Preface

According to "The United Nations World Water Development Report, 2021: Valuing Water, " global freshwater usage has increased by six times in the last 100 years. More than 80% of wastewater is released into the water streams without proper treatment. Conventional wastewater treatment strategies such as the activated sludge process etc., have significantly impacted wastewater treatment on a large scale. However, these treatment processes have disadvantages such as high energy demand during the aeration process, disposal of secondary sludge, and emission of CO_2. Thus, they are not economically viable for a longer period and large-scale. There is a need to develop a sustainable wastewater treatment process to recover waste as a valuable resource, thus strengthening the circular economy. The biological wastewater treatment process mediated by microalgae is a sustainable approach for treating wastewater with simultaneous recovery of nutrients in the form of microalgae biomass. Microalgae growth in the wastewater is highly influenced by various parameters and environmental conditions such as CO_2 content in the inlet air, temperature, initial inoculum level, pH, light intensity, photoperiod, nutrient concentration in wastewater, and much more, together termed as predictor variables. These variables show mutual interaction with each other and cooperatively affect the treatment process. Therefore, in order to enhance the treatment capability and biomass productivity of microalgae, it becomes essential to optimize these variables and provide their right combination. However, it is practically impossible to test all combinations of variables in a single study. Machine learning algorithms can quickly analyze an extensive dataset and determine the best combinations of predictor variables as desired. Machine learning algorithms play a vital role in analyzing large datasets by detecting significant patterns in the data in a minimal time. Machine learning algorithms can quickly process large dataset and detect different patterns of predictor variables for increasing nutrient

removal capability of microalgae, including nitrogen and phosphorus, simultaneously increasing biomass productivity. Moreover, these algorithms can even generate new information from already reported data which is expected from new experiments. Patterns detected from these algorithms can assist in constructing high throuput experimental designs and assist in carrying out efficient wastewater treatment at larger scale.

In the present analysis, decision tree algorithm was used to determine the effects and the best combination of predictor variables, including microalgae class, pre-cultivation stage deciding factors and operating variables, resulting in high biomass productivity and wastewater treatment capability. Decision tree analysis detected 10 different combinations of predictor variables leading to high nitrogen removal efficiency, 10 combinations for high phosphorus removal efficiency and 8 combinations for increased biomass production. These combinations were tested on recently published experimental findings and nearly 80% accuracy was obtained. The results obtained through machine learning analysis can be used in constructing high throughput experimental designs, which may assist in carrying out the efficient wastewater treatment at large scale. In the continuation of the previous analysis, the decision tree was used to analyse the dataset of class *Trebouxiophyceae* and *Chlorophyceae*. Various combinations of cultivation parameters were determined to enhance their performance in wastewater treatment. Nine combinations of cultivation parameters leading to high biomass production and eleven combinations each for high nitrogen removal efficiency and high phosphorus removal efficiency for class *Trebouxiophyceae* were detected by decision tree models. Similarly, eleven combinations for high biomass production, nine for high nitrogen removal efficiency, and eight for high phosphorus removal efficiency were detected for class *Chlorophyceae*. The results obtained through decision tree analysis can provide the optimum conditions of cultivation parameters, saving time in designing new experiments for treating wastewater at a large scale.

In the next objective, one of the data mining tools, association rule mining has been used to find specific conditions of 11 cultivation parameters for enhancing microalgae growth in wastewater. General rules derived from association rule mining showed that biomass productivity and nutrient removal efficiency can be increased by keeping CO_2 content between 0.53-2.53%, light intensity in the range of 200-1500 $\mu mol\ m^{-2}\ s^{-1}$, initial inoculum level from 0.2-0.4 g/L and N/P ratio nearly 15:1 to 50:1. This extracted information can be used to design future experimental runs and will help in implementation of the process at a large scale without wet laboratory experiments. In the subsequent objective, the models obtained from machine learning analysis were verified both computationally and experimentally. Computational validation of the machine learning models was done by testing the rules extracted from them on some recently reported experimental results. The models were able to correctly classify 8 data points out of 11 for nitrogen removal efficiency, 9 data points out of 11 for phosphorus removal efficiency, and 6 data points out of 8 for biomass production. In experimental verification, models were experimentally verified by using two newly isolated strains that predicted a 10% error after experimental verification.

In the last objective, cost-benefit analysis was performed for the application of LEDs for microalgae cultivation. This objective aims to enlist the applications of light-emitting diodes in microalgal cultivation with reference to internally illuminated photobioreactors coupled with evaluation of the cost and energy balance of the artificial lights. The calculation shows that the electrical energy cost incurred during the application of light-emitting diodes for microalgae cultivation is approximately USD 15.19 kg^{-1} DW. The collective fraction of electrical energy transformed into chemical energy (microalgae biomass) is around 6-8%. The cost of the light-emitting diodes can be decreased by the application of Arduino-based automated control system to control the power supply to LEDs, photovoltaic powered photobioreactors and additional light. These techniques of input cost reduction have been also

explored deeply in the present study. As estimated, they can reduce the cost of light-emitting diodes by 50%.