

TOBB ETÜ Bilgisayar Mühendisliği Bölümü
BİL 443/564 Dönem Projesi

2025-26 Bahar Dönemi

Öğrenci Adı Soyadı: Oruç ÇAKIR

Date: 22/01/2026

1. Proje ismi

Çoklu Ajan Destekli Açıklanabilir Yapay Zeka Destekli Metin Sınıflandırma Sistemi
(Multi-Agent Explainable AI-Enhanced Text Classification System)

2. Proje tanımı

Bu projede, metin verilerinin otomatik olarak kategorilere sınıflandırılması hedeflenmektedir. Toplamda aşağıda belirtilen 6 farklı sınıflandırma yöntemi kullanılması hedeflenmektedir.

- 1) Naive Bayes
- 2) Support Vector Machines
- 3) Random Forest
- 4) K-Nearest Neighbors
- 5) Logistic Regression
- 6) Transformer Models

Farklı datasetler uygulanacak ve tüm yöntemler birbiriyle karşılaştırılacaktır. Bu yaklaşım sayesinde hangi modelin hangi veri tipinde binary/multi-class ve hangi dilde Türkçe/İngilizce daha başarılı olduğu analiz edilecektir.

Projenin özgün katkısı olarak 3 ajandan oluşan çoklu ajan sistemi geliştirilecektir.

1. Niyet Sınıflandırıcı Ajanı (Intent Classifier Agent) : Kullanıcıdan gelen metni analiz ederek hangi modelin kullanılacağına otomatik karar verilir.

2. Sınıflandırıcı Ajanı: Niyet sınıflandırıcı ajanın kullanıcının girdisine göre seçtiği 6 farklı yöntemden biri olan sınıflandırıcı model sınıflandırma işlemini gerçekleştirir.

3 Açıklanabilir Yapay Zeka Ajanı: Sınıflandırma kararlarını LIME/SHAP ve LLM kullanarak insan dostu şekilde açıklar.

3. Proje gereksinimleri ve projeden beklenen sonuç

Girdiler:

Veri Seti

Örnek

IMDB	This movie was absolutely fantastic.
Turkish Sentiment	Bu ürün gerçekten harika, kesinlikle tavsiye ederim.
AG News	Apple announces new iPhone with revolutionary AI features.
Turkish News SPOR	Galatasaray Fenerbahçe derbisini kazandı.

Çıktılar:

Veri Seti

Tahmin

Güven

XAI Açıklaması

IMDB	POSITIVE	%94	fantastic ve great kelimeleri olumlu anla taşıyor.
Turkish Sentiment	POZİTİF	%91	harika ve tavsiye kelimeleri memnuniyet gösteriyor.
AG News	TECHNOLOGY	%88	iPhone kelimesi teknoloji kategorisine işaret ediyor.
Turkish News	SPOR	%89	derbi kelimesi spor kategorisine işaret ediyor.

Çalışma Düzenegi:

Önerilen sistem, ham metin verisinin işlenmesinden son kullanıcıya açıklanabilir bir çıktı sunulmasına kadar olan süreci, iş birliği içinde çalışan üç özelleşmiş ajan üzerinden yönetmektedir. Sistemin teknik akışı ve modüllerin detayları aşağıda verilmiştir:

1. Niyet Tespiti ve Yönlendirme (Intent Classifier) Sistemin giriş kapısı olan bu ajan, kullanıcıdan gelen ham metni karşılar. Metnin genel bağlamını (film yorumu, ürün değerlendirmesi veya haber metni) analiz ederek, veriyi ilgili veri seti üzerinde uzmanlaşmış modele yönlendirir. Bu adım, sistemi bağlam farkındalıklı çalışmasını sağlar.

2. Sınıflandırma Uzmanı (Classification Agent) Sınıflandırma işleminin gerçekleştirildiği ana modüldür. Bu ajan, kendi içinde 3 alt aşamadan oluşan bir işlem hattı çalıştırır:

- **A. Veri Ön İşleme (Preprocessing):** Ham veri gürültüden arındırılır. Metinler küçük harfe dönüştürülür, noktalama işaretleri temizlenir. Anlamsal değeri olmayan bağlaçlar ve sık kullanılan kelimeler (stopwords) filtrelenerek veri normalize edilir.
- **B. Öznitelik Çıkarımı (Feature Extraction):** Metinler sayısal vektörlere dönüştürülür. Burada iki yöntem kıyaslamalı olarak kullanılır:
 - TF-IDF: Kelime sıklığına dayalı istatistiksel temsil.
 - Transformer Embeddings (Sentence-BERT): Bağlam ve anlamsal ilişkiyi yakalayan yoğun (dense) vektörler.
- **C. Sınıflandırma (Classification):** Oluşturulan vektörler kullanılarak tahmin yapılır. Geleneksel (Naive Bayes, SVM, RF, KNN, LR) ve modern Transformer modelleri çalıştırılarak sınıf etiketi ve güven skoru üretilir.

