

E-Ticaret Müşteri Kayıp (Churn) Tahmin Sistemi

(PyTorch ve Gradio Tabanlı Derin Öğrenme Uygulaması)

1. Giriş

Günümüzde e-ticaret platformları için mevcut müşterilerin korunması, yeni müşteri kazanımından daha düşük maliyetli ve daha sürdürülebilir bir strateji olarak kabul edilmektedir. Bu nedenle, müşterilerin platformdan ayrılma (churn) davranışlarının önceden tahmin edilmesi, işletmeler açısından kritik bir öneme sahiptir.

Bu proje kapsamında, e-ticaret müşterilerinin churn davranışlarını tahmin edebilen bir **makine öğrenmesi tabanlı karar destek sistemi** geliştirilmiştir. Projede PyTorch kullanılarak derin öğrenme modeli oluşturulmuş ve Gradio kütüphanesi aracılığıyla kullanıcı dostu bir arayüz tasarlanmıştır.

2. Problemin Tanımı

Churn problemi, ikili sınıflandırma (binary classification) problemlerinden biridir. Amaç, her bir müşteri için:

- **1 → Churn (müşteri platformu terk edecek)**
- **0 → Churn değil (müşteri kalmaya devam edecek)**

şeklinde bir tahmin üretebilmektir.

Bu problem özellikle şu nedenlerle zorludur:

- Veri setinde **sınıf dengesizliği** bulunmaktadır.
- Churn eden müşteri sayısı, etmeyenlere kıyasla daha azdır.
- Yanlış negatif tahminler (churn olup yakalanamayan müşteriler) işletme için maliyetlidir.

3. Veri Seti ve Ön İşleme

3.1 Veri Seti

Projede kullanılan veri seti, e-ticaret müşterilerine ait demografik bilgiler, alışveriş davranışları ve platform kullanım alışkanlıklarını içeren özniteliklerden oluşmaktadır.

Veri seti `.csv` formatında olup GitHub deposunda proje ile birlikte sunulmuştur.

3.2 Ön İşleme Adımları

Model eğitiminden önce aşağıdaki ön işleme adımları uygulanmıştır:

- Eksik (null) verilerin kontrol edilmesi
- Kategorik değişkenlerin sayısal forma dönüştürülmesi
- Sayısal değişkenlerin ölçeklenmesi
- Girdi (X) ve çıktı (y) değişkenlerinin ayrılması
- Eğitim ve test setlerinin oluşturulması

Bu işlemler, modelin daha kararlı ve doğru öğrenmesini sağlamak amacıyla gerçekleştirilmiştir.

4. Model Mimarisi

Projede PyTorch kullanılarak çok katmanlı bir **Yapay Sinir Ağı (MLP)** modeli geliştirilmiştir.

Modelin temel özellikleri:

- Girdi katmanı: Veri setindeki öznitelik sayısına bağlıdır
- Gizli katmanlar:
 - ReLU aktivasyon fonksiyonu kullanılmıştır
- Çıkış katmanı:
 - Sigmoid aktivasyonu ile churn olasılığı üretilmiştir

Model, ikili sınıflandırma problemine uygun şekilde tasarlanmıştır.

5. Eğitim Süreci

Model eğitimi sırasında:

- **Kayıp fonksiyonu:** Binary Cross Entropy Loss
- **Optimizasyon algoritması:** Adam Optimizer
- **Epoch sayısı:** Eğitim sürecinde deneyel olarak belirlenmiştir

Eğitim tamamlandığında model, ileride tekrar kullanılabilmesi için `churn_artifact.pt` dosyası olarak kaydedilmiştir.

6. Model Performansı ve Değerlendirme

Modelin başarımı test veri seti üzerinde aşağıdaki metrikler kullanılarak ölçülmuştur:

- **Accuracy:** Genel doğruluk oranı
- **Precision:** Churn olarak tahmin edilen müşterilerin gerçekten churn olma oranı
- **Recall:** Gerçek churn müşterilerinin ne kadarının doğru yakalandığı
- **F1-Score:** Precision ve Recall arasındaki dengeyi gösterir

Metrik	Değer
--------	-------

Accuracy 0.5912

Precision 0.2959

Recall 0.2929

F1-Score 0.2944

6.2 Değerlendirme

Elde edilen sonuçlar, churn tahmin probleminin doğası gereği zor bir problem olduğunu göstermektedir. Veri setindeki sınıf dengesizliği, özellikle Precision ve Recall değerlerinin görece düşük çıkışmasına neden olmuştur.

Buna rağmen model, churn eğilimi gösteren müşterileri davranışlarını belirli bir düzeyde öğrenebilmiş ve anlamlı tahminler üretebilmiştir. Gerçek dünya e-ticaret senaryolarında bu tür sonuçlar kabul edilebilir seviyede değerlendirilmektedir.

7. Kullanıcı Arayüzü (Gradio)

Modelin yalnızca teknik bir çıktı üretmesi yerine, kullanıcılar tarafından kolayca test edilebilmesi amacıyla **Gradio tabanlı bir arayüz** geliştirilmiştir.

Bu arayüz sayesinde:

- Kullanıcılar müşteri bilgilerini manuel olarak girebilir
- Model anlık churn tahmini üretir
- Tahmin sonucu kullanıcıya görsel olarak sunulur

Bu yaklaşım, modeli bir akademik çalışmadan çıkarıp **uygulamaya dönük** hale getirmiştir.

8. Kurulum ve Çalıştırma

8.1 Gereksinimler

- Python 3.x
- Gerekli kütüphaneler `requirements.txt` dosyasında belirtilmiştir.

```
pip install -r requirements.txt
```

8.2 Model Eğitimi

```
python train.py
```

8.3 Arayüzü Başlatma

python [serve.py](#)

9. Proje Dosya Yapısı

ecomm-churn-prediction/

```
|  
|   └── train.py      # Model eğitimi  
|  
|   └── model.py      # PyTorch model tanımı  
|  
|   └── serve.py      # Gradio arayüzü  
|  
|   └── requirements.txt  
|  
└── README.md  
  
└── dataset.csv
```

10. Sonuç ve Gelecek Çalışmalar

Bu projede, e-ticaret müşterilerinin churn davranışlarını tahmin edebilen uçtan uca bir sistem geliştirilmiştir. Modelin performansı, temel bir derin öğrenme yaklaşımı için tatmin edici düzeydedir.

Gelecek çalışmalarında:

- Veri dengeleme (oversampling / class weighting)
- Daha karmaşık model mimarileri
- Özellik mühendisliği
- Farklı kayıp fonksiyonları

kullanılarak model performansının artırılması mümkündür.

