



UTM

UNIVERSITI TEKNOLOGI MALAYSIA

UNIVERSITI TEKNOLOGI MALAYSIA
FACULTY OF COMPUTING
SEMESTER II, SESION 2022/2023

SECB3203: PROGRAMMING FOR BIOINFORMATICS

PROGRESS 1

SECTION: 01

LECTURER'S NAME: DR NIES HUI WEN

STUDENTS	MATRIC NUMBER
MUHAMMAD HARITH ZAEFF BIN KHAIRUL NIZAM	B22EC0007
MEGAT ZARUL AKMAL BIN MOHD ZAHIR	B22EC0070

Contents

Introduction:	2
Problem Background:	4
Aim & Objectives	6
Research Scope	7
Hardware and software requirements	7
Flowchart	8
Summary	9

Introduction:

Colorectal cancer, colloquially referred to as colon cancer, stands as a prevailing global health concern, casting a significant shadow of morbidity and mortality across the world. It's a malady that knows no borders, affecting diverse populations in unique ways. However, when we shift our focus to the context of Malaysia, it becomes evident that the detection and diagnosis of colon cancer constitute an especially formidable challenge within the realm of healthcare. In this dynamic landscape, the critical importance of detecting colon cancer at an early and treatable stage looms large. The timeliness of diagnosis is a matter of paramount concern, for it bears a direct and profound impact on patient outcomes and the allocation of precious healthcare resources.

As degree students immersed in the world of healthcare, we are acutely aware of the pivotal role early detection plays in the prognosis of colon cancer patients. It can spell the difference between life and death, significantly influencing the quality of life and long-term survival prospects. Recognizing the urgency of this issue, we embark on a comprehensive case study that delves into the specific nuances of colon cancer detection within the intricate tapestry of the Malaysian healthcare landscape.

Our mission is clear: to identify the unique factors that exert their influence on early diagnosis and the subsequent effectiveness of medical interventions. By doing so, we aim not only to gain a deeper understanding of the challenges and opportunities within our healthcare system but also to contribute to the ongoing efforts to improve the response to this critical health concern in our country. Through rigorous research and analysis, we hope to shed light on the intricacies of colon cancer detection, ultimately forging a path toward better outcomes and enhanced patient care. In an era where healthcare is a cornerstone of societal well-being, we recognize the significance of our work in the broader context of global health and wellness.

Problem Statement:

The crux of the matter underpinning this case study revolves around the uneven success rate of early colon cancer detection in Malaysia. Early diagnosis is a linchpin in the quest for improved patient outcomes, yet it's a variable that remains frustratingly inconsistent and suboptimal within the Malaysian healthcare framework. This inconsistency in timely detection is a concern of profound significance, with consequences that ripple throughout the healthcare system.

When colon cancer is diagnosed at a later stage, it often necessitates more aggressive and complex treatments, and unfortunately, the odds of survival decrease significantly. These advanced-stage cases also result in a considerable upswing in healthcare costs, imposing additional burdens on both the healthcare system and the affected individuals. As degree students in the healthcare field, we're acutely aware of the multifaceted challenges and ramifications associated with these delayed diagnoses.

Therefore, uncovering the underlying reasons for this variability in early detection success is not merely an academic pursuit but a mission with tangible and far-reaching consequences. Our objective is clear: to gain comprehensive insights into the root causes of this inconsistency, thereby providing a foundation for targeted interventions aimed at enhancing patient outcomes and optimizing the allocation of healthcare resources. The implications of our research are profound, and we are committed to making a meaningful contribution towards a more effective and equitable healthcare system in Malaysia. In a healthcare landscape where time is often the difference between life and death, our work takes on heightened importance, underscoring the urgency of this investigation

Problem Background:

Colorectal cancer in Malaysia is a multifaceted health challenge that is influenced by various factors specific to the country. Several elements contribute to the difficulties in early detection of colon cancer:

