Bu metin bir bankanın "Bireysel- Kredi Kartı(KK), Esnek Hesap(EH) ve Ferdi Kredi(FK) Başvuru Modellerinin Güncellenmesi –Dördüncü Nesil Bireysel Krediler Başvuru Modelleri" projesi ile ilgili detayları içermektedir. Proje, bireysel kredi portföyünde kullanılan başvuru skor kartlarının güncellenmesi amacıyla gerçekleştirilmiştir.

Projede üç ana ürün tipi için ayrı modeller geliştirilmiş: Ferdi Kredi, Kredi Kartı ve Esnek Hesap. Modeller, erişilebilir en güncel temerrüt oranına kalibre edilmiş ve Risk Validasyon Birimi tarafından valide edilmiştir. Proje, ayrıştırma gücü yüksek modellerin geliştirilmesi, risk maliyetlerinin doğru bir şekilde hesaplanması, gerçekleşen temerrüt oranı ile modelin tahmin ettiği oran arasındaki farkın azaltılması, ve müşterinin 12 ay içinde "temerrüde" düşme olasılığının tahmin edilmesi gibi amaçlar doğrultusunda yürütülmüştür.

Proje, 2022 Ocak ayında tamamlanmış olup, süreç veri toplama ve hazırlama, ön-modelleme, modelleme ve kalibrasyon fazlarını içermiştir. Modellerin kalibrasyonu, her bir kredi tipi için ayrı olarak yapılmış, ürün bazında modellerin güncel temerrüt oranları belirlenmiştir.

Veri sözlükleri, geniş bir veri yelpazesini içererek başvuru esnasında temin edilen bilgilerden başvuru sahibinin mevcut müşteri olması durumuna kadar çeşitli veri kaynaklarından türetilen değişkenleri içermektedir.

Projenin amacı ayrıştırma gücü yüksek modeller geliştirerek risk maliyetlerini gerçekçi bir şekilde tahminlemek ve müşteri temerrüt olasılıklarını belirlemektir. Proje, başvuru skor kartlarının otomatize edilmiş bir kredi tahsis sistemi oluşturarak risk yönetimi ve tahsis süreçlerini etkinleştirmeyi amaçlamaktadır.

Ayrıca, metinde projenin tarihçesi, kullanılan modellerin evrimi, ve güncellenen dördüncü nesil başvuru modellerinin detaylarına da yer verilmiştir.

----------------------------------------------------PD------------------------------------------------------------------------

Bankacılıkta "PD," "Probability of Default" ifadesinin kısaltmasıdır. Türkçe'de "Temerrüt Olasılığı" olarak da adlandırılabilir. PD, genellikle bir borçlu veya kredi alıcısının belirli bir dönem içinde iflas etme olasılığını ölçen bir finansal terimdir.

Bankalar, kredi riskini değerlendirmek ve yönetmek amacıyla PD'yi kullanırlar. PD, kredi başvurularını değerlendirirken, kredi verme kararlarını alırken ve kredi portföylerini izlerken önemli bir gösterge olarak kullanılır. Daha yüksek bir PD, borçlunun iflas olasılığının daha yüksek olduğunu gösterir ve bu nedenle kredi riski daha fazladır.

PD'nin belirlenmesi genellikle istatistiksel modeller, kredi derecelendirme ajanslarının değerlendirmeleri veya benzer analitik yöntemler kullanılarak yapılır. Bankalar, kredi riskini etkili bir şekilde yönetmek ve iflas olasılıklarını minimize etmek için PD'yi diğer risk ölçütleriyle birleştirerek kapsamlı bir risk değerlendirmesi yaparlar.

Temerrüte düşmek ve NPL:

**Temerrüte Düşmek (Default):** Temerrüt, bir borçlu veya kredi alıcısının kredi şartlarına uygun olarak ödemelerini gerçekleştirememesi durumunu ifade eder. Temerrüt durumu, genellikle kredi taksitlerinin vadesi geçtikten sonra belirli bir süre içinde yapılmamış olmasıyla ortaya çıkar. Temerrüte düşmek, borçlunun ödemelerini yapma yükümlülüğünü yerine getirememesi ve kredi anlaşmasındaki şartlara uymaması anlamına gelir.

