Test Documentation

Testing Tool Used

- PyTest: Used as the primary testing framework.
- unittest: Provides tools like patch and MagicMock to mock external dependencies such as APIs and models

Testing Strategy

- **Unit Tests**
    - Validates individual functions such as process_image, food_to_nutrition, and transform_data.
    - Ensures correctness of outputs for given inputs.
    - External APIs (e.g., USDA API) and machine learning models are mocked to simulate their behavior during tests.
- **Integration Tests**
    - Verifies that multiple components work together as expected, such as preprocessing an image and running it through the classification pipeline.
- **System Tests**
    - Covers user flows and interactions, including building the container and running all steps in the app.

Test Case Descriptions

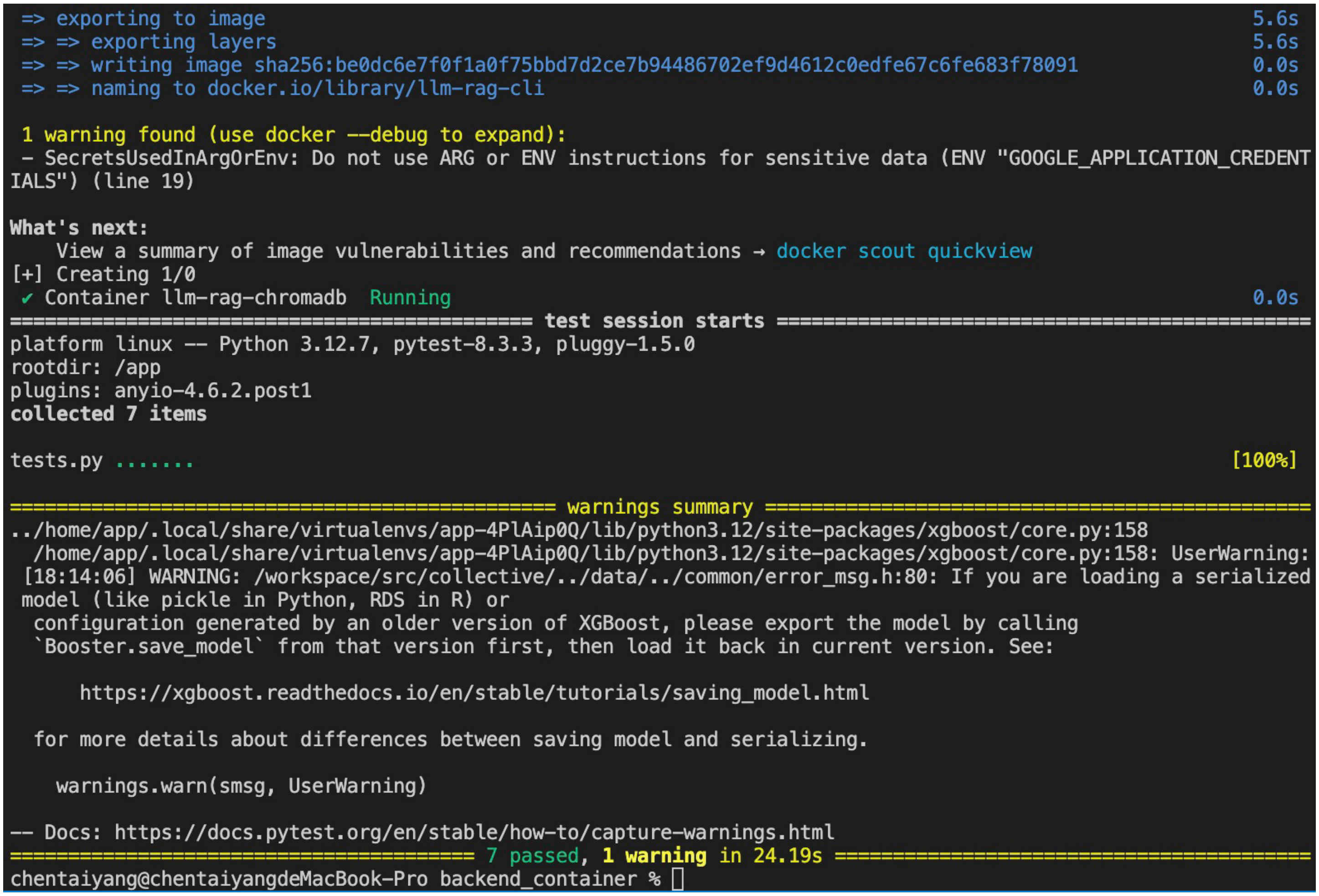
- **Image Processing (process_image)**
    - Goal: Validate the resizing and preprocessing of input images. Ensure compatibility with Keras model input requirements.
    - Input: Dummy images of varying dimensions and content (e.g., all-black, random noise, large, small).
    - Validation: The output shape matches (1, 224, 224, 3).
- **Food Prediction (make_prediction)**
    - Goal: Verify predictions for processed images.
    - Input: Dummy image array.
    - Validation: The returned label is a string, and the returned probability is a non-zero float.
- **Food Nutrition Mapping (food_to_nutrition)**
    - Goal: Test the function that queries the USDA API and calculates nutritional values based on serving size.
    - Input: Food item and serving size.
    - Validation: Output is a Pandas DataFrame with specific columns (Description, Energy, Protein, Fat, Carbohydrate). Nutritional values are correctly scaled based on serving size.
- **Data Transformation (transform_data)**
    - Goal: Validate the transformation of nutrition JSON data into a structured DataFrame.
    - Input: Mock JSON data.
    - Validation: Output is a Pandas DataFrame with renamed columns. Data types and values are correctly formatted.
- **Disease Risk Prediction (nutrition_predict_disease)**
    - Goal: Simulate disease risk prediction using mock machine learning models.
    - Input: Transformed nutrition data.
    - Validation: Mock models produce the correct probabilities for each disease. Output probabilities match the mock behavior.
- **Interaction with LLM (interact_with_llm)**
    - Goal: Ensure correct interaction with a subprocess-based Large Language Model (LLM) pipeline.

