KNN Algorithm to Determine Optimum Agricultural Commodities in Smart Farming

1st Ahmad Cucus
Faculty of Computing
Universiti Malaysia Pahang
Pekan, Malaysia
ahmad.cucus@ubl.ac.id

2nd Al-Fahim Mubarak Ali
Faculty of Computing
Universiti Malaysia Pahang
Pekan, Malaysia
fahim@ump.edu.my

3rd Afrig Aminuddin
Faculty of Computer Science
Universitas Amikom Yogyakarta
Sleman, Indonesia
afrig@amikom.ac.id

4th Yoga Pristyanto
Faculty of Computer Science
Universitas Amikom Yogyakarta
Sleman, Indonesia
yoga.pristyanto@amikom.ac.id

5th Ferian Fauzi Abdulloh
Faculty of Computer Science
Universitas Amikom Yogyakarta
Sleman, Indonesia
ferian@amikom.ac.id

6th Zafril Rizal M. Azmi
Faculty of Computing
Universiti Malaysia Pahang
Pekan, Malaysia
zafril@ump.edu.my

***Abstract*—One of the problems in agriculture is the difficulty of determining the correct type of plant on land with certain conditions. In this study, we tried to apply an algorithm to determine the commodity according to the conditions of the planting area. The idea is to find a match between the variables owned by the environment and the plant profile. The algorithm used in this research is KNN, which was chosen because the variable has a numeric data type to perform the matching. However, sometimes, the matching does not only use variable data types. In some cases, the available variables are string data types. In this study, the researchers tried to improve the matching method on KNN to adjust the string data type variables. At the end of this study, data visualization showed the types of plants that match the case examples from a field. It proves that the proposed method can classify string data types.**

Keywords—Smart farming, KNN Algorithm, Computational intelligence

I. INTRODUCTION

Agriculture is a complicated scientific discipline involving various knowledge, abilities, techniques, and procedures that computerized systems can efficiently assist. There have been numerous attempts to create an automated farming framework that can regulate incoming data and processes appropriately [1]. The world's population and food consumption are expanding quickly, and the impacts of climate change make it more challenging to guarantee food security sustainably [2]. Because developing more innovative agricultural concepts involves methods that use technology and data, data-driven farming is one of the critical concepts and tactics suggested for enhancing production effectively while minimizing its environmental impact. Machine learning is a part of artificial intelligence that works mechanically or gives commands for a selected task. Machine learning aims to understand schooling statistics and suit that training statistics or the data into designs on the way to be beneficial to people. Using a vast quantity of training statistics or data can support making correct decisions and finding the precise outcome [3].

The latest concept in the smart industry in agriculture is smart farming. Smart farming is a concept of agricultural management using modern technology to increase the quantity and quality of agricultural products, a new trend in agricultural technology. The way farmers run their businesses is changing, as in many other sectors. Farmers have more control over cultivating and maintaining crops thanks to machinery, software, and genetic advancements. Agriculture's digitization is an example of a smart industry. Industry 4.0 is a revolution based on digitization, altering production processes, and altering business models to speed up and improve production. It also involves integrating systems employing digitization, from customer requirements to finished goods [4]. Smart farming technologies have greatly aided the agricultural industry's development [5]. The Big Data phenomenon, which refers to an enormous volume of data with various variations that may be recorded, analyzed, and used for decision-making, includes smart farming [6]. Smart farming extends the precision farming concept by adding contextual, situational, and site awareness to real-time management and decision-making tasks [7]. In terms of productivity and sustainability, precision farming (or smart farming) can significantly boost agricultural production [1].

The Key Nearest Neighbor Algorithm (KNN Algorithm) is an algorithm for categorizing data based on learning data acquired from the parameter values owned by the nearest data. In this study, we apply the KNN Algorithm to solve the problem [8]. This approach aims to calculate the separation between the case variables and the variables in the dataset [9]. We add a variable with the data type string in our dataset to test our proposed algorithm.

In this study, we used an example of data on the condition of the land to be planted with one type of plant. In many cases, the matching technique that is carried out using the KNN algorithm uses numeric data type variables because the Euclidean value is obtained from the calculation process with numeric data type variables. However, some data is presented in string data types, such as land and pest types. Euclidean value cannot be calculated when the data type is not numeric. The main contributions of this research are that we try to apply 'string' data type matching techniques, which then we will calculate the Euclidean value between the dataset variable and the case data variable. This technique is used for automatic or semi-automatic data analysis activities, and the data set is then extracted to get interesting patterns, some previously unknown, such as clusters (cluster analysis), anomaly, and dependency detection. The data in this study is a collection of knowledge, and the collection of knowledge in the database is an interdisciplinary field of computer science [10]. This paper presents several sections, such as Introduction: in this section, we write about our research topic, background information, and research objectives. We also clearly explain the research