**NPL (Non-Performing Loan):** NPL, "performans göstermeyen kredi" anlamına gelir. Bu terim, kredinin geri ödenmesinde sorun yaşayan veya vadesi geçmiş olan kredileri ifade eder. Bir kredi, belirli bir süre boyunca ödenmediği veya anlaşma şartlarına uygun olarak geri ödenmediği durumda NPL olarak sınıflandırılır. Finansal kurumlar genellikle kredi portföylerini NPL'ler ve temerrüt durumları açısından izlerler.

Özetle, temerrüte düşmek, bir borçlunun ödeme yükümlülüklerini yerine getirememesi durumunu belirtirken, NPL ise finansal kurumların kredi portföylerindeki geri ödenmeyen veya vadesi geçmiş kredileri tanımlar. Temerrüt durumu genel bir terimken, NPL spesifik olarak kredilerin performansına odaklanan bir kavramdır.

----------------Model Kurulabilmesi için-------------------

Bu metinde, model kurulabilmesi için gerekli olan güvenilir, temiz ve yeterli sayıda veri setine ulaşabilmek için yapılan kontroller ve ön-modelleme süreci açıklanmaktadır.

Model geliştirmenin temel şartlarından biri, iyi ve kötü hedef değişkenleri içeren güvenilir bir veri setine sahip olmaktır. Bu nedenle, veri hazırlama fazının ilk adımı olan kullanıcı kabul testi (KKT) olarak adlandırılan süreçte aşağıdaki kontrol noktaları üzerinde testler yapılmaktadır:

**Veri Kaynağı Testi:** Verinin geldiği kaynakların güvenilirliğini kontrol etmek amacıyla yapılan bir testtir.

**Kayıp Değer ve Veri Dağılım Analizi:** Veri setindeki eksik değerleri ve değişkenlerin dağılımlarını analiz etmek için gerçekleştirilen bir testtir.

**Tutarlılık Testleri:** Veri setindeki değerler arasındaki tutarlılığı kontrol etmek amacıyla yapılan testlerdir.

**Çapraz Kontroller:** Farklı veri kaynakları arasındaki çapraz kontrol testleridir.

Bu testler, tabloların analizlerine ulaşılana kadar devam eden tekrarlı bir süreçtir. Tabloların analizlerine ise bir link üzerinden erişilebilmektedir.

Ön-modelleme süreci, model geliştirecek portföyün özelliklerini belirlemek amacıyla yapılan veri analizi, model geliştirme ve performans dönemi seçimi gibi adımları içerir. Bu aşamada risk gruplarının ayrıştırılması ve alt grupların farklı modellerle yönetilip yönetilemeyeceğine dair ön kararlar alınır. Ancak, segment bazlı model kurmanın ek bir katkı sağlamadığı gözlenmiş ve bu nedenle canlı takip ve kullanım göz önüne alındığında tek bir modelle devam etme kararı alınmıştır.

Veri analizi aşamasında ise, veri sözlüğündeki tüm değişkenler, kayıp değer ve uç değer analizi, dağılım analizi, ve olası değer analizi gibi istatistiksel analizlere tabi tutularak incelenmiştir. Bu analizler, veri setinin karakteristiklerini ve potansiyel sorunlarını anlamak için yapılmıştır.

---------------- Temerrüt ve Hedef Değişken Tanımı

Temerrüt tanımı, müşterinin temerrüde düşme durumunu belirlemek için kullanılan kriterleri içermektedir. Temerrüt, müşterinin belirli şartları yerine getirememesi veya belirli süre içinde borçlarını ödeyememesi durumunu ifade eder. Bu tanım şu durumları kapsamaktadır:

Teminatlara başvurulmaksızın müşterinin bankaya olan borçlarını tamamen ödeyemeyeceğine banka tarafından kanaat getirilmesi.

