

# ### THE-README

**\*\*الحلال\*\***: نظام ذكاء اصطناعي لتصنيف المنتجات الغذائية\*(Yaqaen) يقين

نظام ذكاء اصطناعي متكامل يستخدم معالجة اللغات الطبيعية وتعلم الآلة لتصنيف المنتجات الغذائية تلقائيًا إلى "حلال"، "حرام"، أو "مشتبه فيه" بناءً على قائمة مكوناتها.

"Prosciutto" لقطة شاشة للواجهة التفاعلية للتطبيق تعمل على تصنيف منتج يحتوي على مكون

## نظرة عامة

في ظل تعقيد المنتجات الغذائية الحديثة، يهدف هذا المشروع إلى توفير أداة بسيطة وفعالة لمساعدة المستهلكين المسلمين على اتخاذ قرارات غذائية مستنيرة. يقوم النظام بتحليل النصوص التي تصف المنتج ومكوناته، ويقدم حكمًا شرعيًا متوقعًا مع السبب المحتمل وراء هذا الحكم.

## الميزات الرئيسية

- **تصنيف ثلاثي**: يصنف المنتجات بدقة إلى حلال، حرام، و مشتبه فيه.
- **تفسير النتائج**: يقدم سببًا واضحًا للتصنيفات غير الحلال، بالاعتماد على قاعدة معرفة شرعية.
- **دقة عالية**: تم تدريب النموذج على مجموعة بيانات ضخمة وحقق دقة إجمالية تبلغ 96.69%.
- **منهجية هجينة**: يجمع بين قوة النظام الخبير القائم على القواعد ومرونة التعلم الآلي للوصول إلى أفضل النتائج.
- **واجهة تفاعلية**: تم نشره كواجهة ويب سهلة الاستخدام تم بناؤها باستخدام مكتبة Gradio.

## كيف يعمل النظام؟ (المنهجية)

يعتمد النظام على نهج هجين فريد من ثلاث خطوات رئيسية:

1. **المرحلة الأولى: النظام الخبير**: تم بناء قاعدة معرفة شاملة (قاموس بايثون) تحتوي على مئات المكونات المحرمة والمشبوهة. تم استخدام هذا "العقل المفكر" لتصنيف مجموعة بيانات ضخمة تضم حوالي 2 مليون منتج بشكل آلي ودقيق، مما وفر لنا بيانات تدريب موثوقة.
2. **المرحلة الثانية: تعلم الآلة**: تم استخدام البيانات المصنفة لتدريب نموذج الانحدار تم تحويل النصوص إلى متجهات رقمية باستخدام (Logistic Regression) اللوجستي لتمكين النموذج من "فهم" العلاقة بين الكلمات والحكم الشرعي TfidfVectorizer.
3. **المرحلة الثالثة: التطبيق النهائي**: الواجهة التفاعلية تستخدم هذا النظام الهجين. أولاً، يقوم نموذج تعلم الآلة بإعطاء التنبؤ الأولي. ثانيًا، تقوم قاعدة المعرفة بالبحث في النص لتقديم تفسير واضح ومنطقي للمستخدم.

## التقنيات المستخدمة

- **اللغة:** Python 3.8+
- **تحليل البيانات:** Pandas, NumPy
- **تعلم الآلة:** Scikit-learn
- **الواجهة التفاعلية:** Gradio
- **مكتبات مساعدة:** Joblib (لحفظ النموذج), re (للتعابير النمطية), Matplotlib/Seaborn (لعرض النتائج)

## هيكل المشروع

```
halal-food-classifier/
├── data/
│   └── openfoodfacts_cleaned.csv      # مجموعة البيانات الأولية (اختياري)
├── notebooks/
│   ├── 01_Data_Exploration.ipynb     # كود استكشاف البيانات
│   ├── 02_Rule-Based_Classifier.ipynb # كود بناء وتطبيق النظام الخبير
│   └── 03_ML_Model_Training.ipynb    # كود تدريب وتقييم النموذج
├── app/
│   ├── halal_classifier_model.pkl     # ملف النموذج النهائي المحفوظ
│   └── app.py                        # Gradio الكود الخاص بتشغيل واجهة
├── classified_products.parquet        # البيانات المصنفة والجاهزة للتدريب
└── README.md                          # هذا الملف
```

## الإعداد والتشغيل

اتبع الخطوات التالية لتشغيل المشروع على جهازك المحلي.

### 1. المتطلبات المسبقة

- أو أحدث Python 3.8
- مثبتة على جهازك أداة git

### 2. خطوات الإعداد

#### 1. استنسخ المستودع

```
git clone https://github.com/your-username/halal-food-
classifier.git
cd halal-food-classifier
```

#### 2. أنشئ بيئة افتراضية (مستحسن)

```
python -m venv venv
```

- Windows: `bash venv\Scripts\activate` على نظامي التشغيل
  - macOS/Linux: `bash source venv/bin/activate` على نظامي التشغيل
3. **ثبّت المكتبات المطلوبة:**

```
pip install pandas numpy scikit-learn gradio joblib pyarrow
```

(`parquet`. ضرورة لقراءة ملفات `pyarrow`: ملاحظة)

### 3. كيفية الاستخدام

#### أ. إعادة إنتاج النتائج (التدريب)

إذا كنت ترغب في استعراض عملية التحليل وتدريب النموذج من البداية

1. في المجلد الرئيسي `classified_products.parquet` تأكد من وجود ملف
2. `notebooks/` افتح المجلد.
3. بالتسلسل من 01 إلى 03 (`ipynb`). قم بتشغيل دفاتر الملاحظات

#### ب. تشغيل التطبيق النهائي

لتشغيل واجهة الويب التفاعلية مباشرة

1. `app/:` `bash cd app` انتقل إلى المجلد
2. `app.py:` `bash python app.py` قم بتشغيل ملف
3. سيتم تشغيل التطبيق. افتح متصفح الويب الخاص بك وانتقل إلى العنوان المحلي الذي (`http://127.0.0.1:7860` عادة ما يكون) سيظهر في الطرفية

## فريق العمل والإشراف الأكاديمي

تم تطوير هذا المشروع كمشروع تخرج تطبيقي ضمن متطلبات مادة تنقيب البيانات لطلاب قسم علوم الحاسوب وتقنية المعلومات في جامعة إِب

**إعداد وتطوير المهندسين:**

- علي القواس
- طارق العمري
- عبد الرحمن سنان

**تحت إشراف:**

- م. مالك المصنف

**شكر وتقدير خاص:** نتقدم بجزيل الشكر والتقدير إلى **الدكتور عادل العفيري**، الدكتور الأكاديمي لمادة تنقيب البيانات، على توجيهاته القيمة ودعمه المستمر الذي كان له الأثر الكبير في إنجاز هذا العمل.

## العمل المستقبلي

- تحسين قاعدة المعرفة: إضافة المزيد من المكونات والمترادفات بلغات متعددة.
- تحسين فهم السياق (BERT مثل) نماذج متقدمة: تجربة نماذج التعلم العميق.
- في الواجهة لجمع التصحيحات من "Flag" آلية التغذية الراجعة: تفعيل زر المستخدمين واستخدامها لتحسين النموذج بشكل دوري.
- لمعالجة الفئات النادرة SMOTE معالجة عدم التوازن: تجربة تقنيات متقدمة مثل

## المرحلة 1: استيراد المكتبات الأساسية للمشروع

```
#
=====
#
=====
# الخطوة 1: استيراد المكتبات الأساسية للمشروع
#
=====
=====

# مكتبات التعامل مع البيانات وتحليلها 1.
# -----
import pandas as pd # لاستيراد البيانات ومعالجتها في جداول (DataFrames)
import numpy as np  # للعمليات الرياضية والحسابية المتقدمة

# 2. مكتبات عرض البيانات (Data Visualization)
# -----
import matplotlib.pyplot as plt # لإنشاء الرسوم البيانية الأساسية
import seaborn as sns           # لإنشاء رسوم بيانية أكثر جمالاً وتخصصاً
import plotly.express as px     # لإنشاء رسوم بيانية تفاعلية

# 3. مكتبات معالجة النصوص واللغات الطبيعية (NLP)
# -----
import re # لتنظيف النصوص (Regular Expressions) للتعامل مع التعبيرات النمطية
import nltk # مكتبة شاملة لمعالجة اللغات الطبيعية (مثل تقسيم النص لكلمات)
# nltk.download('stopwords') # قد تحتاج لتنزيل بعض المكونات مثل قائمة الكلمات
# المستبعدة
# nltk.download('punkt') # قد تحتاج لتنزيل مقسم الجمل والنصوص

# 4. مكتبات تعلم الآلة وبناء النماذج
# -----
# scikit-learn: المكتبة الأشهر لتعلم الآلة
from sklearn.model_selection import train_test_split # لتقسيم البيانات إلى
مجموعات تدريب واختبار
from sklearn.feature_extraction.text import TfidfVectorizer # لتحويل
النصوص إلى متجهات رقمية (TF-IDF)
from sklearn.linear_model import LogisticRegression # نموذج الانحدار
اللوجستي (كنموذج أساسي)
```

```

from sklearn.ensemble import RandomForestClassifier # نموذج الغابات العشوائية
from sklearn.metrics import accuracy_score, classification_report,
confusion_matrix # لتقييم أداء النموذج
# import xgboost as xgb # مكتبة XGBoost لأداء قوي وعالي
# 5. (BERT لنماذج متقدمة مثل) مكتبات التعلم العميق
# -----
# import torch # مكتبة PyTorch للتعلم العميق
# from transformers import BertTokenizer,
BertForSequenceClassification # مكتبة Hugging Face لاستخدام نماذج BERT
# 6. (Deployment) مكتبات الواجهة التفاعلية والنشر
# -----
import gradio as gr # لإنشاء واجهة ويب تفاعلية بسهولة لعرض النموذج
from PIL import Image # (إذا أضفنا خاصية) للتعامل مع الصور
# import pytesseract # للتعرف الضوئي على الحروف في الصور (OCR)
# 7. مكتبات مساعدة
# -----
from collections import Counter # لحساب تكرار العناصر (مثل الكلمات)
import warnings # للتحكم في التحذيرات التي قد تظهر
warnings.filterwarnings('ignore') # تجاهل التحذيرات للحفاظ على نظافة
المخرجات
print("تم استيراد جميع المكتبات بنجاح. بيئة العمل جاهزة.")
!تم استيراد جميع المكتبات بنجاح. بيئة العمل جاهزة

```

## (Data Exploration) المرحلة الثانية : الإعداد واستكشاف البيانات

```

# الخطوة 1: استيراد المكتبات اللازمة
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# الخطوة 2: تحديد اسم ملف البيانات
# بناءً على طلبك 'openfoodfacts_cleaned.csv' لقد قمت بتغييره إلى
file_path = 'openfoodfacts_cleaned.csv'

print(f"جاري محاولة تحميل الملف {file_path}")

# الخطوة 3: تحميل البيانات مع معالجة الأخطاء
try:
    df = pd.read_csv(file_path, low_memory=False)

    print("تم تحميل البيانات بنجاح")
    print("-" * 50) # لطباعة خط فاصل للترتيب

```

```

# --- (EDA) بداية قسم الاستكشاف العميق ---

print("1. نظرة عامة على أبعاد البيانات (Shape)")
num_rows, num_cols = df.shape
print(f"    - عدد الصفوف (المنتجات): {num_rows}")
print(f"    - عدد الأعمدة (الخصائص): {num_cols}\n")
print("-" * 50)

print("2. نظرة عامة على أنواع البيانات والقيم غير الفارغة (Info)")
# .info() تعطينا ملخصًا سريعًا لكل عمود: اسمه، عدد القيم غير الفارغة، ونوع البيانات
print("    - سيتم طباعة الملخص أدناه:")
df.info()
print("-" * 50)

print("3. ملخص إحصائي للأعمدة الرقمية (Describe - Numerical)")
# .describe() تحسب الإحصائيات الأساسية (العدد، المتوسط، الانحراف المعياري، إلخ) للأعمدة الرقمية
# يعرض هذا الجدول إحصائيات مثل المتوسط والحد الأدنى والأقصى للقيم
print("الرقمية:")
display(df.describe())
print("-" * 50)

print("4. ملخص إحصائي للأعمدة النصية/الفئوية (Describe - Categorical)")
# بتحليل الأعمدة النصية .describe() تقوم، عند إضافة include='object'،
# (عدد القيم، عدد القيم الفريدة، القيمة الأكثر تكرارًا، وعدد مرات تكرارها)
print("يعرض هذا الجدول إحصائيات مثل القيمة الأكثر تكرارًا للأعمدة النصية:")
display(df.describe(include='object'))
print("-" * 50)

print("5. تحليل القيم المفقودة (Missing Values)")
# .isnull().sum() تحسب عدد القيم الفارغة في كل عمود
missing_values = df.isnull().sum()
missing_percentage = (missing_values / num_rows) * 100
missing_df = pd.DataFrame({'count': missing_values, 'percentage':
missing_percentage})
# نعرض فقط الأعمدة التي تحتوي على قيم مفقودة، مرتبة من الأكثر إلى الأقل
print("الأعمدة التي تحتوي على قيم مفقودة ونسبتها المئوية:")
display(missing_df[missing_df['count'] >
0].sort_values(by='percentage', ascending=False))
print("-" * 50)

print("6. التحقق من الصفوف المكررة (Duplicate Rows)")
duplicate_rows = df.duplicated().sum()
print(f"    - عدد الصفوف المكررة بالكامل في البيانات: {duplicate_rows}\n")
print("-" * 50)

```

```

print("7. عدد القيم الفريدة في كل عمود (Cardinality)")
# تحسب عدد القيم الفريدة، وهو أمر مهم لفهم مدى تنوع البيانات في كل عمود
print("يساعدنا هذا على فهم الأعمدة التي قد تكون فئات (قليلة التنوع) أو - (معرفات عالية التنوع)")
display(df.nunique().sort_values(ascending=False))
print("-" * 50)

# --- قسم الرسوم البيانية (Visualizations) ---

print("8. تصور القيم المفقودة (Missing Values Heatmap)")
# هذا الرسم يساعد على رؤية توزيع القيم المفقودة بشكل بصري
plt.figure(figsize=(15, 8))
sns.heatmap(df.isnull(), cbar=False, cmap='viridis')
plt.title('خريطة حرارية للقيم المفقودة في البيانات')
plt.show()
print("-" * 50)

print("9. تصور الارتباط بين الأعمدة الرقمية (Correlation Heatmap)")
# هذا الرسم يوضح مدى ارتباط الأعمدة الرقمية ببعضها البعض
# نختار الأعمدة الرقمية فقط لتجنب الأخطاء
numeric_df = df.select_dtypes(include=['float64', 'int64'])
if not numeric_df.empty:
    corr_matrix = numeric_df.corr()
    plt.figure(figsize=(12, 10))
    sns.heatmap(corr_matrix, annot=False, cmap='coolwarm') # لتجنب ازدحام الأرقام
    plt.title('خريطة حرارية للارتباط بين الأعمدة الرقمية')
    plt.show()
else:
    print("لا توجد أعمدة رقمية كافية لرسم خريطة الارتباط -")

except FileNotFoundError:
    print(f"❌ خطأ: لم يتم العثور على الملف '{file_path}'.")
    print("يرجى التأكد من أن اسم الملف صحيح وأنه موجود في نفس مجلد الكود.")
except Exception as e:
    print(f"❌ حدث خطأ غير متوقع: {e}")

```

❑ جاري محاولة تحميل الملف openfoodfacts\_cleaned.csv

❑ تم تحميل البيانات بنجاح!

