

```
In [32]: import pandas as pd
import nltk
from nltk.tokenize import word_tokenize, sent_tokenize
from nltk.corpus import stopwords
import csv
import numpy as np
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
import gensim
from gensim.models import Word2Vec
import matplotlib.pyplot as plt
from collections import Counter
import re
```

```
In [33]: # NLTK kaynaklarını indir
nltk.download('punkt', quiet=True)
nltk.download('stopwords', quiet=True)
```

Out[33]: True

```
In [34]: # Türkçe stop words (örnek liste, Zemberek yerine basit bir liste kullanıldı)
tr_stopwords = set(stopwords.words('turkish')) | {
    'bir', 'bu', 'şu', 've', 'ile', 'de', 'da', 'mi', 'mı', 'mu', 'mü',
    'nın', 'nin', 'nun', 'nün', 'ki', 'de', 'da'
}
```

```
In [35]: # Veri setini yükleme
df = pd.read_csv('C:/Users/ademt/Desktop/tarim_problemleri_veriseti.csv', encoding='utf-8')
texts = df['sorun_metin'].tolist()
print(f"Veri seti boyutu: {len(texts)} döküman, ~1 MB")
print("Verinin ilk 500 karakteri:")
print(texts[0][:500])
```

Veri seti boyutu: 200 döküman, ~1 MB

Verinin ilk 500 karakteri:

Bu yaz sezonunda şeftali tarlamızda yapraklarda kahverengileşme ve büzüşme. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, yem takviyesi yapmalı mıyız?

```
In [36]: # Metin ön işleme fonksiyonu
def preprocess_text(text):
    # Küçük harfe çevir
    text = text.lower()
    # Noktalama işaretlerini kaldır
    text = text.translate(str.maketrans('', '', string.punctuation))
    # Tokenize et
    tokens = word_tokenize(text)
    # Stop words'leri kaldır
    tokens = [token for token in tokens if token not in tr_stopwords]
    # Basit stemming (Türkçe için basit bir yaklaşım, son 3 harfi kes)
    tokens = [token[:-3] if len(token) > 3 else token for token in tokens]
    return ' '.join(tokens)
```

```
In [37]: import string
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.tokenize import word_tokenize # Eğer kullanıyorsan

# Örnek ön işleme fonksiyonu
def preprocess_text(text):
    text = text.lower()
    text = text.translate(str.maketrans('', '', string.punctuation))
    tokens = word_tokenize(text)
    return ' '.join(tokens)

# texts adlı metin listesini tanımladığınızı varsayıyoruz
# texts = ["örnek metin1", "örnek metin2", ...]

preprocessed_texts = [preprocess_text(text) for text in texts]

vectorizer = TfidfVectorizer(max_features=1000)
tfidf_matrix = vectorizer.fit_transform(preprocessed_texts)

tfidf_df = pd.DataFrame(tfidf_matrix.toarray(), columns=vectorizer.get_feature_names_out())
tfidf_df.to_csv('tfidf_matrix.csv', index=False)

print("TF-IDF matrisi 'tfidf_matrix.csv' dosyasına kaydedildi.")
```

TF-IDF matrisi 'tfidf_matrix.csv' dosyasına kaydedildi.

```
In [38]: df['cluster'] = kmeans.labels_
df['solution'] = df['cluster'].map(solution_dict)
```

```
In [39]: df[['id', 'sorun_metin', 'cluster', 'solution']].to_csv('clustered_problems_with_solutions.csv', index=False)
```

```
In [40]: # Zipf analizi
```

```
all_tokens = []
for text in preprocessed_texts:
    all_tokens.extend(word_tokenize(text))
word_counts = Counter(all_tokens)
sorted_words = sorted(word_counts.items(), key=lambda x: x[1], reverse=True)
ranks = range(1, len(sorted_words) + 1)
frequencies = [freq for _, freq in sorted_words]
```

```
In [41]: # Zipf grafiği
plt.figure(figsize=(10, 6))
plt.loglog(ranks, frequencies, marker='.')
plt.title("Zipf Yasası - Kelime Frekansları")
plt.xlabel("Sıra (Log)")
plt.ylabel("Frekans (Log)")
plt.grid(True)
plt.savefig('zipf_plot.png')
plt.close()
print("Zipf grafiği 'zipf_plot.png' dosyasına kaydedildi.")
```

