

KOCAELİ ÜNİVERSİTESİ MÜHENDİSLİK FAKÜLTESİ

YAZILIM LABARATUVARI 1.PROJE

190202040

Muhammed Koç

190202064

Osman Şafak

1-)ÖZET:

Projede bizden belirli formattaki tez pdflerden (bitirme projesi araştırma problemleri) veri çekerek anlamlı bazı belirli verileri veri tabanına kaydedilmesi ve bunların bir web sitesinde sunulması istenmiştir. Web sitesinde admin ve kullanıcı bölümü olacaktır.

Admin her yetkiye sahip kullanıcı ise sadece yüklediği pdfleri görüntüleme hakkına sahiptir. Kullanılan PDF için 2 şer sorgu oluşturulmuştur

Sorgu 1: Elde edilen bilgilerden yazara, derse, proje adına, anahtar kelimelere ve döneme göre ayrı ayrı sorgular

Sorgu 2: Belirli dönemde bir kullanıcının bir derste yüklediği projeler (Örn. 2018-2019 Bahar dönemi A. A. kullanıcısının Bitirme dersi için yüklediği projeler)

Proje dil sınırlaması yoktur.

Projede amaç web sitesinin kullanılmasını ve oluşturulmasının öğrenilmesi ve metin madenciliği hakkında tecrübe sahibi olunmasıdır.

2-)GİRİŞ:

Biz projede python ve django tika kütüphanesini metin madenciliği için kullandık. İDE olarak Visual Studio Code Jupyter Notebook faydalandık python kullanmamızın sebebi pratik ve hızlı olması, Django ise kapsamlı ve kuvvetli bir framework olması tika ise Türkçe karakter desteklediği içindir.

Pdfleri farklı kişi sayıları için uzun başlık kısa başlık fazla anahtar kelimesi az anahtar kelimesi opsiyonlarına hata vermeden çalışacak şekilde ayarlamalar yaptık. Ayrıca Türkçe diline göre projeyi hazırladık, bazı kütüphaneler(pypdf2) Türkçe karakterleri desteklemediği için tika kütüphanesinde kullandık.

3-)YÖNTEM:

Verileri TİKA kütüphanesini kullanarak ham metin olarak elde ettikten sonra fazladan oluşan satır atlamak için kullanılan \n ifadelerini temizledikten sonra daha kullanılabilir bir metin elde ettik. Bunun ardından Bazı yerlerde satır sayısına göre bazı yerlerde ise her zaman sabit olan anahtar kelimelere göre veri çekme işlemlerini yaptık her zaman değişme ihtimali olan bazı parametreler için fazladan veya eksik veri gelme ihtimalini göze aldık Buna göre sorgulamalar yaparak pdf'e bağlı olarak bunları kontrol altına aldık ve verileri çektikten sonra mysql kullanarak bunları proje tablosunu aktardık.

Projede danışmanın ve jürilerin ünvanlarını veri çekme esnasında ayırsak da veri tabanında tablolarda çok fazla yer kapladığı için tek bir sütunda bulundurduk.

Web sitesini oluştururken back-end olarak python Django framework'ünü kullandık. Kullanıcıların kaydedilmesi için gereken tablolar Django'da hazır olarak bulunuyordu. Biz de kullanıcı ile alakalı işlemlerde bu tablolardan faydalandık.

Tezlerle alakalı her türlü işlemleri yapabilmek adına tezler isminde bir app oluşturduk. Bu app'in içinde tez isminde bir model oluşturarak veritabanına tezler_tez tablosunu oluşturmasını söylemiş olduk.

Sayfaların görüntülenmesinde MTV yani model,template,view ilişkisini kullandık. Model.py dosyası içinde tablolarımızı oluştururken , view.py de bu modelleri ve templates klasöründe ki html dosyalarını kullanarak sayfanın görünümünü belirledik. Urls.py dosyasında sayfanın hangi uzantıda açılacağını belirledik. Templates klasöründe de html kodlarımızı depoladık.

Kullanıcı veritabanına tez eklemek istediğinde öncelikle dosyayı seçtirdik ve hem dosyayı hem de kullanıcının adını veritabanına kaydettik. Sonrasında kullanıcıyı yönlendirdiğimiz sayfada metin işleme yapılmasının onaylanmasıyla tez üzerinden gerekli bilgileri aldık ve veritabanında yüklenen tezin bulunduğu sütunu güncelledik.

Bizden istenen sorgu1 ve sorgu2'yi hem admin paneline hem de kullanıcı paneline ekledik. Kullanıcılar sadece kendi yükledikleri tabloları görüntüleyebilirken , admin panelinde her tabloya erişim sağlanabilecek şekilde ayarlandı.