  ○ Input: A fixed test prompt.
  ○ Validation: subprocess.run is called with expected arguments.

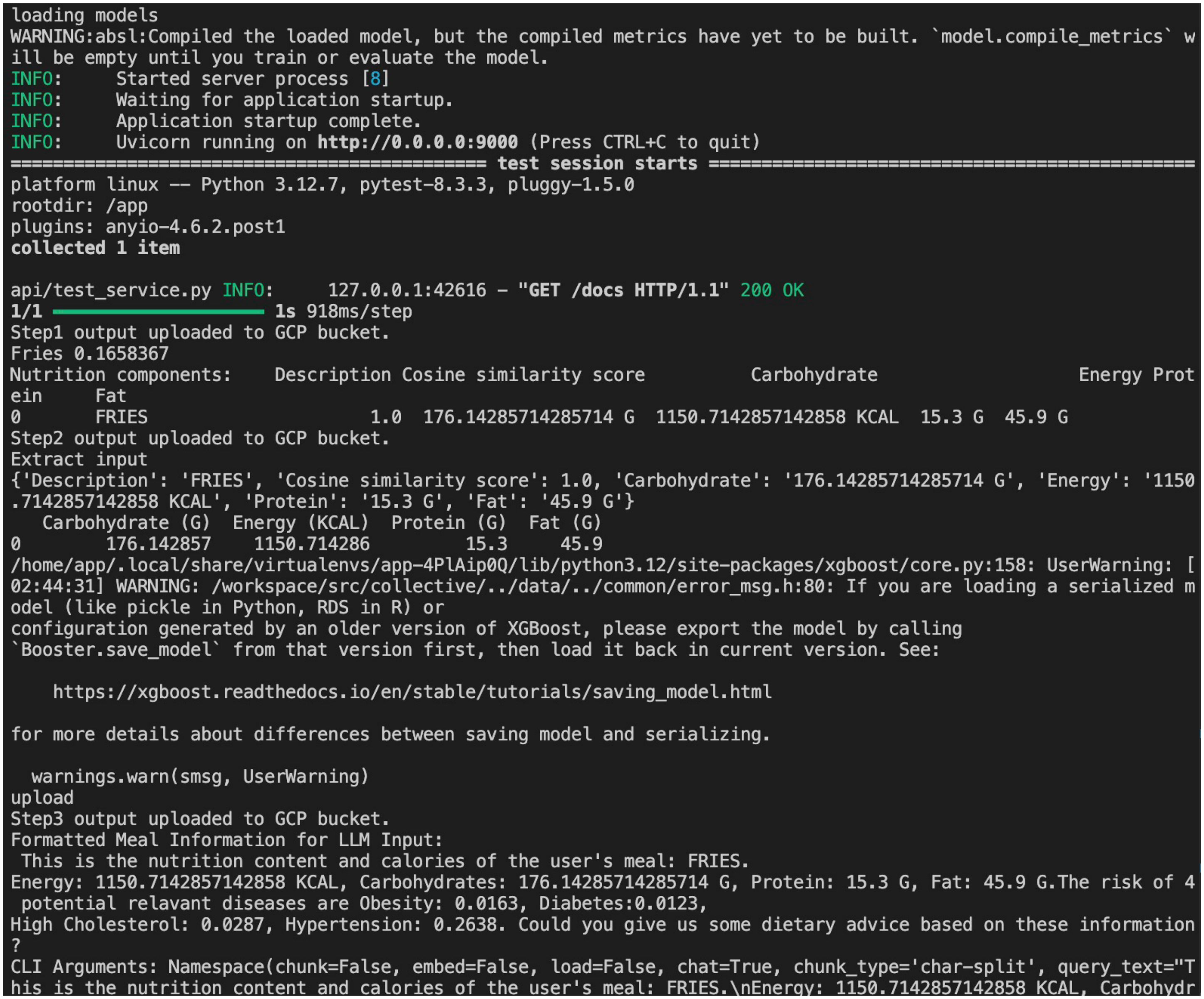
Instructions

- Run the backend container: *sh docker-shell.sh*
- Run all tests using PyTest: *pytest tests.py*

