



Eskişehir Osmangazi Üniversitesi 2024-2025 Güz Dönemi

ÜRETİM VERİ TABANI VERİ YÖNETİM SİSTEMİ

VERİ TABANI YÖNETİM SİSTEMLERİ

Kadriye HARMANCI 152120221094 Abdullah Taha AYDIN 152120211055 Elif Suna GEGİN 15212022105 Umay Ece MANTAR 152120221127

İçindekiler

- **03** Projenin Tanıtımı
- **04** Amaç ve Hedef
- **05** Projenin Eski ve Güncel Halinin Karşılaştırılması
- Of Projenin Genel Değerlendirmesi
- 07-08 Veri Yapısı Tasarımı
- 09-10 Kullanılan Teknolojiler
- 11-13 Karmaşık Sorgular
- 14-16 Saklı Yordamlar

- 17-18 Error Handling
- 19-20 Data Cleaning-Categorizing
- 21 LLM Bağlantısı
- 22-23 Arayüz Tasarımı
- 24 Demo Gösterim
- 25-26 Excel Dosyasının Python'a Aktarımı
- 27 Zaman Olsaydı Eklenecekler



Projenin Tanıtımı

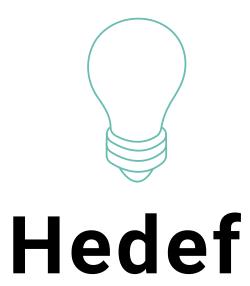
- Projemiz, Ford Otosan'ın üretim bandındaki ürünler ve teknik müdahalelere ilişkin verileri toplayan, düzenleyen ve analiz eden; vektör veritabanı altyapısını baz alan bir uygulamadır.
- Proje, mevcut veri tabanındaki hataları gidermek ve veri yönetim sistemini Vector DB ve Python teknolojileri kullanarak optimize etmeyi hedeflemektedir. PYQT5 ile geliştirilecek bir masaüstü uygulaması olarak tasarlanmış, çalışanlara daha hızlı ve doğru bilgi erişimi sağlayacak bir sistemdir.





Amaç

- Ford Otosan'ın üretim hattındaki verileri analiz ederek chatbot performansını değerlendirmek, mevcut eksiklikleri gidermek ve optimize edilmiş bir veritabanı oluşturmaktır.
- Bu sayede süreçlerdeki verimlilik artırılarak teknik müdahaleler daha hızlı ve doğru şekilde gerçekleştirilecektir.



- Vektör veritabanı ile verilerin doğruluğunu artırarak chatbot performansını iyileştirmek ve karar destek mekanizmalarına daha verimli katkı sağlamaktır.
- Böylece, üretim hattındaki sorunlar daha hızlı tespit edilip çözülerek operasyonel süreçlerin kesintisiz ilerlemesi sağlanacaktır.



Projenin Eski ve Güncel Halinin Karşılaştırılması

- Projede UI için kullanım rahatlığı ve pratikteki kolaylığından dolayı PYQT5 yerine tkinter kütüphanesi kullanılmaya karar verildi.
- ·sentence_transformers kütüphanesi; Tensorflow API, Sckit-learn ve Numpy 'den daha özel bir NLP odaklı çözüm olduğu için tercih edildi.
- ·openpyxl: ChromaDB'de veri saklama Excel yerine DuckDB ve Parquet üzerinden yapılıyor.
- Seaborn ve Matplotlib: Proje veri görselleştirme yerine interaktif sorgu ve yanıtlamaya odaklanmış durumda olduğu için bu kütüphaneler kullanılmıyor.
- ·Arayüzün yönlendirdiği Google hesap açma sayfası ile yetkilendirme üzerinden sisteme erişim sağlanmaktadır.



Projenin Genel Değerlendirmesi

ChromaDB ve Python ile çalışma deneyimi kazanıldı. UI tasarımı, müşteri beklentilerine uygun olarak geliştirildi ve ekip içi iş birliği ile dokümantasyon sağlandı. Ekip, veri yönetimi ve Excel hakkında pratik bilgi edindi.

