

UNIVERSITI TEKNOLOGI MARA

**OPTIMIZING AQUILARIA SPECIES
CLASSIFICATION FOR
AGARWOOD PRODUCTION
THROUGH DATA ANALYSIS AND
KNN MODELLING**

AMIR HUSSAIRI BIN ZAIDI

Thesis submitted in fulfilment
of the requirements for the degree of
Master of Science
(Electrical Engineering)

College of Engineering

November 2024

ABSTRACT

Aquilaria, a valuable plant species renowned for its production of gaharu, commonly known as agarwood, belongs to the Thymelaeaceae family. Among its 21 species, 17 have been identified to yield agarwood, which holds significant commercial value and finds applications in perfumes, fragrances, incense, and traditional medicine. The current classification of *Aquilaria* species relies on human knowledge such as evaluation, leading to inconsistent results due to subjective perception and variability in sensory characteristics. For instance, the same agarwood sample may be classified differently by experts based on its aroma, resulting in unreliable species identification. In response, this study proposes an innovative approach for the identification of *Aquilaria* species using chemical compounds of agarwood oil extracted from GC-FID and classified by KNN classifier. To initiate the process, the identification of significant chemical compounds involves statistical analyses, including boxplot analysis to observe the distribution of *Aquilaria* species. The compounds present across all *Aquilaria* species are then selected as significant, employing the z-score test for further validation. Fourteen compounds are identified to exist universally, with three deemed significant through the z-score test. The KNN algorithm serves as the model classifier for *Aquilaria* species, with all datasets (1st and 2nd datasets) divided into training (80%) and testing (20%) phases. Performance measures, including accuracy, recall, precision, and specificity, are assessed through the confusion matrix. The results demonstrate the KNN algorithm's successful classification of *Aquilaria* species, achieving 100% accuracy, recall, precision, and specificity in both training and testing phases, meeting performance standard criteria.

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and Most Merciful,

I am filled with gratitude, O Allah, for Your boundless blessings that have guided me to this significant milestone in my academic journey. Your grace and mercy have been my companions, helping me overcome challenges and obstacles along the way. I humbly request Your continued guidance, wisdom, knowledge, and understanding.

May Allah's blessings be upon His final Prophet Muhammad (peace be upon him), his family, and his companions.

I express heartfelt appreciation to my supervisor, Assoc. Prof. Ir. Ts. Dr. Nurlaila Ismail, for her unwavering support, guidance, and invaluable advice throughout the research process. Her expertise, patience, and dedication have played a crucial role in the completion of this thesis. I am also immensely grateful to my co-supervisors, Prof. Ir. Ts. Dr. Hj. Mohd Nasir Taib and Ts. Dr. Zakiah Mohd Yusoff, for their support and valuable contributions.

My sincere thanks go to my colleagues in the ASPRG group for their constructive feedback and stimulating discussions, significantly enhancing the quality of this thesis. Their insights have been vital in shaping my ideas and strengthening my arguments.

I dedicate this thesis to my beloved parents, Zaidi bin Mohd Zin and _____, and my family members who have supported and encouraged me throughout my academic journey. Their unwavering love, guidance, and support have been the driving force behind my success. I am truly blessed to have such a wonderful family inspiring me to be my best.

Lastly, I extend my sincere appreciation to Noor Farah Nadira and my dearest friends and housemates for their moral and physical support. Their contributions, both direct and indirect, have been instrumental in the successful completion of this thesis. I am truly blessed to have such wonderful people in my life, and I pray that Allah SWT rewards them abundantly for their kindness and generosity. Alhamdulillah.

TABLE OF CONTENTS

	Page
CONFIRMATION BY PANEL OF EXAMINERS	ii
AUTHOR'S DECLARATION	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
TABLE OF CONTENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	xi
LIST OF SYMBOLS	xiii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1 INTRODUCTION	15
1.1 Research Background	15
1.2 Problem Statement	17
1.3 Research Objectives	18
1.4 Scope of Research	19
1.5 Significance of Study	19
1.6 Thesis Layout	19
CHAPTER 2 LITERATURE REVIEW	21
2.1 Introduction	21
2.2 <i>Aquilaria</i> and Agarwood	21
2.3 Exploration of <i>Aquilaria</i> Realm and Agarwood Grading	23
2.4 The Market Trading of <i>Aquilaria</i> based on their Agarwood.	24
2.5 Agarwood oil	25
2.5.1 The chemical composition of Agarwood oil	25
2.5.2 Analysis techniques for agarwood oil	27
2.5.3 Data collection using GC-FID	28
2.6 Data Visualization and Organization	29
2.6.1 Boxplot Analysis	29

CHAPTER 1

INTRODUCTION

1.1 Research Background

Aquilaria is species of tree in a family of Thymelaeaceae family, that well known for their production of gaharu or famous known as agarwood [1], [2]. This type of tree has a vast number of distributions encompassing the whole world including Malaysia. In Malaysia, *Aquilaria* trees is locally known as depu or engkaras [1] and the product from this tree is classified as a highly valuable non-timber products in West Asia and Middle East [1].

Aquilaria species, also known as agarwood or oud trees, have a rich historical and cultural significance in Southeast Asia and parts of the Indian subcontinent. These trees are highly valued for their resinous wood, which is commonly referred to as agarwood or oud [3]. Agarwood has been used for centuries in the production of perfumes, incense, traditional medicines, and religious ceremonies [1], [4], [5].

The demand for agarwood has consistently outpaced its supply, resulting in its classification as one of the most expensive and sought-after aromatic substances in the world. This high demand, coupled with unsustainable harvesting practices and habitat loss, has led to the endangerment of many *Aquilaria* species [6]–[8]. As a result, the conservation and sustainable management of these species have become imperative to ensure their survival.

Traditionally, the classification of *Aquilaria* species has relied on botanical characteristics such as leaf morphology, bark features, and geographical distribution [9]–[11]. However, these methods have limitations, particularly when dealing with closely related species or hybrids that exhibit subtle variations in their physical attributes. These limitations in accurately identifying *Aquilaria* species can, in turn, affect the quality grading of agarwood derived from these trees, as the agarwood's characteristics are closely linked to the specific *Aquilaria* species it originates from. Consequently, there is a need for more accurate and reliable methods of this species classification.