



Jüri Soru-Cevap Simülasyonu

Soru 1: Bu projede neden geleneksel istatistiksel modeller (ARIMA, Sarima) yerine Derin Öğrenme (Deep Learning) modellerini (GRU/LSTM) tercih ettiniz?

Cevap: "Geleneksel modeller (ARIMA gibi) doğrusal ilişkilerde iyidir ancak karmaşık, doğrusal olmayan (non-linear) ve uzun vadeli bağımlılıkları yakalamakta zorlanır. Stok verilerimiz ani sıçramalar ve karmaşık örüntüler içerdiği için, 'unutma kapıları' (forget gates) sayesinde geçmişteki önemli bilgileri uzun süre hafızada tutabilen LSTM ve GRU modellerini tercih ettik. Ayrıca bu modeller ölçeklenebilirlik ve çok değişkenli veriyi işleme konusunda daha esnektir."

Soru 2: "Seyrek Talep" (Intermittent Demand) problemi stok yönetiminde sıkça görülür. Modeliniz bu durumla nasıl başa çıkıyor?

Cevap: "Grafiklerde de görüldüğü gibi, bazı ürünlerde satışlar sıfıra yapışıyor ve aniden pik yapıyor (sparse data). Bunu aşmak için zaman toplulaştırma (temporal aggregation) stratejisi uyguladık. Haftalık verideki gürültü yerine, 3 Aylık (Quarterly) görünümü ekleyerek modelin 'sıfırları' değil, genel trendi öğrenmesini sağladık. Ayrıca Pinball Loss kullanarak modelin ortalamayı değil, riskli üst sınırları tahmin etmesini sağladık."

Soru 3: Projenin "iş değeri" (business value) nedir? Bu sistemi kullanan bir şirket tam olarak ne kazanacak?

Cevap: "Bu proje iki temel sorunu çözüyor: Satış Kaybı (Lost Sales) ve Stok Maliyeti. Modelimiz sayesinde 'Stok Yok' (Stockout) durumlarını önceden tahmin ederek ciro kaybını engelliyoruz. Aynı zamanda 'Aşırı Stok' uyarısı ile gereksiz depolama maliyetlerini düşürüyoruz. Yani hem geliri artırıyor hem de operasyonel gideri azaltıyoruz."

Soru 4: RFM analizi ile talep tahminini (forecasting) nasıl birleştirdiniz? Bu ikisi birbirini nasıl besliyor?

Cevap: "Forecasting bize 'ne kadar satacağımızı' söylerken, RFM 'kime satacağımızı' söylüyor. Özellikle 'Aşırı Stok' durumundaki ürünler için Dashboard'da o kategoriyi en çok alan RFM segmentlerini (Örn: Sadık Müşteriler) listeliyoruz. Böylece elimizde kalan ürünü rastgele değil, doğru müşteriye kampanya yaparak eritiyoruz."

Soru 5: Stok durumlarını belirlerken (Stok Yok, Aşırı Stok vb.) eşik değerlerini neye göre belirlediniz?

Cevap: "Stok durumlarını belirlerken dinamik bir yaklaşım izledik. Eğer Mevcut Stok < Tahmin ise bu kesin bir krizdir (Stok Yetersiz). Ancak 'Aşırı Stok' için, literatürdeki 'Inventory Turnover' (Stok Devir Hızı) mantığına dayanarak, 6 aylık tahminden fazla stoğu olan ürünleri 'aşırı' olarak etiketledik."

Soru 6: Neden tek bir model yerine hem GRU hem de LSTM modellerini yarıştıyorsunuz?

Cevap: "Her ürünün karakteristiği farklıdır. Kimi ürün son 3 aya (kısa vade) daha duyarlıyken, kimi ürün mevsimseldir. Bu yüzden tek bir modeli dayatmak yerine, her ürün için GRU ve LSTM'i farklı zaman pencerelerinde (aylık, haftalık, 3 aylık) çalıştırıp, validasyon setinde en düşük hatayı (RMSE) veren modeli o ürün için 'Şampiyon' ilan ediyoruz."