ChromaDB, Google Gemini API ve UI entegrasyonu test edilip optimize edildi. Google OAuth ve RAG mimarisi doğrulandı. Sistem, performans ve veri tutarlılığı açısından test edilecek ve kullanıcı beklentileri karşılanacaktır.

Proje, verilerin düzenlenmesiyle operasyonel verimliliği artırdı. Veri tabanı, Ford Otosan'ın ihtiyaçlarına uygun hale getirildi ve sistemin etkin kullanımı sağlandı.



Veri Kaynağı

- Excel dosyasındaki aşağıdaki sütunlardan veriler, sistem veri tabanına aktarılacaktır:
- Açıklama: Kısa problem özeti.
- Uzun Açıklama: Problem ve çözümüne dair detaylar.
- Konum: Problemin yaşandığı ekipmanın yeri.
- Ekipman Numarası: Ekipmanın benzersiz kimlik numarası.



Veri Tabanı Yapısı

Sistem, ChromaDB kütüphanesi kullanarak aşağıdaki yapıda bir veri tabanı oluşturur:

Koleksiyon Adı: equipment_issues

Bu koleksiyon, ekipman arızalarını saklamak için kullanılır. Dokümanlar:

• İçerik: Uzun açıklama, konum ve ekipman numarasından oluşan metin. Örneğin:

"İçerik: Motor çalışmıyor, Konum: A Hattı, Ekipman: EMS256"

• Metadata (Ek bilgiler ve filtreleme için):

Konum: Örneğin, "A Hattı".

Ekipman Numarası: Örneğin, "EMS256".

Açıklama: Problem özeti.

ID: Her kayıt için benzersiz bir kimlik oluşturulur.

Format: Satır indeksine dayalı string (str(index)), vektörler oluşturulurken atanır.

Oluşturulan her vektör 384 boyutludur.



Kullanılan Teknolojiler

Flask

- Hafif bir web çerçevesi olup Google OAuth entegrasyonu ve kullanıcı giriş işlemlerini yönetmek için kullanıldı.
- Kullanıcıların tarayıcı üzerinden chatbot'a erişimini sağlar.

ChromaDB

- Vektör veritabanı olarak kullanıldı; sorguların geçmiş veriyle eşleştirilmesini sağlar.
- Hızlı veri erişimi ve arama işlemleri için tercih edildi.

Sentence Transformers

- Metinleri sayısal vektörlere dönüştürerek benzerlik hesaplamalarında kullanıldı.
- Kullanıcı sorgularının ve verilerin anlamlı eşleştirilmesini sağlar.



Kullanılan Teknolojiler

Google Generative Al (Gemini)

- Kullanıcı sorgularına anlamlı ve teknik çözümler üretmek için kullanıldı.
- LLM modeli sayesinde detaylı ve doğru cevaplar oluşturur.

Tkinter

- Basit bir masaüstü arayüzü sağlamak için kullanıldı.
- Kullanıcıların sorgu girmesi ve cevapları görmesi için grafiksel arayüz oluşturur.

Pandas

- Veri temizleme, düzenleme ve Excel dosyalarından veri okuma işlemleri için kullanıldı.
- Verinin işlenip analiz edilmesini kolaylaştırır.



Karmaşık Sorgular

1. ChromaDB Sorguları (Veritabanı Erişimi):

Kullanıcının girdiği sorguya karşılık gelen verileri collection.query ile alıyor.

```
results = collection.query(query_texts=[user_query], n_results=2)
```

Bu sorgu, ChromaDB'de kaydedilmiş belgelere en yakın olanları bulmak için vektör tabanlı arama kullanıyor.