- 
1. نظرة عامة على أبعاد البيانات (Shape)
- عدد الصفوف (المنتجات): 1982824
  - عدد الأعمدة (الخصائص): 36

- 
2. نظرة عامة على أنواع البيانات والقيم غير الفارغة (Info)

- سيتم طباعة الملخص أدناه :

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1982824 entries, 0 to 1982823  
Data columns (total 36 columns):
```

#	Column	Dtype
0	Unnamed: 0	int64
1	product_name	object
2	generic_name	object
3	quantity	object
4	brands	object
5	categories_en	object
6	origins_en	object
7	countries_en	object
8	traces_en	object
9	additives_n	float64
10	additives_en	object
11	nutriscore_score	float64
12	food_groups_en	object
13	ecoscore_score	float64
14	ecoscore_grade	object
15	main_category_en	object
16	energy-kcal_100g	float64
17	fat_100g	float64
18	saturated-fat_100g	float64
19	monounsaturated-fat_100g	float64
20	polyunsaturated-fat_100g	float64
21	trans-fat_100g	float64
22	cholesterol_100g	float64
23	carbohydrates_100g	float64
24	sugars_100g	float64
25	fiber_100g	float64
26	proteins_100g	float64
27	salt_100g	float64
28	sodium_100g	float64
29	vitamin-a_100g	float64
30	vitamin-c_100g	float64
31	potassium_100g	float64
32	calcium_100g	float64
33	iron_100g	float64
34	fruits-vegetables-nuts-estimate-from-ingredients_100g	float64
35	nutrition-score-fr_100g	float64

```
dtypes: float64(23), int64(1), object(12)
```

```
memory usage: 544.6+ MB
```

-----  
3. ملخص إحصائي للأعمدة الرقمية (Describe - Numerical)

- يعرض هذا الجدول إحصائيات مثل المتوسط والحد الأدنى والأقصى للقيم الرقمية :



Unnamed: 0				
ecoscore_score \				
count	1.982824e+06	727767.000000	1.018967e+06	595222.000000
mean	1.547365e+06	2.032238	8.916433e+00	48.416323
std	9.475745e+05	2.874262	8.848322e+00	24.691730
min	3.000000e+00	0.000000	-1.500000e+01	-30.000000
25%	7.189638e+05	0.000000	1.000000e+00	30.000000
50%	1.486472e+06	1.000000	9.000000e+00	49.000000
75%	2.398193e+06	3.000000	1.600000e+01	69.000000
max	3.239633e+06	46.000000	4.000000e+01	125.000000

energy-kcal_100g			
fat_100g			
saturated-fat_100g \			
count	1.982824e+06	1.982824e+06	1.982824e+06
mean	7.161504e+09	1.507779e+01	5.248085e+00
std	1.008431e+13	8.583571e+02	7.217128e+01
min	0.000000e+00	0.000000e+00	0.000000e+00
25%	1.140000e+02	1.176471e+00	1.000000e-01
50%	2.680000e+02	8.000000e+00	2.000000e+00
75%	4.030000e+02	2.200000e+01	7.300000e+00
max	1.420000e+16	9.999990e+05	9.999900e+04

monounsaturated-fat_100g		
polyunsaturated-fat_100g		
trans-fat_100g \		
count	48155.000000	48119.000000
mean	9.362510	5.401783
std	15.903869	9.537240
min	0.000000	0.000000
25%	0.200000	0.020000
50%	3.500000	2.080000
75%	9.500000	6.670000
max	251.000000	500.000000

...				
proteins_100g				
salt_100g				
sodium_100g				
vitamin-a_100g \				
count	...	1.982824e+06	1.982824e+06	1.982824e+06
				210009.000000

mean	...	9.368445e+00	4.034655e+07	1.613862e+07	0.197146
std	...	7.734714e+01	5.681302e+10	2.272521e+10	17.546780
min	...	-8.330000e+00	0.000000e+00	0.000000e+00	0.000000
25%	...	1.900000e+00	1.000000e-01	4.000000e-02	0.000000
50%	...	6.500000e+00	6.300000e-01	2.520000e-01	0.000000
75%	...	1.300000e+01	1.427500e+00	5.710000e-01	0.000106
max	...	1.000000e+05	8.000000e+13	3.200000e+13	4374.900000

	vitamin-c_100g	potassium_100g	calcium_100g	iron_100g	\
count	2.118320e+05	103513.000000	2.817370e+05	274030.000000	
mean	4.099596e+01	0.701152	7.100931e+02	0.007333	
std	1.645762e+04	17.004889	3.767975e+05	0.332710	
min	0.000000e+00	0.000000	0.000000e+00	-0.000260	
25%	0.000000e+00	0.086000	0.000000e+00	0.000000	
50%	0.000000e+00	0.173000	4.000000e-02	0.001010	
75%	3.900000e-03	0.318000	1.140000e-01	0.002500	
max	7.479000e+06	2000.000000	2.000000e+08	69.000000	

	fruits-vegetables-nuts-estimate-from-ingredients_100g	\
count	727256.000000	
mean	18.644191	
std	37.844359	
min	-4.053013	
25%	0.000000	
50%	0.418750	
75%	22.506744	
max	13461.000000	

	nutrition-score-fr_100g
count	1.018967e+06
mean	8.916433e+00
std	8.848322e+00
min	-1.500000e+01
25%	1.000000e+00
50%	9.000000e+00
75%	1.600000e+01
max	4.000000e+01

[8 rows x 24 columns]

#### 4. (Describe - Categorical) ملخص إحصائي للأعمدة النصية/الفئوية

- يعرض هذا الجدول إحصائيات مثل القيمة الأكثر تكرارًا للأعمدة النصية :

	product_name \
count	1982824
unique	1269709
top	Aceite de oliva virgen extra
freq	1348

	generic_name	quantity
brands \		
count	112704	638615 1144704
unique	83905	35181 169873
top	Pâtes alimentaires de qualité supérieure	500 g Carrefour
freq	375	26026 14705

	categories_en	origins_en	countries_en	traces_en
additives_en \				
count	1051779	85517	1982824	143134
428447				
unique	99222	10698	5551	12055
102860				
top	Snacks	France	France	Nuts E330 - Citric
acid				
freq	33080	17806	642578	14239
24874				

	food_groups_en	ecoscore_grade
main_category_en		
count	877056	1982824
1051779		
unique	46	7
37353		
top	Sugary snacks,Biscuits and cakes	unknown
Groceries		
freq	80379	1371796
50301		

#### 5. (Missing Values) تحليل القيم المفقودة

- الأعمدة التي تحتوي على قيم مفقودة ونسبتها المئوية :

	count
percentage	
polyunsaturated-fat_100g	1934705

97.573209	
monounsaturated-fat_100g	1934669
97.571393	
origins_en	1897307
95.687111	
potassium_100g	1879311
94.779516	
generic_name	1870120
94.315986	
traces_en	1839690
92.781306	
vitamin-a_100g	1772815
89.408591	
vitamin-c_100g	1770992
89.316651	
trans-fat_100g	1709501
86.215468	
iron_100g	1708794
86.179812	
cholesterol_100g	1702905
85.882812	
calcium_100g	1701087
85.791124	
additives_en	1554377
78.392081	
ecoscore_score	1387602
69.981098	
quantity	1344209
67.792653	
fruits-vegetables-nuts-estimate-from-ingredient...	1255568
63.322211	
additives_n	1255057
63.296440	
food_groups_en	1105768
55.767330	
fiber_100g	1058276
53.372160	
nutriscore_score	963857
48.610315	
nutrition-score-fr_100g	963857
48.610315	
main_category_en	931045
46.955504	
categories_en	931045
46.955504	
brands	838120
42.269006	

## 6. التحقق من الصفوف المكررة (Duplicate Rows)

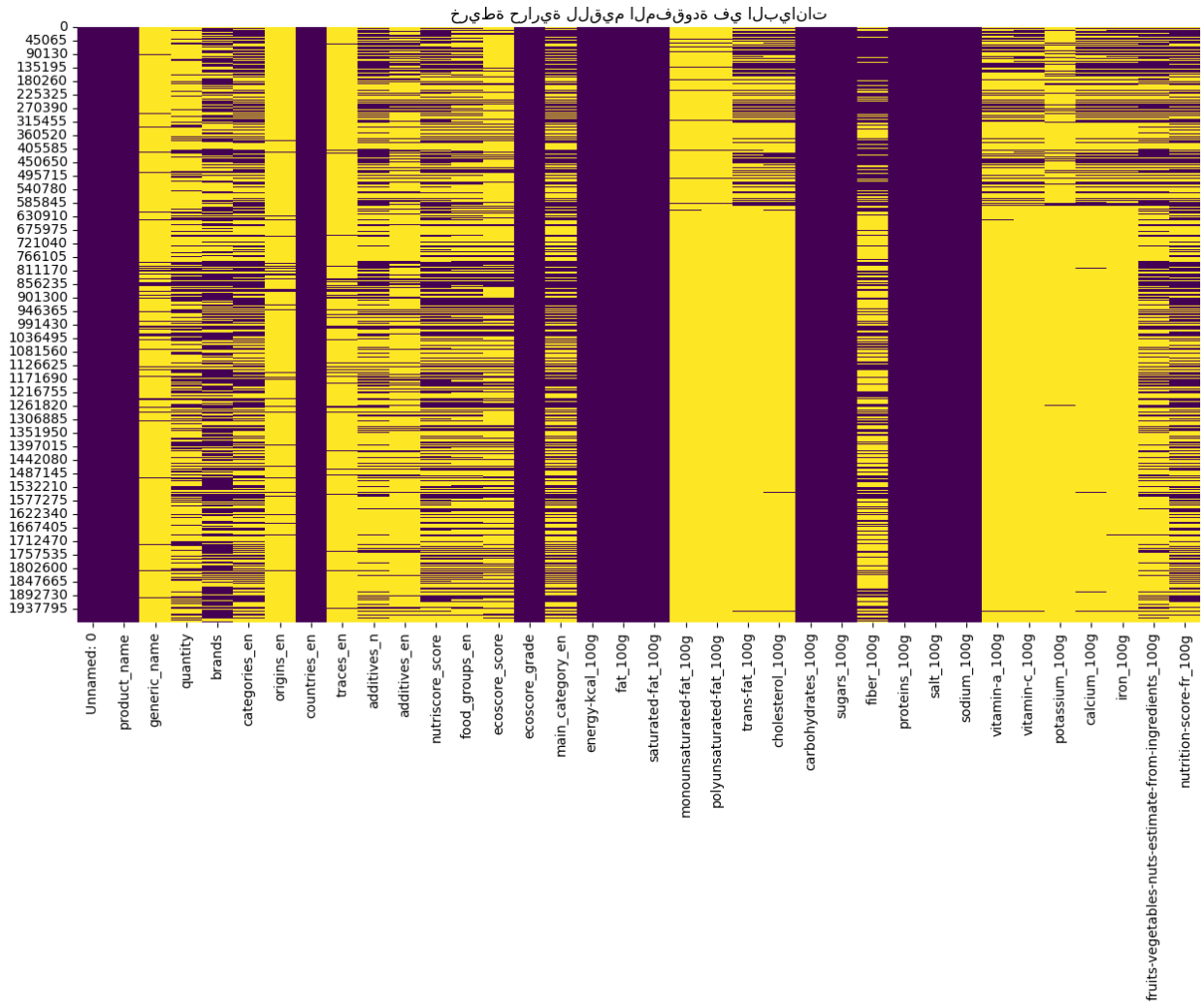
عدد الصفوف المكررة بالكامل في البيانات: 0

## 7. Cardinality (عدد القيم الفريدة في كل عمود)

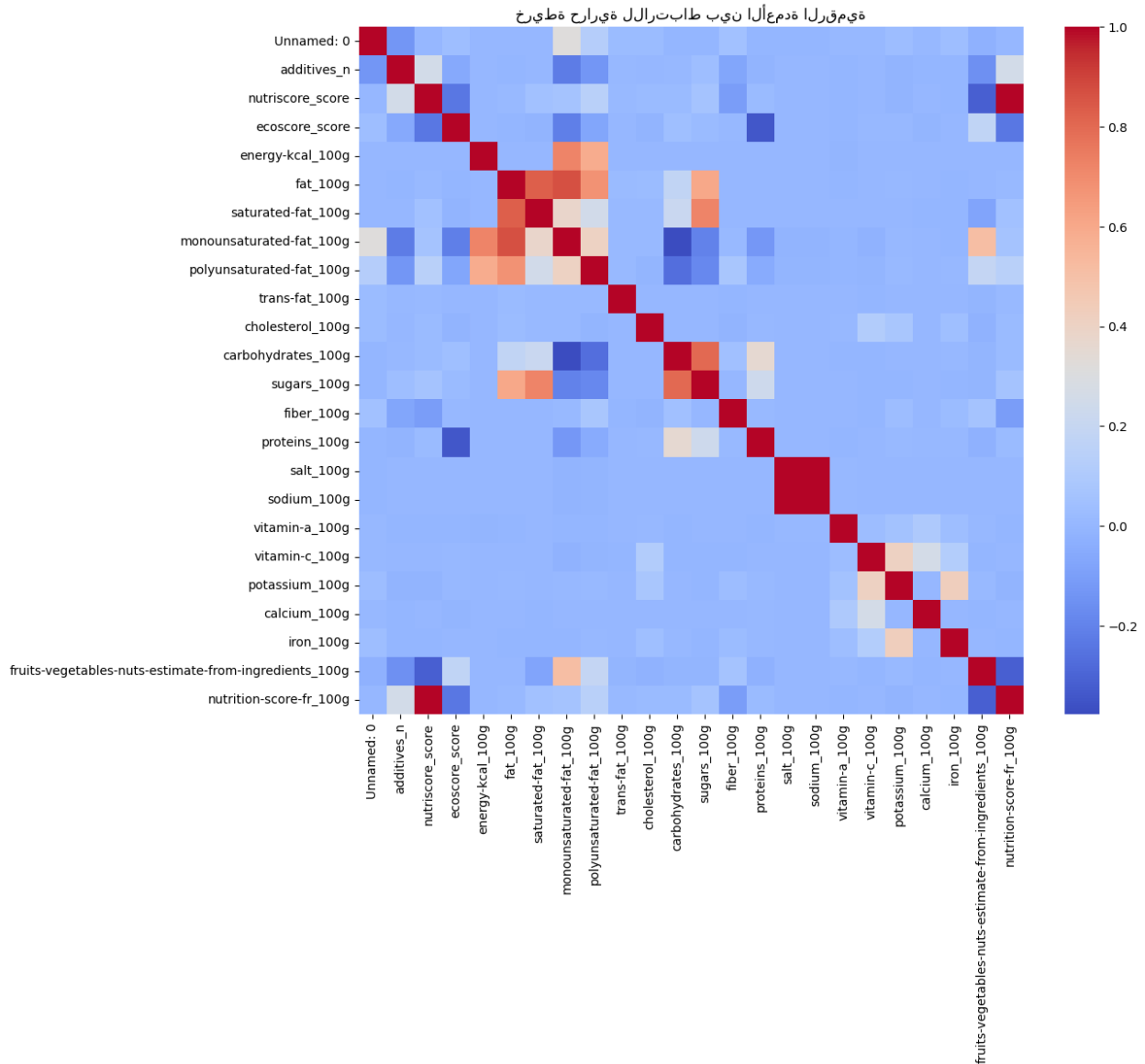
يساعدنا هذا على فهم الأعمدة التي قد تكون فئات (قليلة التنوع) أو معرّفات (عالية التنوع):

Unnamed: 0	1982824
product_name	1269709
brands	169873
fruits-vegetables-nuts-estimate-from-ingredients_100g	122772
additives_en	102860
categories_en	99222
generic_name	83905
main_category_en	37353
sodium_100g	36164
salt_100g	36122
quantity	35181
energy-kcal_100g	34742
carbohydrates_100g	26597
sugars_100g	19813
fat_100g	19300
proteins_100g	17205
saturated-fat_100g	14350
traces_en	12055
origins_en	10698
fiber_100g	7482
countries_en	5551
vitamin-a_100g	3517
calcium_100g	2942
potassium_100g	2904
iron_100g	2737
cholesterol_100g	1834
vitamin-c_100g	1813
monounsaturated-fat_100g	1771
polyunsaturated-fat_100g	1503
trans-fat_100g	894
ecoscore_score	155
nutriscore_score	56
nutrition-score-fr_100g	56
food_groups_en	46
additives_n	39
ecoscore_grade	7
dtype: int64	

## 8. Missing Values Heatmap (تصور القيم المفقودة)



## ----- (Correlation Heatmap) تصور الارتباط بين الأعمدة الرقمية . 9