Soru 7: Modelinizde "Pinball Loss" (Quantile Loss) kullandığınızı görüyoruz. Bu metrik MSE'den (Mean Squared Error) neyi farklı yapıyor ve neden stok yönetiminde önemli?

Cevap: "MSE (Ortalama Kare Hata) modelin 'ortalamayı' bulmasını sağlar. Ancak stok yönetiminde ortalama risklidir; çünkü eksik tahmin (yok satmak), fazla tahminden daha maliyetli olabilir. Pinball Loss ile modele asimetrik bir ceza veriyoruz: 'Eksik tahmin yaparsan daha çok ceza alırsın'. Bu sayede model %90 güven aralığında çalışarak bizi stoksuz kalma riskinden koruyor."

Soru 8: Eğitim sırasında "Early Stopping" ve "Dropout" kullanmanızın sebebi nedir?

Cevap: "Veri setimizdeki dalgalanmalar modelin ezber yapmasına (overfitting) neden olabilirdi. Dropout(0.2) ile nöronların %20'sini rastgele kapatarak modelin daha genel kuralları öğrenmesini sağladık. EarlyStopping ile de model öğrenmeyi tamamladığında (val_loss düşmediğinde) eğitimi durdurarak hem zamandan tasarruf ettik hem de en iyi ağırlıkları koruduk."

Soru 9: Veri setinde 3 aylık (Quarterly) analizi neden ekleme gereği duydunuz? Haftalık ve aylık analiz yetmedi mi?

Cevap: "Haftalık verilerde çok fazla 'sıfır' (hiç satış olmayan günler) vardı. Bu durum modelin sürekli sıfır tahmin etmesine yol açıyordu. Veriyi 3 aylık periyotlara sıkıştırdığımızda trendler ve mevsimsellik (sivri uçlar) daha net ortaya çıktı. Bu, özellikle yavaş hareket eden ürünlerin tahmin başarısını ciddi oranda artırdı."

Soru 10: Modelinizin başarısını ölçmek için neden RMSE (Root Mean Squared Error) metriğini referans aldınız?

Cevap: "RMSE, büyük hataları karesini alarak cezalandırır. Stok yönetiminde küçük hatalar (3 yerine 4 tahmin etmek) tolere edilebilir, ancak büyük hatalar (100 satacakken 10 demek) felakettir. RMSE, bu büyük sapmaları daha sert cezalandırdığı için stok riskini minimize etmede doğru bir metriktir."

Soru 11: Veri ön işleme aşamasında ölçeklendirme (MinMaxScaler) neden gerekiyordu? Yapmasaydık ne olurdu?

Cevap: "Derin öğrenme modelleri (GRU/LSTM), girdilerin 0-1 aralığında olmasını sever. Eğer ölçeklendirme yapmasaydık, satış adedi 10 olan ürünle 10.000 olan ürün arasında modelin ağırlık güncellemesi (gradient descent) dengesiz olurdu ve model yakınsayamazdı (converge edemezdi)."

Soru 12: Modeliniz "Recursive Forecasting" (Özyinelemeli Tahmin) yapıyor. Bu yöntemin avantajı ve riski nedir?

Cevap: "Gelecek 6 ayı tahmin ederken, modelin ilk ay için yaptığı tahmini, ikinci ayın girdisi olarak kullanıyoruz. Bunun avantajı, uzun vadeli tahmin üretebilmesidir. Riski ise, ilk tahminde hata varsa bu hatanın katlanarak (error propagation) ilerlemesidir. Bunu minimize etmek için window boyutunu (geçmiş veri) yeterince geniş tuttuk."

Soru 13: Dashboard'da sunduğunuz "Standart Model" ve "Güvenli Model" arasındaki farkı yöneticiye nasıl açıklarsınız?

Cevap: "Yöneticiye şöyle derim: 'Eğer bütçeniz kısıtlıysa ve ortalama bir stok tutmak istiyorsanız Yeşil Çizgiyi (Standart - MSE) takip edin. Ancak bu ürün "Prestij Ürünü" ise ve müşteriye asla "yok" demek istemiyorsanız, Pembe Çizgiyi (Güvenli - Pinball) dikkate alın.' Bu, maliyet ile müşteri memnuniyeti arasındaki tercihtir."

