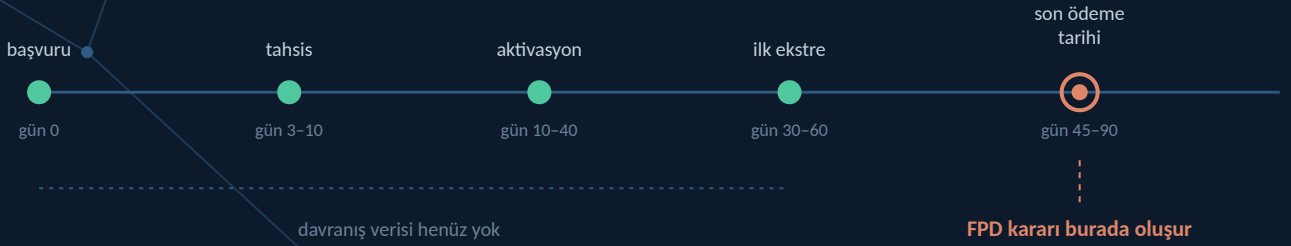


First Payment Default

İlk ödeme temerrüdünü genel temerrütten ayıran şey, hedef tanımı, modellenen popülasyon ve başvuru anında elde gerçekten olan bilgi.



Bu doküman ne anlatıyor?

First payment default (FPD), bir kredi kartı hesabının **ilk ekstresinin asgari ödemesini yapmamasıdır**. Teknik olarak temerrüdün bir alt kümesidir; pratikte ise ayrı bir problemdir — çünkü modelin elinde davranış verisi yoktur, olayın bir kısmı kredi riski bile değildir ve hedefin nasıl tanımlandığı sonucu modelden daha çok etkiler.

Üç cümlede FPD

Nadir ve keskin bir olaydır. Tipik oran %0,3–2 bandındadır; bu yüzden tek başına Gini yanıltıcıdır, üst dilim yakalama ve PR-AUC olmadan model değerlendirilemez.

Saf bir başvuru-anı problemidir. Müşterinin banka nezdinde davranışı yoktur; sinyalin neredeyse tamamı bürodan, beyandan, kanaldan ve ilk kullanım deseninden gelir.

Homojen bir hedef değildir. İçinde kredi riski, başvuru sahtekarlığı ve operasyonel hata bir aradadır. Bunları ayırmadan kurulan model, üç farklı olguyu tek skorla açıklamaya çalışır.

İçindekiler

- 01 FPD nedir, neden ayrı bir problem?
- 02 Hedef tanımı — en kritik karar
- 03 Kartın ilk 90 günü: zaman eksen
- 04 Modellenebilir popülasyon: kim içeride?
- 05 FPD ile genel temerrüt farkı
- 06 Nadir olay: sınıf dengesizliği ve metrik seçimi
- 07 Değişkenler: başvuru anında elde ne var?
- 08 FPD'ye özgü güçlü sinyaller
- 09 İki model, iki an: başvuru ve erken yaşam
- 10 Fraud mu, kredi riski mi?
- 11 Kohort kurgusu ve olgunlaşma
- 12 Modelleme pratiği: nadir sınıfta XGBoost
- 13 Değerlendirme: Gini tek başına yetmez
- 14 Skorun kullanımı: onay, limit, aktivasyon
- 15 İzleme ve erken geri besleme
- 16 Yaygın tuzaklar
- 17 Kontrol listesi ve hızlı referans
- 18 Kaynaklar ve kapsam notu

Kapsam ve ölçüm notu

Anlatım **kredi kartı** ürünü ve bireysel başvuru hattı esas alınarak yazılmıştır. İhtiyaç kredisi ve taksitli ürünlerde FPD mantığı benzerdir, fakat “ilk ekstre” yerine “ilk taksit” tanımı ve farklı bir aktivasyon kurgusu geçerlidir.

Şekillerdeki oranlar ve performans sayıları **temsildir** — tipik davranışın yönünü göstermek içindir, belirli bir portföyün gerçek değerleri değildir.

01 FPD nedir, neden ayrı bir problem?

Tanım basittir: hesap açıldıktan sonra kesilen **ilk ekstrenin** asgari ödemesi, son ödeme tarihini takip eden pencerede yapılmamıştır. Problemi zorlaştıran şey tanım değil, tanımın arkasındaki **bilgi yokluğudur**.

Üç farklı olgu, tek bir etikette toplanıyor

Ödeme kapasitesi	Başvuru sahtekarlığı	Operasyonel hata
Gerçek kredi riski: gelir yetmiyor, borç yükü zaten yüksek, limit ilk ayda dolduruluyor. Skorlanabilir - asıl hedef	Hiç ödeme niyeti yok: sahte/şişirilmiş beyan, kimlik kullanımı, limitin hızla nakde çevrilmesi. Aynı model gerektirir - 10. bölüm	Ekstre ulaşmamış, adres yanlış, otomatik ödeme talimatı kurulmamış, çok küçük bakiye unutulmuş. Gürültü - dışlanmalı

Üçünü ayırmadan kurulan model, üç farklı olguyu tek skorla açıklamaya çalışır.

Temsili dağılım: portföye göre değişmekle birlikte FPD vakalarının azımsanmayacak kısmı ilk kutunun dışındadır.

Şekil 1. FPD homojen bir hedef değildir.

Neden genel PD modeliyle çözülüyor?

Genel bir başvuru skorkartı, 12 aylık bir pencerede temerrüde düşme olasılığını tahmin eder ve hedef kitlesi çok daha geniştir. FPD ise **ilk 30-60 günde** gerçekleşen, çok daha nadir ve çok daha keskin bir olaydır. Genel PD skoru FPD'yi bir dereceye kadar yakalar — kötü müşteri her iki tanımda da kötüdür — fakat üst dilimde ayrışması zayıftır: FPD'ye özgü sinyaller (sorgu patlaması, kanal, aktivasyon hızı, ilk harcama deseni) genel modelin hedefinde seyreler.

FPD'nin sessiz avantajı: hızlı geri besleme

12 aylık PD modelinde bir kohortun olgunlaşmasını beklemek bir yılı bulur. FPD'de etiket **60-90 günde** kesinleşir. Bu, modelin çok daha sık yeniden eğitilebilmesi, kanal ve kampanya bozulmalarının aylar değil haftalar içinde görülmesi demektir. FPD modelinin asıl değeri çoğu zaman ayırım gücünden değil, **bu geri besleme hızından** gelir.

02 Hedef tanımı — en kritik karar

FPD modellerinde performans farkının çoğu, algorithmadan değil hedef tanımından gelir. Aynı portföyde iki farklı tanım, iki farklı model ve iki farklı iş kararı üretir.