4-)DENEYSEL SONUÇLAR:

-----Başlık-----
yarı iletkenler belirgin elektriksel özellikleri
-----isim1-----
halil sari
-----isim2-----
mehmet dal
-----proje türü-----
bitirme projesi
-----hocalar-----
-----danışman-----
prof.dr.
hasan bütün
-----juri1-----
doç.dr.
ahmet yeşim
-----juri2-----
dr. öğr. üyesi
hasan koç
-----tarih-----
2011-2012
güz
-----öğrenci no-----
190201092
1.öğretim
190201091
1.öğretim
-----özet-----
yarı iletkenler belirgin elektriksel özellikleri olan kristal veya amorf katılardır. tipik direnç materyallerinden yüksek bir d
irenç göstermelerine rağmen, dirençleri yalıtkanlar kadar yüksek değildir. metallerin tam tersine dirençleri, sıcaklık arttıkça
azalır. son olarak; istenirse, iletkenlik özellikleri saflıkları “doping” adı verilen yöntem ile bozularak kolaylıkla değiştiril
lebilir. doping yöntemi yarı-iletkenin direncini düşürürken aynı zamanda katkılı yarı-iletkenin farklı düzeyde katkı yapılmış n
oktaları arasında da yarı iletken birleşme noktaları oluşmasına da olanak sağlar. elektronlar, iyonlar ve elektron oyukları gib
i yük taşıyıcıların bu birleşme noktalarındaki davranışları diyotların, transistörlerin ve diğer tüm modern elektronik parçalar
ın temellerini oluşturur.
-----anahtar kelimeler-----
hat arıza,kısa devre,yarı iletkenler,iteratif,biyomedikal,analog hatları,sistem,sesli haberleşme

Giriş												
Ders Ara: bitirme projesi Ara												
BAŞLIK	ÖĞRENCİLER	ÖĞRENCİLERİN NUMARALARI	ÖĞRENCİLERİN ÖĞRETİM TÜRLERİ	DANIŞMAN	JÜRİ1	JÜRİ2	DERS ADI	ANAHTAR KELİMELER	ÖZET	DÖNEM	KULLANICI	TEZ
iletim hatlarında empedans tabanlı arıza yeri tespiti için yeni bir yaklaşım	ali eken	170201025	1.öğretim	prof.dr. nevilhan duru	doç.dr. sevinç ilhan omurca	dr. öğr. üyesi orhan akbulut	bitirme projesi	iletim hatları,mov,pmu,seri kapasitör,arıza yeri bulma algoritmaları	bu çalışmanın amacı, iletim hatlarında arıza yeri tespiti için empedansa dayalı algoritmaları	2018-2019 bahar	k1	/pdf/170201025.pdf
seri kompanze edilmiş iletim hatlarında empedans	hasan şen,ahmet faruk	190201099 ,190201098	1.öğretim,1.öğretim	prof.dr. tulin güneş	doç.dr. sevinç ali omurca	dr. öğr. üyesi orhan koç	bitirme projesi	fiber,yarı iletken,besim,mikro elektromekanik,veri,haber hatları,mühendislik,susuk,manyetik etki,sesli haberleşme	normal durumda yalıtkan olan bu maddeler ısı, ışık, manyetik etki veya elektriksel gerilim gibi dış	2020-2021 bahar	k2	/pdf/SUNUM2.pdf
yarı iletkenler belirgin elektriksel özellikleri	halil sari,mehmet dal	190201092 ,190201091	1.öğretim,1.öğretim	prof.dr. hasan bütün	doç.dr. ahmet yeşim	dr. öğr. üyesi hasan koç	bitirme projesi	hat arıza,kısa devre,yarı iletkenler,iteratif,biyomedikal,analog hatları,sistem,sesli haberleşme	yarı iletkenler belirgin elektriksel özellikleri olan kristal veya amorf katılardır. tipik direnç	2011-2012 güz	k1	/pdf/SUNUM6.pdf

127.0.0.1:8000/admin/sorgu2

Giriş

Dönem : 2020-2021 bahar Kullanıcı : k2 Ders : bitirme projesi Ara

BASLIK	ÖĞRENCİLER	ÖĞRENCİLERİN NUMARALARI	ÖĞRENCİLERİN ÖĞRETİM TÜRLERİ	DANIŞMAN	JÜRİ1	JÜRİ2	DERS ADI	ANAHTAR KELİMELER	ÖZET	DÖNEM	KULLANICI	TEZ
seri kompanze edilmiş iletim hatlarında empedans	hasan gençahmet faruk	190201099 ,190201098	1.öğretim,1.öğretim	prof.dr. tulin güneş	doç.dr. sevinç ali omurca	dr. öğr. üyesi orhan koç	bitirme projesi	fiber,yan iletken,besim,mikro elektromekanik,veri,haber hatları,mühendislik,sıvı,ışık,manyetik etki,sesli haberleşme	normal durumda yalıtılan olan bu maddeler ısı, ışıık, manyetik etki veya elektriksel gerilim gibi dış	2020- 2021 bahar	k2	/pdf/SUNUM2.pdf

