

KURUMLAR İÇİN BÜYÜK DİL MODELİ VE VERİ ANALİZİ AJANI



BURSA TEKNİK
ÜNİVERSİTESİ

EBRAR BEGÜM ŞİPAL

Bilgisayar Mühendisliği

Bursa Teknik Üniversitesi

Lisans Bitirme Tezi Sunumları, Temmuz 2025

- ▶ ERP sistemleri, şirket içi süreçlerin bütünlük bir şekilde yönetilmesini sağlar; ancak kullanım karmaşıklığı, özellikle **yeni kullanıcılar** için ciddi bir engel oluşturmaktadır.
- ▶ Günümüzde kullanıcılar, yazılımlarla **doğal dile** etkileşim kurmayı tercih etmektedir.
- ▶ Bu çalışma, büyük dil modelleri (LLM) ve embedding tabanlı yaklaşımlarla doğal dildeki kullanıcı sorgularını anlamlandırarak ERP veritabanındaki karşılıkları ile eşleştirmeyi ve otomatik SQL sorguları üretmeyi amaçlamaktadır.
- ▶ Böylece, teknik bilgiye ihtiyaç duymadan ERP sistemlerinden veri çekmek mümkün hale gelmektedir.

Literatür Taraması

- ▶ Literatüre hâkim olabilmek amacıyla LLM, NLP, embedding ve ERP sistemlerinde yapay zekâ kullanımı alanlarında akademik ve sektörel kaynak incelenmiştir.
- ▶ Mhaskey (2024), LLM'lerin ERP süreçlerinde operasyonel verimliliği artırma potansiyelini vurgulamaktadır. Gerçek zamanlı veri analizi, doğal dilde sorgulama ve kullanıcı desteği konularında örnekler sunulmuştur.
- ▶ Narne (2022), yapay zekâ ve makine öğrenmesinin ERP sistemlerine entegrasyonunun veri odaklı karar alma süreçlerini nasıl dönüştürdüğünü incelemiştir.
- ▶ Dachepalli (2025), müşteri destek süreçlerinde Transformer tabanlı NLP modellerinin kullanımını önererek, doğal dilin kurumsal sistemlerle etkili şekilde eşleştirilebileceğini göstermiştir.
- ▶ Harshvardhan ve Jayakrishnan (2024), LLM tabanlı çok ajanlı yapılarla ERP süreçlerinin otomasyonunu sağlayan bir mimari önermiştir.

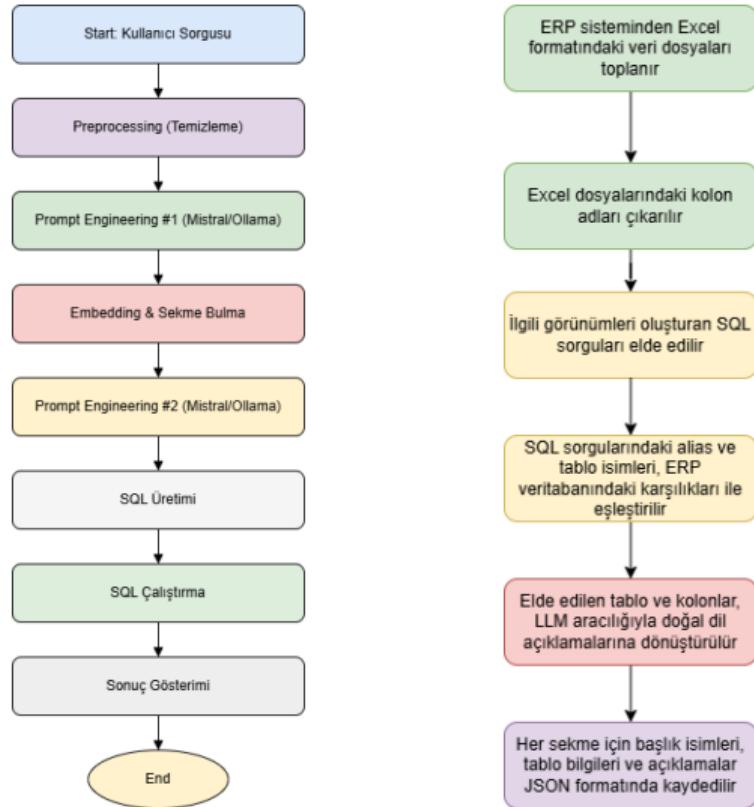
Problem Tanımı

- ▶ Kurumsal sistemlerde doğal dil arayüzü entegrasyonu, genellikle uzun geliştirme ve uylarlama süreleri gerektirmektedir.
- ▶ LLM tabanlı çözümler, genellikle büyük ve temiz veri kümeleriyle çalışmayı gerektirdiğinden; sınırlı ve yapısal olmayan veri ile etkili sonuç almak zordur.
- ▶ Kurum içi veri güvenliği nedeniyle sistemlerin internete açık hale getirilmesi çoğu zaman mümkün değildir; bu da bulut tabanlı LLM servisleriyle entegrasyonu zorlaştırmaktadır.
- ▶ Bu nedenlerle, kurum içinde çalışabilen, az veriyle öğrenebilen ve teknik bilgi gerektirmeden doğal dille ERP verisine ulaşmayı sağlayan bir yapay zekâ ajansı geliştirilmesi hedeflenmiştir.