Karar	Seçenekler	Pratik öneri
Hangi ekstre?	İlk ekstre / ilk iki ekstre / ilk üç ekstre	İlk ekstre = saf FPD. İlk üç ekstre (early default) daha çok gözlem verir, tanımı yumuşatır.
Gecikme eşiği	1+ gün · 30+ gün · 60+ gün	30+ gün en yaygın ve en savunulabilir olanıdır; 1+ gün operasyonel gürültüyü içeri alır.
Ödeme ölçütü	Asgari ödeme / herhangi bir ödeme / tam ödeme	Asgari ödemenin yapılmaması standarttır; "hiç ödeme yok" daha keskin ama daha nadirdir.
Asgari bakiye	Eşik yok / küçük bakiyeler dışlanır	Çok küçük bakiyeleri dışla — unutulmuş 40 TL, kredi riski sinyali değildir.
İyileşme (cure)	Cure sayılır / sayılmaz	FPD'de cure dikkate alınmaz: olay zaten gerçekleşmiştir. Cure'u ayrı bir analiz olarak izle.
Pencere	Son ödeme tarihinden itibaren 30 / 60 gün	Tanımı kohort bazlı kur; takvim ayına göre kesme (3. ve 11. bölüm).

Tanım seçimi bir eşittir, oranı doğrudan belirler

"İlk ekstrede 30+ gün gecikme" tanımıyla FPD oranı tipik olarak **%0,3-1** bandındadır. Tanımı "ilk üç ekstrede 30+ gün" olarak genişletmek oranı iki-üç katına çıkarır, modelin öğrenecek daha çok bad gözlemi olur ve ayırım gücü kağıt üzerinde iyileşir. Bunun bedeli, hedefin saflığını kaybetmesidir: üçüncü ekstrede gecikme artık kısmen davranışsal bir olaydır ve başvuru anındaki bilgiyle açıklanması daha zordur.

Kural: **tanımı iş kararına göre seç, model performansına göre değil.** Skor onay/limit kararında kullanılacaksa saf FPD, erken tahsilat aksiyonunu tetikleyecekse genişletilmiş tanım daha uygundur.

Tanımı dokümente et

Model dokümanında şu beş satır mutlaka bulunmalı: **hangi ekstre, hangi gecikme eşiği, hangi ödeme ölçütü, hangi asgari bakiye eşiği, hangi dışlamalar.** Bağımsız validasyon biriminin ilk sorusu bu olacaktır ve tanımdaki tek satırlık bir muğlaklık, tüm karşılaştırmayı geçersiz kılar.

03 Kartın ilk 90 günü

Hangi bilginin hangi anda var olduğunu bilmeden ne hedef ne de değişken seti doğru kurulabilir. Sızıntının kaynağı neredeyse her zaman bu eksende bir kaymadır.

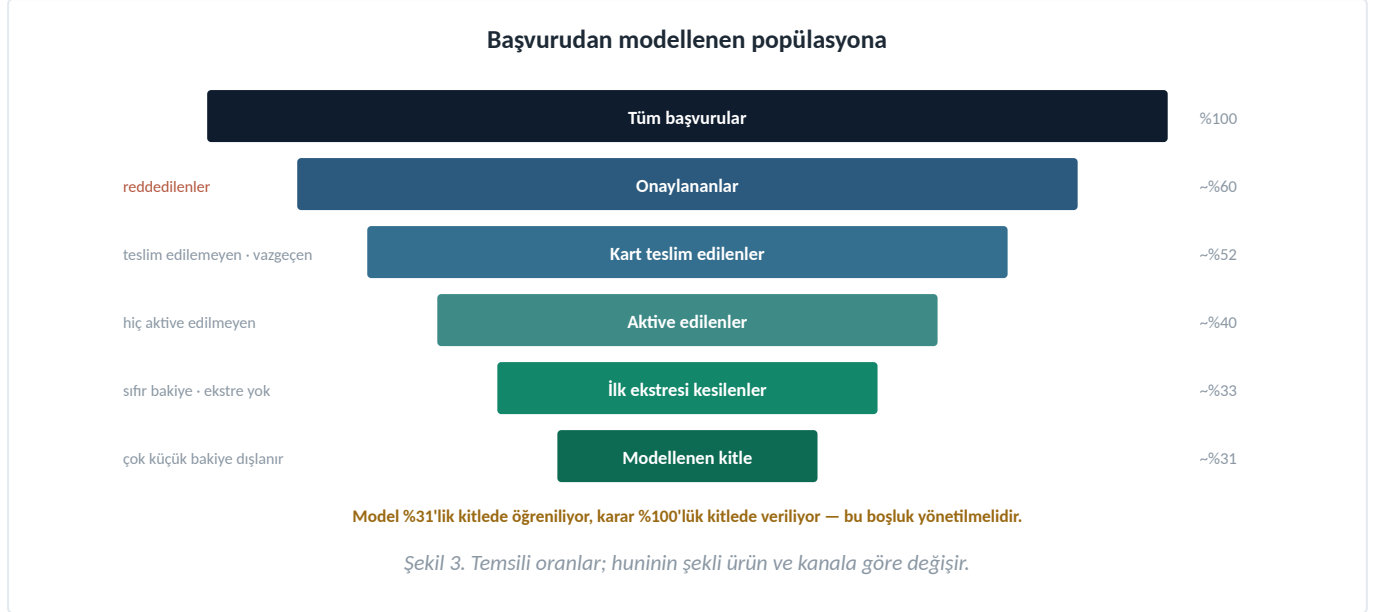


En sık görülen üç sızıntı

- **Tahsis sonrası bilgi başvuru modeline giriyor.** Verilen limit, kart tipi ve kampanya bilgisi karar sonucudur; başvuru skoru bunları kullanamaz. Yüksek limit düşük riskle korele çıkar — çünkü limiti iyi müşteriye banka kendisi vermiştir.
- **İlk ekstre sonrası bilgi kullanılıyor.** Tahsilat aksiyonu, ödeme planı, hesap kapanışı gibi alanlar etiketin sonucudur, nedeni değil.
- **Büro verisi güncel çekilmiş.** Değişkenler başvuru tarihindeki fotoğrafla değil, bugünkü sorguyla üretilirse model gelecekte bilgi alır. Büro alanlarını **başvuru anı kesitinden** üret.

04 Modellenebilir popülasyon

FPD'de en çok tartışılan konu değişken seçimi değil, **kimin modele girdiğidir**. Huninin her adımı örnekleme değiştirir ve son adımda kalan kitle, skorun uygulanacağı kitleyle aynı değildir.