```
df2 = df.copy()
```

## المرحلة الثالثة: بناء النظام الخبير وتصنيف البيانات (Rule-Based Labeling)

```
import pandas as pd
import numpy as np
import re # <-- الخطوة الأولى والمهمة: استيراد مكتبة البحث النمطي

# ما زال محملاً في الذاكرة 'df' الأصلي DataFrame نفترض أن الـ

# --- الخطوة 1: تحميل قاعدة المعرفة الشرعية الشاملة ---
print("--- الخطوة 1: جاري تحميل قاعدة المعرفة الشرعية الشاملة ---")

comprehensive_ingredients_db = {
```

```
#-----  
-----  
#(Haram) القسم الأول: مكونات محرمة بشكل قاطع#  
#-----  
-----  
  
#--- مشتقات الخنزير (Pork Derivatives) ---  
'pork': {'status': 'حرام', 'reason': 'القرآن في النص الصريح'},  
'bacon': {'status': 'حرام', 'reason': 'نوع من لحم الخنزير المقدد والمعالج'},  
'ham': {'status': 'حرام', 'reason': 'لحم فخذ الخنزير، وهو محرم'},  
'lard': {'status': 'حرام', 'reason': 'دهن الخنزير، وهو محرم'},  
'swine': {'status': 'حرام', 'reason': 'اسم آخر للخنزير، وهو محرم'},  
'hog': {'status': 'حرام', 'reason': 'اسم آخر للخنزير، وهو محرم'},  
'boar': {'status': 'حرام', 'reason': 'الخنزير البري، وهو محرم'},  
'porcine': {'status': 'حرام', 'reason': 'كلمة لاتينية تعني أن المصدر من الخنزير'},  
'speck': {'status': 'حرام', 'reason': 'نوع من اللحم المقدد من الخنزير'},  
'schwein': {'status': 'حرام', 'reason': '"كلمة ألمانية تعني خنزير'"},  
'prosciutto': {'status': 'حرام', 'reason': 'نوع من لحم الخنزير الإيطالي المجفف'},  
  
# --- الكحول والمسكرات (Alcohol and Intoxicants) ---  
'alcohol': {'status': 'حرام', 'reason': 'كل مسكر خمر، وكل خمر حرام'},  
'ethanol': {'status': 'حرام', 'reason': 'الاسم الكيميائي للكحول المسكر'},  
'ethyl alcohol': {'status': 'حرام', 'reason': 'الاسم الكيميائي للكحول المسكر'},  
'wine': {'status': 'حرام', 'reason': 'الخمر، وهو محرم بالنص الصريح'},  
'beer': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'liquor': {'status': 'حرام', 'reason': 'مصطلح عام للمشروبات الكحولية'},  
'liqueur': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المحلاة'},  
'brandy': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'cognac': {'status': 'حرام', 'reason': 'نوع من البراندي، وهو مسكر'},  
'gin': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'rum': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'vodka': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'whisky': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'whiskey': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة'},  
'champagne': {'status': 'حرام', 'reason': 'نوع من النبيذ الفوار المسكر'}
```



```
'sake': {'status': 'حرام', 'reason': 'وهو مسكر'},
'vermouth': {'status': 'حرام', 'reason': 'وهو مسكر'},
'sherry': {'status': 'حرام', 'reason': 'وهو مسكر'},
'port': {'status': 'حرام', 'reason': 'وهو مسكر'},
'absinthe': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية'},
'المسكرة'.
```

# --- حيوانات ومواد محرمة أخرى ---

```
'blood': {'status': 'حرام', 'reason': 'الدم المسفوح محرم بالنص الصريح'},
'carrion': {'status': 'حرام', 'reason': 'الميتة (الحيوان الذي لم يذك بالطريقة الشرعية)'},
'carmine': {'status': 'حرام', 'reason': 'تُستخرج من الحشرات، (E120) صبغة'},
'وهي محرمة عند جمهور العلماء',
'cochineal': {'status': 'حرام', 'reason': '(E120) اسم آخر لصبغة كارمين'},
'المستخرجة من الحشرات',
'e120': {'status': 'حرام', 'reason': 'الرمز الأوروبي لصبغة كارمين المستخرجة'},
'من الحشرات',
'e-120': {'status': 'حرام', 'reason': 'الرمز الأوروبي لصبغة كارمين المستخرجة'},
'من الحشرات',
'ci 75470': {'status': 'حرام', 'reason': 'الرمز الدولي لصبغة كارمين'},
'المستخرجة من الحشرات'.
```

#-----  
-----  
# (Doubtful / Mashbooh) القسم الثاني: مكونات مشبوهة  
# هذه المكونات تحتاج إلى تدقيق للمصدر للتأكد من حليتها.  
#-----  
-----

# --- جيلاتين ومشتقاته ---

```
'gelatin': {'status': 'مشتبه فيه (يميل للتحريم)'},
'حيواني (خنزير أو غير مذكي). يتطلب شهادة حلال موثوقة',
'e441': {'status': 'مشتبه فيه (يميل للتحريم)'},
'للجيلاتين. مصدره الغالب حيواني (خنزير أو غير مذكي)',
'e-441': {'status': 'مشتبه فيه (يميل للتحريم)'},
'للجيلاتين. مصدره الغالب حيواني (خنزير أو غير مذكي)'.
```

# --- دهون، شحوم، ومستحلبات ذات مصدر دهني ---

```
'animal fat': {'status': 'مشتبه فيه'},
'أنه من مصدر مذكي شرعاً',
'animal shortening': {'status': 'مشتبه فيه'},
'يجب'. التأكد من أنه من مصدر مذكي شرعاً',
'fat': {'status': 'مشتبه فيه'},
'كلمة عامة للدهن. إذا لم يحدد المصدر'},
'نباتي)، فهو مشبوه',
'tallow': {'status': 'مشتبه فيه'},
'يجب التأكد من أنه'},
'من مصدر مذكي شرعاً'.
```

'shortening': {'status': 'مشتبه فيه', 'reason': 'قد يحتوي على '},  
 'margarine': {'status': 'مشتبه فيه', 'reason': 'قد يحتوي على دهون حيوانية أو '},  
 'glycerin': {'status': 'مشتبه فيه', 'reason': 'قد يكون مصدره دهن حيواني أو '},  
 'glycerine': {'status': 'مشتبه فيه', 'reason': 'اسم آخر للجليسرين، قد يكون '},  
 'glycerol': {'status': 'مشتبه فيه', 'reason': 'قد (E422) اسم آخر للجليسرين'},  
 'e422': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي للجليسرين/الجليسرول'},  
 'e-422': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي للجليسرين/الجليسرول'},  
 'mono and diglycerides': {'status': 'مشتبه فيه', 'reason': 'مستحلبات'},  
 'monoglyceride': {'status': 'مشتبه فيه', 'reason': 'جزء من مستحلبات'},  
 'diglyceride': {'status': 'مشتبه فيه', 'reason': '(E471) جزء من مستحلبات'},  
 'emulsifier': {'status': 'مشتبه فيه', 'reason': 'كلمة عامة للمستحلبات. العديد'},  
 'emulsifiers': {'status': 'مشتبه فيه', 'reason': 'كلمة عامة للمستحلبات. العديد'},  
 'stearic acid': {'status': 'مشتبه فيه', 'reason': 'قد (E570) حمض دهني'},  
 'e570': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لحمض الستريك. قد'},  
 'e-570': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لحمض الستريك. قد'},  
 'calcium stearate': {'status': 'مشتبه فيه', 'reason': 'ملح حمض الستريك'},  
 'magnesium stearate': {'status': 'مشتبه فيه', 'reason': 'ملح حمض'},  
 'e572': {'status': 'مشتبه فيه', 'reason': 'Magnesium الرمز الأوروبي لـ'},  
 'e-572': {'status': 'مشتبه فيه', 'reason': 'Magnesium الرمز الأوروبي لـ'}

# --- بروتينات، إنزيمات، وأحماض أمينية ---

'collagen': {'status': 'مشتبه فيه', 'reason': 'بروتين حيواني المصدر، يخضع'},  
 'hydrolyzed animal protein': {'status': 'مشتبه فيه', 'reason': 'بروتين'},  
 'pepsin': {'status': 'مشتبه فيه', 'reason': 'إنزيم يستخرج غالباً من معدة'}

{}, 'enzymes': {'status': 'مشتبه فيه', 'reason': 'قد تكون من مصدر حيواني محرم', '}', '}', 'enzyme': {'status': 'مشتبه فيه', 'reason': 'قد يكون من مصدر حيواني محرم', '}', '}', 'protease': {'status': 'مشتبه فيه', 'reason': 'نوع من الإنزيمات، قد يكون من مصدر حيواني محرم', '}', '}', 'rennet': {'status': 'مشتبه فيه', 'reason': 'منفحة الجبن، قد تكون من حيوان', '}', '}', 'animal rennet': {'status': 'مشتبه فيه', 'reason': 'منفحة من مصدر حيواني', '}', '}', 'whey': {'status': 'مشتبه فيه', 'reason': 'شرش اللبن، حكمه يعتمد على حكم', '}', '}', 'whey powder': {'status': 'مشتبه فيه', 'reason': 'مسحوق شرش اللبن، حكمه', '}', '}', 'casein': {'status': 'مشتبه فيه', 'reason': 'بروتين الحليب. قد تستخدم في', '}', '}', 'caseinates': {'status': 'مشتبه فيه', 'reason': 'أملاح بروتين الحليب. قد', '}', '}', 'amino acids': {'status': 'مشتبه فيه', 'reason': 'بعض الأحماض الأمينية (مثل', '}', '}', 'l-cysteine': {'status': 'مشتبه فيه', 'reason': 'يمكن (E920) حمض أميني', '}', '}', 'e920': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي ل-ل-سيستين. قد', '}', '}', 'e-920': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي ل-ل-سيستين. قد', '}', '}', 'e921': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لمشتقات ل-سيستين.', '}', '}', 'e-921': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لمشتقات ل-سيستين.', '}', '}', 'e-921': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لمشتقات ل-سيستين.', '}', '}'

# --- نكهات، مرق، ومكونات عامة ---

'stock': {'status': 'مشتبه فيه', 'reason': 'مرق قد يتم تحضيره من لحوم أو', '}', '}', 'broth': {'status': 'مشتبه فيه', 'reason': 'مرق قد يتم تحضيره من لحوم أو', '}', '}', 'bouillon': {'status': 'مشتبه فيه', 'reason': 'مكعبات مرق قد تحتوي على', '}', '}', 'natural flavors': {'status': 'مشتبه فيه', 'reason': 'قد تحتوي على كحول', '}', '}', 'flavouring': {'status': 'مشتبه فيه', 'reason': 'قد تحتوي على كحول كمذيب', '}', '}', 'vanilla extract': {'status': 'مشتبه فيه', 'reason': 'خلاصة الفانيليا التقليدية', '}', '}'

# --- (Additives) مواد مضافة أخرى مشبوهة ---

'shellac': {'status': 'مشتبه فيه', 'reason': 'إفراز صمغي من حشرة اللاك' (E904). '},  
'e904': {'status': 'مشتبه فيه', 'reason': 'إفراز حشري' (E904). '},  
'e-904': {'status': 'مشتبه فيه', 'reason': 'إفراز حشري' (E904). '},  
'e542': {'status': 'مشتبه فيه', 'reason': 'مصدره عظام' (E542). '},  
'e-542': {'status': 'مشتبه فيه', 'reason': 'مصدره عظام' (E542). '},  
'edible bone phosphate': {'status': 'مشتبه فيه', 'reason': 'مصدره عظام' (E542). '},  
'e631': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium inosinate), '},  
'e-631': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium inosinate), '},  
'e635': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium 5\'-ribonucleotide), '},  
'e-635': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium 5\'-ribonucleotide), '},  
'e640': {'status': 'مشتبه فيه', 'reason': 'حمض أميني' (Glycine), '},  
'e-640': {'status': 'مشتبه فيه', 'reason': 'حمض أميني' (Glycine), '},

# --- مجموعة المستحلبات والأملاح الدهنية المشبوهة ---

# السبب العام: قد تكون مشتقة من دهون حيوانية. تتطلب التحقق من المصدر أو شهادة حلال  
'e304': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Ascorbyl palmitate), '},  
'e-304': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Ascorbyl palmitate), '},  
'e322': {'status': 'مشتبه فيه', 'reason': 'مادة الليسيثين' (E322). '},  
'e-322': {'status': 'مشتبه فيه', 'reason': 'مادة الليسيثين' (E322). '},  
'e325': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Sodium Lactate), '},  
'e-325': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Sodium Lactate), '},  
'e326': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Potassium Lactate), '},  
'e-326': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Potassium Lactate), '},

'e327': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة (Calcium Lactate)،  
'},  
'e-327': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة (Calcium Lactate)،  
'},  
'e430': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد  
'},  
'e-430': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد  
'},  
'e431': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد  
'},  
'e-431': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد  
'},  
'e432': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 20) مستحلب،  
'},  
'e-432': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 20) مستحلب،  
'},  
'e433': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 80) مستحلب،  
'},  
'e-433': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 80) مستحلب،  
'},  
'e434': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 40) مستحلب،  
'},  
'e-434': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 40) مستحلب،  
'},  
'e435': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 60) مستحلب،  
'},  
'e-435': {'status': 'مشتبه فيه', 'reason': 'قد (Polysorbate 60) مستحلب،  
'},  
'e436': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد  
'},  
'e-436': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد  
'},  
'e442': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Ammonium  
phosphatides)، قد يكون من مصدر دهني حيواني  
'},  
'e-442': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Ammonium  
phosphatides)، قد يكون من مصدر دهني حيواني  
'},  
'e470': {'status': 'مشتبه فيه', 'reason': 'أملاح الأحماض الدهنية، قد يكون  
'},  
'e-470': {'status': 'مشتبه فيه', 'reason': 'أملاح الأحماض الدهنية، قد يكون  
'},  
'e470a': {'status': 'مشتبه فيه', 'reason': 'أملاح الأحماض الدهنية (صوديوم،  
'},  
'e-470a': {'status': 'مشتبه فيه', 'reason': 'أملاح الأحماض الدهنية (صوديوم،  
'},  
'e470b': {'status': 'مشتبه فيه', 'reason': 'أملاح المغنيسيوم للأحماض الدهنية،  
'}}

'e-470b': {'status': 'مشتبه فيه', 'reason': 'ألاح المغنيسيوم للأحماض الدهنية', 'text': 'قد يكون مصدرها حيوانياً'},  
'e-471': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Mono- and diglycerides)', 'text': 'قد يكون من مصدر دهني حيواني أو نباتي'},  
'e-471': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Mono- and diglycerides)', 'text': 'قد يكون من مصدر دهني حيواني أو نباتي'},  
'e-472': {'status': 'مشتبه فيه', 'reason': 'استرات متنوعة من', 'text': 'قد يكون E471، استرات متنوعة من'},  
'e-472': {'status': 'مشتبه فيه', 'reason': 'استرات متنوعة من', 'text': 'قد يكون E471، استرات متنوعة من'},  
'e-472a': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون مصدرها E471، استرات من'},  
'e-472a': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون E471، استرات من'},  
'e-472b': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون مصدرها E471، استرات من'},  
'e-472b': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون E471، استرات من'},  
'e-472c': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون مصدرها E471، استرات من'},  
'e-472c': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون E471، استرات من'},  
'e-472d': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون مصدرها E471، استرات من'},  
'e-472d': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون E471، استرات من'},  
'e-472e': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون مصدرها E471، استرات من'},  
'e-472f': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون E471، استرات من'},  
'e-472f': {'status': 'مشتبه فيه', 'reason': 'استرات من', 'text': 'قد يكون E471، استرات من'},  
'e-473': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucrose esters)', 'text': 'قد يكون من مصدر دهني حيواني'},  
'e-473': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucrose esters)', 'text': 'قد يكون من مصدر دهني حيواني'},  
'e-474': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucroglycerides)', 'text': 'قد يكون من مصدر دهني حيواني'},  
'e-474': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucroglycerides)', 'text': 'قد يكون من مصدر دهني حيواني'},  
'e-475': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Polyglycerol esters)', 'text': 'قد يكون من مصدر دهني حيواني'},  
'e-475': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Polyglycerol esters)', 'text': 'قد يكون من مصدر دهني حيواني'}



'e476': {'status': 'مشتبه فيه', 'reason': 'مستحلب (PGPR)، قد يكون من مصدر دهني حيواني.'},  
'e-476': {'status': 'مشتبه فيه', 'reason': 'مستحلب (PGPR)، قد يكون من مصدر دهني حيواني.'},  
'e477': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e-477': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e478': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e-478': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e479b': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e-479b': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e481': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e-481': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e482': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e-482': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e483': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e-483': {'status': 'مشتبه فيه', 'reason': 'مستحلب مشتق من أحماض دهنية قد تكون حيوانية المصدر.'},  
'e491': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monostearate)، قد يكون من مصدر دهني حيواني.'},  
'e-491': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monostearate)، قد يكون من مصدر دهني حيواني.'},  
'e492': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan tristearate)، قد يكون من مصدر دهني حيواني.'},  
'e-492': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan tristearate)، قد يكون من مصدر دهني حيواني.'},  
'e493': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monolaurate)، قد يكون من مصدر دهني حيواني.'},  
'e-493': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monolaurate)، قد يكون من مصدر دهني حيواني.'},  
'e494': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monooleate)، قد يكون من مصدر دهني حيواني.'},  
'e-494': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monooleate)، قد يكون من مصدر دهني حيواني.'},  
'e495': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan

```

monopalmitate)}, قد يكون من مصدر دهني حيواني,
'e-495': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan
monopalmitate)}, قد يكون من مصدر دهني حيواني,
}

print("قاعدة المعرفة جاهزة للاستخدام.")
print("-" * 50)

# --- الخطوة 2: المعالجة المسبقة لقاعدة المعرفة (لا تغيير هنا) ---
print("--- الخطوة 2: جاري المعالجة المسبقة لقاعدة المعرفة لزيادة الكفاءة ---")
haram_ingredients = [ing for ing, data in
comprehensive_ingredients_db.items() if data['status'] == 'حرام']
doubtful_ingredients = [ing for ing, data in
comprehensive_ingredients_db.items() if data['status'].startswith('
مشتبه فيه')]
print(f"عدد المكونات المحرمة: {len(haram_ingredients)}")
print(f"عدد المكونات المشبوهة: {len(doubtful_ingredients)}")
print("المعالجة المسبقة تمت بنجاح.")
print("-" * 50)

# --- الخطوة 3: إعداد البيانات ودالة التصنيف المصححة ---
print("--- الخطوة 3: جاري إعداد البيانات وتطبيق دالة التصنيف المصححة ---")
columns_to_keep = ['product_name', 'categories_en', 'additives_en',
'traces_en']
df_focused = df[columns_to_keep].copy()

def classify_product_corrected(row):
    search_text = []
    for col in columns_to_keep: # نبحث في كل الأعمدة التي احتفظنا بها
        if pd.notna(row[col]):
            search_text.append(str(row[col]).lower())

    full_search_text = " ".join(search_text)

    # 1. البحث عن المكونات المحرمة باستخدام مكتبة 're'
    for ingredient in haram_ingredients:
        # re.search يجد النمط في أي مكان في النص
        # re.escape(ingredient) بشكل آمن '-'
        # r'\b' يضمن البحث عن الكلمة كاملة
        pattern = r'\b' + re.escape(ingredient) + r'\b'
        if re.search(pattern, full_search_text):
            status = comprehensive_ingredients_db[ingredient]
            reason = comprehensive_ingredients_db[ingredient]
            return (status, reason)

```



```

# 2. البحث عن المكونات المشبوهة باستخدام مكتبة 're'
for ingredient in doubtful_ingredients:
    pattern = r'\b' + re.escape(ingredient) + r'\b'
    if re.search(pattern, full_search_text):
        status = comprehensive_ingredients_db[ingredient]
['status']
        reason = comprehensive_ingredients_db[ingredient]
['reason']
        return (status, reason)

# 3. الحالة الافتراضية
return ('الم يتم العثور على أي مكونات محرمة أو مشبوهة في قاعدة البيانات '
        'الحالية.')

# تطبيق الدالة المصححة
print("جاري تطبيق الدالة المصححة... قد تستغرق العملية وقتاً أطول الآن بسبب البحث الدقيق.")
results = df_focused.apply(classify_product_corrected, axis=1)
print("اكتمل التصنيف بنجاح")
print("-" * 50)

# --- الخطوة 4: إنشاء الأعمدة وعرض النتائج (لا تغيير هنا) ---
print("--- وعرض النتائج 'status' و 'reason' الخطوة 4: جاري إنشاء عمودي ---")
df_focused[['status', 'reason']] = pd.DataFrame(results.tolist(),
index=results.index)

print("توزيع المنتجات النهائي حسب الحالة")
print(df_focused['status'].value_counts())
print("\nعرض عينة من البيانات النهائية مع الحكم والسبب")
display(df_focused.head(20))

--- الخطوة 1: جاري تحميل قاعدة المعرفة الشرعية الشاملة ---
قاعدة المعرفة جاهزة للاستخدام
-----

--- الخطوة 2: جاري المعالجة المسبقة لقاعدة المعرفة لزيادة الكفاءة ---
- عدد المكونات المحرمة: 38
- عدد المكونات المشبوهة: 138
المعالجة المسبقة تمت بنجاح
-----

--- الخطوة 3: جاري إعداد البيانات وتطبيق دالة التصنيف المصححة ---
. جاري تطبيق الدالة المصححة... قد تستغرق العملية وقتاً أطول الآن بسبب البحث الدقيق
اكتمل التصنيف بنجاح
-----

--- وعرض النتائج 'status' و 'reason' الخطوة 4: جاري إنشاء عمودي ---
توزيع المنتجات النهائي حسب الحالة
status

```

1760931 حلال  
 173020 مشتببه فيه  
 48278 حرام  
 595 مشتببه فيه (يميل للتحريم)  
 Name: count, dtype: int64

: عرض عينة من البيانات النهائية مع الحكم والسبب

	product_name \	categories_en \
0	Skyr	Dairies, Fermented foods, Fermented milk product...
1	Vitória crackers	NaN
2	Hamburguesas de ternera 100%	NaN
3	Blueberry Cobbler Coffee	NaN
4	Sauce Sweet chili 0%	NaN
5	Mini coco	NaN
6	Pur jus de pomme	Plant-based foods and beverages, Beverages, Plan...
7	Salade de carottes râpées	Plant-based foods and beverages, Plant-based fo...
8	Fromage blanc aux myrtilles	NaN
9	Iogurt de cabra	NaN
10	Baguette parisien	NaN
11	"Baguette Lyonnais"	NaN
12	Solène céréales poulet	NaN
13	Tarte noix de coco	Sweet pies, Pies, Coconut pies
14	Caramel au beurre salé	NaN
15	Salade de fruits exotiques	NaN
16	Chouquettes x 30	NaN
17	Fromage blanc pêche	NaN
18	Crème dessert chocolat	NaN
19	Yaourt au chocolat	NaN

18		NaN		
19		NaN		
		additives_en	traces_en	status
\				
0		NaN	NaN	حلال
1		NaN	NaN	حلال
2		NaN	NaN	حلال
3		NaN	NaN	حلال
4		NaN	NaN	حلال
5		NaN	NaN	حلال
6		NaN	NaN	حلال
7		NaN	NaN	حلال
8		NaN	NaN	حلال
9		NaN	NaN	حلال
10		NaN	NaN	حلال
11		NaN	NaN	حلال
12	E150 - Caramel,E160a - Carotene,E202 - Potassi...		NaN	حلال
13		NaN	NaN	حلال
14		NaN	NaN	حلال
15		NaN	NaN	حلال
16		NaN	NaN	حلال
17		NaN	NaN	حلال
18		E406 - Agar	NaN	حلال
19		NaN	NaN	حلال
		reason		
0	... لم يتم العثور على أي مكونات محرمة أو مشبوهة في			
1	... لم يتم العثور على أي مكونات محرمة أو مشبوهة في			

2	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
3	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
4	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
5	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
6	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
7	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
8	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
9	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
10	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
11	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
12	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
13	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
14	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
15	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
16	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
17	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
18	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
19	لم يتم العثور على أي مكونات محرمة أو مشبوهة في

# عرض عينة من المنتجات المحرمة مع سبب التحريم

```
display(df_focused[df_focused['status'] == 'حرام'].head(10))
```

# عرض عينة من المنتجات المشبوهة التي تميل للتحريم

```
display(df_focused[df_focused['status'] == 'مشتبه فيه (يميل للتحريم)'].head(10))
```

	product_name \	categories_en \
62	Rillettes de sanglier au Chouchen	Meats and their products,Spreads,Meats,Prepare...
193	Pizza prosciutto	Meals,Pizzas pies and quiches,Pizzas
207	Root Beer	Beverages,Carbonated drinks,Sodas,Non-alcoholi...
209	A&w - Root Beer - 355ml	Beverages
246	Décors sucrés	NaN
247	Décors sucrés	NaN
252	Chocolat noir Irish whisky	NaN
332	Sweet and tangy pork with spicy sweet potatoes	NaN
393	Babas au rhum	NaN
531	Boneless Pork Chop	Snacks,Sweet snacks,Biscuits and cakes,Pastrie...

531		NaN	
		additives_en	traces_en
status \			
62		NaN	NaN
حرام			
193		NaN	NaN
حرام			
207	E150 - Caramel,E211 - Sodium benzoate,E999 - Q...		NaN
حرام			
209		NaN	NaN
حرام			
246	E120 - Cochineal,E132 - Indigotine,E133 - Bril...		NaN
حرام			
247	E120 - Cochineal,E132 - Indigotine,E133 - Bril...		NaN
حرام			
252		NaN	NaN
حرام			
332		NaN	NaN
حرام			
393		NaN	NaN
حرام			
531		NaN	NaN
حرام			
		reason	
62		الخنزير البري، وهو محرم	
193		نوع من لحم الخنزير الإيطالي المجفف	
207		نوع من المشروبات الكحولية المسكرة	
209		نوع من المشروبات الكحولية المسكرة	
246		المستخرجة من الحشرات (E120) اسم آخر لصبغة كارمين	
247		المستخرجة من الحشرات (E120) اسم آخر لصبغة كارمين	
252		نوع من المشروبات الكحولية المسكرة	
332		لحم الخنزير محرم بالنص الصريح في القرآن الكريم	
393		نوع من المشروبات الكحولية المسكرة	
531		لحم الخنزير محرم بالنص الصريح في القرآن الكريم	
		product_name	
categories_en \			
5007	Gelatin Family Pack		
NaN			
5008	Fruit Gelatin		
NaN			
7240	Gelatin dessert		
NaN			
7242	Gelatin dessert cherry		
NaN			
7243	Lime Gelatin Dessert	Desserts,Food	

additives,Thickeners,Gelatin  
7244 Gelatin dessert

NaN

7245 Gelatin Food

additives,Thickeners,Gelatin

7246 Gelatin Dessert

NaN

7247 Sugar-free gelatin dessert

NaN

7248 Gelatin Dessert

NaN

		additives_en	traces_en	\
5007	E102 - Tartrazine,E110 - Sunset yellow FCF,E12...		NaN	
5008	E110 - Sunset yellow FCF,E129 - Allura red,E13...		NaN	
7240		NaN	NaN	
7242	E129 - Allura red,E133 - Brilliant blue FCF,E2...		NaN	
7243	E102 - Tartrazine,E133 - Brilliant blue FCF,E2...		NaN	
7244		NaN	NaN	
7245		NaN	NaN	
7246	E129 - Allura red,E133 - Brilliant blue FCF,E2...		NaN	
7247		NaN	NaN	
7248		NaN	NaN	

	status	\
5007	مشتبه فيه (يميل للتحريم)	
5008	مشتبه فيه (يميل للتحريم)	
7240	مشتبه فيه (يميل للتحريم)	
7242	مشتبه فيه (يميل للتحريم)	
7243	مشتبه فيه (يميل للتحريم)	
7244	مشتبه فيه (يميل للتحريم)	
7245	مشتبه فيه (يميل للتحريم)	
7246	مشتبه فيه (يميل للتحريم)	
7247	مشتبه فيه (يميل للتحريم)	
7248	مشتبه فيه (يميل للتحريم)	

	reason
5007	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
5008	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7240	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7242	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7243	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7244	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7245	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7246	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7247	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب
7248	... مصدره الغالب حيواني (خنزير أو غير مذكى). يتطلب

```

# حفظ البيانات الجديدة
!pip install pyarrow

Requirement already satisfied: pyarrow in c:\users\ali\anaconda3\lib\
site-packages (14.0.2)
Requirement already satisfied: numpy>=1.16.6 in c:\users\ali\
anaconda3\lib\site-packages (from pyarrow) (1.26.4)

# df_focused النهائي والمصنف موجود باسم DataFrame نفترض أن

print("--- جاري حفظ النتائج النهائية ---")

# --- (الخيار الموصى به) Parquet الخيار 2: الحفظ كملف ---
parquet_filename = 'classified_products.parquet'
print(f"Parquet باسم: {parquet_filename} ...")

# مباشر وبسيط Parquet الحفظ بصيغة
df_focused.to_parquet(parquet_filename, index=False)

print(f"Parquet: {parquet_filename} تم حفظ البيانات بنجاح في ملف")
print("-" * 40)

print("عمل رائع اليوم! بياناتك محفوظة وجاهزة للغد.")

