T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Prediction of Material Properties of Model Organic Biomaterials by Machine Learning
著者(和文)	Kwaria Rudolf Jason
Author(English)	Rudolf Jason Kwaria
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第11580号, 授与年月日:2020年9月25日, 学位の種別:課程博士, 審査員:林 智広,曽根 正人,和田 裕之,柘植 丈治,石田 忠
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第11580号, Conferred date:2020/9/25, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

Powered by T2R2 (Science Tokyo Research Repository)

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

専攻： Department of	物質理工学院材料系 ライフエンジニアリングコース 専攻	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of （理学）
学生氏名： Student's Name	Rudolf Jason Kwaria	指導教員（主）： Academic Supervisor (main)	林　智広　准教授
		指導教員（副）： Academic Supervisor (sub)	教授

要旨（英文 800 語程度）
Thesis Summary（approx.800 English Words ）

In this Ph.D. thesis study, I proposed and discussed a method to optimize the current material design method by employing machine learning algorithm called artificial neural network (ANN) algorithm to predict water contact angle (WCA) and fibrinogen adsorption of self-assembled monolayers (SAMs) as well as making quantitative guide to design SAMs with specific WCA and fibrinogen adsorption. I also discussed the advantage and challenges that needs to be overcome to apply this method.

In case of WCA prediction and construction of quantitative guide, a dataset containing data of WCA and structure descriptors of SAMs were constructed with the purpose of training the ANN algorithm. The dataset was cleaned and preprocessed using various criteria (e.g. same data must not have duplicates, dataset free from outliers, etc.). During the training process, the ANN algorithm itself also underwent some tuning to ensure good prediction accuracy. This involves tuning several parameters of the ANN algorithm. After several iterative process, the ANN algorithm was tasked to make prediction of WCA of SAMs. The prediction of the trained ANN was validated using leave-one-out cross validation (LOOCV) using mean squared error (MSE) and correlation coefficient (R) to measure the prediction accuracy of the ANN algorithm. The validation results show that the ANN can predict WCA with high prediction accuracy. Next, using the trained ANN, the quantified importance of each structure descriptors was calculated to make quantitative guide in which the guide was compared with well-known empirical guide. The result shows that the guide agrees with the well-known empirical guide.

Continued to the next page->

In case of fibrinogen adsorption, I tested the effect of difference experimental approach of each publication towards the prediction accuracy of the ANN algorithm. After making two separate dataset of fibrinogen adsorption and structure descriptors gathered from single laboratory which have standardized measurement procedure (single-lab dataset) and from multiple laboratory which have various measurement procedure (multi-lab dataset). Both datasets underwent several data cleaning and preprocessing using various criteria (e.g. same data must not have duplicates, dataset must be free from outliers, etc.). After the preprocessing step, the parameters for ANN algorithm were tuned before being trained by either dataset. Next, both datasets were used to train the ANN algorithm and the prediction result was validated using LOOCV. The MSE and R of the prediction results shows that ANN trained by single-lab dataset shows higher prediction accuracy than multi-lab dataset. The validation results show that difference of experimental method employed across different laboratories might hinder the predictive performance of the ANN algorithm which means that, whenever possible, the data should be taken under standardized condition. Moreover, the quantitative guide was also created using both datasets. The quantitative guide derived from single-lab dataset shows good agreement with well-known empirical guide to design SAMs with specific fibrinogen adsorption. Interestingly, although the validation result of mixed-lab dataset had low accuracy, the quantitative guide derived from multi-lab dataset also shows several degrees of agreement with well-known empirical guide to design SAMs with specific fibrinogen adsorption. However, as there are still several abnormalities in the quantitative guide, it is still too early to use the quantitative guide derived from mixed-lab dataset.

Lastly, to demonstrate the capability of ANN algorithm for material screening, I made several prediction WCA and fibrinogen adsorption of SAMs outside of the training dataset. The results show good agreement with previously published publication and is also make sense from scientific perspective. Using these results, I can conclude that the ANN algorithm can make material screening which could optimize material design process.

Finally, by looking at the validation results, the quantitative guide constructed using ANN algorithm, as well as the prediction of WCA and fibrinogen adsorption of SAMs outside of the training dataset, it can be concluded that it is possible to optimize the material designing process using ANN algorithm. However, there are challenges in the form of standardizing the experimental method for obtaining data so that it can be used to train the ANN algorithm in the future.

備考 ： 論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。
Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。
Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).