1. **Limited Awareness and Education:** There exists a lack of public awareness and understanding of colon cancer, leading to lower participation in screening programs and reduced knowledge about the disease.
2. **Inadequate Screening Programs:** Colon cancer screening, including colonoscopy and faecal occult blood tests, is not as widely implemented or accessible as necessary to ensure early detection.
3. **Cultural and Language Diversity:** Malaysia's diverse population includes various ethnicities and languages, which can pose communication challenges, potentially affecting healthcare access and awareness of colon cancer risks.
4. **Healthcare Disparities:** Socioeconomic factors and regional disparities affect access to healthcare and resources, further complicating early detection efforts.
5. **Diet and Lifestyle Choices:** The Malaysian diet, characterised by its unique culinary traditions, may contribute to the risk of colon cancer. Additionally, lifestyle factors, such as physical inactivity, can influence disease development.

6. Resource Constraints: Malaysia's healthcare system faces resource constraints, including shortages of medical personnel and equipment, which can hinder early detection and timely interventions.

In summary, the variable success rate in detecting colon cancer at an early stage within Malaysia is a pressing issue, and understanding the underlying factors specific to the country is essential. This case study endeavours to explore these factors comprehensively, with the ultimate aim of proposing strategies and recommendations to enhance early detection rates and improve the chances of successful treatment for colon cancer in Malaysia.

Aim & Objectives

The main objective of this study is to design and implement classification models that demonstrate high levels of accuracy in colon cancer diagnosis using advanced machine learning techniques. This goal is driven by the urgent need for more effective and efficient methods to identify colon cancer at an early stage, which is critical for improving patient outcomes and reducing the burden. common to this disease.

Early detection of colon cancer using machine learning:

- The first goal of this study is to focus on early detection. This involves developing machine learning models capable of identifying the presence of colon cancer at an early stage. By analysing diverse clinical and genetic data, these models aim to identify subtle patterns and markers associated with colon cancer before symptoms start to develop
- This aims to contribute to early intervention and rapid treatment of colon cancer, which can significantly increase a patient's chances of successful recovery.

Evaluate the efficiency of colon cancer classification based on machine learning architecture:

- This evaluation includes rigorous testing, cross-validation, and performance metrics to evaluate the effectiveness of the developed models.
- This goal is necessary to understand how the developed model are applied to the real-world and determine their potential to be use as clinical tools.

Research Scope

This study will be focused on clinical and genetic data related to colon cancer patients and individuals at risk of developing colon cancer. Data for the past 5 years will be taken into account, from 2018 to 2023. The machine learning method is applied for colon cancer classification. Primary data sources will include patient medical records, genetic sequencing data, and demographic information. Area of study will be conducted within the country of Malaysia.