Alt Problemler

- ▶ Kurum içindeki veri şemalarının (sekme, tablo, sütun) doğal dille eşleştirilmesi için uygun embedding yapısının oluşturulması
- ▶ Kullanıcıdan alınan doğal dil sorgusunun doğru bağlama göre işlenmesi (prompt design)
- ▶ En uygun veri alanının embedding uzayında tespit edilmesi (cosine similarity gibi metriklerle)
- ▶ Doğal dilden SQL'e doğru ve güvenli dönüşümün sağlanması
- ▶ Internete çıkmadan lokal çalışabilecek hafif ve verimli bir LLM yapısının kurulması (örneğin: Ollama + Mistral)
- ▶ Az veriyle çalışan sistemin hassasiyetini ve başarımını artırmak için test süreci

- ▶ Yazılım Yöntemleri:
 - ▶ Prompt Engineering
 - ▶ Semantic Textual Similarity (Embedding ve cosine similarity)
 - ▶ Doğal Dilden SQL'e Dönüştürme (LLM tabanlı)
- ▶ Kullanılan Kütüphane ve Araçlar:
 - ▶ Ollama (Mistral 7B modeli ile lokal LLM çalışma)
 - ▶ Sentence Transformers (paraphrase-multilingual-MiniLM-L12-v2 modeli)
 - ▶ SQLite (Veritabanı testi ve sorgulama)
 - ▶ Python (Veri işleme, istem hazırlama, embedding karşılaştırmaları)
- ▶ Veri Hazırlık Süreci:
 - ▶ ERP sistemine ait sekme, tablo ve sütun adlarının manuel olarak çıkarılması
 - ▶ Excel dosyalarının vektörel karşılıklarının embedding uzayına aktarılması
 - ▶ Doğal dil sorgularının eşleştirme başarımını ölçmek için test sorguları oluşturulması



Uygulama – Veri Hazırlama Aşaması

- ▶ **Excel İşleme:** Belirli bir klasördeki tüm Excel dosyaları taranır.
- ▶ **Başlık Okuma:** Her dosyanın ilk satırından kolon başlıklarları çıkarılır.
- ▶ **SQL Eşleşmesi:** SQL sorgularındaki alias ve kolon isimleri ile eşleştirilir.
- ▶ **Veri Yapısı:** JSON formatında, sekme-adı, tablo, alias, kolon ve açıklama bilgileri ile yapılandırılır.

Uygulama – LLM ve Arayüz Bileşenleri

- ▶ **Embedding ile Sekme Bulma:** Kullanıcı sorgusu SentenceTransformer ile vektörleştirilir.
- ▶ **En Benzer Sekme Seçimi:** Önceden oluşturulan sekme cümleleriyle benzerlik karşılaştırması yapılır.
- ▶ **OpenRouter API ile Sorgu Üretilimi:** Uygun sekmeye göre LLM, çalıştırılabilir SQL sorgusu üretir.
- ▶ **Streamlit Arayüzü:** Kullanıcıdan doğal dilde sorgu alır, sonucu anlık olarak arayüzde gösterir.

Sonuçlar – Doğruluk ve Performans

- ▶ LLM destekli SQL üretimi genel olarak kullanıcı sorgularına yüksek doğrulukla cevap verebilmektedir.
- ▶ Sekme eşleşmesi SentenceTransformer modeli sayesinde anlamlı bir şekilde gerçekleştirilmiştir.
- ▶ Sistem Streamlit üzerinden gerçek zamanlı etkileşim sağlamaktadır.
- ▶ Ancak sorgu üretiminde bazı özel durumlarda alias uyumsuzlukları veya eksik WHERE filtreleri gözlenmiştir.

Sonuçlar – Geliştirilebilir Yönler

- ▶ Daha hassas sonuçlar için daha güçlü embedding modelleri (örn. BGE, Instructor) kullanılabilir.
- ▶ Prompt engineering süreci kullanıcı sorgularına göre daha dinamik ve esnek hale getirilebilir.
- ▶ SQL üretiminde tablo ilişkileri ve JOIN mantıkları daha derin şekilde modellenebilir.
- ▶ Sorgu hızını artırmak için veritabanı sorgu optimizasyonları eklenebilir.

Sonuçlar – Güçlü Yönler

- ▶ Kullanıcı dostu doğal dil arayüzü ile teknik olmayan kullanıcılar da veri sorgulayabilir.
- ▶ Genişletilebilir yapı: Yeni sekmeler kolaylıkla sisteme entegre edilebilir.
- ▶ Excel'den gelen yapılar otomatik işlenerek esnek veri eşleşmesi sağlanmıştır.
- ▶ Gelişmiş prompt design sayesinde LLM çıktılarında yorum eklenmemesi ve çalıştırılabilirlik korunmuştur.

Sorularınız ?

- ▶ Dinlediğiniz için teşekkürler