HOKKAIDO UNIVERSITY

Title	Study on Multi-armed Bandit Algorithm for Sequential Experiments to Predict the Best Molecule with Dynamic Feature Selection [an abstract of dissertation and a summary of dissertation review]
Author(s)	Md Menhazul Abedin
Degree Grantor	北海道大学
Degree Name	博士(総合化学)
Dissertation Number	甲第16136号
Issue Date	2024-09-25
Doc URL	https://hdl.handle.net/2115/93524
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	doctoral thesis
File Information	Md_Menhazul_abstract.pdf, 論文内容の要旨

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称　博士（総合化学）　氏名　ムハンマド　メンハズル　アブディン

学 位 論 文 題 名

Study on Multi-armed Bandit Algorithm for Sequential Experiments to Predict the Best Molecule with Dynamic Feature Selection
（動的特徴選択による最良分子を推定する逐次実験のための多腕バンディットアルゴリズムの研究）

The objective of this research is to find the molecule having the highest/lowest quantity of the desired property by chemists such as hydration free energy, and reaction enantioselectivity, solvation affinity etc., by fewer experiments as much as possible via bridging chemistry and information science. I develop the theoretical framework of efficient sequential experiments based on one of reinforcement learning, linear bandit (optimism in the face of uncertainty linear bandit-OFUL) combined with dynamic feature selection along the experiment. As follows I present the outline of the thesis.
In Chapter 1, the purpose of the dissertation and background of the study are described.
In Chapter 2, This chapter is designed to introduce the data we utilized in this research endeavor. The aim of this thesis work is to develop a new algorithm for sequential optimization of molecular property. We develop a modified version of Optimistic in the Face of Uncertainty Linear (OFUL) bandit algorithm along with automatic stopping condition which enables to stop the algorithm after getting the desired molecule immediately. In this regard we consider three real data set and two synthetic data set to validate our developed algorithm. The inclusion of both synthetic and real data aims to demonstrate the method' s versatility.
In Chapter 3, I explain the methods for sequential optimization for molecular property prediction. Chemists find the best property molecule by sequentially testing the molecules usually from simple molecules or available ones from company to more complicated molecules. This is very time consuming, costly in terms of money and labor both, sometimes it is hard to design the sequential experiments optimally if the set of candidate molecules is very large. To solve this issue, sequential optimization is one of the promising approaches (for computational search) in identifying the optimal candidate(s) (molecules, reactants, drugs etc.) with desired properties (reaction yield, selectivity, efficacy etc.) from a large set of potential candidates, while minimizing the number of experiments required.
Molecular property is predicted using the molecular fragment features calculated from the molecular structures. However, the high dimensionality of the feature space (e.g., molecular descriptors) makes it often difficult to utilize the relevant features during the process of updating the set of candidates to be examined.
In this thesis, I developed a new sequential optimization algorithm for molecular problems based on a reinforcement learning, multi-armed linear bandit framework, and on-line, dynamic feature selections in which relevant molecular descriptors are updated along with the experiments. I also designed a stopping condition aimed to guarantee the reliability of the chosen candidate from the data set pool.
In Chapter 4, I show the results and the discussions. The developed algorithm was examined by comparing with Bayesian optimization, using two synthetic datasets and three real datasets in which one dataset includes hydration free energy of molecules, free energy difference between enantiomer products in chemical reaction and wavelength of molecules. We found that the dynamic feature selection in representing the desired properties along the experiments provides a better performance (e.g., time required to find the best candidate and stop the experiment) as the overall trend, and that our multi-armed linear bandit approach with dynamic feature selection scheme outperforms the standard Bayesian optimization (BO) with fixed feature variables. The comparison of our algorithm to BO with dynamic feature selection is also addressed.
In Chapter 5, all the results are summarized, and future plans have been described.