King Saud University

College of Computer and Information Sciences

Department of Information Technology

IT 362 Course Project

Semester-1, 1447H

Phase 1

# Saudi food cuisine multilabel classification

Prepared by

| Section No. | Name | ID |
|---|---|---|
| 56594 | Walah Alsaeed | 444201689 |
| | Norah Alshamsan | 444200817 |
| | Joud Alzahrani | 444200862 |
| | Norah Alkathiri | 444200439 |
| | Aljudi Altorkistani | 444203157 |

Supervised by

Khulood Alyahya

IT 362 Course Project

Semester-1, 1447H

# 1. Introduction

In today's world, food delivery platforms, restaurant menus, and recipe websites all serve as valuable resources for information about local cuisines. Global cuisines such as American, Italian, Chinese are all extensively researched in large datasets, while the Saudi cuisine remains underexplored in that department. Dishes like Jareesh, Kabsa, and Mandi are some examples of the most prevalent dishes in Saudi Arabia, yet we lack structured datasets and tools that are able to properly distinguish these dishes in terms of identification and classification.

Our goal in this project is to address this problem and develop a structured dataset along with a computer vision model that can correctly recognize these dishes and their components. We can achieve this by using scraping techniques on recipe websites and online menus to get images and their descriptions, this process helps in developing a model that not only detects the Saudi dish but is also able to identify the main components of it (e.g. rice, chicken, salad). This model can be beneficial in developing nutritional-based applications and food delivery systems.

The main research question in this project is:

**How can we utilize computer vision techniques by applying them to images of Saudi food to automatically identify the dish and its components?**

# 2. Data Sources

**Data Sources:**

The dataset has been collected from many sources (e.g. Restaurant menus, Recipe websites), these sources were the best for collecting our data since it provided an image along with a textual description of it. Specific examples of scraped sources include:

Restaurant websites: Shawaya House, Najd Village, Al-Saudi

Recipe Websites: Sayidaty, Fufu's kitchen

Our intention behind selecting a wide range of sources was to include the greatest number of resources to ensure diversity in our dataset in terms of executing Saudi dishes in commercial or home-cooking context.

**Features:**

Item name (Categorical\Text): the name of the dish.

Ingredients (Text): that were used in the making of the dish.

Scraping Date (Timestamp): The date of when the scraping technique was applied.

Image url

**Data types:**

The dataset is multimodal in general; it contains images which will be used as input during the computer vision process and textual metadata which will be used for classification and labeling since it describes the scraped data.

**Evaluation of dataset biases:**

1. Cultural Bias: Many sources (especially restaurants) tend to focus on modernized versions of Saudi dishes, which can be misleading and unauthentic.
2. Image Bias: Photos are highly styled, meaning the background can include many decorations and bright lighting. Which highly differs from reality, affecting the model's generalization to "real-world" food images.
3. Class imbalance: Certain ingredients are extremely prevalent in Saudi dishes (e.g. Rice, Chicken), while others are rare to find. This imbalance could bias the model toward frequent dominant classes

## 3. Objectives

**Insights:**

1. Dish ingredients trend: Understand which ingredients are used the most in the Saudi cuisine.
2. Dish category popularity: Identify which category of dishes dominate restaurant menus and recipe websites the most.

# IT 362 Course Project

Semester-1, 1447H

**Questions:**

1. Main research question: **How can we utilize computer vision techniques by applying them to images of Saudi food to automatically identify the dish and its components?**
2. Category question: Which food categories (e.g. rice-based, dessert, meat-based) appear the most in Saudi cuisine dishes?

**Tasks:**

1. Automatic dish classification: Train and test models the differentiate between different Saudi dishes.
2. Dish ingredient detection: build multi-label classifying model that can correctly identify the ingredients of a dish.

## 4. Method

To achieve the objectives stated above, we plan to combine computer vision and textual analysis within a multimedia dataset. The process will follow these steps:

1. **Data Preprocessing**
   - Prepare dish images for model training (resizing, normalization, augmentation).
   - Align each image with its corresponding textual description for easier cross-reference during analysis.
2. **Automatic Dish Classification**
   - Train a supervised model to distinguish between different Saudi dishes.
   - Evaluate performance using accuracy and confusion matrix to identify strengths and weaknesses.

3. **Dish Ingredient Detection (Multi-Label Classification)**
   a. Develop a multi-label model capable of detecting multiple components within a single dish image (e.g., rice + chicken + salad).
   b. Evaluate model performance to measure the accuracy

4. **Insights Extraction**
   a. **Dish Ingredient Trends**: Aggregate predictions to identify the most common ingredients in Saudi cuisine.
   b. **Dish Category Popularity**: Group dishes into broader categories (rice-based, meat-based, desserts, etc.) and compare their frequencies.

5. **Bias and Generalization Analysis**
   a. Assess how dataset imbalances (e.g., overrepresentation of rice/chicken dishes) affect model performance.
   b. Conduct controlled tests to check whether the model generalizes well across different image styles (e.g., restaurant photography vs. home-cooked photos).

## 5. Challenges Faced

1. **Web scraping of dynamic websites:** Many of the data sources providing Saudi dishes load their websites dynamically by using JavaScript, which created a challenge for us since the only way was to use tools like Selenium or Playwright, which we weren't fully familiar with. However, with more practice on these tools, we could become more comfortable handling dynamic content and automate the extraction process more effectively in future projects.

2. **Limited dish information:** some sources provide the name of the dish by itself without any listings of ingredients; to solve this issue we will have to manually classify each dish image which is very time consuming

3. **Image quality and styling:** as mentioned before, scraped images were highly stylized with decoration and lighting, which would add an extra step where we need to crop this styling to get the cleanest version of the image

4. **Dish class imbalance:** Certain ingredients like rice and chicken appear frequently in Saudi dishes, this can cause the data to be skewed and raises the risk of model biases.