2. Gemini API Kullanımı (LLM Yanıt Üretimi):

generate_LLM_answer fonksiyonu aracılığıyla Google **Gemini API'yi çağırarak, kullanıcının** sorgusu ve bulunan belge içeriğiyle cevap oluşturuyor.

```
response = chat.send_message( prompt + context)
```



Karmaşık Sorgular

3. Google OAuth İşlemleri (Kimlik Doğrulama):

Kullanıcı girişini sağlamak için Google OAuth kullanılıyor, verilerin doğruluğu kontrol ediliyor

```
token = oauth.google.authorize_access_token()
nonce = session.pop('nonce', None)
if not nonce:
    return "Nonce eksik!", 400
user_data = oauth.google.parse_id_token(token, nonce=nonce)
```

4. Vektör Temsili (Sentence Transformer):

Kullanıcıdan alınan sorgunun vektör temsili oluşturuluyor ve veritabanında sorgu için kullanılıyor.

```
query_vector = model.encode(user_query)
```



2.3

Karmaşık Sorgular

5. Çoklu İstemci ve Sunucu Arayüzleri (Flask + Tkinter):

Flask, OAuth kimlik doğrulamasını sağlarke<mark>n, Tkinter üzerinden kullanıcı arayüzü</mark> sağlanıyor. İki bileşen arasında veri akışı s<mark>ağlanıyor</mark>.

6. Threading (Paralel İşlemler):

Flask sunucusu arka planda çalışırken Tkinter uygulaması aktif tutuluyor:

threading.Thread(target=start_flask).start()



Saklı Yordamlar

1. handle_query():

Kullanıcı sorgusunu alır, Sentence Transformers ile vektörleştirir, ChromaDB'de arama yapar ve Gemini API kullanarak cevap üretir. Tüm bu adımları tek bir işlemde gerçekleştirir.

```
f handle query():
 user_query = query_input.get().strip() # Kullanıcı sorgusunu alın ve boşlukları temizleyin
 if not user query:
     update_response("Lütfen bir sorgu girin.")
     return
 update_response("Cevaplaniyor...") # "Cevaplaniyor..." mesajini göster
 root.update() # GUI'yi güncelle
 try:
     query vector = model.encode(user query)
     results = collection.query(query_texts=[user_query], n_results=2)
     if results['documents'][0]:
         context = "\n\n".join(results["documents"][0])
         prompt = f"Sorgu: {user query}"
         gemini_response = generate_LLM_answer(prompt, context, RAG_LLM)
         response = "CHATBOT YANITI:\n" + gemini_response if gemini_response else "Gemini API'den sonuç alınamadı."
         save_to_xml(user_data['email'], query_input.get(), response) # Sorguyu ve yanıtı XML dosyasına kaydedin
     else:
         prompt = f"Sorgu: {user_query}\n\nDaha önce böyle bir sorun yaşanmadı, bunu belirt ve genel bir çözüm öner."
         gemini response = generate LLM answer(prompt, "", RAG LLM)
         response = "ChromaDB'den sonuç bulunamadı.\n" + (gemini_response if gemini_response else "Gemini API'den sonuç alınamadı.
         save to xml(user data['email'], query input.get(), response) # Sorguyu ve yanıtı XML dosyasına kaydedin
 except Exception as e:
     response = f"Hata oluştu: {str(e)}"
 update response(response)
```



Saklı Yordamlar

2- LLM Özelleştirmesi

```
def build chatBot(system instruction):
    model = genai.GenerativeModel('gemini-1.5-flash-latest', system instruction=system instruction)
    chat = model.start chat(history=[])
    return chat
def generate LLM answer(prompt, context, chat):
    response = chat.send_message( prompt + context)
    return response.text
def generateRAG LLM(prompt):
    RAG LLM = build chatBot(system prompt)
    return RAG LLM
system_prompt= """ You are a technical support assistant.
Your primary role is to provide concise, actionable solutions to technical problems based on provided context and historical problem records. Here's how to re
1. You will receive a user query alongside two historical problem records that are similar to the current issue. Analyze the query and the records to provide
2. Based on the information provided, deliver a response in bullet points with clear steps (maximum 5-6 points) on how the problem can be solved.
3. If no historical problem records are provided, or if the query does not match any known issues, respond with: "Bu sorunla daha önce hic karşılaşılmamış.
4. Ensure responses are concise and fully in Turkish.
5. In some cases, you might be asked questions about the chat session itself (e.g., summarizing, listing questions). For these, do not refer to the user query
```



