CHAPTER 1 INTRODUCTION

1.1 Research Background

One of the most common problems in society is health problems. Among them, nutritional disorders are a common concern. Obesity is one of the major nutritional problemsfaced by the global community today. It is a leading cause of non-communicable diseases (NCDs) such as stroke, coronary heart disease, and diabetes mellitus. Because obesity is a major risk factor for preventable and treatable NCDs, addressing it is critical. Obesity refers to excess body weight caused by an excessive accumulation of fat in the body[1]. Obesity caused by a high intake of foods that are high sugar and fat. If the energy from these foods is not burned or eliminated, it accumulates, which can have toxic effects on the body and also lead to diseases that carry significant risks[2].

Obesity is a multifactorial disease, with evidence suggesting that the interaction between environmental and genetic factors interact to cause most cases. These factors include physical activity, lifestyle, and nutrition factors, such as eating habits and the early introduction of solid foods in infants. Obesity can occur when a person's diet isunbalanced or when they frequently consume unhealthy foods such as junk food, fast food, and high-calorie foods[3]. Obesity can be analyzed through classification in data mining, which use variables such as weight, height, age, and Body Mass Index (BMI) to cluster data and identify obesity levels or categories in a more structured way. One of the classification methods used in data mining is Naïve Bayes. The Naïve Bayes is a classification method based on the popular Bayes probability theorem, commonly

---

[1] Patonah Patonah, Lia Marliani, and Yani Mulyani, "Edukasi Pola Hidup Sehat Kepada Masyarakat Di Kelurahan Manjahlega Kota Bandung Dalam Menanggulangi Obesitas Sebagai Faktor Resiko Penyakit Kardiovaskular," *Amaliah: Jurnal Pengabdian Kepada Masyarakat* 3, no. 2 (2019): 354–61, https://doi.org/10.32696/ajpkm.v3i2.290.

[2] Syifa Arifani and Zulia Setiyaningrum, "Faktor Perilaku Berisiko Yang Berhubungan Dengan Kejadian Obesitas Pada Usia Dewasa Di Provinsi Banten Tahun 2018," *Jurnal Kesehatan* 14, no. 2 (2021): 160–68, https://doi.org/10.23917/jk.v14i2.13738.

[3] Noviasty Reny Hanani Retno, Badrah Siti, "Pola Makan , Aktivitas Fisik Dan Genetik Mempengaruhi Kejadian Obesitas" 14 (2021): 120–29.

used to create simple models, especially in document classification and disease prediction. The Naïve Bayes algorithm is a machine learning algorithm. The Naïve Bayes is often recommended over other algorithms due to its high speed and the fact that the results of data analysis using Naïve Bayes are very effective and accurate[4]. In addition to Naïve Bayes, Kolmogorov-Arnold Networks (KANs) refer to the application of the Kolmogorov-Arnold theorem for data representation and processing in various data mining tasks. KANs is a concept in neural network theory that derives from the function representation theorem known as the Kolmogorov-Arnold theorem. This theorem relates to the ability of neural networks to represent continuous functions more efficiently[5].Due to the many factors and impacts caused by obesity, society must maintain a healthy diet and lifestyle consistently. As time progresses, it has become easier to obtain unhealthy food, yet as a healthy and reasonable human being, it is not acceptable to indulge excessively, especially in eating and drinking. In addition to causing several serious diseases, overindulgence in eating and drinking is also something that is disliked by Allah SWT, as stated in His words in Surah Al-A'raf, verse 31:

۞ يٰبَنِيْٓ اٰدَمَ خُذُوْا زِيْنَتَكُمْ عِنْدَ كُلِّ مَسْجِدٍ وَّكُلُوْا وَاشْرَبُوْا وَلَا تُسْرِفُوْاۚ اِنَّهٗ لَا يُحِبُّ الْمُسْرِفِيْنَ

"O children of Adam, take your adornment at every masjid and eat and drink, but be not excessive. Indeed, He likes not those who commit excess." (Al-A'raf 7:31)

In the verse above, Allah SWT has explained that in the context of worship, He has provided food and drink. Therefore, eat and drink whatever you desire from the halal, good, and nutritious foods and beverages. However, do not be excessive in anything,

[4] Achmad Ridwan, "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus," *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)* 4, no. 1 (2020): 15–21, https://doi.org/10.47970/siskom-kb.v4i1.169.

[5] Ziming Liu et al., "KAN: Kolmogorov-Arnold Networks," 2024, 1–48, http://arxiv.org/abs/2404.19756.

whether in worship by adding methods or proportions or in eating and drinking. Indeed, Allah does not like excessive, meaning He does not bestow His mercy and rewards upon those who overdo anything[6].

This research applies data mining to classify the likelihood of obesity sufferers into four categories: normal weight, overweight, obesity, and underweight. The data used in this study is sourced from the Kaggle platform. This classification not only helps determine an individual's level of obesity but can also serve as a reference for decision-making in future research. In line with the objective of this study, the goal is to compare the accuracy results of the Naïve Bayes method and Kolmogorov-Arnold Networks (KANs) to achieve better and more reliable outcomes.

1.2 Research Problem Formulation

Based on the background problem explanation presented, the following problem statementsareformulated: The high risk of obesity in various community groups leads to several serious diseases, such as coronary heart disease, stroke, and diabetes mellitus, and can even result in death. Therefore, this study aims to evaluate and compare the accuracy of the Naïve Bayes and Kolmogorov Arnold Networks methods.

1.3 Problem Limitations

To keep the formulated problem focused, the following boundaries have been defined for the present research :

- The data utilized in this study is exclusively sourced from the Kaggle platform.

[6] Nutrigenomik Berdasarkan, Tafsiran Ayat, and Surah Al- A Raf, "[THEMATIC ANALYSIS OF NUTRITION AND HEALTHY LIFESTYLE BASED ON NUTRIGENOMIC IN VERSE 31 OF SURAH AL- A ' RAF] ANALISIS TEMATIK PEMAKANAN DAN GAYA HIDUP SIHAT BERASASKAN Ilmu Nutrigenomik Semakin Mendapat Perhatian Sejak Dua Dekad Kebelakangan Ini , Khus," 2021, 152–65.

- The classification of data will be limited to a comparison of the methods of Kolmogorov-Arnold Networks (KANs) and Naïve Bayes.
- The programming language employed is Python, with the use of Jupyter Notebook tools.

1.4 Research Objectives

Based on the research problem formulated above, the objectives of this study are as follows:

To classify obesity data using the Kolmogorov-Arnold Networks (KANs) and Naïve Bayes methods.

1.5 Research Benefits

The expected benefits of this research are as follows:

a. For the Readers

- Can serve as a reference and knowledge source for consideration in future research.
- This can encourage further research on methods that can be used in data mining classification.

b. For the Author

- Able to contribute positively to understanding the health field, which is rich in information.
- Able to provide insight into public understanding of a disease.

c. For the University

- Can showcase the student's ability to master the material learned during their academic studies.
- Can demonstrate the students' capability in applying their knowledge to evaluate their academic progress.