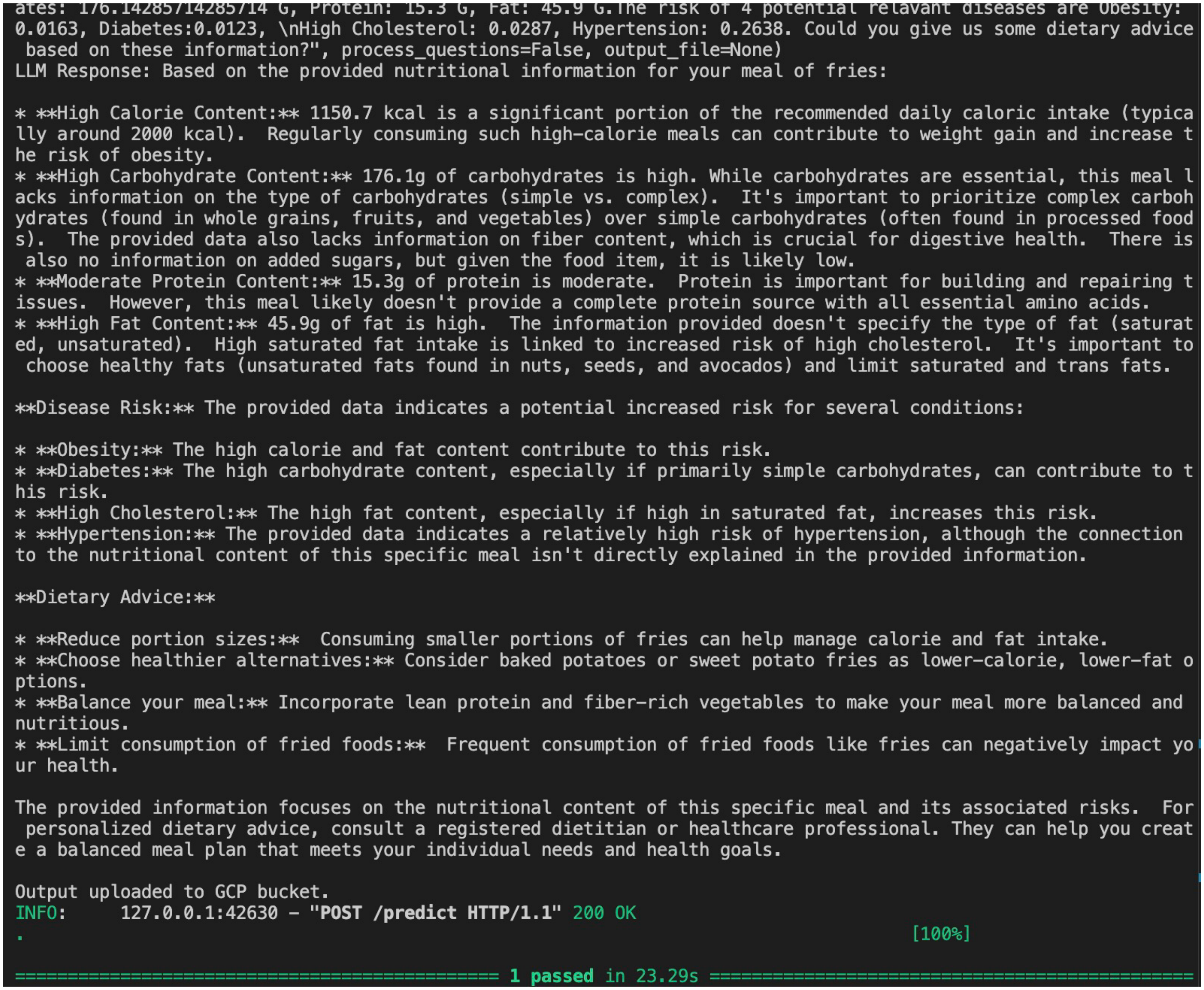
Expected Results



```
 => exporting to image                                                                              5.6s
 => => exporting layers                                                                             5.6s
 => => writing image sha256:be0dc6e7f0f1a0f75bbd7d2ce7b94486702ef9d4612c0edfe67c6fe683f78091        0.0s
 => => naming to docker.io/library/llm-rag-cli                                                      0.0s

 1 warning found (use docker --debug to expand):
 - SecretsUsedInArgOrEnv: Do not use ARG or ENV instructions for sensitive data (ENV "GOOGLE_APPLICATION_CREDENT
IALS") (line 19)

What's next:
    View a summary of image vulnerabilities and recommendations → docker scout quickview
[+] Creating 1/0
 ✔ Container llm-rag-chromadb  Running                                                              0.0s
============================================= test session starts ==============================================
platform linux -- Python 3.12.7, pytest-8.3.3, pluggy-1.5.0
rootdir: /app
plugins: anyio-4.6.2.post1
collected 7 items

tests.py .......                                                                                      [100%]

=============================================== warnings summary ===============================================
../home/app/.local/share/virtualenvs/app-4PlAip0Q/lib/python3.12/site-packages/xgboost/core.py:158
  /home/app/.local/share/virtualenvs/app-4PlAip0Q/lib/python3.12/site-packages/xgboost/core.py:158: UserWarning:
 [18:14:06] WARNING: /workspace/src/collective/../data/../common/error_msg.h:80: If you are loading a serialized
 model (like pickle in Python, RDS in R) or
  configuration generated by an older version of XGBoost, please export the model by calling
  `Booster.save_model` from that version first, then load it back in current version. See:

      https://xgboost.readthedocs.io/en/stable/tutorials/saving_model.html

  for more details about differences between saving model and serializing.

    warnings.warn(smsg, UserWarning)

-- Docs: https://docs.pytest.org/en/stable/how-to/capture-warnings.html
========================================= 7 passed, 1 warning in 24.19s =========================================
chentaiyang@chentaiyangdeMacBook-Pro backend_container % ▯
```



```
loading models
WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built. `model.compile_metrics` w
ill be empty until you train or evaluate the model.
INFO:     Started server process [8]
INFO:     Waiting for application startup.
INFO:     Application startup complete.
INFO:     Uvicorn running on http://0.0.0.0:9000 (Press CTRL+C to quit)
============================================= test session starts ==============================================