Saklı Yordamlar

3. assign_category():

Veriyi analiz eder, açıklamaların kategori benzerliklerini ölçer ve en uygun kategoriye otomatik olarak atar. Tekrarlanabilir ve sistematik bir sınıflandırma sağlar.

```
def assign_category(row):
    similarities = {cat: cosine_similarity([row['aciklama_vector']], [vec]).item() for cat, vec in category_vectors.items()}
    return max(similarities, key=similarities.get) #
```

4. generate_LLM_answer():

Kullanıcı sorgusu ve bağlamı kullanarak Gemini API üzerinden anlamlı ve teknik çözümler üretir. Tekrarlayan cevap oluşturma işlemini tek bir yapıda yönetir.

```
def generate_LLM_answer(prompt, context, chat):
    response = chat.send_message( prompt + context)
    return response.text
```



Error Handling

Kullanıcı Sorguları ve Hata Kontrolü

Kullanıcıdan gelen sorguların doğru şekilde işlenebilmesi için, geçersiz veya boş sorgulara karşı bir hata kontrolü yapılmıştır. Eğer kullanıcı boş bir sorgu gönderirse, sistem bir hata mesajı verir ve işlem yapılmaz.

```
def handle_query():
    user_query = query_input.get().strip()
    if not user_query:
        update_response("Lütfen bir sorgu girin.")
        return
```



Error Handling

Google OAuth Giriş Hatası (Flask İle Entegre)

Google OAuth entegrasyonu sırasında, kullanıcı kimlik doğrulama ve erişim hataları için hata yönetimi yapılmaktadır. Eğer bir hata oluşursa, hata mesajı kullanıcıya gösterilir.

```
@app.route('/authorize')
def authorize():
    global user_data
    try:
        token = oauth.google.authorize_access_token()
        nonce = session.pop('nonce', None)
        if not nonce:
            return "Nonce eksik!", 400
        user_data = oauth.google.parse_id_token(token, nonce=nonce)
        session['user'] = user_data
        return redirect('/success')
    except Exception as e:
        return f"Bir hata oluştu: {e}", 500
```



Data Cleaning-Categorizing

Veri kategorizasyonu

Mevcut verilerde, açıklamalar sütununda yazım hataları veya benzeri nedenlerle kategori tekrarları bulunmaktadır. Ayrıca, genel bir kategorilendirme sütununun eksikliği, verilerin düzenlenmesini ve analiz edilmesini zorlaştırmaktadır.

- ✓ (Tümünü Seç)
 ✓ atc arızası
 ✓ atc arızası takım değiştirmiyor
 ✓ atc arızası tezgah baslatılamıyor
 ✓ atc kapağı çıktı
 ✓ ATC KAPAĞI KIRIK
 ✓ Atc kapı arızası.
- ✓ (Tümünü Seç)

 ✓ Atc Sorunları

 ✓ Diğer

 ✓ Kapı Sorunları

 ✓ Motor Sorunları

 ✓ Takım Sorunları
- atc arızası takım değiştirmiyor

 atc arızası tezgah baslatılamıyo

 atc kapağı çıktı

 atc kapağı kırık

 atc kapı arızası.

 atc kapı swich arızası

Eski ve hatalı açıklamalar sütunu

Yeni kategori sütunu

Düzenlenmiş açıklamalar



Data Cleaning-Categorizing

Bu işlemde, veri temizleme adımında metinlerdeki yazım hataları düzeltilir, harfler küçültülür ve metinler standart bir forma getirilir. Temizlenen veriler, kategorize etme işlemleri için sonraki adımlarda kullanıma hazır hale getirilir.

```
def clean_text(text):
    text = text.replace("I", "a")
    text = re.sub(r'arizasi', 'arazasi', text, flags=re.IGNORECASE)
    text = text.lower()
    return text
```

```
# Temizleme işlemi

data['Açıklama'] = data['Açıklama'].apply(clean_text)

# SentenceTransformer modeli yükleme

model = SentenceTransformer('all-MiniLM-L6-v2')

# Açıklamaları vektöre dönüştürme

data['aciklama_vector'] = data['Açıklama'].apply(model.encode)
```



