

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Generalization Analysis on dependency of Model Complexity in Bayesian Deep Neural Network
著者(和文)	永安修也
Author(English)	Shuya Nagayasu
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12656号, 授与年月日:2024年3月26日, 学位の種別:課程博士, 審査員:渡邊 澄夫,金森 敬文,山下 真,中野 張,高邊 賢史
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第12656号, Conferred date:2024/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	永安 修也		
論文審査 審査員		氏名	職名		氏名	職名
	主査	渡邊 澄夫	教授		高邊 賢史	准教授
	審査員	金森 敬文	教授	審査員		
		山下 真	教授			
		中野 張	准教授			

論文審査の要旨 (2000 字程度)

本論文は、**Generalization Analysis on dependency of Model Complexity in Bayesian Deep Neural Network** (ベイズ深層神経回路網における汎化へのモデル複雑さ依存性の解析)と題し、英文5章より構成されている。

第1章「**Introduction**」では、深層神経回路網とベイズ学習に関する従来の研究をまとめ、本研究の目的は深層学習のベイズ汎化誤差がモデルの複雑さにどのように依存するかを二つのケースを通して解明することであると述べている。第1のケースとしてデータ生成分布がモデルで実現不可能でありモデルの中で最適な確率分布が複数個存在するときには、サンプルサイズの増加に伴って平均汎化誤差が増大することを示すと述べている。第2のケースとして畳込み神経回路網がデータ生成分布に対して過剰パラメータを持つときスキップ接続を持つ場合と持たない場合において平均汎化誤差の上界が異なることを示すと述べている。

第2章「**Bayesian learning theory**」では、ベイズ学習理論に基づく事後分布と予測分布の定義を説明し、平均汎化誤差と平均自由エネルギーの理論的な性質を説明している。

第3章「**Learning theory for non-identifiable optimal probability distributions**」では、データ生成分布から学習モデルまでのカルバック・ライブラ情報量を最小にするという意味で最適となる複数の異なる確率分布が存在する場合には、対数尤度関数が相対的に有限な分散であるという条件が成り立たないために従来理論が適用できないことを示している。さらに最適な分布の個数が有限であるとき、パラメータ集合を有限個の共通部分をもたない部分集合に分割することで各部分集合の上で周辺尤度の漸近挙動を調べて、その最大値を比較することで平均自由エネルギーと平均汎化誤差の性質が導出できると述べている。平均自由エネルギーは従来理論とは異なりサンプルサイズの平方根の項を含むこと、および平均汎化誤差がサンプルサイズの増大に伴って増大することを示している。

第4章「**Free Energy of Bayesian Convolutional Neural Networks**」では、画像認識等において広く用いられている畳込み神経回路網において、スキップ接続を持つものとスキップ接続を持たないものとの理論的な比較を行っている。尤度関数が定数となるパラメータの部分集合の事後確率のサンプルサイズに対するオーダーを調べることでベイズ

学習の推測精度を調べる方法を構成し、この方法に基づいてスキップ接続を持つ畳み込み神経回路網は、データ生成分布に対して冗長なモデルであっても平均自由エネルギーと平均汎化誤差が大きくなること、およびスキップ接続を持たないときには冗長な度合に応じてそれらの上界が増大することを示している。

第5章「**Conclusion**」では、本研究で得られた結果についてまとめ、深層学習における意義を考察している。

以上を要するに本論文は、深層神経回路網のベイズ学習における理論的な性質を解析し、複数の最適なモデルが存在するという条件とスキップ接続の有無に関する条件が平均自由エネルギーと平均汎化誤差に与える影響を解明したものであり、理学上、貢献するところが極めて大きい。よって本論文は博士（理学）にふさわしいものと認められる。

注意：「論文審査の要旨及び審査員」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。