Zipf grafiği 'zipf_plot.png' dosyasına kaydedildi.

```
In [42]: from sklearn.cluster import KMeans

# K-Means kümeleme
num_clusters = 5 # Varsayılan küme sayısı
kmeans = KMeans(n_clusters=num_clusters, random_state=42)
kmeans.fit(tfidf_matrix)

# Sonuçları TF-IDF DataFrame'ine ekleyelim
tfidf_df['cluster'] = kmeans.labels_
```

C:\Users\ademt\anaconda3\Lib\site-packages\sklearn\cluster_kmeans.py:1412: FutureWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly to suppress the warning
super()._check_params_vs_input(X, default_n_init=10)

```
In [43]: # Orijinal metinleri TF-IDF DataFrame'e ekle
tfidf_df['text'] = texts # orijinal metinler

# Kümelendiği metinleri yazdırma
print("\nKümelendiği Sorun Metinleri:")
for cluster_id in range(num_clusters):
    print(f"\nKüme {cluster_id}:")
    cluster_texts = tfidf_df[tfidf_df['cluster'] == cluster_id]['text'].tolist()
    for text in cluster_texts[:5]: # Her kümeden ilk 5 metni göster
        print(f"- {text}")
```

Kümelendirilmiş Sorun Metinleri:

Küme 0:

- mısır tarlamızda genç fidelerde çekiş kaybı yaşanması. aktarma organlarında sorun kaynaklı olabilir mi?
- Son hasatta traktörümüzün çekiş kaybı yaşanması. gübreleme programını yeniden düzenlemeli miyim?
- kabak tarlamızda başakların dolmadan sararması. Tarlanın drenaj sistemi motor gücü yetersizliği. sulama sıklığını değiştirmeli miyim?
- domates tarlamızda aşırı su birikimi nedeniyle kök çürümesi. Tarlanın drenaj sistemi azot eksikliği. sulama sıklığını değiştirmeli miyim?
- marul ağaçlarımızda çekiş kaybı yaşanması. drenaj eksikliği söz konusu olabilir mi?

Küme 1:

- patates tarlamızda genç fidelerde genç sürgünlerde kurumalar ve yaprak dökümü. azot eksikliği kaynaklı olabilir mi?
- Yeni aldığımız ilaçlama makinesi kök bölgesinde yumuşama ve siyahlaşma. aktarma organlarında sorun olabilir mi?
- Yeni aldığımız traktör genç sürgünlerde kurumalar ve yaprak dökümü. kök çürüklüğü olabilir mi?
- Seradaki kabaklarda genç sürgünlerde kurumalar ve yaprak dökümü. Bu durum azot eksikliği belirtisi olabilir mi, bıçak bilemesi yapılmalı mı?
- Seradaki patateslerde meyve dökülmesi aşırı arttı. Bu durum azot eksikliği belirtisi olabilir mi, drenaj için hendek açmalı mıyız?

Küme 2:

- üzüm ağaçlarımızda yapraklarda kahverengileşme ve büzüşme. motor gücü yetersizliği söz konusu olabilir mi?
- Son hasatta traktörümüzün meyve dökülmesi aşırı arttı. toprak analizi yaptırmalı mıyım?
- Yeni aldığımız biçerdöver meyve dökülmesi aşırı arttı. toprakta yoğun su birikimi olabilir mi?
- Seradaki üzümlerde aşırı su birikimi nedeniyle kök çürümesi. Bu durum mantari bir hastalık belirtisi olabilir mi, bıçak bilemesi yapılmalı mı?
- kabak ağaçlarımızda meyve dökülmesi aşırı arttı. motor gücü yetersizliği söz konusu olabilir mi?