3.3.1

LLM Bağlantısı

Gemini (gemini-1.5-flash-latest) adlı LLM modeline API üzerinden erişim sağlıyoruz. Eğer kullanıcının yaşadığı problem daha önce veri tabanında kayıtlıysa, bu veriler veri tabanından alınarak LLM'e gönderilir ve model bu bilgilere dayanarak bir çözüm üretir. Eğer problem veri tabanında bulunmuyorsa, Gemini problemin daha önce yaşanmadığını belirtir ve kullanıcının bir yetkiliyle iletişime geçmesini önerir.

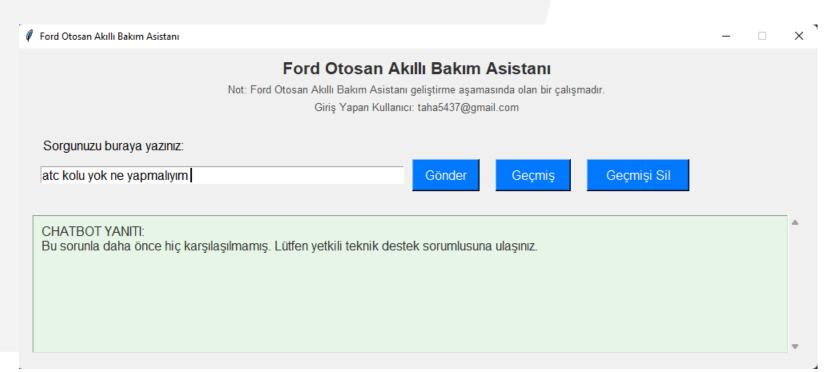
```
try:
    query_vector = model.encode(user_query)
    results = collection.query(query_texts=[user_query], n_results=2)

if results['documents'][0]:
    context = "\n\n".join(results["documents"][0])
    prompt = f"Sorgu: {user_query}"
    gemini_response = generate_LLM_answer(prompt, context, RAG_LLM)

    response = "CHATBOT YANITI:\n" + gemini_response if gemini_response else "Gemini API'den sonuc alinamadi."
    save_to_xml(user_data['email'], query_input.get(), response) # Sorguyu ve yanıtı XML dosyasına kaydedin
    else:
        prompt = f"Sorgu: {user_query}\n\nDaha önce böyle bir sorun yaşanmadı, bunu belirt ve genel bir cözüm öner."
        gemini_response = generate_LLM_answer(prompt, "", RAG_LLM)

        response = "ChromaDB'den sonuc bulunamadı.\n" + (gemini_response if gemini_response else "Gemini API'den sonuc alinamadı.")
        save_to_xml(user_data['email'], query_input.get(), response) # Sorguyu ve yanıtı XML dosyasına kaydedin

except Exception as e:
    response = f"Hata oluştu: {str(e)}"
```

