İSTANBUL BİLGİ AKADEMİSİ

AI Model Explainability & Interpretability Eğitimi

YZ2512

Eğitim Amacı

Katılımcıların yapay zekâ ve makine öğrenmesi modellerinin karar mekanizmalarını anlamalarını sağlayarak, model çıktılarının açıklanabilirliğini ve yorumlanabilirliğini artırmalarını sağlamaktır.

Eğitim Detayları

- 🏢 Kurumsal Eğitim
- 👥 Kapalı Grup / Bire-bir
- 💻 Online 🏫 Yüz Yüze
- 🕒 Eğitim Süresi: 12 Saat / 2 Gün
- 📝 Sertifika: Katılım Sertifikası
- 📆 Eğitim Takvimi: Hİ / HS

Eğitim İçeriği

Explainability ve Interpretability Kavramları

- Model açıklanabilirliği ve yorumlanabilirlik tanımları
- İş süreçlerinde model açıklanabilirliğinin önemi
- Açıklanabilirlik ile regülasyon uyumu
- Model güvenilirliği ve etik kullanımı

Temel Yöntemler ve Teknikler

- Feature importance ve SHAP değerleri
- LIME ve lokal açıklama yöntemleri
- Model-agnostic ve model-specific teknikler
- Görselleştirme ile model davranışı analizi

Black-box Modellerin Açıklanması

- Derin öğrenme modelleri ve açıklanabilirlik zorlukları
- Neural network yorumlama yöntemleri
- Sensitivity analysis ve partial dependence plots
- Gerçek dünya örnekleri ile model analizi

Uygulamalı Çalışmalar

- Veri seti üzerinde SHAP ve LIME uygulamaları
- Model çıktılarının yorumlanması ve raporlanması
- Açıklanabilirlik ile iş kararlarının desteklenmesi
- Kurumsal senaryolarda vaka çalışmaları

Model Güvenilirliği ve Regülasyon

- Açıklanabilirlik ve etik ilişkisi
- Model bias tespiti ve düzeltme
- Regülasyon ve veri uyumluluğu
- Model izleme ve sürekli değerlendirme

Hedef Kitle

Veri bilimciler, makine öğrenmesi mühendisleri, iş analistleri, ürün yöneticileri, proje ekipleri ve yapay zekâ modellerini daha şeffaf ve güvenilir kullanmak isteyen tüm katılımcılar.