Küme 3:

- Bu yaz kurak geçtiği için sığırların genç sürgünlerde kurumalar ve yaprak dökümü. gübreleme programını yeniden düzenlemeli miyim?
- Bu yaz kurak geçtiği için sığırların genç sürgünlerde kurumalar ve yaprak dökümü. bıçak bilemesi yapılmalı mı?
- Bu yaz kurak geçtiği için sığırların kök bölgesinde yumuşama ve siyahlaşma. acil müdahale gerekli mi?
- Bu yaz kurak geçtiği için atların genç sürgünlerde kurumalar ve yaprak dökümü. bıçak bilemesi yapılmalı mı?
- Bu yaz kurak geçtiği için keçilerin çekiş kaybı yaşanması. sulama sıklığını değiştirmeli miyim?

Küme 4:

- Bu yaz sezonunda şeftali tarlamızda yapraklarda kahverengileşme ve büzüşme. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, yem takviyesi yapmalı mıyız?
- Bu yaz sezonunda şeftali tarlamızda yapraklarda kahverengileşme ve büzüşme. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, bıçak bilemesi yapılmalı mı?
- Bu yaz sezonunda patates tarlamızda aşırı su birikimi nedeniyle kök çürümesi. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, fungusit uygulaması gerekli mi?
- Bu yaz sezonunda domates tarlamızda meradaki ot tüketimi azalması. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, bıçak bilemesi yapılmalı mı?
- Bu yaz sezonunda domates tarlamızda meradaki ot tüketimi azalması. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, gübreleme programını yeniden düzenlemeli miyim?

```
In [44]: # Elbow yöntemi ile optimum küme sayısını belirleme
inertia = []
for k in range(2, 10):
    kmeans_temp = KMeans(n_clusters=k, random_state=42)
    kmeans_temp.fit(tfidf_matrix)
    inertia.append(kmeans_temp.inertia_)
```

```
C:\Users\ademt\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: FutureWarning: The default value of
'n_init' will change from 10 to 'auto' in 1.4. Set the value of 'n_init' explicitly to suppress the warning
super()._check_params_vs_input(X, default_n_init=10)
```

```
In [45]: # Elbow grafiği
plt.figure(figsize=(10, 6))
plt.plot(range(2, 10), inertia, marker='o')
plt.xlabel('Küme Sayısı')
plt.ylabel('Inertia')
plt.title('Elbow Yöntemi ile Küme Sayısı Seçimi')
plt.grid(True)
plt.savefig('elbow_plot.png')
plt.close()
print("Elbow grafiği 'elbow_plot.png' dosyasına kaydedildi.")
```

Elbow grafiği 'elbow_plot.png' dosyasına kaydedildi.

```
In [46]: from sklearn.metrics import silhouette_score

# Silhouette skoru
silhouette_avg = silhouette_score(tfidf_matrix, kmeans.labels_)
print(f"Silhouette Skoru: {silhouette_avg:.4f}")
```

Silhouette Skoru: 0.0799

```
In [47]: # Çözüm önerileri tanımlama (örnek)
solution_dict = {
    0: "Toprak analizi yaptırarak azot eksikliğini kontrol edin ve uygun gübreleme yapın.",
    1: "Drenaj sistemini kontrol edin ve su birikimini önlemek için hendek açmayı değerlendirin.",
    2: "Fungisit uygulaması yaparak mantari hastalıkları kontrol altına alın.",
    3: "Sulama sıklığını optimize edin ve sulama sistemini kontrol edin.",
}
```

```
4: "Bıçak bilemesi yaparak makine performansını artırın ve tıkanıklıkları önleyin."
}
```

```
In [48]: # Kümelere çözüm önerileri ekleme
tfidf_df['solution'] = tfidf_df['cluster'].map(solution_dict)
```

```
In [49]: # TF-IDF DataFrame oluşturma
tfidf_df = pd.DataFrame(tfidf_matrix.toarray(), columns=vectorizer.get_feature_names_out())

# Orijinal metinleri ekle
tfidf_df['sorun_metin'] = texts # veya df['sorun_metin']

# Kümeleri ekle
tfidf_df['cluster'] = kmeans.labels_

# Çözüm önerileri ekle
tfidf_df['solution'] = tfidf_df['cluster'].map(solution_dict)
```

```
In [50]: print("\nKümeler ve Çözüm Önerileri:")
for cluster_id in range(num_clusters):
    print(f"\nKüme {cluster_id} Çözüm Önerisi: {solution_dict[cluster_id]}")
    cluster_texts = tfidf_df[tfidf_df['cluster'] == cluster_id]['sorun_metin'].tolist()
    for text in cluster_texts[:3]: # Her kümeden ilk 3 metni göster
        print(f"- Sorun: {text}")
```