--- جاري حفظ النتائج النهائية ---
Parquet باسم: classified_products.parquet ... جاري الحفظ كملف
Parquet: classified_products.parquet تم حفظ البيانات بنجاح في ملف
-----
عمل رائع اليوم! بياناتك محفوظة وجاهزة للغد.

import pandas as pd
df_final = pd.read_parquet('classified_products.parquet')
print("تم تحميل البيانات النهائية بنجاح")
display(df_final.head())

!تم تحميل البيانات النهائية بنجاح

```

	product_name \		categories_en	additives_en
0	Skyr			
1	Vitória crackers			
2	Hamburguesas de ternera 100%			
3	Blueberry Cobbler Coffee			
4	Sauce Sweet chili 0%			

	traces_en \		categories_en	additives_en
0	Dairies, Fermented foods, Fermented milk product...			None
	None			
1		None		None
	None			

2	None	None
None		
3	None	None
None		
4	None	None
None		

	status	reason
0	...	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
1	...	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
2	...	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
3	...	لم يتم العثور على أي مكونات محرمة أو مشبوهة في
4	...	لم يتم العثور على أي مكونات محرمة أو مشبوهة في

# للتعامل مع البيانات *pandas* استيراد مكتبة

```
import pandas as pd
```

# تحديد اسم ومسار الملف الذي حفظناه

# (نفترض أنه في نفس مجلد الكود)

```
file_path = 'classified_products.parquet'
```

```
print(f"❑ محاولة تحميل الملف : '{file_path}'")
```

try:

# *DataFrame* لقراءة الملف مباشرة إلى *read\_parquet* استخدام دالة

# *df\_final* الجديد *DataFrame* سنسمي الـ

```
df_final = pd.read_parquet(file_path)
```

```
print(f"\n❑ تم تحميل البيانات النهائية بنجاح")
```

```
print(f"    - إجمالي عدد المنتجات المحملة : {len(df_final)}")
```

```
print(f"    - الأعمدة المتاحة هي : {list(df_final.columns)}")
```

# عرض أول 10 صفوف من البيانات للتأكد من أنها سليمة

```
print("\n❑ عرض عينة من البيانات التي تم تحميلها")
```

```
display(df_final.head(10))
```

# لعرض 15 صفًا عشوائيًا من البيانات

# في كل مرة تشغل هذا الكود، ستري عينة مختلفة

```
display(df_final.sample(15))
```

except *FileNotFoundError*:

```
print(f"\n❑ خطأ: لم يتم العثور على الملف '{file_path}'.")
```

```
print("    - يرجى التأكد من أن اسم الملف صحيح وأنه موجود في نفس مجلد الكود.")
```

```
print("    - تأكد من أنك قمت بتنشغيل كود الحفظ في الجلسة السابقة بنجاح")
```

except *Exception* as e:

```
print(f"\n❑ حدث خطأ غير متوقع أثناء القراءة : {e}")
```

```
❑ محاولة تحميل الملف : 'classified_products.parquet'
```



❑ تم تحميل البيانات النهائية بنجاح !  
 - إجمالي عدد المنتجات المحملة: 1982824  
 - الأعمدة المتاحة هي: ['product\_name', 'categories\_en', 'additives\_en', 'traces\_en', 'status', 'reason']

❑ عرض عينة من البيانات التي تم تحميلها :

	product_name \	categories_en	additives_en
0	Skyr		
1	Vitória crackers		
2	Hamburguesas de ternera 100%		
3	Blueberry Cobbler Coffee		
4	Sauce Sweety chili 0%		
5	Mini coco		
6	Pur jus de pomme		
7	Salade de carottes râpées		
8	Fromage blanc aux myrtilles		
9	Iogurt de cabra		

	traces_en \	categories_en	additives_en
0	Dairies,Fermented foods,Fermented milk product...		None
1	None	None	None
2	None	None	None
3	None	None	None
4	None	None	None
5	None	None	None
6	Plant-based foods and beverages,Beverages,Plan...		None
7	Plant-based foods and beverages,Plant-based fo...		None
8	None	None	None
9	None	None	None

	status	reason
0	حلال ... لم يتم العثور على أي مكونات محرمة أو مشبوهة في	
1	حلال ... لم يتم العثور على أي مكونات محرمة أو مشبوهة في	
2	حلال ... لم يتم العثور على أي مكونات محرمة أو مشبوهة في	
3	حلال ... لم يتم العثور على أي مكونات محرمة أو مشبوهة في	
4	حلال ... لم يتم العثور على أي مكونات محرمة أو مشبوهة في	
5	حلال ... لم يتم العثور على أي مكونات محرمة أو مشبوهة في	

6	حلال... لم يتم العثور على أي مكونات محرمة أو مشبوهة في
7	حلال... لم يتم العثور على أي مكونات محرمة أو مشبوهة في
8	حلال... لم يتم العثور على أي مكونات محرمة أو مشبوهة في
9	حلال... لم يتم العثور على أي مكونات محرمة أو مشبوهة في

	product_name \
494226	Red Chicken Pakora
1095389	Fresh bio
936302	La pintade sauce au riesling
1954886	6 Kaisersemeln
859671	Crottin Chavignol X2 Dbf,
770362	Pechuga de pollo marinada al ajillo
1810573	Sardina aceite oliva rr-125
649532	Andouillette de Troyes
398830	Vanilla Chocolate Chip Keto Ice Cream
748470	Saucisses à rôtir - diots
372612	Éclair panache foit six
445438	French bread pizza
515343	iberico
452565	Pecan pie energy bars, pecan pie
1126221	Chouchous pralines amandes piment doux

	categories_en \
494226	None
1095389	None
936302	None
1954886	Plant-based foods and beverages,Plant-based fo...
859671	Dairies,Fermented foods,Fermented milk product...
770362	None
1810573	Seafood,Fishes and their products,Canned foods...
649532	None
398830	Desserts,Frozen foods,Frozen desserts,Ice crea...
748470	Meats and their products,Meats,Prepared meats,...
372612	None
445438	Meals,Pizzas pies and quiches,Pizzas
515343	None
452565	Snacks
1126221	None

	additives_en	
traces_en \		
494226	None	None
1095389	None	None
936302	None	None
1954886	E262 - Sodium acetates,E262ii - Sodium diaceta...	None
859671	None	None

770362	None	None
1810573	None	None
649532	None	None
398830	E407 - Carrageenan,E410 - Locust bean gum,E412...	None
748470	None	None
372612	None	None
445438	E202 - Potassium sorbate,E250 - Sodium nitrite...	None
515343	None	None
452565	E322 - Lecithins,E322i - Lecithin,E341 - Calci...	None
1126221	None	None
status		reason
494226	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
1095389	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
936302	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
1954886	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
859671	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
770362	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
1810573	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
649532	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
398830	...قد يكون مصدره دهن حي ، (E422) مشتبه فيه اسم آخر للجليسرين	...
748470	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
372612	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
445438	... يمكن استخراجه من شعر الإنسان (E920) مشتبه فيه حمض أميني	...
515343	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال
452565	... مشتبه فيه مادة الليسيتين. غالباً من الصويا أو البيض (حلا	...
1126221	لم يتم العثور على أي مكونات محرمة أو مشبوهة في	...حلال

## المرحلة الرابعة : تدريب وتقييم النموذج

الخطوة 0: استيراد المكتبات الضرورية #

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.pipeline import Pipeline
from sklearn.metrics import classification_report,
```

```

ConfusionMatrixDisplay
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score

print("تم استيراد جميع المكتبات بنجاح.")

# --- الخطوة 1: تحميل البيانات وهندسة الميزات ---

# تحميل البيانات المصنفة مسبقًا من ملف parquet
# هذا الملف هو نتاج عملك الشاق في المرحلة السابقة
print("classified_products.parquet...")
df = pd.read_parquet('classified_products.parquet')
print(f"تم تحميل البيانات بنجاح. عدد المنتجات: {len(df)}")

# تحديد الأعمدة النصية التي ستستخدم كميزات للنموذج
text_columns = ['product_name', 'categories_en', 'additives_en',
                'traces_en']

# معالجة القيم الفارغة (NaN)
# نقوم بتحويل أي قيمة فارغة في الأعمدة النصية إلى نص فارغ
# هذا يمنع حدوث أخطاء عند محاولة دمج النصوص
print("جاري معالجة القيم الفارغة...")
for col in text_columns:
    df[col] = df[col].fillna('')

# إنشاء عمود الميزات المدمج 'text_features'
# ندمج محتوى الأعمدة النصية في عمود واحد ليكون مدخل النموذج
print("text_features...")
df['text_features'] = df[text_columns].agg(' '.join, axis=1)

# عرض مثال على الصفوف الأولى للتأكد من أن كل شيء يعمل كما هو متوقع
print("\nمثال 'text_features' على البيانات بعد إنشاء:")
print(df[['product_name', 'status', 'reason',
          'text_features']].head())

# --- الخطوة 2: تحديد المدخلات والمخرجات وتقسيم البيانات ---

# (y) والمخرجات (X) تحديد متغيرات الإدخال
X = df['text_features'] # الميزات التي سيتعلم منها النموذج
y = df['status']        # الهدف الذي نريد التنبؤ به

# تقسيم البيانات إلى مجموعات تدريب واختبار
# للتدريب، 20% للاختبار 80%
# stratify=y يضمن أن نسبة الفئات (حلال، حرام، مشبوه) تبقى كما هي في مجموعتي
# التدريب والاختبار
print("\nجاري تقسيم البيانات إلى مجموعات تدريب واختبار...")
X_train, X_test, y_train, y_test = train_test_split(

```

```

X, y,
test_size=0.2,      # من البيانات للاختبار 20%
random_state=42,    # لضمان الحصول على نفس التقسيم في كل مرة نشغل فيها الكود
stratify=y          # للحفاظ على توازن الفئات
)
print(f"حجم مجموعة التدريب: {len(X_train)}")
print(f"حجم مجموعة الاختبار: {len(X_test)}")

# --- الخطوة 3: بناء وتدريب النموذج (التجربة الأكثر نضجاً) ---
print("\nجاري بناء وتدريب النموذج بالتجربة الأكثر نضجاً (50 ألف ميزة مع تحسينات إضافية)...")
model_pipeline = Pipeline([
    ('tfidf', TfidfVectorizer(
        max_features=50000,      # نعم، لنطبق استنتاجك الصحيح <--
        min_df=3,                # إجراء وقائي للنظافة <--
        ngram_range=(1, 2)      # لفهم السياق <--
    )),
    ('clf', LogisticRegression(
        random_state=42,
        solver='saga',
        multi_class='multinomial',
        class_weight='balanced' # لإجبار النموذج على الاهتمام بالأخطاء <--
    ))
])

# تدريب النموذج بالكامل على بيانات التدريب
# قد تستغرق هذه الخطوة بعض الوقت بناءً على حجم البيانات
model_pipeline.fit(X_train, y_train)
print("اكتمل تدريب النموذج بنجاح!")

# --- الخطوة 4: تقييم أداء النموذج ---
# عمل تنبؤات على بيانات الاختبار
print("\n...جاري تقييم النموذج على بيانات الاختبار")
y_pred = model_pipeline.predict(X_test)

# عرض تقرير التصنيف المفصل
# لكل فئة (recall) والاستدعاء (precision) يوضح الدقة
print("\n--- (Classification Report) تقرير التصنيف ---")
print(classification_report(y_test, y_pred))

# --- الجزء الجديد: حساب وعرض الدقة الشاملة كنسبة مئوية ---
# accuracy_score نحسب الدقة باستخدام دالة

```

```

overall_accuracy = accuracy_score(y_test, y_pred)
# نطبعها بشكل منسق وواضح كنسبة مئوية
print("\n--- الدقة الشاملة (Overall Accuracy) ---")
print(f"نسبة دقة النموذج الإجمالية هي: {overall_accuracy * 100:.2f}%")
print("=====")

# (Confusion Matrix) عرض مصفوفة الارتباك
# توضح بشكل مرئي أين ينجح النموذج وأين يخطئ
print("\n--- مصفوفة الارتباك (Confusion Matrix) ---")
fig, ax = plt.subplots(figsize=(8, 6))
ConfusionMatrixDisplay.from_predictions(y_test, y_pred, ax=ax,
cmap='Blues')
plt.title('مصفوفة الارتباك للنموذج')
plt.show()

# --- الخطوة 2: تحليل الأخطاء بعمق ---
# بشكل كامل 'text_features' لضمان عرض النصوص الطويلة في عمود
pd.set_option('display.max_colwidth', 300)

print("--- بداية تحليل الأخطاء ---")

# الخطوة 2.1: تجميع بيانات الاختبار ونتائج التنبؤ في إطار واحد
# التي تم df للوصول إلى الصفوف الأصلية الكاملة في loc[X_test.index]. نحن نستخدم
استخدامها في مجموعة الاختبار
# وليس فقط (product_name, reason) مثل هذا يعطينا كل الأعمدة الأصلية
'text_features'
analysis_df = df.loc[X_test.index].copy()
analysis_df['predicted_label'] = y_pred

print(f"منتجاً {len(analysis_df)} تم إنشاء إطار بيانات التحليل بنجاح ويحتوي على (مجموعة الاختبار).")

# الخطوة 2.2: عزل الأخطاء الأكثر خطورة (حرام -> تم توقعه ك حلال)
# نحن نبحث عن الحالات التي كان فيها الحكم الحقيقي هو 'حرام' ولكن النموذج تنبأ بأنه 'حلال'
haram_as_halal_errors = analysis_df[
    (analysis_df['status'] == 'حرام') &
    (analysis_df['predicted_label'] == 'حلال')
]

print(f"\n>>> تم العثور على {len(haram_as_halal_errors)} منتجاً 'حرام' تم تصنيفه بالخطأ ك 'حلال'")

# الخطوة 2.3: عرض وتحليل النتائج
# نعرض الأعمدة المهمة التي تساعدنا في فهم الخطأ
# product_name: لتعرف على المنتج
# status: الحكم الحقيقي

```

```

# predicted_label: الحكم الخاطئ
# reason: سبب الحكم الحقيقي (من نظامك الخبير الأصلي)، هذا العمود ثمين جداً للمقارنة
# text_features: النص الكامل الذي رآه النموذج لاتخاذ قراره
display(haram_as_halal_errors[['product_name', 'status',
'predicted_label', 'reason', 'text_features']].head(20))

# --- قالب لتحليل أنواع أخرى من الأخطاء ---
print("\n--- تحليل نوع آخر من الأخطاء (مشتبه فيه -> تم توقعه ك حلال) ---")

mashbooh_as_halal_errors = analysis_df[
    (analysis_df['status'] == 'مشتبه فيه') &
    (analysis_df['predicted_label'] == 'حلال')
]

print(f"\n>>> {len(mashbooh_as_halal_errors)} منتجاً 'مشتبه فيه' تم العثور على ('تم تصنيفه بالخطأ ك 'حلال')")
display(mashbooh_as_halal_errors[['product_name', 'status',
'predicted_label', 'reason', 'text_features']].head(10))

# يمكنك استخدام نفس النمط لعزل أي خطأ تراه في مصفوفة الارتباك
# مثال: حرام -> مشتبه فيه
# (analysis_df['status'] == 'حرام') & (analysis_df['predicted_label']
== 'مشتبه فيه')

.تم استيراد جميع المكتبات بنجاح
classified_products.parquet...
تم تحميل البيانات بنجاح. عدد المنتجات: 1982824
...جاري معالجة القيم الفارغة
'text_features'...جاري إنشاء عمود الميزات المدمج

'text_features': مثال على البيانات بعد إنشاء
product_name status \
0 Skyr حلال
1 Vitória crackers حلال
2 Hamburguesas de ternera 100% حلال
3 Blueberry Cobbler Coffee حلال
4 Sauce Sweet chili 0% حلال

reason \
0 لم يتم العثور على أي مكونات محرمة أو مشبوهة في قاعدة البيانات الحالية
1 لم يتم العثور على أي مكونات محرمة أو مشبوهة في قاعدة البيانات الحالية
2 لم يتم العثور على أي مكونات محرمة أو مشبوهة في قاعدة البيانات الحالية
3 لم يتم العثور على أي مكونات محرمة أو مشبوهة في قاعدة البيانات الحالية
4 لم يتم العثور على أي مكونات محرمة أو مشبوهة في قاعدة البيانات الحالية

```

```

text_features
0   Skyr Dairies,Fermented foods,Fermented milk
products,Cheeses,Desserts,Dairy desserts,Fermented dairy
desserts,Skys,Cream cheeses
1
Vitória crackers
2
Hamburguesas de ternera 100%
3
Blueberry Cobbler Coffee
4
Sauce Sweet chili 0%

