

Doğal Dil İşlemeye Giriş - Ödev 3

Doğal Dil İşleme (DDİ) alanındaki ön araştırmalarımız sonucunda, Türkçenin sondan eklemeli yapısının ve zengin morfolojisinin yarattığı "veri seyreltisi" ile "morfolojik belirsizlik" sorunlarını aşmayı hedefleyen bir uygulama geliştirilmesi planlanmaktadır. Bu kapsamda sunulan proje önerisi, finansal piyasalardaki karmaşık metin verilerini anlamlandırabilen ve kullanıcılara anlık piyasa analizi sunan bir Akıllı Yatırımcı Asistanı (Chatbot) projesidir.

Projenin temel motivasyonu, yatırımcıların sosyal medya veya forum dillerinde kullandığı "gürültülü metinlerin" örneğin "yrn borsa ne olr?" standart dile dönüştürülmesi ve bu metinlerden finansal niyet tespiti yapılmasıdır. Proje, sadece bir soru-cevap mekanizması değil, aynı zamanda finansal metinlerden duygu analizi yapabilen bir karar destek sistemi olarak kurgulanmıştır.

Önerilen projenin dilbilimsel ve teknolojik temelleri, aşağıda belirtilen dört temel akademik ve teknik çalışma üzerinden detaylandırılmıştır.

Sistemin ilk ve en kritik aşamasını oluşturan metin ön işleme süreci, Zemberek-python kütüphanesinin sunduğu morfolojik analiz yetenekleri üzerine inşa edilmiştir. Zemberek-python; metni cümlelere ve kelimelere ayırma, kök bulma (lemmatization), yazım denetimi ve metin normalleştirme gibi temel görevlerde Python ekosisteminde yüksek başarımlı sağlamaktadır. Bu kütüphanenin sunduğu modüller sayesinde, kullanıcıdan gelen "hisseler dussede alsak" gibi bozuk metinler önce normalize edilecek, ardından morfolojik analiz ile finansal terimlerin kökleri belirlenerek sistemin ilerleyen aşamalarındaki anlamsal karmaşa minimize edilecektir.

Sistemin anlamsal analiz ve karar verme katmanında ise Schweter (2020) tarafından geliştirilen BERTurk modeli temel alınacaktır. BERT mimarisini temel alan bu model, Türkçe metinler üzerinde önceden eğitilmiş olup metin sınıflandırma ve soru-cevaplama gibi karmaşık görevlerde yüksek başarımlı sunmaktadır. Projede kullanılacak olan BERTurk modeli, kullanıcının niyetini (intent) tespit etmekle kalmayacak, aynı zamanda finansal metinlerin anlamsal bileşenlerini analiz ederek kullanıcının sorusuna en uygun yanıtın üretilmesini sağlayacaktır. Bu teknolojik dayanak, asistanın sadece kelime eşleşmesi yapmasının ötesine geçerek gerçek bir doğal dil anlayışı kazanmasına olanak tanıyacaktır.

Finansal alana özgü duygu analizi ve risk değerlendirmesi süreçlerinde ise "Customer Credit Risk Scoring Using NLP: A News Analysis Approach (2025)" başlıklı güncel çalışma referans alınmıştır. Bu araştırma, Türkçe finansal haber metinlerinde BERTurk tabanlı modeller kullanarak otomatik duygu analizi ve varlık ismi tanıma (NER) işlemlerinin nasıl gerçekleştirileceğini deneysel verilerle ortaya koymaktadır. Geliştirilecek Akıllı Yatırımcı Asistanı, bu çalışmadaki metodolojiyi takip ederek haber kaynaklarından gelen verileri pozitif, negatif ve nötr olarak sınıflandıracak ve yatırımcıya proaktif risk uyarıları sunacaktır.

Sistemin bütünsel mimarisi ve kullanıcı etkileşim modeli için AVA (Financial Service Chatbot) çalışması vizyoner bir dayanak olarak seçilmiştir. AVA projesi, derin öğrenme tabanlı transformatör modellerinin finansal hizmetlerde nasıl bir diyalog yönetimiyle kullanılabileceğini teknik bir çerçeveye sunmaktadır. Bu çalışma, bizim asistanımızda kullanıcıya sunulacak "kişiselleştirilmiş yanıt üretimi" ve "bağlamsal diyalog akışı" süreçlerine akademik bir rehberlik sağlayacaktır.

Bu dört kaynağın bütünleştirilmesiyle, Türkçenin morfolojik engellerini Zemberek-python ile aşan ve BERTurk ile finansal anlamlandırma yapabilen kapsamlı bir chatbot mimarisi oluşturulacaktır.

Uygulama Aşamaları

2-3 haftalık süreçte gerçekleştirilecek uygulama aşaması şu temel adımlardan oluşacaktır:

1. **Veri Temizleme:** Zemberek-python kullanılarak kullanıcı sorgularındaki yazım hatalarının düzeltilmesi ve kök bulma işlemleri.
2. **Niyet Tanıma:** Hazır BERTurk modelleri kullanılarak kullanıcının sorusunun kategorize edilmesi.
3. **Cevap Havuzu:** Tespit edilen niyet ve duygu tonuna göre (AVA çalışmasında önerildiği gibi) uygun cevabın kullanıcıya iletilmesi.

Bu araştırma raporu, Türkçeye özgü yapısal zorlukların akademik yaklaşımlarla nasıl çözülebileceğini ve bu çözümlerin finansal teknolojiler alanında nasıl uygulanabileceğini detaylandırmaktadır. Sonuç olarak önerilen sistem, hem akademik literatürdeki morfolojik analiz yöntemlerini hem de modern derin öğrenme modellerini bir araya getirerek özgün bir DDİ uygulaması meydana getirecektir.

Eren Atasoy 2303013123

Kaynakça

1. **Zemberek-Python:** Python library for Turkish Natural Language Processing. <https://pypi.org/project/zemberek-python/>
2. **Schweter, S. (2020).** *BERTurk - A BERT model for Turkish (dbmdz/bert-base-turkish-cased)*. Hugging Face. <https://huggingface.co/dbmdz/bert-base-turkish-cased>
3. **DergiPark (2025).** *Customer Credit Risk Scoring Using Natural Language Processing (NLP): A News Analysis Approach*. DOI: [10.47933/ijeir.1814441](https://doi.org/10.47933/ijeir.1814441)
4. **Yu, S., et al. (2021).** *AVA: A Financial Service Chatbot Based on Deep Bidirectional Transformers*. Frontiers in Applied Mathematics and Statistics. DOI: [10.3389/fams.2021.604842](https://doi.org/10.3389/fams.2021.604842)

