# CHAPTER 5

# DISCUSSION AND CONCLUSION

## 5.1 INTRODUCTION

In this chapter, we discuss the main findings and implications of the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water using Python. We also provide the limitations and recommendations of the study, and the conclusion and future directions of the research.

## 5.2 DISCUSSION

The main objective of this study was to develop and evaluate the predictive models of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water using Python. The study used the data obtained from Etuk’s library, whose study was based synthesized a nanocellulosic composite from cellulose nanofibrils (CNF) and modified cellulose nanofibrils (MF-CNF) and iron oxide nanoparticles (Fe3O4) and tested its adsorption performance for cobalt II ion in aqueous solution. The study applied three machine learning algorithms, namely ridge regression (RR), Support Vector Regressor (SVR) and random forest regression (RFR).

This project also implemented six kinetic models, namely Adams-Bohart, Thomas, Weibull, Wolborska, Yan, and Yoon-Nelson, to the experimental data to determine the concentration of absorbate on absorbent, and compared their performance and accuracy using various metrics, such as the coefficient of determination (R2), the root mean square error (RMSE), and the mean absolute error (MAE), Hybrid, Marquardt and others.

The main findings of the study are summarized as follows:

* The RR model had the highest performance and accuracy among the machine learning algorithms, with an R2 value of 0.8436, a RMSE value of 0.4052, and a MAE value of 0.276. This indicates that the RR model was able to overcome the overfitting problem of the MLR model by introducing a regularization term that penalized the large coefficients of the model. The RR model was able to fit the data better and generalize better to the unseen data.
* The RFR model had a slightly lower performance and accuracy than the RR model, with an R2 value of 0.8096, a RMSE value of 0.447, and a MAE value of 0.2706. This indicates that the RFR model was able to capture the nonlinear and complex relationship between the input variables and the output variable by using an ensemble of decision trees that learned from different subsets of the data. However, the RFR model was also prone to overfitting and high variance, as it depended on the number and the depth of the trees, and the randomness of the data splitting.
* The Yan model had the highest performance and accuracy among the kinetic models, with an R2 value of 0.99, a SSE value of 0.004346, and a SAE value of 0.169300. This indicates that the Yan model was able to describe the adsorption kinetics of cobalt II ion on the nanocellulosic composite very well, as it considered both the adsorption and the desorption processes, and had two parameters that could be adjusted to fit the data. The Yan model had the lowest errors and the highest correlation among the kinetic models.
* The Wolborska and the Adams-Bohart models had the lowest performance and accuracy among the kinetic models, with R2 value of 0.84, SSE value of 0.302615 and SAE value of 1.826538, respectively. This indicates that the Wolborska and the Adams-Bohart models were not able to describe the adsorption kinetics of cobalt II ion on the nanocellulosic composite very well, as they assumed a constant adsorption rate and a linear relationship between the adsorption efficiency and the contact time, which were not valid for the experimental data. The Wolborska and the Adams-Bohart models had the highest errors and the lowest correlation among the kinetic models.

The main implications of the study are as follows:

* The study demonstrated the feasibility and the usefulness of using Python as a tool for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. Python is a powerful and versatile programming language that has many libraries and modules that can facilitate the data collection, data analysis, and data visualization of the study. Python is also an open-source and free software that can be easily installed and used by anyone who is interested in the study.
* The study provided a comprehensive and comparative evaluation of the machine learning algorithms and the kinetic models for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. The study used various metrics and methods to measure and compare the performance and the accuracy of the models, and to identify the strengths and the weaknesses of the models. The study also provided the graphical and numerical representations of the models and their results, which can help the visualization and the interpretation of the models and their implications.
* The study contributed to the knowledge and the understanding of the adsorption process and the adsorption kinetics of cobalt II ion on the nanocellulosic composite. The study revealed the nonlinear and complex relationship between the input variables (initial concentration and contact time) and the output variable (adsorption efficiency), and the influence and the significance of the model parameters on the adsorption efficiency. The study also revealed the best-fitting and the worst-fitting models for the adsorption kinetics of cobalt II ion on the nanocellulosic composite, and the reasons and the explanations for their performance and accuracy.