```

...جاري تقسيم البيانات إلى مجموعات تدريب واختبار

حجم مجموعة التدريب: 1586259

حجم مجموعة الاختبار: 396565

...جاري بناء وتدريب النموذج بالتجربة الأكثر نضجاً (50 ألف ميزة مع تحسينات إضافية)

```

C:\Users\Ali\anaconda3\Lib\site-packages\sklearn\linear_model\
_sag.py:350: ConvergenceWarning: The max_iter was reached which means
the coef_ did not converge
  warnings.warn(

```

!اكتمل تدريب النموذج بنجاح

...جاري تقييم النموذج على بيانات الاختبار

--- (Classification Report) تقرير التصنيف ---

			precision	recall	f1-score	support
		9656	0.88	0.98	0.80	حرام
		352186	0.98	0.97	1.00	حلال
		34604	0.89	0.98	0.82	مشتبه فيه
119	0.07	0.99	0.04	(يميل للتحريم)		
		accuracy			0.97	396565
		macro avg	0.66	0.98	0.71	396565
		weighted avg	0.98	0.97	0.97	396565

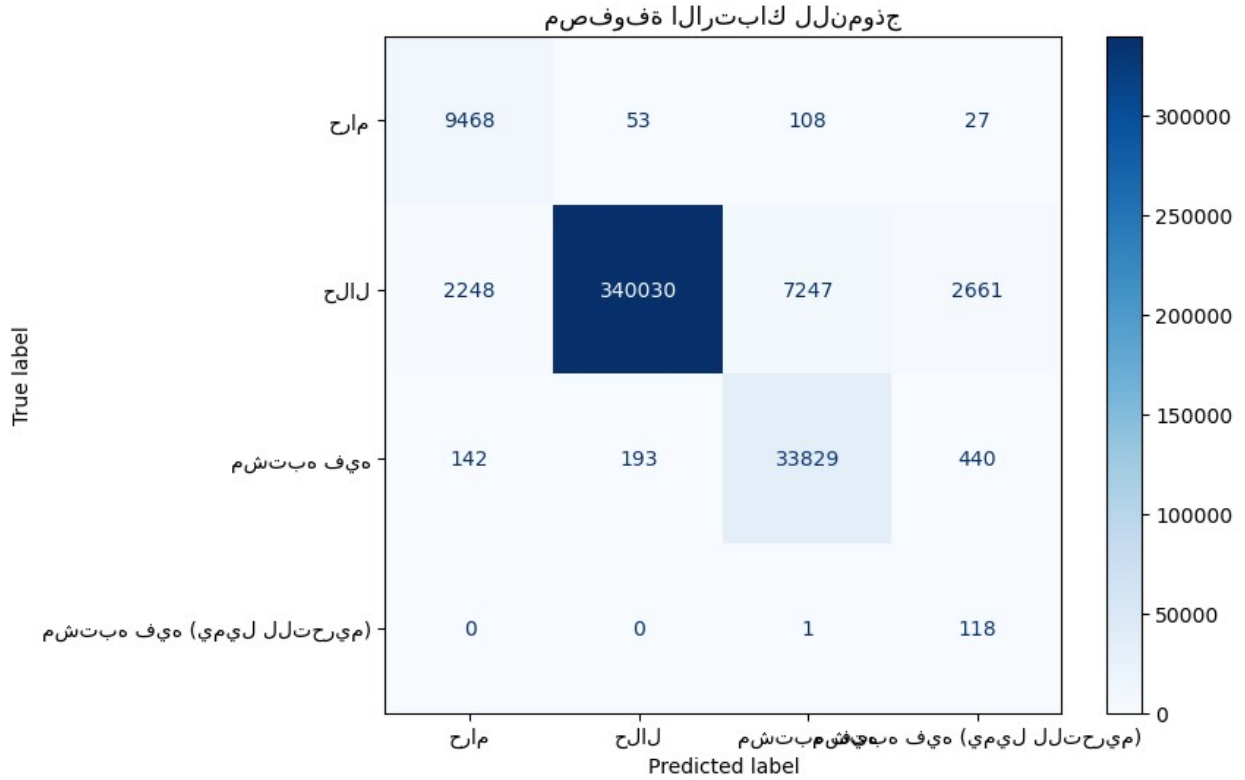
--- (Overall Accuracy) الدقة الشاملة ---

%نسبة دقة النموذج الإجمالية هي: 96.69

=====

--- (Confusion Matrix) مصفوفة الارتباك ---





--- بداية تحليل الأخطاء ---

تم إنشاء إطار بيانات التحليل بنجاح ويحتوي على 396565 منتجاً (مجموعة الاختبار)

: 'تم العثور على 53 منتجاً 'حرام' تم تصنيفه بالخطأ كـ 'حلال' >>

```
product_name \
509269 Perky
Penguins Gummi Candy
1503678 Gin de
Grande Bretagne
496972 Vista d'oro, the preservatory preserve, pineapple with sake &
candied ginger
1290747
Sake Chazuke
561538 Chipotle
Barbecue Chicken
569857 Tart cherry &
port preserves
1958622
Sugar Free Pasito
1556706 gli sfiziosi - 2 mini panpizzato
speck e scamorza
1262828
Bio Wiener
```

1322498	
Ginger Beer	
1257284	
EAA Blood Orange	
1054229	Confiture
cerise et champagne	
121439	
Cola Champagne	
1391971	Breaded
chicken ham cheese	
1480072	
Sake Mei Kwei Lu	
1053305	
Cancoillote Absinthe	
695145	
Sake Box	
207568	Pepperoni +
Bacon Pizza Snack	
634257	
Oeufs Confiseurs	
801679	Tonnelets en chocolat noir
fourré liqueur	

	status	predicted_label \
509269	حلال	حرام
1503678	حلال	حرام
496972	حلال	حرام
1290747	حلال	حرام
561538	حلال	حرام
569857	حلال	حرام
1958622	حلال	حرام
1556706	حلال	حرام
1262828	حلال	حرام
1322498	حلال	حرام
1257284	حلال	حرام
1054229	حلال	حرام
121439	حلال	حرام
1391971	حلال	حرام
1480072	حلال	حرام
1053305	حلال	حرام
695145	حلال	حرام
207568	حلال	حرام
634257	حلال	حرام
801679	حلال	حرام

	reason \
509269	المستخرجة من الحشرات (E120) اسم آخر لصبغة كارمين.

1503678 . نوع من المشروبات الكحولية المسكرة  
 496972 . نبيذ الأرز الياباني، وهو مسكر  
 1290747 . نبيذ الأرز الياباني، وهو مسكر  
 561538 . لحم الخنزير محرم بالنص الصريح في القرآن الكريم  
 569857 . نوع من النبيذ المقوى، وهو مسكر  
 1958622 . المستخرجة من الحشرات (E120) اسم آخر لصبغة كارمين  
 1556706 . نوع من اللحم المقدد من الخنزير  
 1262828 . 'كلمة ألمانية تعني 'خنزير'  
 1322498 . نوع من المشروبات الكحولية المسكرة  
 1257284 . الدم المسفوح محرم بالنص الصريح  
 1054229 . نوع من النبيذ الفوار المسكر  
 121439 . نوع من النبيذ الفوار المسكر  
 1391971 . لحم فخذ الخنزير، وهو محرم  
 1480072 . نبيذ الأرز الياباني، وهو مسكر  
 1053305 . نوع من المشروبات الكحولية المسكرة  
 695145 . نبيذ الأرز الياباني، وهو مسكر  
 207568 . نوع من لحم الخنزير المقدد والمعالج  
 634257 . المستخرجة من الحشرات (E120) اسم آخر لصبغة كارمين  
 801679 . نوع من المشروبات الكحولية المحلاة

#### text\_features

509269 Perky Penguins Gummi  
 Candy Snacks, Sweet snacks, Confectioneries E102 - Tartrazine, E110 -  
 Sunset yellow FCF, E120 - Cochineal, E129 - Allura red, E133 - Brilliant  
 blue FCF, E171 - Titanium dioxide, E330 - Citric acid, E428 -  
 Gelatine, E440 - Pectins, E903 - Carnauba wax  
 1503678  
 Gin de Grande Bretagne Snacks, Sweet snacks, Cocoa and its  
 products, Chocolates, Dark chocolates, Dark chocolate bar  
 496972  
 Vista d'oro, the preservatory preserve, pineapple with sake & candied  
 ginger Condiments, Sauces, Dips, Groceries  
 1290747  
 Sake Chazuke Beverages, Condiments, Groceries E306 - Tocopherol-rich  
 extract, E330 - Citric acid, E331 - Sodium citrates, E33liii - Trisodium  
 citrate, E621 - Monosodium glutamate  
 561538  
 Chipotle Barbecue Chicken Frozen foods E260 - Acetic acid, E330 -  
 Citric acid, E340 - Potassium phosphates, E340ii - Dipotassium  
 phosphate, E412 - Guar gum, E415 - Xanthan gum, E440 - Pectins, E460 -  
 Cellulose, E460ii - Powdered cellulose Pork  
 569857  
 Tart cherry & port preserves Plant-based foods and beverages, Plant-  
 based foods, Breakfasts, Spreads, Plant-based spreads, Sweet spreads, Fruit  
 and vegetable preserves  
 1958622 Sugar Free Pasito Plant-based foods and  
 beverages, Beverages, Plant-based foods, Fruits and vegetables based

foods, Fruits based foods, Carbonated drinks, Fruits, Waters, Tropical  
 fruits, Carbonated soft drinks without fruit juice, Carbonated  
 waters, Carbonated soft drink without fruit juice without sugar ...  
 1556706 gli sfiziosi - 2  
 mini panpizzato speck e scamorza Plant-based foods and  
 beverages, Plant-based foods, Snacks, Cereals and potatoes, Salty  
 snacks, Breads E202 - Potassium sorbate, E252 - Potassium nitrate, E500 -  
 Sodium carbonates, E500ii - Sodium hydrogen carbonate  
 1262828  
 Bio Wiener Schwein, Wiener-wurstchen E250 - Sodium nitrite, E331 -  
 Sodium citrates Celery, Mustard  
 1322498  
 Ginger Beer Beverages, Carbonated drinks, Carbonated soft drinks without  
 fruit juice, Carbonated soft drink without fruit juice without sugar  
 and with artificial sweeteners  
 1257284  
 EAA Blood Orange Dietary supplements E330 - Citric acid, E641 -  
 Leucine, E950 - Acesulfame K, E955 - Sucralose  
 1054229  
 Confiture cerise et champagne Plant-based foods and beverages, Plant-  
 based foods, Breakfasts, Spreads, Plant-based spreads, Sweet spreads, Fruit  
 and vegetable preserves, Jams, Cherry jams  
 121439  
 Cola Champagne Beverages, Carbonated drinks, Sodas E102 -  
 Tartrazine, E110 - Sunset yellow FCF, E330 - Citric acid  
 1391971 Breaded chicken ham cheese Meats and their products, Meat  
 preparations, Meats, Chicken and its products, Poultry, Chicken  
 preparations, Turkey and its products, Breaded  
 products, Chickens, Turkeys, Turkey preparations, Cordons bleus, Turkey  
 cordons bleus, fr:chicken-cordons-bleus E14XX - Modified Starch, E2...  
 1480072  
 Sake Mei Kwei Lu  
 1053305  
 Cancoillote Absinthe Dairies, Fermented foods, Fermented milk  
 products, Cheeses, French cheeses, Processed cheese, Cancoillottes, Melted  
 cheese E452 - Polyphosphates, E452i - Sodium polyphosphate  
 695145  
 Sake Box  
 207568 Pepperoni + Bacon Pizza Snack Frozen foods E171 - Titanium  
 dioxide, E202 - Potassium sorbate, E296 - Malic acid, E330 - Citric  
 acid, E339 - Sodium phosphates, E339ii - Disodium phosphate, E339iii -  
 Trisodium phosphate, E341 - Calcium phosphates, E341i - Monocalcium  
 phosphate, E412 - Guar gum, E461 - Methy...  
 634257 Oeufs Confiseurs Snacks, Sweet snacks, Confectioneries, Candies  
 E101 - Riboflavin, E101i - Riboflavin, E120 - Cochineal, E141 - Copper  
 complexes of chlorophylls and chlorophyllins, E141i - Copper complexes  
 of chlorophylls, E141ii - Copper complexes of chlorophyllins, E300 -  
 Ascorbic acid, E406 - Agar, E414...  
 801679

Tonnelets en chocolat noir fourré liqueur Snacks, Sweet snacks, Cocoa and its products, Chocolates, Dark chocolates

--- تحليل نوع آخر من الأخطاء (مشتبه فيه -> تم توقعه ك حلال) ---

: 'تم العثور على 193 منتجاً 'مشتبه فيه' تم تصنيفه بالخطأ ك 'حلال' >>>

	product_name	status \
118943	Caffeine free pineapple soda	مشتبه فيه
689261	Hagelslag mix	مشتبه فيه
1933026	Blackcurrant flavored drink with nata de coco	مشتبه فيه
1941080	Mountain Dew	مشتبه فيه
1317968	Creame Egg Single	مشتبه فيه
223495	Juice monster	مشتبه فيه
296071	Chocolate tree cookie kit, chocolate	مشتبه فيه
1868276	Sandwich de nata sin azúcares	مشتبه فيه
581785	Candy dispenser	مشتبه فيه
1781014	Preparado soluble al cacao	مشتبه فيه

	predicted_label \
118943	حلال
689261	حلال
1933026	حلال
1941080	حلال
1317968	حلال
223495	حلال
296071	حلال
1868276	حلال
581785	حلال
1781014	حلال

	reason \
118943	قد يكون مصدره (E422) اسم آخر للجليسرين
689261	دهن حيواني أو نباتي . مادة الليسيتين. غالباً من الصويا أو البيض (حلال)، ولكن يوضع في الشبهة للاحتياط التام.
1933026	قد (Calcium Lactate) مادة مضافة
1941080	قد يكون من مصدر حيواني
1317968	قد (Ammonium phosphatides) مستحلب
223495	يكون من مصدر دهني حيواني
296071	قد يكون مصدره (E422) اسم آخر للجليسرين
	دهن حيواني أو نباتي .
	قد (Polysorbate 60) مستحلب

. يكون من مصدر دهني حيواني

1868276 مادة الليسيثين. غالباً من الصويا أو البيض (حلال)، ولكن يوضع في الشبهة للاحتياط التام.

581785 قد يكون (E572) ملح حمض الستياريك

. مصدر الحمض دهن حيواني

1781014 مادة الليسيثين. غالباً من الصويا أو البيض (حلال)، ولكن يوضع في الشبهة للاحتياط التام.

text\_features

118943

Caffeine free pineapple soda Beverages, Carbonated drinks, Sodas E102 - Tartrazine, E110 - Sunset yellow FCF, E211 - Sodium benzoate, E330 - Citric acid, E445 - Glycerol esters of wood rosin

689261

Hagelslag mix Bread coverings, Cooking helpers, Food decorations, Pastry helpers, Sprinkles, Chocolate sprinkles E322 - Lecithins, E414 - Acacia gum

1933026

Blackcurrant flavored drink with nata de coco Beverages, Waters, Flavored waters, Bottled water flavoured with sugar, Sweetened beverages E133 - Brilliant blue FCF, E211 - Sodium benzoate, E327 - Calcium lactate, E330 - Citric acid

1941080

Mountain Dew Beverages, Carbonated drinks, Waters, Sodas, Carbonated waters, Energy drinks, Sweetened beverages E100 - Curcumin, E102 - Tartrazine, E211 - Sodium benzoate, E330 - Citric acid, E331 - Sodium citrates, E445 - Glycerol esters of wood rosin

1317968

Creme Egg Single Snacks, Sweet snacks, Cocoa and its products, Chocolates, Milk chocolates E160c - Paprika extract, E442 - Ammonium phosphatides

223495

Juice

monster Beverages, Carbonated drinks, Sodas, Sweetened beverages E129 - Allura red, E202 - Potassium sorbate, E211 - Sodium benzoate, E330 - Citric acid, E331 - Sodium citrates, E414 - Acacia gum, E445 - Glycerol esters of wood rosin, E955 - Sucralose

296071 Chocolate tree cookie kit, chocolate Snacks, Sweet snacks, Biscuits and cakes, Biscuits E102 - Tartrazine, E110 - Sunset yellow FCF, E129 - Allura red, E132 - Indigotine, E133 - Brilliant blue FCF, E1400 - Dextrin, E171 - Titanium dioxide, E202 - Potassium sorbate, E330 - Citric acid, E331 - Sodium citrates...

