CHAPTER 1 INTRODUCTION

1.1 Research Background

Drinking water is a basic necessity that plays an essential role in maintaining human health and supporting various daily activities. Drinking water can be obtained from various sources, including bottled drinking water, refill drinking water, piped water, protected wells, and protected springs. Protected wells and springs should be located at least 10 meters from the nearest waste, garbage, or sewage disposal site (Susanti, 2020). Among these options, refill drinking water has become one of the preferred choices because it is practical and relatively more affordable (Suryani and Kusumayati, 2022). In Indonesia, the mandatory drinking water parameters are regulated by the Regulation of the Minister of Health of the Republic of Indonesia No. 2 of 2023. According to this regulation, the permissible Total Dissolved Solids (TDS) concentration is less than 300 mg/L or ppm, while the pH value must remain within the range of 6.5 to 8.5 (Kemenkes, 2023). Despite these standards, users of refill drinking water do not have continuous information on whether the TDS or pH values remain within the permissible range or not. Monitoring drinking water parameters is important to provide users with information about quality of the drinking water they consume daily. The Internet of Things (IoT) enables drinking water parameters to be monitored automatically and in real time. IoT allows sensor measurement data to be transmitted to and displayed on digital platforms, making the information accessible anytime and anywhere (Hartawan and Sudiarsa, 2019).

Water parameter monitoring systems have been developed by previous researchers. Lestari and Zafia (2022) has developed water quality monitoring on the PAMSIMAS (Penyedia Air Minum dan Sanitasi Berbasis Masyarakat) system in Cendana Village using turbidity, TDS and pH sensors. Monitoring data is displayed on the ThingSpeak website and measurement history is stored using a data logger. The system achieved a performance success rate of 86,6%. The developed system only provided information after the water quality parameters had exceeded the permissible limits. It was not equipped with a prediction feature to estimate when the parameters would exceed the permissible limits, nor with a notification feature to provide early warnings to users. Prasetyo et al. (2025) has developed a water quality monitoring system using three sensors, namely TDS, temperature and pH sensors. Water quality was classified using the Sugeno Fuzzy Logic method. The measurement results were displayed on a self developed website, and notifications were sent via Telegram when water quality deteriorated. The system remained focused on reporting the current water quality condition and did not include a prediction feature to estimate when the water parameters would exceed the permissible limits.

Previous studies have only provided information about water quality at the time the measurements were taken. The information is useful for determining the current condition of the water but does not provide an estimate of when the water parameters will reach the permissible limits. As artificial intelligence advances, machine learning has been increasingly adopted. It can identify relationships among parameters and temporal patterns that are difficult to detect from instantaneous measurements. This capability opens up opportunities to develop systems that not only monitor but also predict future conditions. Pang et al. (2025) developed a time series based prediction model for four water quality parameters in a drinking water treatment plant. The study employed the Extreme Gradient Boosting (XGBoost) model. The results demonstrated that the machine learning model combined with time feature engineering achieved better performance than the deep learning models. The developed model predicted the values of four water quality parameters by incorporating a 12-hour time lag to capture the delayed effects of reagent addition. The study focused on predicting the values of water quality parameters at specific points in time. It did not estimate the time required for the parameters to reach the permissible limits, nor with a notification feature to provide early warnings to users. Information on the estimated time until a parameter exceeds the permissible limits is expected to support more informative early warning systems and facilitate more effective to making decision.

Based on the previous problems, this study developed a prototype of an IoT and machine learning based system to predict the time until TDS $\geq$ 300 ppm or pH falls outside the permissible range of 6.5 to 8.5 in refill drinking water. The prediction utilizes patterns in temperature, TDS, and pH observed during the monitoring process. The prediction results are integrated into a mobile application, enabling users to receive information about changes in parameter conditions before they exceed permissible limits anytime and anywhere. The developed system is not intended to determine the suitability of drinking water for consumption or to replace laboratory testing. The novelty of this study lies in the development of a system capable of predicting the time until pH or TDS exceeds the permissible limits while providing early warning notifications.

1.2 Research Purpose

This study aims to develop an IoT and machine learning based prototype system to predict the time until TDS $\geq$ 300 ppm or pH falls outside the 6.5 to 8.5 range in refill drinking water.

1.3 Research Benefits

The benefits of this study are as follows:

1. Provide refill drinking water users with information on the predicted time until TDS $\geq$ 300 ppm or pH falls outside the 6.5 to 8.5 range.
2. Contributes to the advancement of instrumentation physics through the implementation of an integrated monitoring system incorporating sensors, IoT, machine learning, cloud databases, and mobile application.

1.4 Research Scope and Limitations

The scope and limitations of this research are as follows:

1. The object of this research is refill drinking water used in indoor household environment.
2. Machine learning is used to predict the remaining time until TDS $\geq$ 300 ppm or pH falls outside the range of 6.5 to 8.5.

3. The machine learning models are developed using temperature, TDS, and pH data as input features.
4. The machine learning models evaluated in this study are Random Forest Regressor, XGB Regressor, and Support Vector Regression (SVR).
5. The sensors used are the TDS SEN0244 sensor, the DS18B20 temperature sensor and the pH 4502C sensor.
6. The system only monitors three parameters, namely temperature, TDS, and pH.
7. The system does not evaluate all drinking water quality parameters required by Minister of Health Regulation No. 2 of 2023.
8. The system is not intended to determine the overall suitability of drinking water or to replace laboratory testing.
9. The prediction results presented in this study are based on the monitoring data collected under the experimental conditions of the developed prototype. Therefore the observed time required for pH or TDS to exceed the permissible limit should not be interpreted as a general lifetime of refill drinking water.