Müşterinin bankaya olan yükümlülüklerini ifa etmede 90 günden fazla gecikmesi.

Bu temerrüt tanımına göre, 2009 yılında Temerrüt Tanımlama Sistemi (DIS) oluşturulmuştur. DIS, müşterinin herhangi bir kredisinin temerrüt etmesi durumunda müşterinin seyyal ve takipteki tüm kredilerini temerrüt etmiş olarak gösterir. Dönüşüm çalışmalarıyla CRD1 - CRD2 dönüşümü gerçekleştirilmiş ve geçmişe dönük temerrüt olayları CRD2 yöntemi ile hesaplanmıştır.

4-2) Hedef Değişken Tanımı:

Hedef değişken, modelin temelini oluşturan bir müşterinin temerrüt durumunu gösteren değişkendir. Temerrüt tanımına dayanarak, müşterinin sahip olduğu ürünlerden herhangi birinde temerrüte düşmesi durumunda hedef değişken 1 (bir) değerini alır. Hedef değişkenin oluşturulmasında "İYİ" ve "KÖTÜ" gözlemler ayrıştırılır.

"KÖTÜ" tanımına göre, müşterinin herhangi bir ürününde temerrüte düşmesi durumunda hedef değişken 1 değerini alır.

"İYİ" tanımına göre, performans dönemi içerisinde 3 ay veya daha fazla aktif durumda olan ve temerrüte düşmeyen müşteriler için hedef değişken 0 değerini alır.

Hedef değişkeni, "İYİ" ve "KÖTÜ" gözlemleri ayırt etmek için temerrüt ayraçları kullanılarak oluşturulmuştur. Bu hedef değişkeni, Esnek Hesap Başvuru Skor Kartı modelinin temelini oluşturmaktadır. Performans dönemi 12 ay olarak belirlenmiş ve hedef değişkeni "TARGET\_FLG" olarak adlandırılmıştır.

Bu aşamadan sonra, metin Esnek Hesap Başvuru Modeli'nin geliştirilme sürecini detaylı bir şekilde ele alacak şekilde devam edecektir.

**Performans dönemi, başvuru tarihinden başlayarak temerrüdün gözlemlenebilmesi amacıyla belirlenmesi gereken zaman dilimini ifade etmektedir. Basel II uzlaşısına göre içsel derecelendirme yaklaşımlarında, 12 aylık performans döneminin, kredi borçlusunun temerrüde düşme olasılığının hesaplanması için uygun olduğu kabul edilmektedir. Buna dayanarak, Esnek Hesap Başvuru Skor kartı için hedef değişkeninde 12 aylık performans dönemi kullanılmış ve hedef değişkeni TARGET\_FLG olarak belirlenmiştir.**

Kullanılan eleme kriterlerinin detayları aşağıda paylaşılmıştır;

* Personel başvuruları,
* Yaşı 18’den küçük olan müşteri başvuruları,
* Başvuru yaptığı tarih itibariyle hali hazırda temerrüt durumunda olan başvurular,
* Başvuru tarihinden sonraki 12 aylık süreç içerisinde Esnek Hesap ürününde herhangi bir risk bakiyesi olmayan başvurular,
* Başvuru sırasında 3 aydan eski KKB sorgusu kullanılan başvurular,
* \*\*\*\*\*\*\*\*\*
* **Çapraz ürün olarak satılan Esnek Hesap başvuruları,**
* **Merged CIF**

