



HOKKAIDO UNIVERSITY

Title	The Analysis of Infectious Diseases via Machine Learning [an abstract of dissertation and a summary of dissertation review]
Author(s)	TESSMER, HEIDI LYNN
Degree Grantor	北海道大学
Degree Name	博士(獣医学)
Dissertation Number	甲第13258号
Issue Date	2018-06-29
Doc URL	https://hdl.handle.net/2115/71244
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/4.0/
Type	doctoral thesis
File Information	Heidi_Lynn_TESSMER_review.pdf, 審査の要旨



学位論文審査の要旨

博士の専攻分野の名称：博士（獣医学）

氏名：Heidi Lynn Tessmer

審査委員	主査	教授	鈴木	定彦
	副査	教授	高田	礼人
	副査	准教授	小柳	香奈子（情報科学研究科）
	副査	准教授	瀧川	一学（情報科学研究科）
	副査	助教	大森	亮介

学位論文題名

The Analysis of Infectious Diseases via Machine Learning
(機械学習による感染症の解析)

本論文ではバイオインフォマティクス分野における機械学習の活用に関する二つの研究成果を紹介した。第一章は感染症疫学データ解析、特に回帰問題における機械学習の活用について、第二章はウイルスの宿主の分類問題における機械学習の活用についてである。

呼吸器感染症の伝播過程を推定し予測するには、基本再生産数を推定することが一般的である。基本再生産数をはじめとした疫学パラメータの推定の困難さは、従来の統計手法で必要とする陽な尤度関数の導出の困難さに帰着する。近年、陽な尤度関数を必要としない基本再生産数をはじめとした疫学パラメータの推定法として Approximate Bayesian Computation が使われている。本論文では、様々な機械学習の手法、Multilayer Perceptron、Convolutional Neural Network、Long term short memory を用い、その性能を Approximate Bayesian Computation と比較した。比較には、伝播過程シミュレーションデータだけでなく実際のインフルエンザ A H1N1pdm09、流行性耳下腺炎、麻疹の疫学データも使用し以下の知見を得た。機械学習は推定の時間効率は Approximate Bayesian Computation を上回るが、Approximate Bayesian Computation は異なる病原体でより普遍的に正確な推定をすることが可能である。

多種の宿主を持つ感染症は特にその制御が困難であり、自然宿主の同定はその制御の第一歩である。しかしながら、新規および再興感染症の自然宿主の同定は困難となっている。そこで、フルーツバット、カモを宿主とする一本鎖 RNA ウイルスの塩基配列データを主成分分析により解析し、自然宿主の分類を試みた。歴史的に、塩基配列データ上でのウイルスと宿主の関係は G+C コンテンツと CpG の比率の比較により解析されてきた。しかしながら、多くのデータでこれら二つの指標だけでは説明できないウイルスと宿主の関係性も多く存在する。そこで、これら二つ

の指標に代わる新しい塩基の組み合わせを提唱し、実塩基配列データでその性能を検証した。この塩基の組み合わせを用いて塩基配列データを解析すれば、新規感染症流行における自然宿主を同定することが可能となる。

よって、審査委員一同は、上記学位論文提出者 Heidi Lynn Tessmer 氏の学位論文は、北海道大学大学院獣医学研究科規程第6条の規定による本研究科の行う学位論文の審査等に合格と認めた。