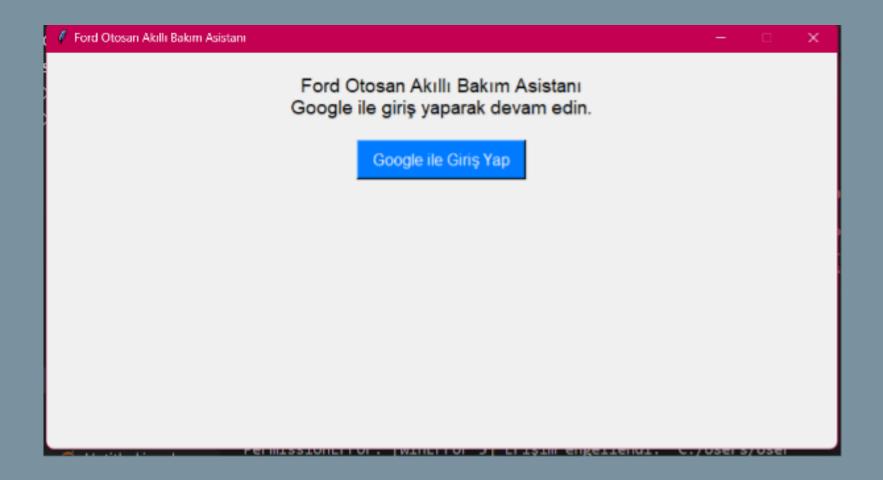




Arayüz Tasarımı

Giriş Ekranı

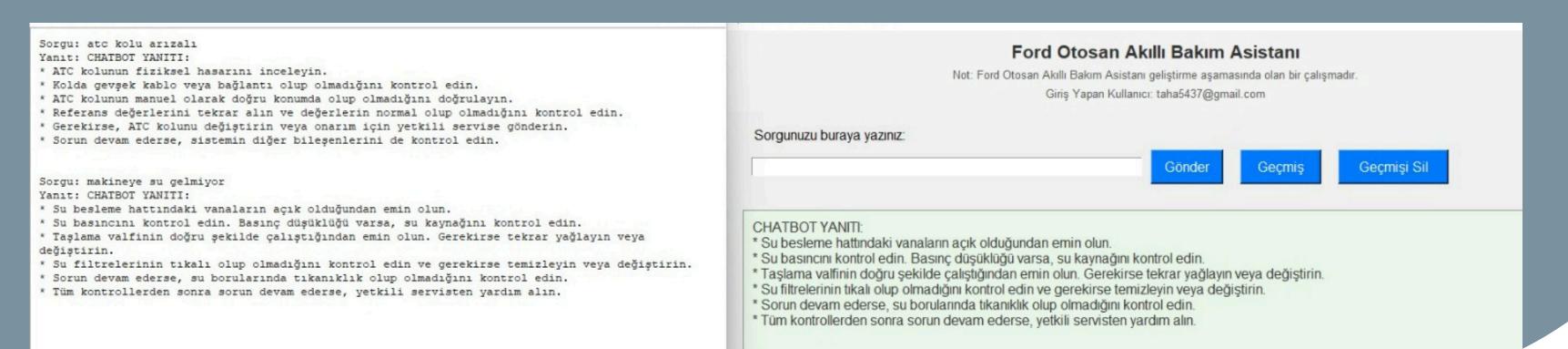
- Google OAuth kullanarak kullanıcı girişini sağlar.
- Kullanıcı giriş yaptıktan sonra chatbot ekranına geçiş yapar.



