlerini konularına göre ayırma
- Kullanıcı davranışları belirleme

Sınıflandırma İşlemi

Sınıflandırma işlemi üç aşamadan oluşur:

1. Model oluşturma
2. Model değerlendirme
3. Modeli kullanma

Sınıflandırma İşlemi: Model Oluşturma

1. Model Oluşturma:

- ❑ Her nesnenin sınıf etiketi olarak tanımlanan niteliğinin belirlediği bir sınıfta olduğu varsayılır
- ❑ Model oluşturmak için kullanılan nesnelerin oluşturduğu veri kümesi öğrenme kümesi olarak tanımlanır
- ❑ Model farklı biçimlerde ifade edilebilir
 - IF – THEN – ELSE kuralları ile
 - Karar ağaçları ile
 - Matematiksel formüller ile

Sınıflandırma İşlemi: Model Değerlendirme

2. Model Değerlendirme:

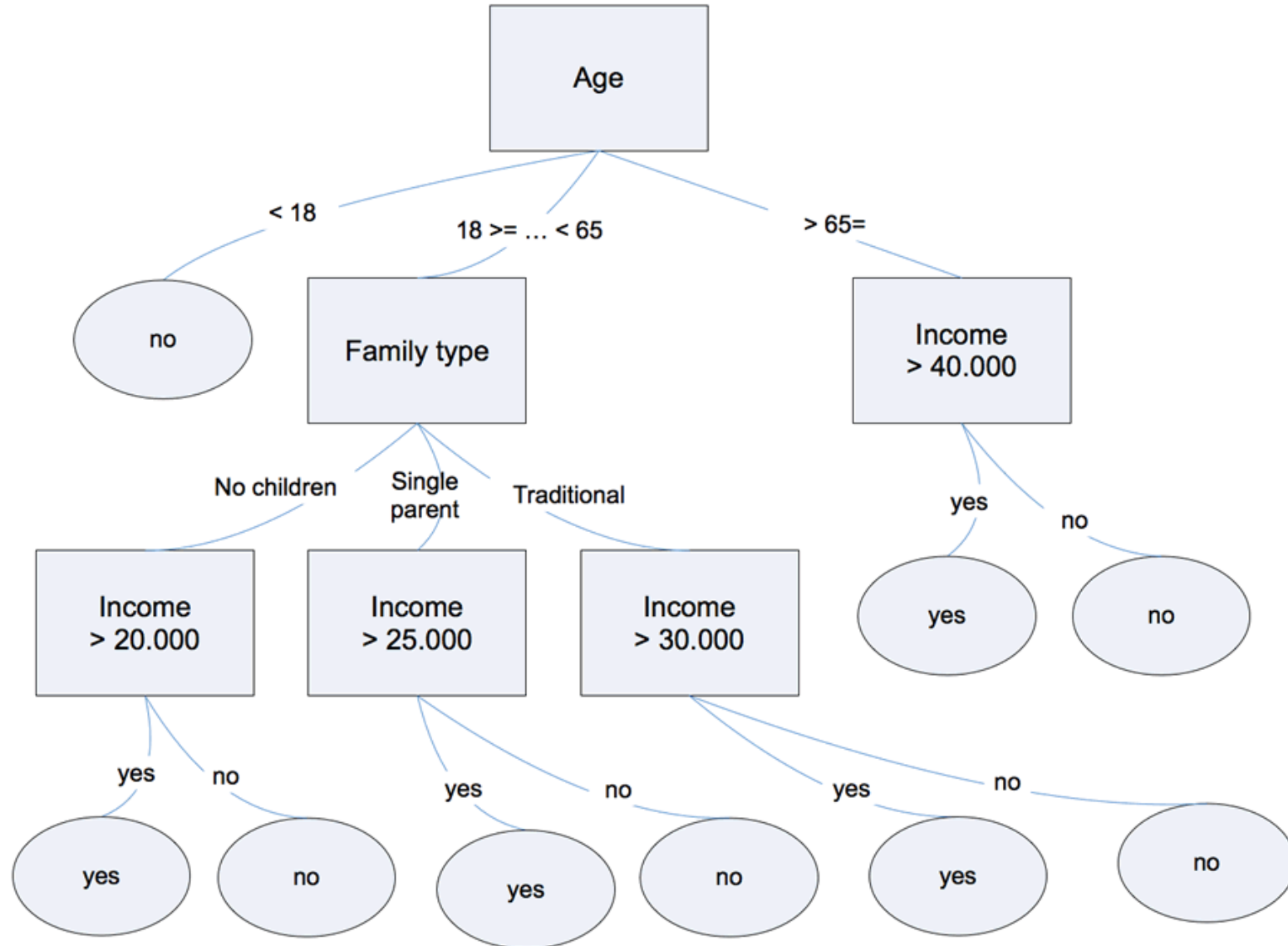
- ❑ Modelin başarımı (doğruluğu) sınıama kümesi örnekleri kullanılarak belirlenir.
- ❑ Sınıf etiketi bilinen bir sınıama kümesi örneği model kullanılarak belirlenen sınıf etiketiyle karşılaştırılır.
- ❑ Modelin doğruluğu, doğru sınıflandırılmış sınıama kümesi örneklerinin toplam sınıama kümesi örneklerine oranı olarak belirlenir.
- ❑ Sınıama kümesi model öğrenirken kullanılmaz.