platform linux -- Python 3.12.7, pytest-8.3.3, pluggy-1.5.0
rootdir: /app
plugins: anyio-4.6.2.post1
collected 1 item

api/test_service.py INFO:     127.0.0.1:42616 - "GET /docs HTTP/1.1" 200 OK
1/1 ━━━━━━━━━━━━━━━━━━━━ 1s 918ms/step
Step1 output uploaded to GCP bucket.
Fries 0.1658367
Nutrition components:    Description Cosine similarity score          Carbohydrate                   Energy Prot
ein      Fat
0       FRIES                      1.0  176.14285714285714 G  1150.7142857142858 KCAL  15.3 G  45.9 G
Step2 output uploaded to GCP bucket.
Extract input
{'Description': 'FRIES', 'Cosine similarity score': 1.0, 'Carbohydrate': '176.14285714285714 G', 'Energy': '1150
.7142857142858 KCAL', 'Protein': '15.3 G', 'Fat': '45.9 G'}
   Carbohydrate (G)  Energy (KCAL)  Protein (G)  Fat (G)
0        176.142857    1150.714286          15.3     45.9
/home/app/.local/share/virtualenvs/app-4PlAip0Q/lib/python3.12/site-packages/xgboost/core.py:158: UserWarning: [
02:44:31] WARNING: /workspace/src/collective/../data/../common/error_msg.h:80: If you are loading a serialized m
odel (like pickle in Python, RDS in R) or
configuration generated by an older version of XGBoost, please export the model by calling
`Booster.save_model` from that version first, then load it back in current version. See:

    https://xgboost.readthedocs.io/en/stable/tutorials/saving_model.html

for more details about differences between saving model and serializing.

  warnings.warn(smsg, UserWarning)
upload
Step3 output uploaded to GCP bucket.
Formatted Meal Information for LLM Input:
 This is the nutrition content and calories of the user's meal: FRIES.
Energy: 1150.7142857142858 KCAL, Carbohydrates: 176.14285714285714 G, Protein: 15.3 G, Fat: 45.9 G.The risk of 4
 potential relavant diseases are Obesity: 0.0163, Diabetes:0.0123,
High Cholesterol: 0.0287, Hypertension: 0.2638. Could you give us some dietary advice based on these information
?
CLI Arguments: Namespace(chunk=False, embed=False, load=False, chat=True, chunk_type='char-split', query_text="T
his is the nutrition content and calories of the user's meal: FRIES.\nEnergy: 1150.7142857142858 KCAL, Carbohydr
```



```
ates: 176.14285714285714 G, Protein: 15.3 G, Fat: 45.9 G.The risk of 4 potential relavant diseases are Obesity:
0.0163, Diabetes:0.0123, \nHigh Cholesterol: 0.0287, Hypertension: 0.2638. Could you give us some dietary advice
 based on these information?", process_questions=False, output_file=None)
LLM Response: Based on the provided nutritional information for your meal of fries:

* **High Calorie Content:** 1150.7 kcal is a significant portion of the recommended daily caloric intake (typica
lly around 2000 kcal).  Regularly consuming such high-calorie meals can contribute to weight gain and increase t
he risk of obesity.