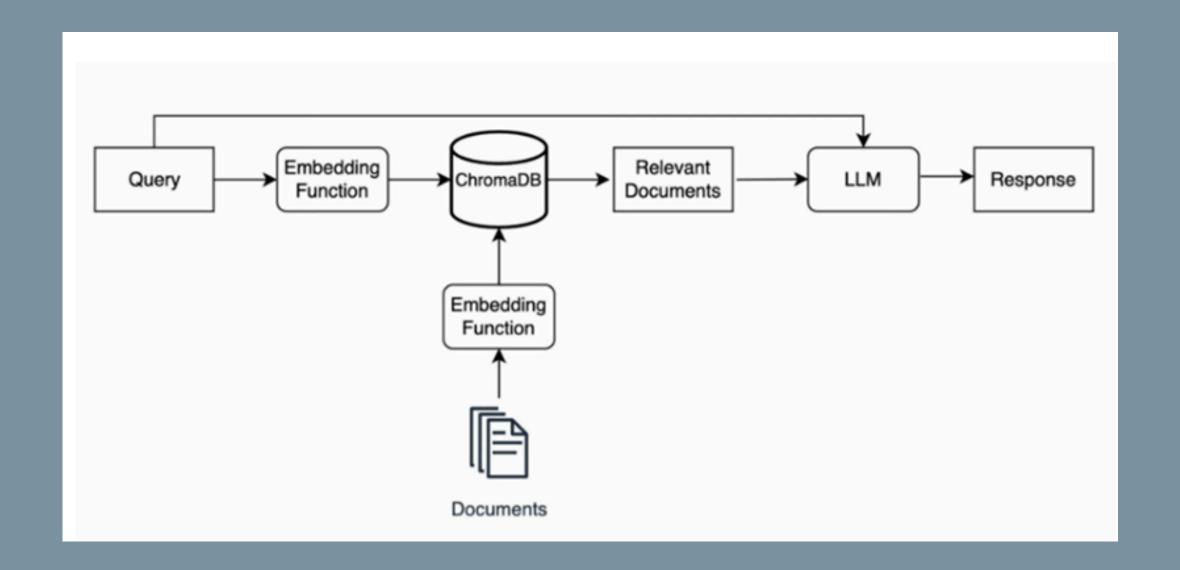
Arayüz Tasarımı

Chatbot Ekrani

- Sorgu Girişi: Kullanıcının yazdığı soruları alır.
- Gönder Butonu: Sorguyu işlemesi için arka uca iletir.
- Cevap Alanı: Gelen yanıtları kullanıcıya gösterir.
- Geçmiş Sorgulara Erişim: Kullanıcılar, daha önce verdikleri sorgulara ve alınan yanıtlara kolayca ulaşabilirler.



DEMO GÖSTERIM



Excel Dosyasını Python'a Aktarma

- Excel dosyasının yolu belirtilir
- Dosya yolu, Excel dosyasının bilgisayardaki tam konumudur.
- Pandas kütüphanesi kullanılır
- Pandas, Excel dosyalarını okumak için read_excel() fonksiyonunu sağlar.

```
import pandas as pd
from sentence transformers import SentenceTransformer
import chromadb
from chromadb.config import Settings
client = chromadb.PersistentClient(path="C:/Users/User/Desktop/Database/chroma_db")
collection_name = "equipment_issues"
collection = client.get_or_create_collection(name=collection_name)
file_path = 'C:/Users/User/Desktop/Deneme/ÖrnekVeri.xlsx' #Verilerin bulunduğu excel dosyasının dizin yolunu gir.
data = pd.read excel(file path)
model = SentenceTransformer('all-MiniLM-L6-v2')
# 7. Vektörlere Dönüştür ve Depola
for index, row in data.iterrows():
    combined_text = f"{row['Uzun Açıklama']}"
    vector = model.encode(combined_text)
    collection.add(
        documents=[combined_text],
       metadatas=[{
            "Konum": row["Konum"],
            "Ekipman Numarası": row["Ekipman Numarası"],
            "Açıklama": row["Açıklama"]
        ids=[str(index)]
print(f"Koleksiyonda toplam {collection.count()} kayıt var.")
```



4.1.4

Diğer Entegrasyonlar (XML, XSLSX)