Aktive edilmeyen kart bad değildir, good da değildir

Hiç aktive edilmemiş ya da ilk ekstresi sıfır bakiyeyle kesilmiş bir hesabın ödeme davranışı hakkında hiçbir bilgi yoktur. Bunları **bad=0** olarak modele koymak en sık yapılan hatadır: iyi ödeyicilerin yanına, aslında **gözlenmemiş** bir kitle eklenir. Model bu kitleyi ayırt eden özellikleri (düşük ilgi, pasif müşteri profili) düşük riskle ilişkilendirir. Doğrusu bu hesapları **indeterminate** olarak işaretleyip modelden çıkarmak, oranlarını ayrıca izlemektir — aktivasyon oranının kendisi bir kanal kalitesi göstergesidir.

Reject inference'ı görmezden gelme

Model yalnız onaylananlar üzerinde öğrenir; oysa skor tüm başvuruya uygulanacaktır. Mevcut politika sıkıysa reddedilen kitle model tarafından hiç görülmez ve skorun o bölgedeki davranışı ekstrapolasyondur. Asgari düzeyde yapılabilecek şey: reddedilenlerin büro tabanlı sonraki performansını izlemek, **politika kuralıyla reddedilenleri** skor kararından ayrı tutmak ve kabul oranı değiştiğinde modelin bu bölgede test edilmediğini model dokümanına yazmaktır.

05 FPD ile genel temerrüt farkı

İki model de "temerrüt" tahmin eder; fakat gözlem penceresi, veri kaynağı, olay oranı ve kullanım amacı farklıdır. Aynı hattın iki ayrı aracıdır, birbirinin yerine geçmez.

Boyut	FPD modeli	Genel PD / davranış modeli
Hedef	İlk ekstrede asgari ödemenin yapılmaması	12 ay içinde 90+ gün gecikme
Gözlem penceresi	30-90 gün	12-24 ay
Olay oranı	%0,3-2 — çok nadir	%2-8 — görece bol
Veri kaynağı	Büro, beyan, kanal, ürün; davranış yok	Banka içi davranış verisi ağırlıklı
Etiket olgunlaşması	2-3 ay — çok hızlı	12-18 ay — yavaş
Yeniden eğitim sıklığı	6 ayda bir mümkün	Yılda bir veya daha seyrek
Fraud bileşeni	Belirgin — ayrıştırılmalı	Sınırlı
Tipik kullanım	Onay eşiği, ilk limit, aktivasyon kontrolü	Limit yönetimi, karşılık, fiyatlama
Tipik ayırım gücü	Gini 0,45-0,60 (nadir olay, keskin sinyal)	Gini 0,40-0,55

İkisi rakip değil, tamamlayıcıdır

Doğru kurgu genellikle şudur: **genel başvuru skoru** ana kabul kararını verir, **FPD skoru** bu kararın üzerine ikinci bir filtre olarak biner. Genel skorda kabul bölgesinde olup FPD skorunda en riskli dilime düşen başvurular, reddedilmek yerine **düşük ilk limit, nakit avans kısıtı** veya **ek doğrulama** ile yönetilir. Bu, iki modelin de kendi güçlü olduğu yerde kullanılmasını sağlar.

FPD skorunu genel PD skorunun yerine koyma

FPD skoru yalnız ilk ekstreyi tahmin eder; portföyün ikinci yılındaki davranışı hakkında sistematik bir şey söylemez. Karşılık hesabında, beklenen kayıp modellerinde veya IFRS 9 aşama geçişlerinde FPD skorunu doğrudan PD olarak kullanmak yanlıştır. FPD bir **kabul ve erken müdahale aracıdır**.

06 Nadir olay ve metrik seçimi

%1'lik bir bad oranında, hiçbir şey öğrenmeyen bir model %99 doğruluk verir. FPD'de metrik seçimi model seçiminden daha belirleyicidir.

Aynı model, dört farklı metrikle dört farklı hikaye



Metrik	Değer	Ne söylüyor?
Accuracy	%98,8	hiçbir şey
Gini	0,52	genel sıralama
PR-AUC	0,09	nadir sınıfta gerçek
Üst %1 yakalama	×11	operasyonel değer
Brier	0,011	kalibrasyon

Aynı model. Accuracy'ye bakan "mükemmel", PR-AUC'ye bakan "zayıf" der.

Temsili değerler; Gini yüksekken PR-AUC'nin düşük görünmesi nadir olayda normaldir.

Şekil 4. Nadir olayda tek metrikli değerlendirme yanıltır.

Hangi metrikler birlikte raporlanmalı?

- **Gini / AUC** — genel sıralama gücü. Zorunlu ama tek başına yetersiz.
- **PR-AUC** — nadir sınıfta gerçek performansı gösterir; taban çizgisi bad oranının kendisidir (%1,2 bad için rastgele model 0,012 verir).
- **Üst dilim yakalama (lift)** — en riskli %1 ve %5'te ortalamanın kaç katı bad yakalandığı. Karar eşiğinin çalışacağı bölge burasıdır, toplam Gini bu bölgedeki farkı gizleyebilir.
- **KS** — sektörde yaygın; dağılımların en çok ayrıştığı noktayı verir.
- **Brier / kalibrasyon eğrisi** — skor bir olasılık olarak kullanılacaksa zorunlu.

Sınıf ağırlığı: çözüm değil, araç

`scale_pos_weight = neg/pos` vermek nadir sınıfta öğrenmeyi kolaylaştırır ama **çıktı artık olasılık değildir** — model riski sistematik olarak abartır. FPD'de bunun iki sonucu vardır: sıralama bozulmaz, dolayısıyla Gini ve lift güvenilir kalır; buna karşılık Brier ve kalibrasyon eğrisi anlamsızlaşır. Skoru yalnız eşik/dilim kararında kullanacaksan ağırlık verebilirsin. Olasılık gerekiyorsa ayrı bir dilimde kalibrasyon katmanı öğren.

07 Başvuru anında elde ne var?

FPD modelinin tüm sinyali beş kaynaktan gelir. Değişken üretme çabasını bu beş başlığa bölmek, dağınık bir "her şeyi at" yaklaşımından çok daha verimlidir.

Kaynak	Tipik değişkenler	FPD'deki gücü
Büro (KKB)	Skor, son 1/3/6 ay sorgu sayısı, yeni açılan hesap sayısı, mevcut gecikme, toplam limit ve kullanım oranı, kapalı hesap geçmişi	En güçlü blok. Özellikle sorgu ve yeni hesap yoğunluğu FPD'ye genel PD'den daha duyarlıdır.
Başvuru beyanı	Gelir, çalışma süresi, meslek, eğitim, medeni durum, ikamet süresi	Orta. Asıl değer beyanın kendisinde değil, tutarlılığında : beyan-büro uyumsuzluğu.
Kanal ve satış	Şube / dijital / aracı / mağaza, aracı kodu, kampanya, satış temsilcisi	Yüksek ve dalgalı. Kanal bazlı FPD oranı zaman içinde en hızlı bozulan boyuttur.
Ürün ve teklif	Talep edilen limit, kart tipi, aidat yapısı, taksit kampanyası	Orta. Verilen limit değil, talep edilen limit kullanılabilir (endojenite — 16. bölüm).
Banka içi (varsa)	Müşteri kıdemi, maaş müşterisi mi, mevduat/yatırım bakiyesi, diğer ürün davranışı	Çok yüksek — ama yalnız mevcut müşteride vardır; yeni müşteride tamamen boştur.