* **High Carbohydrate Content:** 176.1g of carbohydrates is high. While carbohydrates are essential, this meal l
acks information on the type of carbohydrates (simple vs. complex).  It's important to prioritize complex carboh
ydrates (found in whole grains, fruits, and vegetables) over simple carbohydrates (often found in processed food
s).  The provided data also lacks information on fiber content, which is crucial for digestive health.  There is
 also no information on added sugars, but given the food item, it is likely low.
* **Moderate Protein Content:** 15.3g of protein is moderate.  Protein is important for building and repairing t
issues.  However, this meal likely doesn't provide a complete protein source with all essential amino acids.
* **High Fat Content:** 45.9g of fat is high.  The information provided doesn't specify the type of fat (saturat
ed, unsaturated).  High saturated fat intake is linked to increased risk of high cholesterol.  It's important to
 choose healthy fats (unsaturated fats found in nuts, seeds, and avocados) and limit saturated and trans fats.

**Disease Risk:** The provided data indicates a potential increased risk for several conditions:

* **Obesity:** The high calorie and fat content contribute to this risk.
* **Diabetes:** The high carbohydrate content, especially if primarily simple carbohydrates, can contribute to t
his risk.
* **High Cholesterol:** The high fat content, especially if high in saturated fat, increases this risk.
* **Hypertension:** The provided data indicates a relatively high risk of hypertension, although the connection
to the nutritional content of this specific meal isn't directly explained in the provided information.

**Dietary Advice:**

* **Reduce portion sizes:**  Consuming smaller portions of fries can help manage calorie and fat intake.
* **Choose healthier alternatives:** Consider baked potatoes or sweet potato fries as lower-calorie, lower-fat o
ptions.
* **Balance your meal:** Incorporate lean protein and fiber-rich vegetables to make your meal more balanced and
nutritious.
* **Limit consumption of fried foods:**  Frequent consumption of fried foods like fries can negatively impact yo
ur health.

The provided information focuses on the nutritional content of this specific meal and its associated risks.  For
 personalized dietary advice, consult a registered dietitian or healthcare professional. They can help you creat
e a balanced meal plan that meets your individual needs and health goals.

Output uploaded to GCP bucket.
INFO:     127.0.0.1:42630 - "POST /predict HTTP/1.1" 200 OK
.                                                                                          [100%]

============================================== 1 passed in 23.29s ==============================================
```