## 5.3 LIMITATIONS AND RECOMMENDATIONS

The study had some limitations and challenges that should be acknowledged and addressed. The limitations and recommendations of the study are as follows:

* The study used a limited and specific data set that was obtained from Etuk’s library, the study synthesized a nanocellulosic composite from cellulose nanofibrils (CNF) and iron oxide nanoparticles (Fe3O4) and tested its adsorption performance for cobalt II ion in aqueous solution. The data set had only two input variables (initial concentration and contact time) and one output variable (adsorption efficiency), and had only 10 data points. The data set also had some outliers and noise that could affect the quality and the reliability of the data. Therefore, the study results and conclusions may not be generalizable and applicable to other data sets and other adsorbents and adsorbates. The study recommends using more and diverse data sets that can cover a wider range of input variables and output variables, and that can represent different adsorbents and adsorbates. The study also recommends using more and robust data cleaning and data validation methods that can improve the quality and the reliability of the data.
* The study used only three machine learning algorithms, namely ridge regression (RR), and random forest regression (RFR), Support Vector Regressor(SVR) and six kinetic models, namely Adams-Bohart, Thomas, Weibull, Wolborska, Yan, and Yoon-Nelson, for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. The study did not consider other machine learning algorithms and other kinetic models that could also be suitable and effective for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. Therefore, the study results and conclusions may not be comprehensive and optimal, and may miss some important and relevant aspects and features of the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. The study recommends using more and diverse machine learning algorithms and kinetic models that can capture and explain the nonlinear and complex relationship between the input variables and the output variable, and that can provide more accurate and reliable predictions of the adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water.
* The study used only the coefficient of determination (R2), the root mean square error (RMSE), and the mean absolute error (MAE), Hybrid, Marquardt, Sum of Squared Error (SSE), Non-Linear Chi-Square, and Average Relative Error (ARE) as the metrics for measuring and comparing the performance and the accuracy of the machine learning algorithms and the kinetic models for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. The study did not consider other metrics and methods that could also be useful and informative for measuring and comparing the performance and the accuracy of the machine learning algorithms and the kinetic models for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. Therefore, the study results and conclusions may not be sufficient and valid, and may overlook some significant and critical aspects and factors of the performance and the accuracy of the machine learning algorithms and the kinetic models for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. The study recommends using more and diverse metrics and methods that can evaluate and assess the performance and the accuracy of the machine learning algorithms and the kinetic models for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water.

## 5.4 CONCLUSION AND FUTURE DIRECTIONS

The conclusion of this study is that the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water using Python is a feasible and useful approach that can provide valuable insights and implications for the adsorption process and the adsorption kinetics of cobalt II ion on the nanocellulosic composite. The study showed that the ridge regression model and the Yan model were the best-fitting and the most accurate models among the machine learning algorithms and the kinetic models, respectively, and that they could capture and explain the nonlinear and complex relationship between the input variables and the output variable, and provide reliable and precise predictions of the adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. The study also showed that Python was a powerful and versatile tool that could facilitate the data collection, data analysis, and data visualization of the study, and that it could offer many libraries and modules that could enhance the performance and the quality of the study.

The future directions of this study are as follows:

* The study can be extended and improved by using more and diverse data sets that can cover a wider range of input variables and output variables, and that can represent different adsorbents and adsorbates. This can help to increase the generalizability and the applicability of the study results and conclusions, and to test the robustness and the validity of the study models and methods.
* The study can be expanded and enriched by using more and diverse machine learning algorithms and kinetic models that can capture and explain the nonlinear and complex relationship between the input variables and the output variable, and that can provide more accurate and reliable predictions of the adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. This can help to provide a more comprehensive and comparative evaluation of the study models and methods, and to identify the strengths and the weaknesses of the study models and methods.
* The study can be refined and optimized by using more and diverse metrics and methods that can evaluate and assess the performance and the accuracy of the machine learning algorithms and the kinetic models for the predictive modelling of adsorption efficiency of nanocellulosic composite for the removal of cobalt II ion from waste water. This can help to provide a more sufficient and valid measurement and comparison of the study models and methods, and to address and resolve the issues and problems of the study models and methods.