

KOCAELİ ÜNİVERSİTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ
YAZILIM LAB. II
PROJE 3

Proje İlan Tarihi: 26/04/2024

Proje Teslim Tarihi: 17/05/2024

Akademik Makale Öneri Sistemi

Proje, kullanıcıların ilgi alanlarına ve okuma geçmişlerine uygun olarak kişiselleştirilmiş makale önerileri sunan bir sistemdir. Kullanıcılara ilgi duydukları doğru makaleleri yönlendirmek projenin ana hedefidir.

Proje İsterleri

1. Veri Kaynağı

- Hazır makale veri kümesi kullanılacaktır. Örnek hazır veri kümesi: Inspec ve Krapivin2009

2. Ön İşleme ve Veri Analizi

- NLP (Doğal Dil İşleme) teknikleri ile makalelerin ön işleme.
- NLP araçları ile metin ön işleme: Python ile NLTK veya spaCy kütüphanelerinden biri kullanılarak yapılacaktır.
- Yapılacak ön işleme yöntemleri: İngilizce stopwords (etkisiz kelimeler) temizlenmesi (bunun için hazır stop words listeleri kullanılacaktır), noktalama işaretleri çıkarılması, kelime köklerinin bulunması.
- FastText ve SCIBERT modelleri kullanılarak makaleler ve kullanıcı profili için vektör temsillerinin (vector embeddings) oluşturulması.
- Kullanıcı profili için vektör temsili, bir kullanıcının tüm ilgi alanlarının vektör temsillerinin ortalaması alınarak oluşturulmalıdır.
- İlgi alanları vektörleri benzerlik yönünden değerlendirilerek, FastText ve SCIBERT vektör temsilleri için ayrı ayrı 5 öneri sunulması beklenmektedir.
- Makale-kullanıcı benzerliği Cosine Similarity metriği ile hesaplanacaktır.
- Yapılan önerilerin performans değerlendirmesi: Precision ve Recall değerleri hesaplanarak yapılmalıdır. Değerlendirme sonuçları arayüzde gösterilmelidir.

3. Kullanıcı Profil Oluşturma ve Yönetimi

- Kullanıcı Girişi: Kullanıcıların sisteme kayıt olabilmesi için bir arayüz tasarlanmalıdır.

- Arayüzde öncelikle sisteme giriş için üyelik işlemi yapılmalıdır.
- Üyelik formu aracılığıyla kullanıcı demografik bilgilerinin yanı sıra kullanıcının akademik ilgi alanları alınmalıdır.
- Profil Yönetimi: Kullanıcı profillerinin oluşturulması, güncellenmesi ve yönetilmesi.

4. Öneri Motoru

- Başlangıç Önerileri: Kullanıcının üyelik oluştururken belirttiği ilgi alanlarına göre oluşturulmalıdır. Daha sonraki öneriler, kullanıcının ilgi alanı ve okuma geçmişi göz önünde bulundurularak dinamik olarak güncellenmelidir.
- Algoritma Geliştirme: Kullanıcı profilleri ve okuma alışkanlıklarına dayanarak önerilerde bulunacak algoritmalar geliştirilmelidir.
- Kişiselleştirilmiş Öneriler: Kullanıcının geçmiş etkileşimlerine göre önerilerin dinamik olarak güncellenmelidir.
- Geri Bildirim Sistemi: Kullanıcı geri bildirimlerinin toplanması ve modelin buna göre güncellenmesi: Kullanıcıya FastText ve SCIBERT ile sunulan 5'er öneriden her birine uygun veya değil olarak şeklinde sisteme geri dönüş yapması beklenmektedir.
- Bir sonraki öneri kullanıcının ilgi alanına uygun olarak işaretlediği makale/makaleler de dahil edilerek belirlenmelidir.

5. Kullanıcı Arayüzü

- Frontend Geliştirme: Kullanıcı dostu bir web arayüzü için dil kısıtı bulunmamaktadır.
- Arama Fonksiyonları: Kullanıcılara aradıkları makaleleri kolayca bulabilecekleri arama çubuğu ve filtreleme seçenekleri oluşturulmalıdır.

6. Kullanılacak Teknolojiler

- Uygulama bir web projesi olacağından verilerin senkronizasyonu ve algoritmaların çalışabilmesi için uygun bir veri tabanı kullanmanız gerekmektedir. Örnek veri tabanı MongoDB.
- Yapay Zeka ve Makine Öğrenmesi: Python

Ödev Teslimi

- Proje raporu IEEE formatında (önceki yıllarda verilen formatta) 4 sayfa uzunluğunda olmalıdır. Rapor; akış diyagramı veya yalancı kod içermeli, özet, giriş, yöntem, deneysel sonuçlar, sonuç, kaynakça ve ER (veri tabanı için) diyagramından oluşmalıdır.
- Dersin takibi projenin teslimi dâhil edestek2.kocaeli.edu.tr sistemi üzerinden yapılacaktır. edestek2.kocaeli.edu.tr sitesinde belirtilen tarihten sonra teslim edilen projeler kabul edilmeyecektir.
- Proje ile ilgili sorular edestek2.kocaeli.edu.tr sitesindeki forum üzerinden **Arş. Gör. Kübra Erat** veya **Arş. Gör. Dilara Görmez Açık**'a sorulabilir. Proje teslimine 2 gün kala sorulan hiçbir soruya cevap verilmeyecektir.

- Sunum tarihleri daha sonra duyurulacaktır.
- Sunum sırasında;
 - Algoritma, geliřtirdiđiniz kodun eřitli kısımlarının ne amaçla yazıldıđı ve geliřtirme ortamı hakkında sorular sorulabilir.
 - Kullandıđınız herhangi bir satır kodu açıklamanız istenebilir.
- Projenin tanıtım toplantısı detayları e-destek üzerinden duyurulacaktır.
- Proje grupları **en fazla 2 kiřiden** oluřmalıdır. Proje grup bilgileri e-destekte paylařılacak link üzerinden en ge **30 Nisan Cuma** gnne kadar girilmelidir. Bu tarihten sonra gruplarda herhangi bir deđiřiklik yapılmayacaktır.