Kullanıcının chatbot ile olan sohbet geçmişi XML formatında saklanır. Chatbot'un çalıştırıldığı klasörde kaydedilir. Kullanıcı "Geçmişi Sil" butonuna bastığında bu dosya silinerek geçmiş temizlenir.

```
EXPLORER
                                                     ħ taha5437_gmail.com.xml X
DATABASEPROJECT
                                     taha5437_gmail.com.xml
                                       1 <?xml version='1.0' encoding='utf-8'?>
> .dist
                                       2 <history><entry><query>ATC kolu arızalandı </query><response>CHATBOT YANITI:
proje1.py
                                           * ATC kolunun fiziksel durumunu kontrol edin. Hasar veya gevşek bağlantı olup olmadığını inceleyin.
taha5437_gmail.com.xml
                                           * ATC kolunun manuel olarak hareket ettirilebildiğini doğrulay<mark>ı</mark>n. Hareket kı<mark>sıtlılığı</mark> varsa, mekanik bir sorun olabilir.
TahAydn03_outlook.com.xml
                                           * Kapı açık switchinin bağlantısını kontrol edin. Gevşek veya hasarlı bağlantı ATC kolunun yanlış çalışmasına neden olabilir.
                                           * ATC kolunun referans değerlerini kontrol edin. Yanl<mark>ı</mark>ş referans değerleri ar<mark>ı</mark>zaya neden olabilir.  Bu değerleri yeniden ayarlay<mark>ı</mark>n.
                                           * Sorun devam ederse, ilgili kablolar<mark>ı</mark> ve bağlant<mark>ı</mark>lar<mark>ı</mark> inceleyin. Kopuk veya k<mark>ı</mark>sa devre yapm<mark>ı</mark>ş kablolar olabilir.
                                           * Eğer sorun çözülmezse, yetkili servisi arayın.
                                          </response></entry><entry><query>su gelmiyor </query><response>CHATBOT YANITI:
                                      10 * Taşlama vanas<mark>ının açık olduğundan emin olun. Tamamen açı</mark>k konumda olduğundan emin olmak için manuel olarak kontrol edin.
                                      * Vananın elektrik bağlant<mark>ısını</mark> kontrol edin. Geçici bir güç kesintisi veya gevşek bir bağlant<mark>ı</mark> olabilir.
                                      * Aktivatör vananın doğru çalıştığından emin olun. Eğer muadil vana ile denenmişse ve sorun devam ediyorsa, vananın kendisi arızalı olabilir.
                                      * Su giriş hatt<mark>ı</mark>ndaki bas<mark>ıncı</mark> kontrol edin. Su bas<mark>ıncı</mark> yetersizse, su tesisat<mark>ı</mark>nda başka bir sorun olabilir.
                                      14 * Sistemde tıkanıklık olup olmadığını kontrol edin. Tıkalı bir filtre veya boru, su akışını engelleyebilir.
                                           * Yukar<mark>ı</mark>daki adımları denedikten sonra sorun devam ederse, yetkili servisi aray<mark>ı</mark>n.
                                      16 </response></entry><entry><query>makine başlatılamadı </query><response>CHATBOT YANITI:
                                           * Su basınç sensörünün doğru çalıştığından emin olun. Sensörün doğru bir değer okuduğunu doğrulayın.
                                      18 * Su basınç ayarını kontrol edin. Ayarlanan basınç değerinin makine gereksinimlerini karşıladığından emin olun.
                                           * Su beslemesindeki basıncı kontrol edin. Yetersiz su basıncı, makineyi başlatmayı engelleyebilir.
                                           * Işık bariyerinin (uşık perdesinin) doğru hizaland<mark>ığı</mark>ndan ve temiz olduğundan emin olun. Toz veya kir, alg<mark>ı</mark>lamayı engelleyebilir.
                                           * Işık perdesinin ayarlarını kontrol edin. Hassasiyet ayarı doğru şekilde yapılmış mı kontrol edilmeli.
                                           * Yukarıdaki kontrollerden sonra sorun devam ederse, makineyi kontrol eden ana kontrol panelini ve elektrik bağlantıları inceleyin.
                                      25 </response></entry></history>
```



İlerde Yapılması Planlananlar

- Her kullanıcının geçmiş sohbetleri kullanıcının mail adresiyle aynı isimdeki xml dosyasında tutulur. Geçmiş sohbetlere şuan her kullanıcı xml dosyasını açarak ulaşabilir. Bunu şifrelenmiş bir şekilde tutarak diğer kullanıcıların başka kullanıcıların geçmiş sohbetlerinin bulunduğu xml dosyasına erişiminin engellenmesi planlanmaktadır.
- Chatbot'a şuan için sadece google hesabı ile giriş yapılabilmektedir. Bir kayıt oluşturma sistemi yoktur. Chatbot'a bir e-mail adresi veya telefon numarasıyla kayıt olabilme özelliği eklenmesi planlanmaktadır



TEŞEKKÜRLER