İki popülasyon, tek model olmayabilir

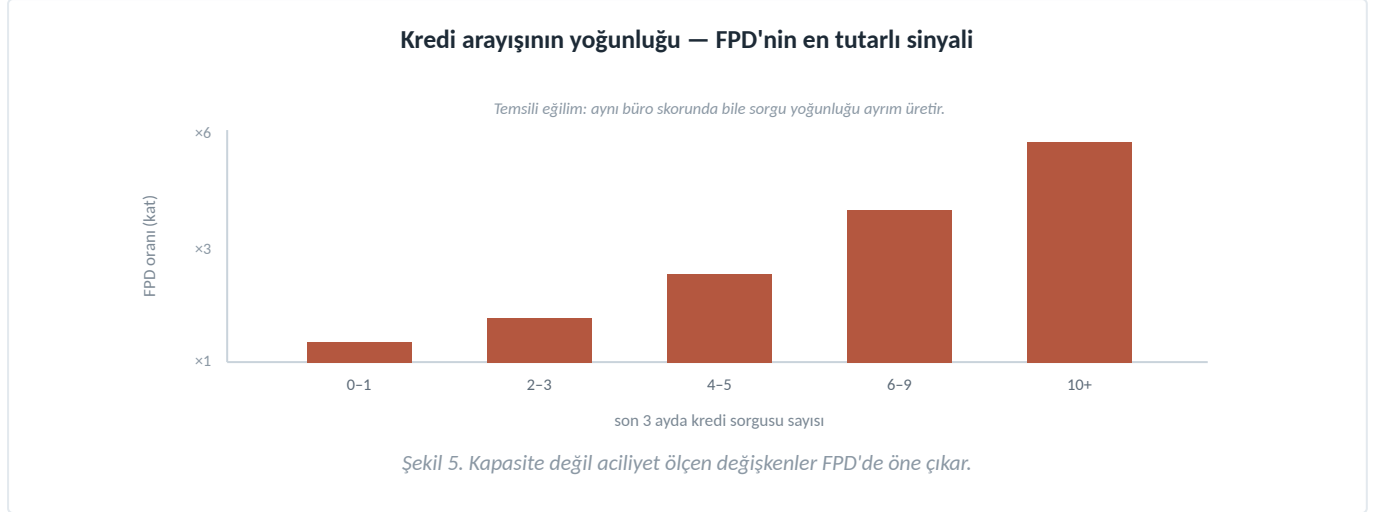
Mevcut banka müşterisi ile hiç ilişkisi olmayan yeni müşteri, aynı modelde çok farklı bilgi setiyle temsil edilir: birinde onlarca içsel davranış değişkeni doludur, diğerinde hepsi eksiktir. Ağaç modelleri bunu bir dereceye kadar tolere eder, fakat pratikte **iki ayrı model** (veya en azından popülasyonu ayıran açık bir bayrak) genellikle daha temiz sonuç verir ve model dokümanında savunulması çok daha kolaydır.

Oran değişkenleri ham tutarlardan güçlüdür

Talep edilen limit / beyan edilen gelir, büro toplam borç / gelir, sorgu sayısı / hesap yaşı gibi **oranlar** FPD'de ham tutarlardan istikrarlı biçimde daha güçlüdür ve enflasyonist bir ortamda çok daha stabildir: nominal gelir her yıl kayarken oranın dağılımı yerinde kalır. Model dokümanında da savunması kolaydır — "gelirinin 4 katı limit talep eden müşteri" iş tarafının anladığı bir cümledir.

08 FPD'ye özgü güçlü sinyaller

Genel PD modelinde ikinci planda kalan bazı değişkenler, FPD'de ilk sıraya çıkar. Nedeni ortaktır: hepsi **aciliyet** ve **niyet** ölçer, kapasiteyi değil.



- **Sorgu ve yeni hesap yoğunluğu.** Kısa sürede çok sayıda kredi başvurusu, nakit sıkışıklığının en erken ve en temiz göstergesidir. Büro skoru bunu kısmen içerir, ama FPD için ham sorgu sayıları ayrıca değerlidir.
- **Limit talebi / gelir oranı.** Gelirin katları düzeyinde limit talebi hem kapasite hem niyet sinyalidir.
- **Kanal ve aracı kimliği.** Belirli bir aracının, mağazanın veya kampanyanın FPD oranı portföy ortalamasının kat kat üzerinde olabilir; bu genellikle satış baskısının veya beyan kalitesinin göstergesidir.
- **Beyan-büro tutarsızlığı.** Beyan edilen gelire büro borç yükünün uyumsuzluğu, tek başına bir değişken olarak üretilmeye değer.
- **Aktivasyon hızı ve ilk kullanım deseni.** Kartın aktive edilmesiyle limitin büyük kısmının kullanılması arasındaki sürenin çok kısa olması, nakit avans ağırlıklı kullanım ve tek seferde yüksek tutarlı harcama klasik erken risk deseni. Bu grup yalnız **erken yaşam modelinde** kullanılabilir — 9. bölüm.

Delta değişkenleri: FPD'de ne işe yarar, ne işe yaramaz

Harcama artışı, son 3 ayda toplam risk değişimi gibi **değişim/delta** değişkenleri kredi kartı modellemesinin en güçlü aletlerindendir — fakat FPD'nin başvuru modelinde **yeni müşteride tanımsızdır**. Kullanılabilecekleri iki yer vardır: (a) mevcut banka müşterisinde içsel davranış deltaları, (b) büro tarafında toplam risk ve limit kullanımının son aylardaki değişimi. İkincisi çoğu portföyde yeterince kullanılmayan bir kaynaktır ve saf başvuru modelinde bile mevcuttur.

09 İki model, iki an

Aynı hedefi tahmin eden iki ayrı model kurmak fazlalık değil, doğru kurgudur: karar anları farklıdır, elde olan bilgi farklıdır, tetikledikleri aksiyon farklıdır.

