

AR-GE MERKEZİ PROJE KARTI

Proje Durumu:

Proje Kodu:

Proje Adı:

Proje Özeti:

Projenin Konusu:

Proje Konusunu Belirleyen İhtiyaçlar: (Maks.5.000 karakter)

Şirket Tekilleştirme Projesi, müşteri ilişkileri, aday arama ve öneri sistemlerinde kısmi ya da bulanık çoğaltmaların oluşması sorununu çözmek amacıyla başlamıştır. Manuel olarak girilen veri setleri nedeniyle veri duplikasyonu ve veri kalitesinde düşüş gibi sorunlarla karşılaşmaktadır. Özellikle, farklı kişiler tarafından girilen aynı şirket isimleri küçük farklılıklar gösterebilir, bu da hem müşteri ilişkilerinde hem de iş süreçlerinde hatalara yol açabilir.

Bu proje, yapay zeka ve dil modellerini kullanarak verilerin tekilleştirilmesini sağlayacak ve bu sayede verilerin daha tutarlı ve güvenilir hale gelmesine yardımcı olacaktır. Veri işleme sürecinde, doğal dil işleme (NLP) metotları ile beraber, Üretken Yapay Zeka ve embedding yönetimleri kullanılacaktır.

Projenin bir diğer önemli hedefi, müşteri verilerinin ve aday kayıtlarının doğruluğunu artırarak, şirketin müşteri ilişkilerini ve operasyonel süreçlerini daha verimli hale getirmektir.

Proje Kapsamında Yapılan Faaliyetler: (Maks.5.000 karakter)

Arge ve Galatasaray Üniversitesi Akademik Üyeleri ile beraber bir embedding model üzerinde çalışılıyor:

Veri Kümelerinin İncelenmesi ve Yöntem Belirlenmesi:

Verilerin ilk gözden geçirilmesiyle veri türleri, eksik veri oranları ve veri dağılımları hakkında bilgi edinilir. Veri keşfi için görselleştirme araçları kullanılacaktır. Verilerin gruplandırılması amacıyla clustering uygulanır.

Uygun model seçimi (veri dağılımına göre):

Embedding ve model eğitimi gibi, modellerin başarıları ve zorlukları tartışılır. Veri seti eğitim, doğrulama ve test kümelerine ayrılacaktır. Embedding, yüksek boyutlu verilerin (örneğin, kelimeler veya resimler) daha düşük boyutlu vektörlere indirgenmesi anlamına gelir. Bu, veri analizinde verimli hale gelmeleri için kullanılacaktır. Embedding teknikleri üzerine yapılan akademik çalışmalarda kullanılan yöntemler (örneğin, BERT, GPT) araştırılır ve analiz edilir.

Precision (doğruluk) ve Recall (duyarlılık) arasında bir denge kuran bir metrikler, model performansını değerlendirmek için kullanılabilir.

PoC (proof of concept): Başarı metriklerden sonra ürün ekibi ile beraber model ve sonuçlar paylaşılır ve canlı geçişi yapılır.

Projenin Yenilikçi Yönü ve Ar-Ge Niteliği: (Maks.5.000 karakter)

Projenin en temel yenilikçi yönü, müşteri verilerinin doğruluğunu sağlamak amacıyla geleneksel veri temizleme yöntemlerinin ötesine geçerek, yapay zeka temelli yaklaşımlar kullanmasıdır. Geleneksel yöntemler genellikle manuel ve kural tabanlı süreçlere dayanırken, bu proje derin öğrenme modelleri, doğal dil işleme (NLP) teknikleri ve Embedding algoritmaları kullanarak çok daha dinamik, esnek ve akıllı bir çözüm sunmaktadır. Bu işlem ile cümleler, belirli boyutlara sahip vektörlere dönüştürülür. Kullanılan Embedding modeli, bu vektörlerin boyutlarını belirliyor.

Projenin Beklenen Çıktıları ve Faydaları: (Maks.5.000 karakter)

Şirket Tekilleştirme Projesi'nin en önemli çıktılarından biri, veri kalitesinin önemli ölçüde artması olacaktır. Tekilleştirilmiş veri, müşteri ilişkileri yönetimi (CRM), aday takibi, raporlama ve iş zekası gibi kritik alanlarda daha doğru ve tutarlı veri sağlanmasına olanak tanıyacaktır. Manuel veri girişi sırasında meydana gelen hatalar, veri tabanlarında çoğaltmalar ve tutarsızlıklar yaratır. Bu durum, veri arama ve doğrulama süreçlerini karmaşık hale getirir.

Şirket Tekilleştirme Projesi'nin bir diğer önemli faydası, aday arama süreçlerinin iyileştirilmesidir. İş arama süreçlerinde aday aramada sıklıkla benzer şirket isimlerle aranmaktadır. Bu durum, aday arama karmaşıklığına yol açarak, adayın değerlendirme sürecinde gözden kaçmasına veya yanlış eşleşmelere neden olabilir. Şirket tekileştirmede şirketlerin değeri tekil olup, arama sonuçlarının doğruluğu artırmaktadır.