**----------------------------Gözlem Dönemi Hakkında-----------------------------------------**

* Dokümanın bu kısmında, Esnek Hesap (EH) Başvuru Skor Kartı'nın geliştirme süreci ve kullanılan değişken azaltım teknikleri üzerine bilgiler bulunmaktadır. İlk olarak, Esnek Hesap başvurularından oluşan tahmin kitlesi 12 aylık performans dönemine dayanarak oluşturulmuştur. Model geliştirme dönemi için Ocak 2012 - Ağustos 2019 tarihleri arasındaki 86 aylık bir süre belirlenmiştir. Bu karar, Basel II Uzlaşısı ve İDDY yönetmeliğine uyum gereklilikleri doğrultusunda alınmıştır.
* Dokümanın devamında, Geliştirme, Validasyon, Fazladan Zaman ve Güncel dönemler olmak üzere dört farklı dönem belirlenmiştir. Bu dönemlerin gözlem sayıları ve temerrüt oranları Tablo.3'te ayrıntılı bir şekilde sunulmuştur. Geliştirme, Validasyon ve Fazladan Zaman dönemlerinin temerrüt oranları sırasıyla 3.01%, 3.01% ve 1.13% olarak görülmektedir. Güncel dönem için temerrüt oranı belirtilmemiştir, çünkü performans dönemi tamamlanmamış gözlemler bu aşamada "İYİ" ya da "KÖTÜ" olarak kategorize edilememiştir.
* Şekil.7'de Geliştirme, Validasyon ve Fazladan Zaman dönemlerindeki aylık popülasyon dağılımı ve temerrüt oranları gösterilmektedir.
* Daha sonra, Esnek Hesap başvurularının model geliştirme sürecinde kullanılacak değişkenler üzerinde yapılan eleme tekniklerine odaklanılmıştır. Değişken azaltımı için iki seviyeli bir süreç uygulanmıştır. İlk aşamada, iş kararı, kayıp değer oranları ve trend analizi kullanılarak değişken sayısı azaltılmıştır. İkinci aşamada ise düşük Gini değerleri, değişken kümeleme tekniği ve popülasyon istikrarlılık göstergesi PSI kullanılarak değişken sayısı daha da daraltılmıştır.
* Son olarak, PSI değerleri kullanılarak popülasyonun stabilitesini değerlendiren bir kontrol yapılmıştır. PSI değeri 0.1'in altındaki değişkenler modelden çıkarılmış, ancak 0.1 - 0.25 arasındakiler manuel olarak değerlendirilmiştir. Gini değişimine odaklanılarak, modelin performansını olumsuz etkileyebilecek değişkenler elenmiştir.
* Bu süreç sonunda, 869 adet değişkenin kaldığı belirtilmiştir. Dokümanın devamında, bu değişkenlerin regresyon analizine tabi tutulduğu ve modelin geliştirme aşamasının detaylarına geçilecektir.

**Varyans Arttırıcı Faktör (VIF) Kullanılarak Yapılan Değişken Elemesi:**

* Çoklu doğrusal bağlantı problemi, regresyon modellerinde parametre varyansının yüksek çıkması gibi sorunlara neden olabilir.
* VIF, çoklu doğrusal bağlantıyı belirlemek için yaygın bir yöntemdir. VIF, parametre tahminlerinin ve varyanslarının gerçek değerlerinden ne kadar uzaklaştığını belirler.
* VIF değeri 2.5'ten büyükse, çoklu bağlantının var olduğunu gösterir. Bu durumdaki değişkenler elemeye tabi tutulur.
* Çalışmada, VIF değeri 2 ve üzeri olan değişkenlerin regresyon eleme adımına geçildiği belirtilmiştir.

**Regresyon Kullanılarak Yapılan Değişken Seçim Süreci:**

* Geriye doğru seçim işlemi, regresyon modellerinde kullanılan bir değişken seçim yöntemidir.
* Başlangıçta tüm değişkenler modele alınır, p<0.05 anlamlılık seviyesini geçmeyen değişkenler birer birer çıkarılarak anlamlı değişkenler belirlenir.
* Değişken azaltımı seviye 2 sonucunda 47 değişken elenmiştir.
* WOE katsayıları kontrol edilerek, değişkenin katsayısının işaretinin pozitif olmaması durumunda değişkenler tekrar eleme kriteri olarak kullanılmıştır.
* Modelde sağladıkları GINI katkısına bakılarak fazla değişkenin modeli verimsiz hale getirebileceği göz önüne alınarak ek bir eleme uygulanmıştır.
* Final modelde 21 adet değişken kullanılmıştır.