Patika.dev & NewMind AI Bootcamp Bitirme Projesi

Business Reviews Analysis and Conclusion Generation

## **Hedef:**

Bu case'in amacı, çeşitli hizmetler ve ürünleriyle ilgili kullanıcı yorumlarını analiz etmek, konuları ve görüşleri (positive, neutral, negative) sınıflandırmak ve sonrasında bu yorumlardan anlamlı sonuçlar (conclusions) üretmektir. Bu işlemi, sınıflandırıcı ve/veya cümle benzerliği (sentence similarity) modelleri kullanarak yapacağız. Sonuçları ise LLM (Large Language Model) kullanarak oluşturacağız.

## **Girdi Dosyaları:**

**1. Reviews (yelp\_data.json)**

Bu dosya, her bir firma hakkında kullanıcıların yaptığı yorumları ve bu yorumların pozisyonlarını içerir.

* review\_id: Her yoruma özgü benzersiz kimlik numarası.
* business\_id: Yorumun yapıldığı restoranın ID’si (bu kolon, ilgili hizmetle bağlantı kurar).
* review\_text: Kullanıcı yorumunun metni. Bu, hizmet hakkında olumlu, olumsuz veya nötr bir geri bildirim olabilir.
* roberta\_scores: Yorumun içeriğiyle ilgili pozisyonu belirler. Bu pozisyonlar şu şekilde olabilir:
  + Positive: Yorumun olumluluğu.
  + Neutral: Yorumun netice bildirmemesi.
  + Negative: Yorumun olumsuzluğu.

## **Adımlar:**

**1. Konular ve Görüşlerin Sınıflandırılması**

Yorumlar, belirli bir pozisyonla (positive, neutral, negative) sınıflandırılmıştır. Ancak, bu yorumları daha anlamlı hale getirmek için aşağıdaki işlemleri gerçekleştireceğiz:

* Positive: Yorumun olumluluğu.
* Neutral: Yorumun netice bildirmemesi.
* Negative: Yorumun olumsuzluğu.

Bu adımda, yorumları doğru bir şekilde sınıflandırmak ve uygun pozisyonlarda gruplayabilmek için bir **sınıflandırıcı** veya **cümle benzerliği** modeli kullanabilirsiniz.

**ÖNEMLİ**: Kümeleme ve sınıflandırma algoritmaları accuracy, precision, recall, f1 gibi metrikler ile karşılaştırılmalıdır.

**2. Sonuç (Conclusion) Üretimi**

Sonuç üretirken, LLM (Large Language Model) kullanarak tüm görüşler (positive, neutral, negative) birleştirilir ve anlamlı bir sonuç oluşturulur. LLM, bu görüşleri analiz ederek, doğruluğu yüksek bir sonuç çıkarmalıdır.

LLM dışında başka doğal dil işleme (NLP) modelleri de kullanılabilir. Özellikle BERT ya da GPT-3/4 gibi modellerin de çözüm sürecine dahil edilmesi faydalı olabilir. Bununla birlikte, kullanılan sınıflandırıcıların doğruluğu (örneğin, doğruluk oranı, F1 skoru vs.) ve sonuç üretiminde açıklık gereklidir.

**ÖNEMLİ**: Kaggle veya Colab gibi platformlarda kullanılabilecek büyüklükte open source LLM’ler ile inference yapılabilmeli ve üretilen metinler Rouge, Bleu gibi metrikler ile karşılaştırabilmelidir. GPT 3/4 yerine küçük LLM kullanımı daha iyi olacaktır. Uygun ve iyi cevap veren 0.5 B/1B/3B tarzı LLM modellerinin araştırılması ve implementasyonun yapılması, prompt engineering, uygun model seçimi yetenekleriniz değerlendirilecektir.

## **Entegrasyon:**

Model ve uygulama arasındaki entegrasyon, event-driven mimaride yapılmaktadır. Platforma yapılan her yeni yorum, bir "event" mesajı olarak modele iletilir ve bu yorumlar analiz edilip sonuçlar başka bir "event" mesajı olarak geri gönderilir. Model, HealthPulse API'si üzerinden çalışarak bu entegrasyonu sağlar.

## **Sonuçların Üretimi:**

Model, konularla ilgili tüm görüşleri (positive, neutral, negative) birleştirerek anlamlı bir sonuç oluşturacaktır. Aşağıda örnek bir conclusion verilmektedir:

**Örnek Conclusion:**

* Sonuç: "Yorumlar, ... firmasının oldukça iyi olduğunu gösteriyor. Ancak bazı kullanıcılar, oradaki deneyimlerinin hızlı ve profesyonel olduğunu belirtirken, diğerleri bu hizmetin daha iyi olabileceğini savunuyor. Yine de, genel olarak ... firması önerilmektedir."

## **Değerlendirme:**

Bu case, firma hizmetlerinin farklı yönlerini anlamak için hem sınıflandırma hem de anlamlı metin işleme tekniklerinin birleşimini kullanır. Yorumların doğru bir şekilde sınıflandırılması ve sonuçların mantıklı bir şekilde üretilmesi, firma hizmetlerinin kalitesini ve kullanıcı algısını anlamak adına kritik önem taşır. Probleme nasıl yaklaştığınız, düşünme süreciniz ve sunabileceğiniz çözümler ile ilgili puanlama yapılacaktır.