Sınıflandırma İşlemi: Modeli Kullanma

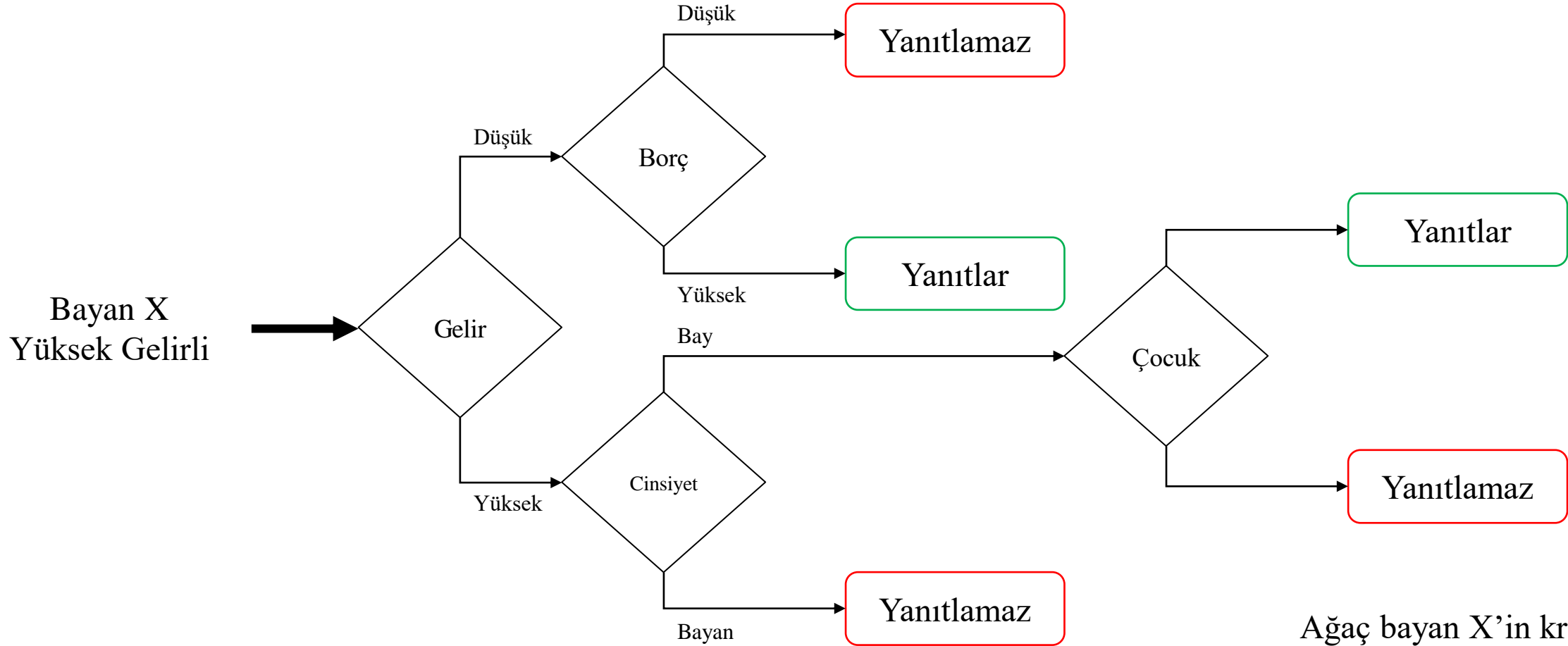
3. Modeli kullanma:

- Model daha önce görülmemiş örnekleri sınıflandırmak için kullanılır
 - Örneklerin sınıf etiketlerini tahmin etme
 - Bir niteliğin değerini tahmin etme

Örnek: Bir kredi talebini sınıflandırmak için karar ağacı



Örnek: Bir Kredi Kartı Kampanyasında Kişinin Yanıtı



Ağaç bayan X'in kredi kartı kampanyasına yanıt vermeyeceğini öngörür.

BİTTİ 😊

Kaynaklar:
Yrd.Doç.Dr. Kadriye ERGÜN Ders Notları Balıkesir Üniversitesi MMF Endüstri Mühendisliği Bölümü

Dr.Günay TEMÜR