

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC
RESEARCH

FERHAT ABBAS UNIVERSITY OF SETIF

FACULTY OF SCIENCES
DEPARTMENT OF COMPUTER SCIENCE

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE
DEGREE OF MASTER IN CYBERSECURITY

Theme

**Aklee, AI-Based Mobile App for Nutrition
Label Classification**

Presented by:

Bourouba Mohamed El Khalil

Supervisor:

Pr. Toumi Lyazid

June 2025

Contents

Acknowledgments

Dedication

Abstract

1 Introduction

1.1 Context and Motivation 7

1.2 Objectives 8

1.3 Methodology 9

1.3.1 Image Acquisition 9

1.3.2 Text Extraction using OCR 9

1.3.3 Data Parsing and Cleaning 9

1.3.4 Nutritional Analysis using AI Model 9

1.3.5 Output Generation 10

1.3.6 Feedback and Continuous Learning (Optional) 10

1.4 Related Work 10

1.5 Optical Character Recognition in Mobile Apps
11

2 Machine Learning in Nutrition Analysis 12

2.1 Dataset Preparation 12

2.2 Label Encoding and Data Splitting 13

2.3 Comparative Analysis of Machine Learning Algorithms
... 13

2.3.1 Performance Overview 13

2.3.2 Detailed Algorithm Analysis 13

2.3.3 Why CatBoost Outperforms Other Algorithms 14

2.3.4 Model Selection: CatBoost Classifier 15

2.3.5 Prediction and Evaluation 16

2.3.6 Confusion Matrix 16

2.3.7 Feature Importance Analysis 16

2.4 Step 1: Data Loading and Renaming 17

2.5 Step 2: Feature and Target Separation 18

2.6 Step 3: Label Encoding 18

2.7 Step 4: Splitting the Dataset 18

2.8 Step 5: Model Training using CatBoost
19

2.9 Step 6: Model Prediction 19

2.10 Step 7: Model Evaluation	20
2.11 Step 8: Confusion Matrix Visualization	20
2.12 Step 9: Feature Importance	20
2.13 Step 10: Model and Encoder Serialization	21
2.14 Model Export and Deployment	21
2.15 Conclusion	21
3 Methodology	22
3.1 Architecture Overview	22
3.2 Class and Sequence Diagram Explanation: From Login to Nutrition Prediction	22
3.2.1 Authentication Component	23
3.2.2 OCR Component	23
3.2.3 Comment Component	24
3.2.4 Class Relationships Summary	25
3.3 Sequence Diagram Explanation	25
3.3.1 Detailed Steps	25
3.3.2 Purpose of the Diagram	26
3.4 Followup diagrams	27
3.4.1 User Interaction Workflow (Use Cases)	29
3.4.2 System Integration Points	30
3.5 Tesseract OCR	39
3.6 OCR Accuracy Analysis: Nutritional Label Variations	42
3.6.1 Key Differences and OCR Challenges	42
3.6.2 Integration with Tesseract OCR Pipeline	44
3.6.3 OCR Performance Analysis	44
3.7 Library Imports and Initialization	46
3.8 Model Loading and App Setup	47
3.9 Nutrition Input Schema	47
3.10 Prediction Endpoint	48
3.11 Test Endpoint	49
3.12 CORS Configuration	50
3.13 Conclusion	50
4 Dockerization	53

4.1 Overview of Tools	53
4.2 Docker Compose Services Description	54
4.3 Networks and Volumes	56
5 Conclusion	57

Abstract

With the growing interest in healthy eating and digital solutions, the need for accessible tools that assist users in evaluating food products has increased. In this work, we introduce Aklee, a mobile application that leverages artificial intelligence and optical character recognition (OCR) to classify food products based on their nutritional labels.

Our system uses Tesseract OCR to extract text from scanned product images and applies a Catboost classifier to determine the nutritional quality. The application is built with a modern architecture using React Native for the mobile frontend, Laravel and FastAPI for backend services, and containerized using Docker.

The goal is to provide a user-friendly and efficient tool that aids consumers in making healthier food choices through real-time analysis.

Keywords: Nutrition Label, Classification, OCR, CatBoost, Mobile Application, React Native, FastAPI, Docker.