1868276 Sandwich de nata sin azúcares Desserts, Frozen foods, Frozen desserts, Ice creams and sorbets, Ice creams, Ice cream sandwiches, Congelados, Helado, Helados-y-sorbetes, Postres, Postres-helados, Sandwich-de-nata, Sandwiches-de-helado E101 - Riboflavin, E101i - Riboflavin, E1200 - Polydextrose, E322 - Lecithins...

581785

Candy dispenser Snacks, Sweet snacks, Confectioneries E102 -

Tartrazine,E110 - Sunset yellow FCF,E129 - Allura red,E133 - Brilliant blue FCF,E330 - Citric acid,E572 - Magnesium stearate 1781014

Preparado soluble al cacao Cocoa and its products,Cocoa and chocolate powders,Cocoa powders E322 - Lecithins,E341 - Calcium phosphates,E341ii - Dicalcium phosphate Milk

# حفظ النموذج

import joblib

print("جاري حفظ النموذج النهائي...")

# بالكامل في ملف *pipeline* نقوم بحفظ الـ

joblib.dump(model\_pipeline, 'halal\_classifier\_model.pkl')

print("تم حفظ النموذج بنجاح في ملف 'halal\_classifier\_model.pkl'")

...جاري حفظ النموذج النهائي

'halal\_classifier\_model.pkl' تم حفظ النموذج بنجاح في ملف

## المرحلة الخامسة : بناء الواجهة التفاعلية (Final Application)

```
pip install gradio
```

Requirement already satisfied: gradio in c:\users\ali\anaconda3\lib\site-packages (5.27.0)

Requirement already satisfied: aiofiles<25.0,>=22.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (24.1.0)

Requirement already satisfied: anyio<5.0,>=3.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (4.2.0)

Requirement already satisfied: fastapi<1.0,>=0.115.2 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.115.12)

Requirement already satisfied: ffmpy in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.5.0)

Requirement already satisfied: gradio-client==1.9.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (1.9.0)

Requirement already satisfied: groovy~=0.1 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.1.2)

Requirement already satisfied: httpx>=0.24.1 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.27.0)

Requirement already satisfied: huggingface-hub>=0.28.1 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.30.2)

Requirement already satisfied: jinja2<4.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (3.1.4)

Requirement already satisfied: markupsafe<4.0,>=2.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (2.1.3)

Requirement already satisfied: numpy<3.0,>=1.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (1.26.4)

Requirement already satisfied: orjson~=3.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (3.10.16)

Requirement already satisfied: packaging in c:\users\ali\anaconda3\

lib\site-packages (from gradio) (23.2)  
Requirement already satisfied: pandas<3.0,>=1.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (2.2.2)  
Requirement already satisfied: pillow<12.0,>=8.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (10.3.0)  
Requirement already satisfied: pydantic<2.12,>=2.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (2.10.3)  
Requirement already satisfied: pydub in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.25.1)  
Requirement already satisfied: python-multipart>=0.0.18 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.0.20)  
Requirement already satisfied: pyyaml<7.0,>=5.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (6.0.1)  
Requirement already satisfied: ruff>=0.9.3 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.11.7)  
Requirement already satisfied: safehttpx<0.2.0,>=0.1.6 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.1.6)  
Requirement already satisfied: semantic-version~=2.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (2.10.0)  
Requirement already satisfied: starlette<1.0,>=0.40.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.46.2)  
Requirement already satisfied: tomlkit<0.14.0,>=0.12.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.13.2)  
Requirement already satisfied: typer<1.0,>=0.12 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.15.2)  
Requirement already satisfied: typing-extensions~=4.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (4.12.2)  
Requirement already satisfied: uvicorn>=0.14.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio) (0.34.2)  
Requirement already satisfied: fsspec in c:\users\ali\anaconda3\lib\site-packages (from gradio-client==1.9.0->gradio) (2024.3.1)  
Requirement already satisfied: websockets<16.0,>=10.0 in c:\users\ali\anaconda3\lib\site-packages (from gradio-client==1.9.0->gradio) (15.0.1)  
Requirement already satisfied: idna>=2.8 in c:\users\ali\anaconda3\lib\site-packages (from anyio<5.0,>=3.0->gradio) (3.7)  
Requirement already satisfied: sniffio>=1.1 in c:\users\ali\anaconda3\lib\site-packages (from anyio<5.0,>=3.0->gradio) (1.3.0)  
Requirement already satisfied: certifi in c:\users\ali\anaconda3\lib\site-packages (from httpx>=0.24.1->gradio) (2025.1.31)  
Requirement already satisfied: httpcore==1.\* in c:\users\ali\anaconda3\lib\site-packages (from httpx>=0.24.1->gradio) (1.0.2)  
Requirement already satisfied: h11<0.15,>=0.13 in c:\users\ali\anaconda3\lib\site-packages (from httpcore==1.\*->httpx>=0.24.1->gradio) (0.14.0)  
Requirement already satisfied: filelock in c:\users\ali\anaconda3\lib\site-packages (from huggingface-hub>=0.28.1->gradio) (3.13.1)  
Requirement already satisfied: requests in c:\users\ali\anaconda3\lib\site-packages (from huggingface-hub>=0.28.1->gradio) (2.32.2)



Requirement already satisfied: tqdm>=4.42.1 in c:\users\ali\anaconda3\lib\site-packages (from huggingface-hub>=0.28.1->gradio) (4.66.4)

Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\ali\anaconda3\lib\site-packages (from pandas<3.0,>=1.0->gradio) (2.9.0.post0)

Requirement already satisfied: pytz>=2020.1 in c:\users\ali\anaconda3\lib\site-packages (from pandas<3.0,>=1.0->gradio) (2024.1)

Requirement already satisfied: tzdata>=2022.7 in c:\users\ali\anaconda3\lib\site-packages (from pandas<3.0,>=1.0->gradio) (2023.3)

Requirement already satisfied: annotated-types>=0.6.0 in c:\users\ali\anaconda3\lib\site-packages (from pydantic<2.12,>=2.0->gradio) (0.6.0)

Requirement already satisfied: pydantic-core==2.27.1 in c:\users\ali\anaconda3\lib\site-packages (from pydantic<2.12,>=2.0->gradio) (2.27.1)

Requirement already satisfied: click>=8.0.0 in c:\users\ali\anaconda3\lib\site-packages (from typer<1.0,>=0.12->gradio) (8.1.7)

Requirement already satisfied: shellingham>=1.3.0 in c:\users\ali\anaconda3\lib\site-packages (from typer<1.0,>=0.12->gradio) (1.5.0)

Requirement already satisfied: rich>=10.11.0 in c:\users\ali\anaconda3\lib\site-packages (from typer<1.0,>=0.12->gradio) (13.3.5)

Requirement already satisfied: colorama in c:\users\ali\anaconda3\lib\site-packages (from click>=8.0.0->typer<1.0,>=0.12->gradio) (0.4.6)

Requirement already satisfied: six>=1.5 in c:\users\ali\anaconda3\lib\site-packages (from python-dateutil>=2.8.2->pandas<3.0,>=1.0->gradio) (1.16.0)

Requirement already satisfied: markdown-it-py<3.0.0,>=2.2.0 in c:\users\ali\anaconda3\lib\site-packages (from rich>=10.11.0->typer<1.0,>=0.12->gradio) (2.2.0)

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in c:\users\ali\anaconda3\lib\site-packages (from rich>=10.11.0->typer<1.0,>=0.12->gradio) (2.15.1)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\ali\anaconda3\lib\site-packages (from requests->huggingface-hub>=0.28.1->gradio) (2.0.4)

Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\ali\anaconda3\lib\site-packages (from requests->huggingface-hub>=0.28.1->gradio) (2.2.2)

Requirement already satisfied: mdurl~=0.1 in c:\users\ali\anaconda3\lib\site-packages (from markdown-it-py<3.0.0,>=2.2.0->rich>=10.11.0->typer<1.0,>=0.12->gradio) (0.1.0)

Note: you may need to restart the kernel to use updated packages.

```
import gradio as gr
import joblib
import re
```

```
# --- الخطوة 0: استرجاع قاعدة المعرفة (القاموس الخبير) ---
# يجب أن يكون هذا القاموس متاحاً لديك من المرحلة الأولى للمشروع
# هذا مجرد مثال مصغر، استخدم القاموس الكامل الخاص بك
```

```

comprehensive_ingredients_db = {

#-----
#(Haram) القسم الأول: مكونات محرمة بشكل قاطع#
#-----

#--- مشتقات الخنزير (Pork Derivatives) ---
'pork': {'status': 'حرام', 'reason': 'لحم الخنزير محرم بالنص الصريح في القرآن الكريم.'},
'bacon': {'status': 'حرام', 'reason': 'نوع من لحم الخنزير المقدد والمعالج.'},
'ham': {'status': 'حرام', 'reason': 'لحم فخذ الخنزير، وهو محرم.'},
'lard': {'status': 'حرام', 'reason': 'دهن الخنزير، وهو محرم.'},
'swine': {'status': 'حرام', 'reason': 'اسم آخر للخنزير، وهو محرم.'},
'hog': {'status': 'حرام', 'reason': 'اسم آخر للخنزير، وهو محرم.'},
'boar': {'status': 'حرام', 'reason': 'الخنزير البري، وهو محرم.'},
'porcine': {'status': 'حرام', 'reason': 'كلمة لاتينية تعني أن المصدر من الخنزير.'},
'speck': {'status': 'حرام', 'reason': 'نوع من اللحم المقدد من الخنزير.'},
'schwein': {'status': 'حرام', 'reason': '"كلمة ألمانية تعني خنزير'."},
'prosciutto': {'status': 'حرام', 'reason': 'نوع من لحم الخنزير الإيطالي المجفف.'},

# --- الكحول والمسكرات (Alcohol and Intoxicants) ---
'alcohol': {'status': 'حرام', 'reason': 'كل مسكر خمر، وكل خمر حرام.'},
'ethanol': {'status': 'حرام', 'reason': 'الاسم الكيميائي للكحول المسكر.'},
'ethyl alcohol': {'status': 'حرام', 'reason': 'الاسم الكيميائي للكحول المسكر.'},
'wine': {'status': 'حرام', 'reason': 'الخمر، وهو محرم بالنص الصريح.'},
'beer': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},
'liquor': {'status': 'حرام', 'reason': 'مصطلح عام للمشروبات الكحولية المقطرة.'},
'liqueur': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المحلاة.'},
'brandy': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},
'cognac': {'status': 'حرام', 'reason': 'نوع من البراندي، وهو مسكر.'},
'gin': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},
'rum': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},
'vodka': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},
'whisky': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},
'whiskey': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية المسكرة.'},

```

```
'champagne': {'status': 'حرام', 'reason': 'نوع من النبيذ الفوار المسكر'},
'sake': {'status': 'حرام', 'reason': 'نبيذ الأرز الياباني، وهو مسكر'},
'vermouth': {'status': 'حرام', 'reason': 'نبيذ مقوى ومنكه، وهو مسكر'},
'sherry': {'status': 'حرام', 'reason': 'نوع من النبيذ المقوى، وهو مسكر'},
'port': {'status': 'حرام', 'reason': 'نوع من النبيذ المقوى، وهو مسكر'},
'absinthe': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية'},
'المسكرة': {'status': 'حرام', 'reason': 'نوع من المشروبات الكحولية'}
```

#### # --- حيوانات ومواد محرمة أخرى ---

```
'blood': {'status': 'حرام', 'reason': 'الدم المسفوح محرم بالنص الصريح'},
'carrion': {'status': 'حرام', 'reason': 'الميتة (الحيوان الذي لم يذك بالطريقة الشرعية)'},
'carmine': {'status': 'حرام', 'reason': 'تُستخرج من الحشرات، (E120) صبغة'},
'cochineal': {'status': 'حرام', 'reason': 'وهي محرمة عند جمهور العلماء (E120) اسم آخر لصبغة كارمين'},
'e120': {'status': 'حرام', 'reason': 'المستخرجة من الحشرات'},
'e-120': {'status': 'حرام', 'reason': 'الرمز الأوروبي لصبغة كارمين المستخرجة من الحشرات'},
'ci 75470': {'status': 'حرام', 'reason': 'الرمز الدولي لصبغة كارمين المستخرجة من الحشرات'}
```

```
#-----
#
# (Doubtful / Mashbooh) القسم الثاني: مكونات مشبوهة
# هذه المكونات تحتاج إلى تدقيق للمصدر للتأكد من حليتها.
#-----
#
```

#### # --- جيلاتين ومشتقاته ---

```
'gelatin': {'status': 'مشتبه فيه (يميل للتحريم)', 'reason': 'مصدره الغالب'},
'e441': {'status': 'مشتبه فيه (يميل للتحريم)', 'reason': 'حيواني (خنزير أو غير مذكي). يتطلب شهادة حلال موثوقة'},
'e-441': {'status': 'مشتبه فيه (يميل للتحريم)', 'reason': 'الرمز الأوروبي القديم'},
'tallow': {'status': 'مشتبه فيه (يميل للتحريم)', 'reason': 'لليجلاتين. مصدره الغالب حيواني (خنزير أو غير مذكي)'}
```

#### # --- دهون، شحوم، ومستحلبات ذات مصدر دهني ---

```
'animal fat': {'status': 'مشتبه فيه', 'reason': 'دهن حيواني. يجب التأكد من أنه من مصدر مذكي شرعاً'},
'animal shortening': {'status': 'مشتبه فيه', 'reason': 'دهن حيواني. يجب التأكد من أنه من مصدر مذكي شرعاً'},
'fat': {'status': 'مشتبه فيه', 'reason': 'كلمة عامة للدهن. إذا لم يحدد المصدر (نباتي)، فهو مشبوه'},
'tallow': {'status': 'مشتبه فيه', 'reason': 'الشحم الحيواني. يجب التأكد من أنه من مصدر مذكي شرعاً'}
```

{}, من مصدر مذكى شرعاً  
 'shortening': {'status': 'مشتبه فيه', 'reason': 'قد يحتوي على '},  
 'دهون حيوانية. يجب التأكد من المصدر',  
 'margarine': {'status': 'مشتبه فيه', 'reason': 'قد يحتوي على دهون حيوانية أو '},  
 'مستحلبات مشبوهة المصدر',  
 'glycerin': {'status': 'مشتبه فيه', 'reason': 'قد يكون مصدره دهن حيواني أو '},  
 'نباتي. يجب التأكد من المصدر النباتي',  
 'glycerine': {'status': 'مشتبه فيه', 'reason': 'اسم آخر للجليسرين، قد يكون '},  
 'مصدره دهن حيواني أو نباتي',  
 'glycerol': {'status': 'مشتبه فيه', 'reason': 'اسم آخر للجليسرين (E422)، قد '},  
 'يكون مصدره دهن حيواني أو نباتي',  
 'e422': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي للجليسرين/الجليسرول.'},  
 'قد يكون مصدره دهن حيواني أو نباتي',  
 'e-422': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي '},  
 'للجليسرين/الجليسرول. قد يكون مصدره دهن حيواني أو نباتي',  
 'mono and diglycerides': {'status': 'مشتبه فيه', 'reason': 'مستحلبات '},  
 'قد تكون من مصدر دهني حيواني أو نباتي. يجب التحقق (E471)',  
 'monoglyceride': {'status': 'مشتبه فيه', 'reason': 'جزء من مستحلبات '},  
 'قد يكون مصدره دهني حيواني أو نباتي (E471)',  
 'diglyceride': {'status': 'مشتبه فيه', 'reason': '(E471) جزء من مستحلبات '},  
 'قد يكون مصدره دهني حيواني أو نباتي',  
 'emulsifier': {'status': 'مشتبه فيه', 'reason': 'كلمة عامة للمستحلبات. العديد '},  
 'منها مشبوه المصدر (حيواني)',  
 'emulsifiers': {'status': 'مشتبه فيه', 'reason': 'كلمة عامة للمستحلبات. '},  
 'العديد منها مشبوه المصدر (حيواني)',  
 'stearic acid': {'status': 'مشتبه فيه', 'reason': 'قد (E570) حمض دهني '},  
 'يكون من مصدر حيواني أو نباتي',  
 'e570': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لحمض الستريك. قد '},  
 'يكون من مصدر حيواني أو نباتي',  
 'e-570': {'status': 'مشتبه فيه', 'reason': 'الرمز الأوروبي لحمض الستريك. قد '},  
 'يكون من مصدر حيواني أو نباتي',  
 'calcium stearate': {'status': 'مشتبه فيه', 'reason': 'ملح حمض الستريك. '},  
 'قد يكون مصدر الحمض دهن حيواني',  
 'magnesium stearate': {'status': 'مشتبه فيه', 'reason': 'ملح حمض '},  
 'قد يكون مصدر الحمض دهن حيواني (E572) الستريك',  
 'e572': {'status': 'مشتبه فيه', 'reason': 'Magnesium الرمز الأوروبي لـ '},  
 'قد يكون مصدره دهن حيواني stearate.',  
 'e-572': {'status': 'مشتبه فيه', 'reason': 'Magnesium الرمز الأوروبي لـ '},  
 'قد يكون مصدره دهن حيواني stearate.'