Windows taskbar: 0°C, 12:49, 21.12.2021

5-)KARŞILAŞILAN PROBLEMLER:

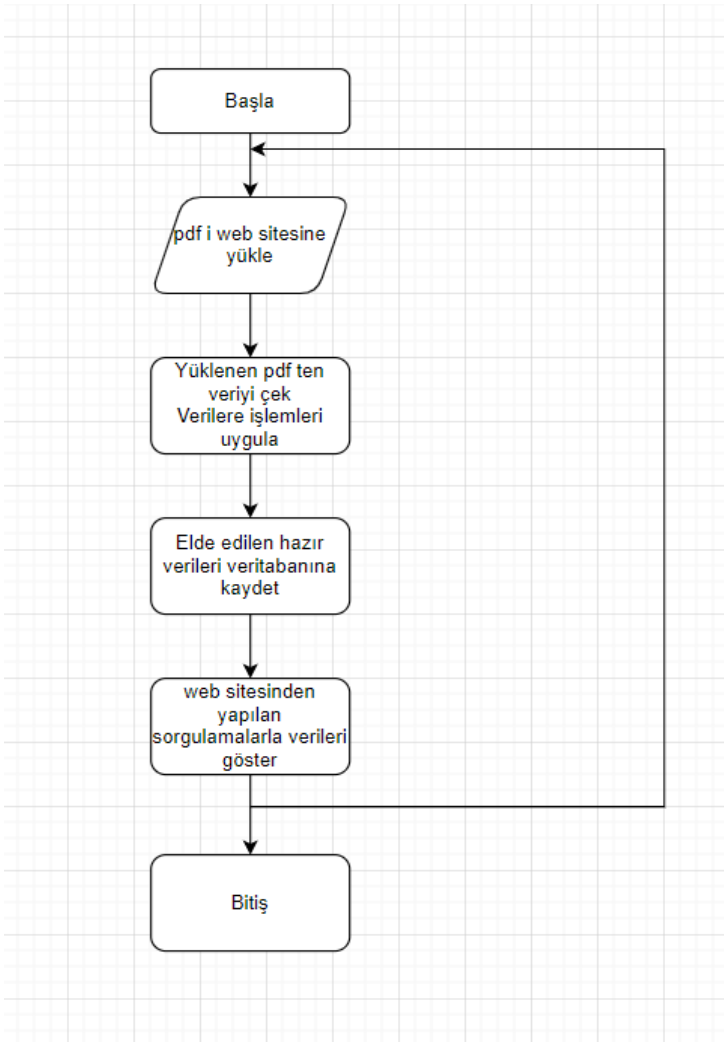
Projede Türkçe karakter zorunlu olduğu için en çok bilinen farklı hediye iki kütüphanesini denemek ama Türkçe karakter sıkıntısından ötürü sorunlar yaşadık bu sorunu çözmek için internette çeşitli aramalar yapsak da herhangi bir yöntem bulamadık bunun üzerine farklı kütüphaneler denedik en son tika'yı kullandık.

Veri çekme esnasında her Pdfin formatı değişebileceği bir noktanın yanına konulan boşluğa göre veri çekme işleminin değişebileceği için çok hassas ve değişken olmasından ötürü zorluklar yaşadık ve bir projede bir, iki veya üç kişi olmasının sayfa sayısını düzenini değiştirmesinden ötürü zorluklar yaşadık.

6-)SONUÇ:

Pythonla djangonun kullanımını öğrendik ve pekiştirdik. Metin madenciliğinin ne olduğu ve nasıl yapıldığı hakkında fikir sahibi olduk. Pdf teki ham metinden veri tabanına veri nasıl aktarılabilir hakkında kendimizi geliştirdik. Daha karmaşık sql sorgulamalarında mesafe kat ettik. Temel düzeyde framework nasıl kullanıldığını öğrendik.

7-)AKIŞ ŞEMASI:



8-)KAYNAKÇA:

-->[How to Make and Use an ERD/EER Diagram in MYSQL Workbench - YouTube](#)

--> [Remove \n From the String in Python | Delft Stack](#)

--> [python — Bir PDF dosyasından metin nasıl çıkarılır? \(it-swarm-tr.com\)](#)

--> [Python ile PDF'den Veri Çıkararak Metin Ön İşleme Teknikleri, Çeviri ve Word Cloud | by Yiğit Şener | Data Runner | Medium](#)

--> [python - how to add new fonts to pyfpdf - Stack Overflow](#)

[1 VİDEODA DJANGO DERSLERİ | PYTHON ile WEB SİTESİ YAPMAK - YouTube](#)