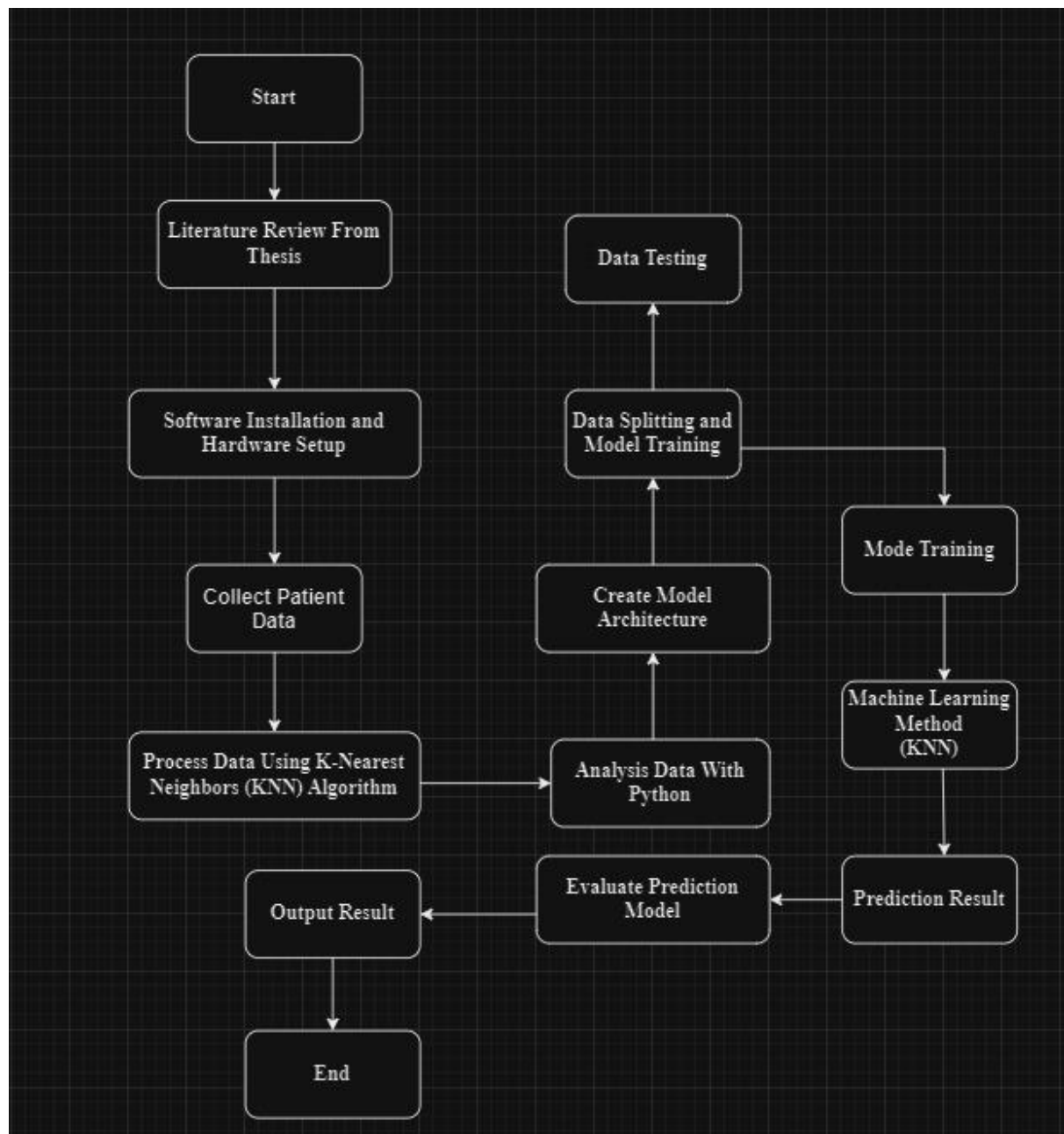
The research will use advanced data analysis techniques including machine learning algorithms that are k-Nearest-Neighbors (KNN). Method that will be used include data preprocessing, feature selection, model development, cross-validation, and performance evaluation. It will follow a systematic process to create, test and validate machine learning models for colon cancer classification. The approach of this research is to apply machine learning algorithms, especially KNN, to analyse and classify the collected data. The goal is to create models that can accurately identify colon cancer cases.

Hardware and software requirements

The hardware that will be used in this research is matching the requirements needed and is important in order to ensure the process of designing and developing the K- Nearest Neighbours (KNN) for colon cancer classification. Hardware requirements for this research is a laptop with Intel Core i5 9th gen processor. The operating systems are Windows 11 with 8GB RAM memory

Besides, the software required are XLSTAT program, a software that can run (KNN). Its a very simple, easy-to-understand, and versatile machine learning algorithm. It's used in many different areas, such as handwriting detection, image recognition, and video recognition. KNN is most useful when labelled data is too expensive or impossible to obtain, and it can achieve high accuracy in a wide variety of prediction-type problems. Besides that, Microsoft Word 2019 are used for documentation and Microsoft Excel 2019 used for data management.

Flowchart



Summary

Colorectal cancer, commonly known as colon cancer, is a pressing health problem in Malaysia, with uneven success rates in early detection. The main problem is difficulties in detecting colon cancer at early stage, leading to aggressive treatments, reduced survival rates, and increased health care costs. Contributing factors include limited awareness, inadequate screening programs, cultural diversity, health care disparities, dietary choices, and limited resources. The main goal of the research is to develop accurate machine learning models for early diagnosis of colon cancer, with the aim of improving early intervention and treatment. Focusing on data from the past five years and using advanced data analysis techniques, this study aims to improve early detection rates and treatment outcomes for patients in Malaysia.