3. Açıklanabilir Yapay Zeka (XAI Agent) Sistemi bir kara kutu olmaktan çıkaran bu ajan, Agent 2'den gelen tahmin sonucunu alır ve iki aşamalı bir analiz uygular:

- **A. LIME ve SHAP Analizi:** Modelin verdiği kararda hangi kelimelerin ne kadar etkili olduğu matematiksel olarak hesaplanır.
- **B. LLM Tabanlı Açıklama Üretimi:** Hesaplanan sayısal etki değerleri, Büyük Dil Modeliaracılığıyla işlenir. Son kullanıcıya sadece etiket değil, doğal dilde bir neden-sonuç ilişkisi sunulur.

Çıktı (Output): Süreç sonunda kullanıcıya sınıf etiketi, güven skoru ve XAI modülü tarafından üretilen yorumlu açıklama raporlanır.

4. Projenin basit yol haritası (belirlenen tarihlere göre öngörülen plan)

Dönem	Tarih	Hedef ve Görevler
Hafta 1-2	23 Ocak - 5 Şubat	Veri seti seçimi ve ön işleme adımları
Hafta 3-4	6 - 19 Şubat	TF-IDF implementasyonu ve model eğitimi
Hafta 5-6	20 Şubat - 1 Mart	Transformer embeddings implementasyonu
Durum Raporu	2 Mart 2026	4-5 sayfalık ilerleme raporu teslimi
Hafta 7-8	3 - 16 Mart	Niyet Tespiti ajani entegrasyonu
Hafta 9-10	17 - 30 Mart	LLM Agent entegrasyonu ve demo hazırlığı
Hafta 11	31 Mart - 7 Nisan	K-fold validation, analizler ve IEEE rapor yazımı
Final Teslimatı	8 Nisan 2026	IEEE formatında final rapor ve demo sunumu

5. Projede kullandığınız varsayımlar.

Veri seti yeterli büyüklükte olacak (10.000+ örnek)
Sınıflar arasında aşırı dengesizlik olmayacak
Metinler okunabilir ve temiz formatta olacak
Programlama Dili: Python 3.x
Kütüphaneler: scikit-learn, sentence-transformers, LIME, SHAP, LangChain, huggingface.
Donanım: Intel I7 CPU, 32Gb RAM. NVIDIA GeForce RTX 4060 8GB VRAM.

Niyet Sınıflandırma (Intent Classifier) Veri Kısıtı ve Çözüm Stratejisi

- Veri Seti Erişilebilirliği Kısıtı:** Proje kapsamında geliştirilecek olan Niyet Sınıflandırıcı Agent için, doğrudan bu amaca hizmet eden etiketli ve temiz bir veri seti bulmanın zorluğu öngörülmektedir.
- Aşamalı Model Stratejisi:** Bu veri kısıtını aşmak için iki aşamalı bir strateji benimsenmiştir:

Başlangıç: İlk aşamada eğitim verisine ihtiyaç duymayan Zero-Shot Learning yeteneği sayesinde bir llm kullanılacaktır.

Optimizasyon: Sistem çalıştıkça biriken veriler veya LLM yardımıyla etiketlenen veriler kullanılarak, maliyeti düşürmek ve hızı artırmak amacıyla raporda belirtilen 6 denetimli öğrenme yönteminden en uygun olanı eğitilecek ve ilerleyen süreçte LLM'in yerini alacaktır.

6. Diğer proje önerileri

Karakter tanıma (OCR), bir yazılı metnin hangi dilde yazılmış olduğunun tespiti
Yüz ifadesi tespiti
Hareket ve hareketli nesne tanıma
Otomatik görsel denetleme
Optimum proje çizelgesi veya ders programı oluşturmak