Kümeler ve Çözüm Önerileri:

Küme 0 Çözüm Önerisi: Toprak analizi yaptırarak azot eksikliğini kontrol edin ve uygun gübreleme yapın.
- Sorun: mısır tarlamızda genç fidelerde çekiş kaybı yaşanması. aktarma organlarında sorun kaynaklı olabilir mi?
- Sorun: Son hasatta traktörümüzün çekiş kaybı yaşanması. gübreleme programını yeniden düzenlemeli miyim?
- Sorun: kabak tarlamızda başakların dolmadan sararması. Tarlanın drenaj sistemi motor gücü yetersizliği. sulama sıklığını değiştirmeli miyim?

Küme 1 Çözüm Önerisi: Drenaj sistemini kontrol edin ve su birikimini önlemek için hendek açmayı değerlendirin.
- Sorun: patates tarlamızda genç fidelerde genç sürgünlerde kurumalar ve yaprak dökümü. azot eksikliği kaynaklı olabilir mi?
- Sorun: Yeni aldığımız ilaçlama makinesi kök bölgesinde yumuşama ve siyahlaşma. aktarma organlarında sorun olabilir mi?
- Sorun: Yeni aldığımız traktör genç sürgünlerde kurumalar ve yaprak dökümü. kök çürüklüğü olabilir mi?

Küme 2 Çözüm Önerisi: Fungisit uygulaması yaparak mantari hastalıkları kontrol altına alın.
- Sorun: üzüm ağaçlarımızda yapraklarda kahverengileşme ve büzüşme. motor gücü yetersizliği söz konusu olabilir mi?
- Sorun: Son hasatta traktörümüzün meyve dökülmesi aşırı arttı. toprak analizi yaptırmalı mıyım?
- Sorun: Yeni aldığımız biçerdöver meyve dökülmesi aşırı arttı. toprakta yoğun su birikimi olabilir mi?

Küme 3 Çözüm Önerisi: Sulama sıklığını optimize edin ve sulama sistemini kontrol edin.
- Sorun: Bu yaz kurak geçtiği için sığırların genç sürgünlerde kurumalar ve yaprak dökümü. gübreleme programını yeniden düzenlemeli miyim?
- Sorun: Bu yaz kurak geçtiği için sığırların genç sürgünlerde kurumalar ve yaprak dökümü. bıçak bilemesi yapılmalı mı?
- Sorun: Bu yaz kurak geçtiği için sığırların kök bölgesinde yumuşama ve siyahlaşma. acil müdahale gerekli mi?

Küme 4 Çözüm Önerisi: Bıçak bilemesi yaparak makine performansını artırın ve tıkanıklıkları önleyin.
- Sorun: Bu yaz sezonunda şeftali tarlamızda yapraklarda kahverengileşme ve büzüşme. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, yem takviyesi yapmalı mıyız?
- Sorun: Bu yaz sezonunda şeftali tarlamızda yapraklarda kahverengileşme ve büzüşme. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, bıçak bilemesi yapılmalı mı?
- Sorun: Bu yaz sezonunda patates tarlamızda aşırı su birikimi nedeniyle kök çürümesi. Sulama artırılmasına rağmen bitkilerde düzelme olmadı, fungusit uygulaması gerekli mi?

```
In [51]: # Sonuçları CSV dosyasına kaydetme
df[['id', 'sorun_metin', 'cluster', 'solution']].to_csv('clustered_problems_with_solutions.csv', index=False)
print("Kümelenmiş sorunlar ve çözüm önerileri 'clustered_problems_with_solutions.csv' dosyasına kaydedildi.")
```

Kümelenmiş sorunlar ve çözüm önerileri 'clustered_problems_with_solutions.csv' dosyasına kaydedildi.

```
In [ ]:
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js