--- بروتينات، إنزيمات، وأحماض أمينية --- #

'collagen': {'status': 'مشتبه فيه', 'reason': 'بروتين حيواني المصدر، يخضع '},  
 'لنفس حكم الجيلاتين',  
 'hydrolyzed animal protein': {'status': 'مشتبه فيه', 'reason': 'بروتين '},  
 'حيواني متحلل. يجب التأكد من مصدر الحيوان وطريقة ذبحه',

إنزيم يستخرج غالباً من معدة '، 'reason': 'مشتبه فيه'، 'status': 'pepsin':  
 '، 'enzymes': '، 'reason': 'مشتبه فيه'، 'status': 'قد تكون من مصدر حيواني محرم،'  
 '، 'enzyme': '، 'reason': 'مشتبه فيه'، 'status': 'قد يكون من مصدر حيواني محرم،'  
 '، 'protease': '، 'reason': 'مشتبه فيه'، 'status': 'نوع من الإنزيمات، قد يكون من'  
 '، 'rennet': '، 'reason': 'مشتبه فيه'، 'status': 'منفحة الجبن، قد تكون من حيوان'  
 '، 'animal rennet': '، 'reason': 'مشتبه فيه'، 'status': 'منفحة من مصدر حيواني،'  
 '، 'whey': '، 'reason': 'مشتبه فيه'، 'status': 'شرش اللبن، حكمه يعتمد على حكم'  
 '، 'whey powder': '، 'reason': 'مشتبه فيه'، 'status': 'مسحوق شرش اللبن، حكمه'  
 '، 'casein': '، 'reason': 'مشتبه فيه'، 'status': 'بروتين الحليب. قد تستخدم في'  
 '، 'caseinates': '، 'reason': 'مشتبه فيه'، 'status': 'أملاح بروتين الحليب. قد'  
 '، 'amino acids': '، 'reason': 'مشتبه فيه'، 'status': 'بعض الأحماض الأمينية (مثل'  
 '، 'l-cysteine': '، 'reason': 'مشتبه فيه'، 'status': 'يمكن (E920) حمض أميني'  
 '، 'e920': '، 'reason': 'مشتبه فيه'، 'status': 'استخراجه من شعر الإنسان (حرام) أو ريش الدواجن أو مصادر صناعية'  
 '، 'e-920': '، 'reason': 'مشتبه فيه'، 'status': 'الرمز الأوروبي ل-ل-سيستين. قد'  
 '، 'e921': '، 'reason': 'مشتبه فيه'، 'status': 'الرمز الأوروبي لمشتقات ل-سيستين. قد'  
 '، 'e-921': '، 'reason': 'مشتبه فيه'، 'status': 'الرمز الأوروبي لمشتقات ل-سيستين. قد'  
 '، 'e920': '، 'reason': 'مشتبه فيه'، 'status': 'الرمز الأوروبي لمشتقات ل-سيستين. قد'

# --- نكهات، مرق، ومكونات عامة ---

'stock': '، 'reason': 'مشتبه فيه'، 'status': 'مرق قد يتم تحضيره من لحوم أو'  
 '، 'broth': '، 'reason': 'مشتبه فيه'، 'status': 'مرق قد يتم تحضيره من لحوم أو'  
 '، 'bouillon': '، 'reason': 'مشتبه فيه'، 'status': 'مكعبات مرق قد تحتوي على'  
 '، 'natural flavors': '، 'reason': 'مشتبه فيه'، 'status': 'قد تحتوي على كحول'  
 '، 'flavouring': '، 'reason': 'مشتبه فيه'، 'status': 'قد تحتوي على كحول كمذيب'  
 '، 'vanilla extract': '، 'reason': 'مشتبه فيه'، 'status': 'خلاصة الفانيليا التقليدية'

'}, تحتوي على نسبة من الكحول كمذيب

# --- (Additives) مواد مضافة أخرى مشبوهة ---

'shellac': {'status': 'مشتبه فيه', 'reason': 'إفراز صمغي من حشرة اللاك' (E904)},  
'e904': {'status': 'مشتبه فيه', 'reason': 'إفراز حشري' (E904)},  
'e-904': {'status': 'مشتبه فيه', 'reason': 'إفراز حشري' (E904)},  
'e542': {'status': 'مشتبه فيه', 'reason': 'مصدره عظام' (E542)},  
'e-542': {'status': 'مشتبه فيه', 'reason': 'مصدره عظام' (E542)},  
'edible bone phosphate': {'status': 'مشتبه فيه', 'reason': 'مصدره عظام' (E542)},  
'e631': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium inosinate)},  
'e-631': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium inosinate)},  
'e635': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium 5'-ribonucleotide)},  
'e-635': {'status': 'مشتبه فيه', 'reason': 'مُحسِّن نكهة' (Sodium 5'-ribonucleotide)},  
'e640': {'status': 'مشتبه فيه', 'reason': 'حمض أميني' (Glycine)},  
'e-640': {'status': 'مشتبه فيه', 'reason': 'حمض أميني' (Glycine)},

# --- مجموعة المستحلبات والأملاح الدهنية المشبوهة ---

# السبب العام: قد تكون مشتقة من دهون حيوانية. تتطلب التحقق من المصدر أو شهادة حلال.  
'e304': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Ascorbyl palmitate)},  
'e-304': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Ascorbyl palmitate)},  
'e322': {'status': 'مشتبه فيه', 'reason': 'مادة الليسيثين. غالباً من الصويا أو البيض (حلال)' (E322)},  
'e-322': {'status': 'مشتبه فيه', 'reason': 'مادة الليسيثين. غالباً من الصويا أو البيض (حلال)' (E322)},  
'e325': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Sodium Lactate)},  
'e-325': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Sodium Lactate)},  
'e326': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Potassium Lactate)},  
'e-326': {'status': 'مشتبه فيه', 'reason': 'مادة مضافة' (Potassium Lactate)},



Lactate)، قد تكون من مصدر حيواني ،  
'e327': {'status': 'مشتبه فيه' , 'reason': ' مادة مضافة (Calcium Lactate) ،  
قد تكون من مصدر حيواني  
'e-327': {'status': 'مشتبه فيه' , 'reason': ' مادة مضافة (Calcium Lactate) ،  
قد تكون من مصدر حيواني  
'e430': {'status': 'مشتبه فيه' , 'reason': ' مشتق من أحماض دهنية قد  
.تكون حيوانية المصدر  
'e-430': {'status': 'مشتبه فيه' , 'reason': ' مشتق من أحماض دهنية قد  
.تكون حيوانية المصدر  
'e431': {'status': 'مشتبه فيه' , 'reason': ' مشتق من أحماض دهنية قد  
.تكون حيوانية المصدر  
'e-431': {'status': 'مشتبه فيه' , 'reason': ' مشتق من أحماض دهنية قد  
.تكون حيوانية المصدر  
'e432': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 20) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e-432': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 20) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e433': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 80) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e-433': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 80) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e434': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 40) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e-434': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 40) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e435': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 60) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e-435': {'status': 'مشتبه فيه' , 'reason': ' (Polysorbate 60) مستحلب، قد  
.يكون من مصدر دهني حيواني  
'e436': {'status': 'مشتبه فيه' , 'reason': ' مشتق من أحماض دهنية قد  
.تكون حيوانية المصدر  
'e-436': {'status': 'مشتبه فيه' , 'reason': ' مشتق من أحماض دهنية قد  
.تكون حيوانية المصدر  
'e442': {'status': 'مشتبه فيه' , 'reason': ' (Ammonium  
phosphatides)، قد يكون من مصدر دهني حيواني ،  
'e-442': {'status': 'مشتبه فيه' , 'reason': ' (Ammonium  
phosphatides)، قد يكون من مصدر دهني حيواني ،  
'e470': {'status': 'مشتبه فيه' , 'reason': ' أملاح الأحماض الدهنية، قد يكون  
.مصدرها حيوانياً  
'e-470': {'status': 'مشتبه فيه' , 'reason': ' أملاح الأحماض الدهنية، قد يكون  
.مصدرها حيوانياً  
'e470a': {'status': 'مشتبه فيه' , 'reason': ' أملاح الأحماض الدهنية (صوديوم،  
.بوتاسيوم، كالسيوم)، قد يكون مصدرها حيوانياً  
'e-470a': {'status': 'مشتبه فيه' , 'reason': ' أملاح الأحماض الدهنية (صوديوم،  
.بوتاسيوم، كالسيوم)، قد يكون مصدرها حيوانياً

'e470b': {'status': 'مشتبه فيه', 'reason': 'أملاح المغنيسيوم للأحماض الدهنية', 'قد يكون مصدرها حيوانياً'},  
'e-470b': {'status': 'مشتبه فيه', 'reason': 'أملاح المغنيسيوم للأحماض الدهنية', 'قد يكون مصدرها حيوانياً'},  
'e471': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Mono- and diglycerides)', 'قد يكون من مصدر دهني حيواني أو نباتي'},  
'e-471': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Mono- and diglycerides)', 'قد يكون من مصدر دهني حيواني أو نباتي'},  
'e472': {'status': 'مشتبه فيه', 'reason': 'E471 استرات متنوعة من', 'قد يكون مصدرها دهني حيواني'},  
'e-472': {'status': 'مشتبه فيه', 'reason': 'E471 استرات متنوعة من', 'قد يكون مصدرها دهني حيواني'},  
'e472a': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e-472a': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e472b': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e-472b': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e472c': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e-472c': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e472d': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e-472d': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e472e': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e-472e': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e472f': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e-472f': {'status': 'مشتبه فيه', 'reason': 'E471 استرات من', 'قد يكون مصدرها دهني حيواني'},  
'e473': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucrose esters)', 'قد يكون من مصدر دهني حيواني'},  
'e-473': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucrose esters)', 'قد يكون من مصدر دهني حيواني'},  
'e474': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucroglycerides)', 'قد يكون من مصدر دهني حيواني'},  
'e-474': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sucroglycerides)', 'قد يكون من مصدر دهني حيواني'},  
'e475': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Polyglycerol esters)', 'قد يكون من مصدر دهني حيواني'},  
'e-475': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Polyglycerol esters)', 'قد يكون من مصدر دهني حيواني'}



esters)، قد يكون من مصدر دهني حيواني،  
'e476': {'status': 'مشتبه فيه', 'reason': 'مستحلب (PGPR) مصدر،  
'e-476': {'status': 'مشتبه فيه', 'reason': 'مستحلب (PGPR)، قد يكون من  
'e477': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e-477': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e478': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e-478': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e479b': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e-479b': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e481': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e-481': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e482': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e-482': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e483': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e-483': {'status': 'مشتبه فيه', 'reason': 'مشتحلب مشتق من أحماض دهنية قد  
'e491': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
monostearate)، قد يكون من مصدر دهني حيواني،  
'e-491': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
monostearate)، قد يكون من مصدر دهني حيواني،  
'e492': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
tristearate)، قد يكون من مصدر دهني حيواني،  
'e-492': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
tristearate)، قد يكون من مصدر دهني حيواني،  
'e493': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
monolaurate)، قد يكون من مصدر دهني حيواني،  
'e-493': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
monolaurate)، قد يكون من مصدر دهني حيواني،  
'e494': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monooleate)،  
'e-494': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan  
monooleate)، قد يكون من مصدر دهني حيواني،

```

'e495': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monopalmitate)، قد يكون من مصدر دهني حيواني'},
'e-495': {'status': 'مشتبه فيه', 'reason': 'مستحلب (Sorbitan monopalmitate)، قد يكون من مصدر دهني حيواني'},
}

# --- الخطوة 1: تحميل النموذج المدرب ---
model = joblib.load('halal_classifier_model.pkl')

# --- الخطوة 2: تعريف دالة التنبؤ والشرح (النسخة الهجينة) ---
def predict_and_explain(text_input):
    if not model or not text_input.strip():
        return {}, ""

    # الجزء الأول: التنبؤ باستخدام نموذج التعلم الآلي
    prediction = model.predict([text_input])[0]

    # الجزء الثاني: البحث عن السبب باستخدام قاعدة المعرفة
    reason = "لم يتم تحديد سبب مباشر من قاعدة المعرفة." # سبب افتراضي

    if prediction in ['حرام', 'مشتبه فيه']:
        # نبحث في النص عن المكونات من قاموسنا الخبير
        for ingredient, details in comprehensive_ingredients_db.items():
            # استخدام تعابير نمطية للبحث عن الكلمة كاملة
            if re.search(r'\b' + re.escape(ingredient) + r'\b',
text_input, re.IGNORECASE):
                # إذا وجدنا مكوناً، نستخدم سببه ونخرج من الحلقة
                reason = details['reason']
                break # نكتفي بأول سبب نجده
    elif prediction == 'حلال':
        reason = "لم يتم العثور على مكونات محرمة أو مشبوهة واضحة."

    return prediction, reason

# --- الخطوة 3: بناء الواجهة التفاعلية ---
iface = gr.Interface(
    fn=predict_and_explain,
    inputs=gr.Textbox(lines=5, label="أدخل اسم المنتج والمكونات"),
    outputs=[
        gr.Textbox(label="الحكم الشرعي المتوقع (Predicted Status)"),
        gr.Textbox(label="السبب المحتمل (Potential Reason)"),
    ],
    title="نظام خبير لتصنيف المنتجات الغذائية",
    description="أدخل معلومات المنتج لتحصل على الحكم الشرعي المتوقع مع السبب"
)

```

```
المحتمل.",  
    examples=[["Prosciutto crudo"], ["Cheddar cheese, water, salt,  
rennet"]]  
)
```

```
iface.launch()
```

```
* Running on local URL:  http://127.0.0.1:7860
```

```
To create a public link, set `share=True` in `launch()`.
```

```
<IPython.core.display.HTML object>
```