	Model 1 • Başvuru FPD	Model 2 • Erken yaşam
Skorlama anı	Başvuru — karar öncesi	Kart açıldıktan 15–30 gün sonra
Girdi	Büro, beyan, kanal, ürün talebi	Aynı değişkenler + aktivasyon, ilk harcama, nakit avans, MCC deseni
Hedef	İlk ekstre FPD	İlk ekstre FPD (aynı hedef, daha çok bilgi)
Tipik ayırım gücü	Gini 0,45–0,55	Gini 0,60+ — ilk kullanım deseni çok güçlüdür
Aksiyon	Ret, referral, düşük ilk limit, ek doğrulama	Limit dondurma, nakit avans kapatma, erken iletişim
Kısıt	Kararı verdiği için politika ve adalet incelemesi ağır	Müşteri ilişkisi zedelenmemeli; aksiyon ölçülü olmalı

İkinci modelin cazibesine dikkat

Erken yaşam modeli çok daha yüksek Gini verir ve bu, birinci modelin "kötü" olduğu izlenimi yaratır. Oysa ikisi aynı yarışta değildir: ikinci model, riskin bir kısmı **zaten gerçekleşmişken** tahmin yapar — limit ilk hafta nakde çevrilmişse FPD'yi tahmin etmek kolaydır. Karşılaştırma yalnız **aynı bilgi setiyle** anlamlıdır. İkinci modelin değeri ayırım gücünde değil, **kayıbı önleyecek zamanda** aksiyon üretmesindedir.

Aksiyonun geri besleme etkisi

Erken yaşam skoru limit dondurma tetikliyorsa, o hesapların bir kısmı artık FPD'ye düşmeyecektir — model kendi etiketini değiştirir. Bir sonraki eğitimde bu hesaplar "good" görünür ve model kendi başarısını öğrenemez. Çözüm: aksiyon alınan hesapları **bayraklayıp** eğitimde ayrı ele almak ve mümkünse küçük bir kontrol grubunu aksiyonsuz bırakmak. Bu, model riski tarafında da beklenen bir kontroldür.

10 Fraud mu, kredi riski mi?

FPD vakalarının bir bölümü ödeyememe değil, **ödememe niyetidir**. Bu iki olgu farklı değişkenlerle açıklanır ve farklı aksiyon gerektirir; tek etikette birleştirmek her ikisini de zayıflatır.



Neden ayırmak gerekir?

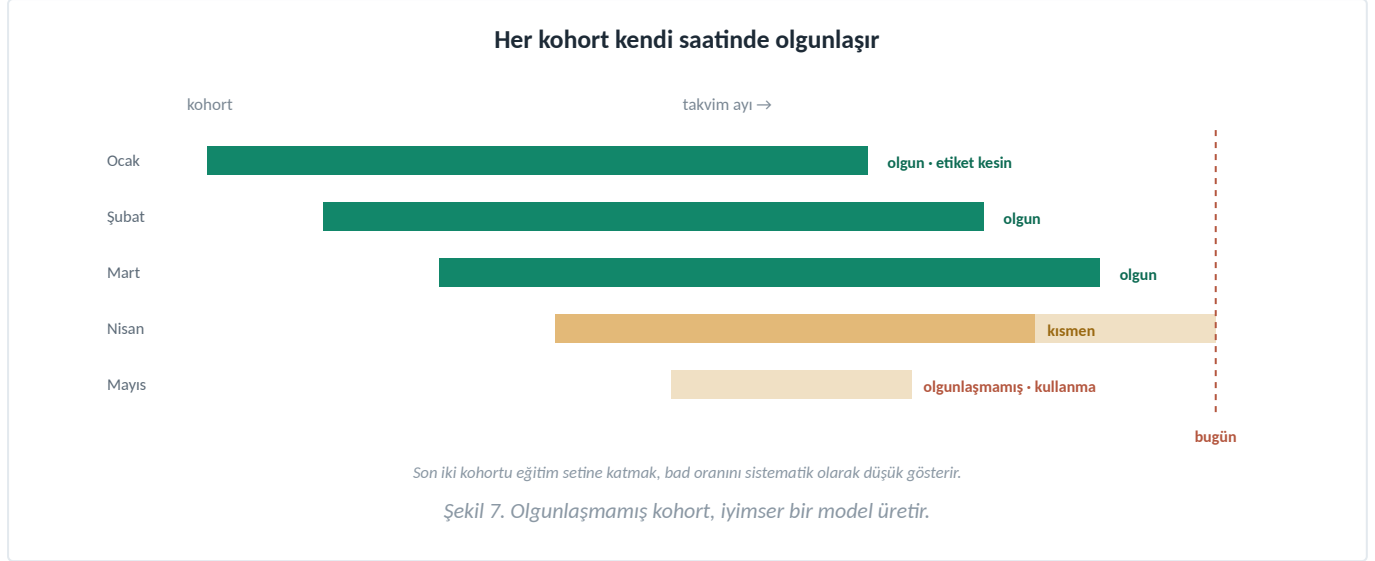
Fraud vakaları, kredi riski değişkenleriyle **ters** desenler taşır: sahte beyan genellikle temiz bir profil gösterir, büro skoru iyi olabilir, borç yükü düşük görünür. Bu vakalar kredi modeline karıştığında model, "temiz profil" ile FPD arasında zayıf ve yanıltıcı bir ilişki öğrenmeye zorlanır. Sonuç, iki tarafta da bulanıklaşan bir skordur.

Pratik ayrıştırma kuralı

Etiket kütüğüne üç bayrak ekle: **hiç ödeme yapılmadı, iletişim kurulamadı, limitin %80'inden fazlası ilk 15 günde nakde çevrildi**. Üçü birden sağlanan vakalar fraud yönlü olarak işaretlenir. Bunları kredi modelinden çıkarmak yerine **ayrı bir hedef değişken** olarak tutmak en iyisidir: aynı veriden hem FPD hem başvuru fraud modeli eğitilebilir, ikisi üretimde ayrı eşiklerle çalışır.

11 Kohort kurgusu ve olgunlaşma

FPD'de örneklem takvim ayına göre değil, **hesap açılış kohortuna** göre kurulur. Bu, konunun en sık atlanan ve en pahalıya mal olan teknik detaydır.



Neden kohort, neden takvim değil?

Bir hesabın FPD'ye düşüp düşmediği, açılış tarihinden 60–90 gün sonra belli olur. Örneklemi "2026 ilk yarısı" gibi takvim aralığıyla kesersen, son iki ayın hesapları etiketlenmeden içeri girer. Bunlar veriye $bad=0$ olarak yazılırsa **bad oranı sistematik olarak düşer** ve model, dönemin sonuna doğru riskin azaldığını öğrenir. Doğrusu: kohortu açılış ayına göre kur, yalnız **tam olgunlaşmış** kohortları eğitime al, kısmen olgunlaşmışları izleme raporunda ayrı göster.