Soru 14: Bir kullanıcı "Haftalık" yerine "Aylık" seçtiğinde arka planda model eğitim stratejisi nasıl değişiyor?

Cevap: "Kullanıcı periyodu değiştirdiğinde sadece grafik değişmiyor; veri seti ve modelin geçmişe bakış açısı (window) da değişiyor. Haftalıkta son 12 haftaya (3 ay) bakarken, Aylıkta son 6 aya bakıyoruz. Bu da modelin kısa vadeli dalgalanmalara mı yoksa genel trende mi odaklanacağını belirliyor."

Soru 15: Dashboard üzerindeki "Kritik Stok Listesi" neye göre sıralanıyor ve önceliklendiriliyor?

Cevap: "Listeyi, 'Stok Yok' ve 'Stok Yetersiz' durumlarına göre filtreliyoruz. Sistem, tahmin edilen satışın (NextMonthForecast) mevcut stoktan (RemainingStock) büyük olduğu durumları en acil olarak işaretliyor."

Soru 16: Bu sistemi gerçek zamanlı (real-time) bir e-ticaret sitesine entegre etmek istesek karşılaştığımız en büyük zorluk ne olurdu?

Cevap: "En büyük zorluk Eğitim Süresi (Training Latency) olurdu. Her yeni sipariş geldiğinde modeli baştan eğitmek imkansızdır. Çözüm olarak; modeli günde 1 kez gece (batch processing) eğitip, gün içinde sadece önceden eğitilmiş ağırlıklarla tahmin (inference) yaptırırdım."

Soru 17: Dış faktörleri (promosyonlar, bayramlar, hava durumu) modele eklemek isteseydiniz mimariyi nasıl değiştirdiniz?

Cevap: "Mevcut modelimiz 'Univariate' (Tek Değişkenli) çalışıyor. Dış faktörleri eklemek için mimariyi 'Multivariate' (Çok Değişkenli) yapıya çevirdim. Input Shape'i (Window, 1) yerine (Window, Feature_Sayısı) yapardım. Fiyat, indirim oranı, özel gün bayrağı gibi özellikleri input olarak verirdim."

Soru 18: Ürün sayısı 18.000'den 1 milyona çıksaydı, run_analysis.py scriptiniz performans açısından nasıl etkilenirdi ve bunu nasıl çözerdiniz?

Cevap: "Şu anki for döngüsü 1 milyon ürünle günlerce sürer. Çözüm olarak; 1. Paralel İşleme (Multiprocessing): joblib kullanarak işlemi çekirdek sayısına bölmek. 2. Kümeleme: Benzer satış karakteristiğine sahip ürünleri gruplayıp (clustering) tek bir model eğitmek."

Soru 19: RFM segmentasyonu statik bir analizdir. Bunu dinamik hale getirmek için ne önerirsiniz?

Cevap: "Şu an analiz sabit bir tarihe göre yapılıyor. Bunu dinamik yapmak için, her gün çalışan bir script ile son 30/60/90 günlük pencereler üzerinden RFM skorlarını günceller ve müşterinin segment değişimini (Örn: 'Sadık'tan 'Uykuda'ya geçişini) takip eden bir 'Müşteri Yaşam Döngüsü' (CLV) modülü eklerdim."

Soru 20: Projenin en zayıf yönü veya "keşke vaktim olsaydı da şunu da yapsaydım" dediğiniz kısmı nedir?

Cevap: "En büyük geliştirme fırsatı, 'Tedarik Süresi' (Lead Time) parametresini sisteme eklemek olurdu. Şu an sadece 'Stok Yetersiz' diyoruz ama ürünün depoya gelmesi 30 gün sürüyorsa, siparişi 30 gün önce vermemiz gerektiğini sisteme öğretmek, sistemi tam bir tedarik zinciri çözümüne dönüştürdü."