

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | |
|  | | FİNAL RAPORU | | | | |  | |
|  |  | | | | | | |  |
|  | | | |  |  | | | |
|  | | | | Sude Gül ÜZÜM |  | | | |
|  | | | | IZV701Veri Analitiğinin Temelleri—Kemal İLTER—28.05.2025 |  | | | |
|  | | |  | | |  | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1. Özellik Mühendisliği: Esra’nın kullandığı değişkenlere ek olarak hangi özellikleri önerirsiniz? Neden? | | | | |  |
|  | |  |  |  |  | |
|  | Özel günlere denk gelen haftalarda örnek olarak resmi tatiller, dini bayramlar, anneler günü gibi özel günlere denk gelen haftalarda değişkenliği görmek için bir sütun eklenerek gözlemler yapılabilir. Sütun değeri 0-1 değerlerini alarak o hafta özel gün içeren bir hafta mı olup olmadığı anlaşılabilir ve satış dalgalanmalarını anlamdırılabilir.  Mağazalara özgü değerleri göz önüne alınarak mağazanın büyüklüğü, konumu, rakip mağazalara yakınlığı ve yakınındaki rakip mağaza sayıları, mağazanın bulunduğu konumdaki nüfus yoğunluğu gibi demografik değişkenler müşteri segmenti gibi detaylarda göz önüne alınarak değerlendirmeler yapmak çok daha anlamlı sonuçlara varmamıza olanak sağlar. | | | | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 2. Model Seçimi: Regresyon modeli yerine kullanabileceğiniz iki alternatif yöntem önererek olası avantajlarını tartışın. | | | | |  |
|  | |  |  |  |  | |

Aslında benim yukarıda belirttiğim değişkenleri gibi değişken sütunları oluşturup kategorik değişkenler olarak tanımlarsak hangi faktörlerin mağaza satışlarını etkileyen faktörler olduğunu öğrenmek için sınıflandırma algoritmalarından lojistik regresyon ve karar ağacı modelini kullanmanın mantıklı olacağını düşündüm aynı soruyu birkaç yapay zeka araçlarına sorduğumda random forest’ı mevsimsellik, gelir ve promosyon gibi faktörler için ve XGBoost algoritmasını tahmin performansı için uygun olacağı cevaplarını aldım. Bahsettiğim 4 algortima da denererek performans karşılaştırılması yapılabilir.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 3. İletişim: Yönetici özetinin daha ikna edici ve görsel açıdan daha güçlü olması için neler eklenebilir? | | | | |  |
|  | |  |  |  |  | |

Sunumu yaparken Esra’nın bahsettiği promosyon etkisi, sıcaklık etkisi gibi sonuçları grafik ve ısı haritaları kullanarak yazılardan çok orta ve üst seviye yöneticinin baktığında anlam çıkaracabileceği ve aksiyon almasına yardımcı olacak seneryo analizi grafikleri kullanarak, kuralcı analitik uygulamaları dersinde gördüğümüz what-if analizlerini kullanarak dashboard tasarımı ile anlatımını hikayeleştirme yönetimini kullanarak riskleri ve fırsatları vurgulama yaparak etkili bir iletişimde bulunmuş olur.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 4. Etik Boyut: Hava koşullarına dayalı promosyon stratejisi farklı gelir düzeylerindeki bölgeler için uygun mu? Bu nasıl ölçülebilir ve takip edilebilir? | | | | |  |
|  | |  |  |  |  | |

Hava koşullarına dayalı promosyon stratejisi farklı gelir düzeylerindeki bölgeler için eşitsizlik yaratma riski vardır tabii ki. Bunun için müşteri segmentine uygun promosyonlar uygulabilir. Ama bir şirket için önemli olan etik olması değil en fazla karı nasıl elde edeceği olduğu için bununla ilgilenmeyeceklerini düşünüyorum.

Bu durumu ölçebilmek için pilot uygulamalar yapılarak her gelir seviyesine uygun mağazaları pilot mağaza seçerek ölçeklendirip takibini yapabiliriz. Ya da anketler, satış sonrasındaki iade oranları gibi ölçekler ile takibi yapılabilir.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 5. Tekrarlanabilirlik: R kodunu proje ekibinin kolayca sürdürebileceği bir yapıya dönüştürmek için hangi iyileştirmeleri önerirsiniz? | | | | |  |
|  | |  |  |  |  | |

R da satır satır çalıştırdığımızda hangi satırları seçeceğimiz bazen kafa karışıklığına yol açtığı için ve çalıştırırken onda neyi çalıştırdığımı çıktıyı görmeden anlayamıyorum o yüzden en başına yorum olarak ekliyorum kodun başlangıç ve bitişinin belirli olması çalıştırma işleminde büyük kolaylık sağlıyor ve bir başkasının da anlayıp devamında işlem yapması işlevsellik açısından iyi oluyor. R markdown dosyasında aslında kodun gözükmesinden ziyade chunk’lar kullanılarak sonucun gözükmesi, başlık ve yorum satırlarından oluşması sürdürebilirlik açısından daha iyi olur.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 6. Bakış Açısı: Bu raporu hazırlayan siz olsaydınız yukarıdaki konuların dışında neleri farklı yapmak isterdiniz? | | | | |  |
|  | |  |  |  |  | |

Zaten yukarıda 1,2 ve 3 sorularda neleri yapacağımından bahsettim. Eklemek istediğim sütunlardan, kullanmak istediğim modellerden, sunum yaparken daha çok grafiklerden oluşan ve farklı seneryoları analizleri what-if analizleri kullanarak bir dashboard tasarlamak istediğimi belirttim. R dosyamı da Rmarkdown dosyası oluşturarak her bir kod bloğunu chunkların içine yazarak çıktı da kodları göstermemeye özen gösterirdim korelasyon analizi kullanarak faktörler arasındaki ilişkileri gözlemlerdim. Eksik değerleri silmek yerine bir model geliştirerek en doğru şekilde doldurmak isterdim ve Tahmin yaparken her bir tahminin güven derecelerini de analize eklerdim.