Zaman bazlı bölme, rastgele bölme değil

Eğitim ve OOT ayrımı da kohort eksenindedir: eski kohortlar eğitim, sonraki kohortlar OOT. Rastgele bölme, aynı kampanyanın ve aynı kanal dönemlerinin her iki tarafa dağılmasına yol açar; model kampanya kimliğini ezberler, OOT performansı iyimser çıkar. FPD'de kampanya ve kanal etkisi çok güçlü olduğu için bu hata burada genel PD modellerinden daha ağır sonuç verir.

Mevsimsellik ve kampanya etkisi

Kart satışının yoğunlaştığı dönemler (yılbaşı, okul dönemi, büyük kampanyalar) hem müşteri karışımını hem de FPD oranını kaydırır: agresif satış dönemlerinde kabul edilen kitle sistematik olarak daha risklidir. En az **bir tam yılı** kapsayan kohort seti kullan ve OOT penceresini mümkünse farklı bir kampanya rejiminden seç — modelin yalnız bir satış dönemine göre ayarlanmadığını böyle görürsün.

12 Modelleme pratiği

Algoritma seçimi bu problemin en kolay kısmıdır. Nadir sınıf ve az sayıda bad gözlem, ayarın tamamını **muhafazakar** tarafa çeker.

Konu	FPD'de öneri	Neden
Kapasite	<code>max_depth = 3-4</code>	Bad sayısı az; derin ağaç birkaç gözlemlik desen ezberler.
Yaprak alt sınırı	<code>min_child_weight</code> yüksek tut	Hessian toplamı üzerinden çalışır; nadir sınıfta etkili bir fren.
Öğrenme oranı	0,02-0,03 + çok tur	Yavaş öğrenme, nadir sınıfta varyansı düşürür.
Düzenleştirme	<code>reg_lambda</code> ve <code>gamma</code> cömert	Gürültülü değişkenlerin ağaca sızmasını engeller.
Sınıf dengesizliği	<code>scale_pos_weight</code> veya ağırlıksız + eşik	Sıralama ikisinde de benzer; ağırlık kalibrasyonu bozar.
Monotonluk	Büro skoru, gecikme, sorgu sayısı için kısıt koy	Denetimde savunulabilirlik ve OOT stabilitesi artar.
Erken durdurma	Ayrı validation, metrik olarak AUC veya PR-AUC	OOT asla erken durdurmada kullanılmaz.

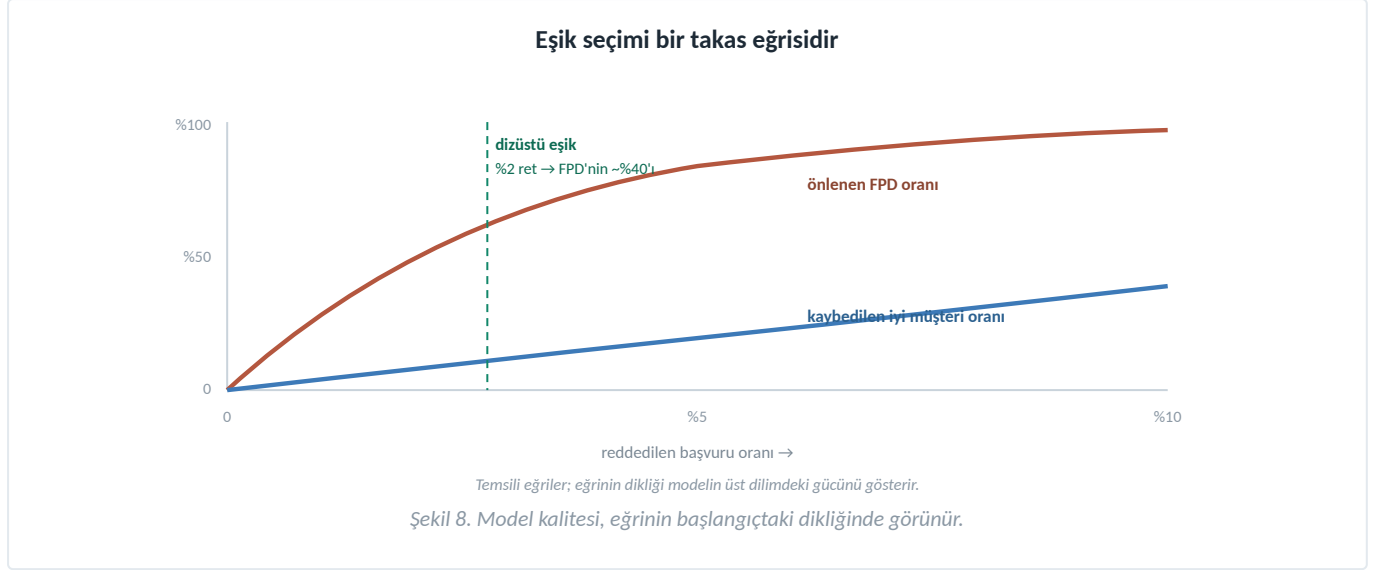
```
// FPD başlangıç parametre bloğu
{
  "objective": "binary:logistic",
  "eval_metric": ["auc", "aucpr"],
  "tree_method": "hist",
  "max_depth": 3,
  "min_child_weight": 30,
  "learning_rate": 0.02,
  "gamma": 0.2,
  "reg_lambda": 10.0,
  "subsample": 0.8,
  "colsample_bytree": 0.7,
  "scale_pos_weight": 80,           // neg/pos ~ %1,2 bad oranı için
  "n_estimators": 5000,           // üst sınır – erken durdurma belirler
  "early_stopping_rounds": 150,
  "monotone_constraints": "(-1,1,1,0,0)", // büro skoru ↓risk, gecikme/sorgu ↑risk
  "random_state": 42
}
```

Bad sayısı, satır sayısından önemlidir

500 bin satırlık bir veri kulağa bol gelir; %0,5 bad oranında bu **2.500 bad gözlem** demektir ve model fiilen bu sayıyla öğrenir. Pratik eşik: 1.000'in altında bad gözlemle kurulan bir FPD modeli kırılgandır — tanımları genişletmek (ilk üç ekstre), daha çok kohort eklemek veya modeli daha da sığ tutmak gerekir. Değişken sayısını da bu sayıya göre sınırla: 2.500 bad gözlem için 300 değişkenlik bir havuz fazladır.

13 Değerlendirme: karar eşiği

FPD skorunun değeri, ortalama performansta değil **en riskli dilimde** ortaya çıkar. Doğru soru "Gini kaç?" değil, "bu eşikte kaç FPD önleniyor, karşılığında kaç iyi müşteri kaybediliyor?"dur.



Swap-set analizi

Yeni skoru mevcut politikayla karşılaştırırken toplam Gini yerine **swap-set** tablosu kur: yeni modelin reddedip eskisinin kabul ettiği başvurular (swap-out) ile tersi (swap-in). Kararı bu iki kümenin gerçekleşmiş FPD oranları verir. İyi bir model, swap-out kümesinde portföy ortalamasının çok üzerinde, swap-in kümesinde çok altında bir FPD oranı gösterir. Bu tablo, iş tarafına Gini'den çok daha anlamlı gelir ve model dokümanının en ikna edici sayfasıdır.

Raporlama şablonu

Her model versiyonu için tek satır: **medyan Gini ± std · PR-AUC · KS · üst %1 ve %5 yakalama · Brier · %2 ret eşiğinde önlenen FPD · swap-out FPD oranı · PSI**. Kararı bu tabloya bakarak ver; tek bir Gini sayısına bakarak değil.

14 Skorun kullanımı

FPD skorunun en verimli kullanımı ret değildir. Ret, hem gelirden hem müşteriden vazgeçmek demektir; oysa riskin çoğu **koşullu kabul** ile yönetilebilir.

Skor bandı	Beklenen FPD	Önerilen aksiyon
En riskli %1	Ortalamanın ~10 katı	Ret veya zorunlu ek doğrulama; fraud hattına yönlendirme.
%1-5	~4-6 kat	Düşük ilk limit, nakit avans kapalı, gelir belgesi talebi.
%5-20	~2 kat	Kademeli limit: ilk 3 ay düşük, ödeme performansına göre otomatik artış.
%20-80	Ortalama civarı	Standart akış.
En iyi %20	Ortalamanın altında	Yüksek ilk limit, hızlı onay, kampanya hedeflemesi.

Kademeli limit: FPD'ye en uygun araç

FPD'nin doğrudan kaybı, kullanılan limitle sınırlıdır. Riskli bandı reddetmek yerine ilk 3 ay boyunca düşük limitle çalıştırmak, kaybı büyüklük olarak kırpar ve ödeme davranışını gözleme fırsatı verir. Üç düzgün ödeme sonrası limit otomatik yükseltilir. Bu yapı, ret ile tam kabul arasındaki geniş orta bölgeyi kullanılabilir hale getirir ve FPD skorundan alınan iş değerini belirgin biçimde artırır.

Nakit avans kısıtı ve aktivasyon kontrolü

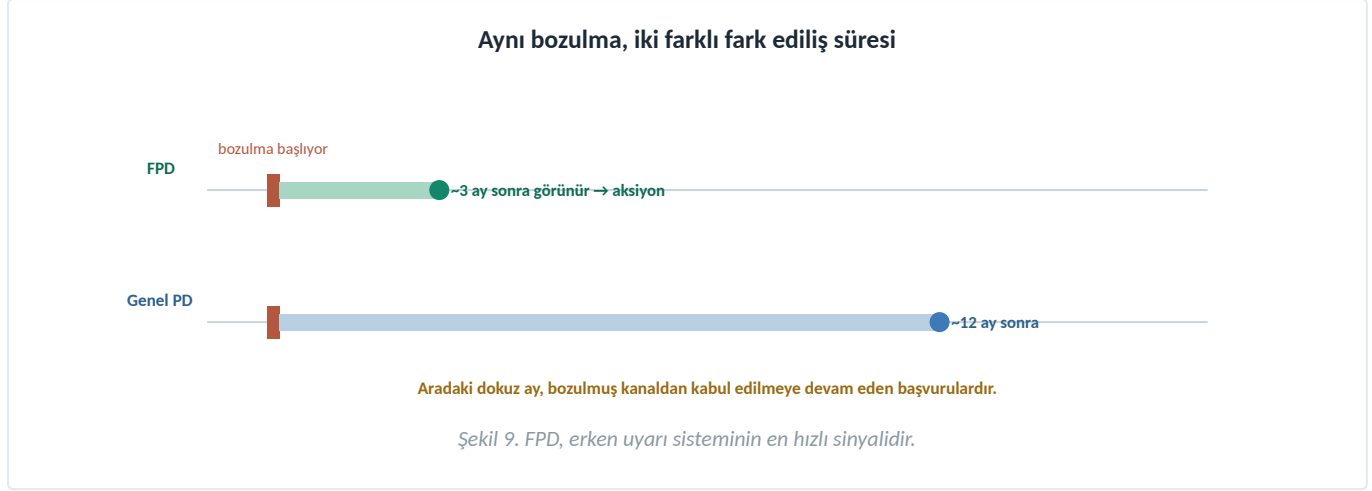
FPD ve başvuru sahtekarlığı vakalarının ortak deseni, limitin hızla nakde çevrilmesidir. Riskli bantta ilk 30-60 gün nakit avansı kapatmak, kredi kararını değiştirmeden kaybın en hızlı gerçekleştiği kanalı kapatır — düşük maliyetli ve müşteri deneyimini en az zedeleyen müdahaledir.

Skor eşiği bir politika kararıdır

Eşiği model ekibi tek başına belirlemez. Kabul oranı, hedeflenen portföy büyümesi, kampanya taahhütleri ve risk iştahı eşiği doğrudan etkiler. Model ekibinin görevi **eğriyi** vermektir: her eşikte kaç FPD önlenir, kaç iyi müşteri kaybedilir, hangi kanal ne kadar etkilenir. Kararı bu eğriye bakarak iş ve risk birlikte verir.

15 İzleme ve erken geri besleme

FPD modelinin en büyük üstünlüğü izlemede ortaya çıkar: etiket 2-3 ayda kesinleştiği için bozulma, aylar değil haftalar içinde görülebilir.



Ne izlenmeli?

- **Kohort bazlı FPD oranı** — açılış ayına göre, olgunlaşma durumu belirtilerek.
- **Kanal ve aracı kırılımı** — FPD'de en hızlı bozulan boyut budur; aracı bazında oranı ayrı izle.
- **PSI (skor dağılımı) ve CSI (değişken dağılımı)** — kabul politikası veya kampanya değiştiğinde ilk bunlar kayar.
- **Aktivasyon ve ilk kullanım oranları** — modelin dışında ama popülasyon tanımını doğrudan etkiler.
- **Eksik oranları** — bir büro alanı veya kaynak sistem devre dışı kaldığında skor sessizce kayar.
- **Swap-set gerçekleştirmeleri** — modelin reddettiği kitle gerçekten kötü çıkıyor mu?

Makro etkiyi modelden değil, izlemeden yönet

Faiz, işsizlik veya enflasyondaki değişim FPD oranını topluca yukarı çeker; fakat bunlar bir başvurunun diğerinden daha riskli olduğunu söylemez — **seviyeyi** kaydırır, **sıralamayı** değil. Makro değişkenleri başvuru modeline değişken olarak koymak, eğitim döneminin makro rejimini modele sabitler ve rejim değiştiğinde skor yanlış yönde kayar. Doğru yer: **kalibrasyon katmanı ve izleme** — beklenen FPD oranını makro senaryoya göre güncelle, sıralamayı modele bırak.

Yeniden eğitim ritmi

Etiket hızlı olgunlaştığı için FPD modeli 6 ayda bir yeniden eğitilebilir. Fakat her sapmada modeli yenilemek doğru refleks değildir: bozulmanın kaynağı çoğu zaman modelde değil **kanal karışımındadır**. Önce kırılımlara bak; oran belirli bir aracı veya kampanyada patlamışsa çözüm yeni model değil, politika müdahalesidir.

16 Yaygın tuzaklar

Modeli ya da karşılaştırmayı sessizce bozan hatalar.

Tuzak	Ne olur?	Doğrusu
Aktive edilmeyen kartları bad=0 saymak	Gözlenmemiş kitle iyi ödeyicilerin arasına karışır; pasiflik düşük riskle ilişkilendirilir.	Indeterminate olarak modelden çıkar, oranını ayrıca izle.
Verilen limiti değişken olarak kullanmak	Yüksek limit düşük riskle korele çıkar — çünkü kararın sonucudur.	Talep edilen limiti kullan; verilen limit yalnız erken yaşam modelinde.
Olgunlaşmamış kohortu eğitime katmak	Bad oranı düşük görünür; model dönemin sonunda risk azalıyor sanır.	Yalnız tam olgunlaşmış kohortlar; diğerlerini raporda ayrı göster.
Rastgele train/test bölmesi	Kampanya ve kanal dönemleri iki tarafa dağılır; OOT iyimser çıkar.	Zaman/kohort bazlı bölme.
Fraud vakalarını ayırmamak	Model "temiz profil" ile FPD arasında yanıltıcı ilişki öğrenir.	Ayrı bayrak, mümkünse ayrı model.
Büro verisini bugünkü kesitten çekmek	Gelecekte bilgi sızar; validasyon mükemmel görünür.	Başvuru anı kesiti (as-of date) zorunlu.
Yalnız Gini raporlamak	Üst dilim farkı ve kalibrasyon görünmez.	PR-AUC, KS, üst dilim yakalama ve Brier birlikte.
Çok küçük bakiyeleri bad saymak	Unutulmuş 40 TL, kredi riski sinyali gibi öğrenilir.	Asgari bakiye eşiği koy ve dokümente et.
Makro değişkeni modele koymak	Eğitim döneminin rejimi modele sabitlenir.	Seviye kaymasını kalibrasyon ve izleme ile yönet.
Erken yaşam modelini başvuru modeliyle kıyaslamak	Farklı bilgi setleri; kıyas anlamsız.	Yalnız aynı bilgi setiyle karşılaştır.
Aksiyon alınan hesapları bayraklamamak	Model kendi başarısını "good" olarak öğrenir.	Aksiyon bayrağı tut; küçük bir kontrol grubu bırak.
Reject inference'ı yok saymak	Skorun reddedilen bölgedeki davranışı test edilmemiştir.	Sınırı dokümente et, büro takibiyle izle.

17 Kontrol listesi

Bir FPD modeli kurmadan önce ve teslim etmeden önce cevaplanması gereken sorular.

Kurmadan önce

- ☐ Hedef tanımı beş bileşeniyle yazıldı mı? (ekstre · gecikme eşiği · ödeme ölçütü · asgari bakiye · dışlamalar)
- ☐ Modellenabilir popülasyon hunisi çıkarıldı mı, her adımdaki kayıp oranı biliniyor mu?
- ☐ Aktive edilmeyen ve sıfır bakiyeli hesaplar indeterminate olarak ayrıldı mı?
- ☐ Fraud yönlü vakalar bayraklandı mı?
- ☐ Kohortlar açılış ayına göre kuruldu ve olgunlaşma kontrol edildi mi?
- ☐ Toplam bad gözlem sayısı model karmaşıklığını taşıyor mu?
- ☐ Tüm değişkenler başvuru anı kesitinden mi üretildi?

Teslim etmeden önce

- ☐ OOT en az iki farklı kohort penceresinde ölçüldü mü?
- ☐ En az 5 seed ile medyan $\pm$ std raporlandı mı?
- ☐ PR-AUC, KS ve üst dilim yakalama raporda var mı?
- ☐ Skor olasılık olarak kullanılacaksa kalibrasyon katmanı kuruldu ve OOT'de doğrulandı mı?
- ☐ Swap-set analizi mevcut politikaya karşı yapıldı mı?
- ☐ Kanal, ürün ve segment kırılımlarında performans ayrı ayrı bakıldı mı?
- ☐ Monotonluk kısıtları ve iş mantığına aykırı yön taşıyan değişkenler kontrol edildi mi?
- ☐ Sürüm, seed, veri kesiti ve parametreler model dokümanına yazıldı mı?

18 Kaynaklar ve kapsam notu

FPD, akademik literatürden çok sektör pratiğiyle şekillenmiş bir konudur; aşağıdakiler çerçeveyi kuran temel başvuru kaynaklarıdır.

- 1 **Siddiqi — Intelligent Credit Scoring / Credit Risk Scorecards**
Hedef tanımı, gözlem-performans penceresi ve reject inference bölümleri
- 2 **Thomas, Edelman & Crook — Credit Scoring and Its Applications**
Vintage analizi, kohort kurgusu ve kabul eşiği ekonomisi
- 3 **Baesens, Roesch & Scheule — Credit Risk Analytics**
Dengesiz sınıf, model doğrulama ve performans ölçütleri
- 4 **XGBoost Documentation — Parameters, Imbalanced Data, Monotonic Constraints**
xgboost.readthedocs.io/en/stable/
- 5 **Saito & Rehmsmeier — The Precision-Recall Plot Is More Informative than the ROC Plot (PLOS ONE, 2015)**
Nadir olayda metrik seçiminin gerekçesi
- 6 **BCDataWorks Çalışma Notları N° 04 — Random Forest • XGBoost**
Model seçimi, ayar felsefesi ve adil karşılaştırma protokolü

Kapsam notu

Bu doküman **bireysel kredi kartı** ürününe ve başvuru hattına odaklanır. Taksitli krediler, kredili mevduat hesabı ve ticari kartlarda FPD mantığı benzerdir; fakat "ilk ekstre" yerine "ilk taksit" tanımı, farklı bir aktivasyon kurgusu ve farklı bir fraud profili geçerlidir. Oranlar ve tipik performans aralıkları portföye, kabul politikasına ve satış kanalı karışımına göre belirgin biçimde değişir.

Son söz

FPD modellemesinde başarıyı belirleyen sıralama şudur: **hedef tanımı, modellenenebilir popülasyonun doğru kurulması, fraud ayrıştırması, kohort disiplini** — ve ancak bunlardan sonra algoritma. İlk dördü doğruysa basit bir model bile işe yarar; yalnızca hiçbir algoritma toparlamaz.