様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月　 １ 日現在

機関番号：１２１０２
研究種目：研究活動スタート支援
研究期間：2019～2022
課題番号：１９Ｋ２４３３７

研究課題名（和文）機械学習モデル多様化による機械学習応用システムの高信頼化設計および評価

研究課題名（英文）Reliability evaluation and design for machine learning application systems using diverse machine learning models

研究代表者
町田　文雄（Machida, Fumio）

筑波大学・システム情報系・准教授

研究者番号：５０８４２２０９

交付決定額（研究期間全体）：（直接経費）　　2,200,000円

研究成果の概要（和文）：本研究は機械学習を組み込んだ情報システムを高信頼化する技術およびそのシステムを評価するモデル化と解析手法を確立することを目的として実施した．高信頼化する手段として，一つのシステムに複数の機械学習モデルを用い，モデル間の多様性を高める手法に着目した．研究の結果，多様な機械学習モデルに加え，多様な入力を用いることで，システムの誤った出力を抑制できることを明らかにした．さらに，機械学習モデルと入力データの多様性を測る指標を定義し，多様性指標とシステムの信頼性を関係づける評価モデルを提案して有効性を示した．本研究成果により機械学習システムを高信頼化するNバージョン機械学習システムと呼ぶ手法を確立した．

研究成果の学術的意義や社会的意義
現在機械学習は様々なシステムで利用され，自動運転など高度な安全性が求められるシステムへの応用も進んでいる．本研究で確立したNバージョン機械学習システムは機械学習応用システムの安全性や信頼性向上に役立てることができる．特に，入力データを多様化して用いる手法は従来手法と比較してより低コストで実現できる高信頼化手法として期待できる．また，本研究では機械学習モデルや入力データの多様性とシステムの出力の信頼性の関係をモデル化して明らかにした．機械学習システムの信頼性がその構成要素や入力の多様性によって特徴づけられるという興味深い結果は信頼性工学の学術会議でも認められ，当該研究分野の発展に貢献した．

研究成果の概要（英文）：This research aimed to develop a technique for improving the reliability of information systems employing machine learning components and to establish the models and analysis methods for evaluating such systems. As a means to enhance reliability, the study focused on a system approach using multiple machine learning models for leveraging the diversity of the models. The outcome of the research unveiled that incorrect system outputs can be reduced by using diverse input data in addition to diverse machine learning models. Moreover, this study proposed and showed the effectiveness of the reliability model that formulates the relation between system output reliability and the diversity metrics defined for machine learning models and input data. This study established a new technique for improving the reliability of machine learning systems that is named the N-version machine learning system.

研究分野：システムディペンダビリティ

キーワード：機械学習システム　信頼性　多様性

１．研究開始当初の背景

近年，深層学習等の機械学習アルゴリズムの発展により，機械学習モデルによる予測や分類の精度が飛躍的に向上した．しかし，サンプルからの学習によって得られる機械学習モデルは，任意の入力データに対して必ず期待通りの正しい出力結果が保証されるわけではない．明確に仕様が定められるソフトウェアプログラムとは異なり，機械学習モデルの出力は学習の際に用いたデータやアルゴリズムの性質に左右される．機械学習モデルを高い信頼性や安全性が求められる情報システムで利用する際には，機械学習モデルが誤った出力をすることを想定してシステム全体としての信頼性を確保する方法が求められる．

２．研究の目的

本研究は機械学習を組み込んだ情報システムを高信頼化する技術およびそのシステムを評価するモデル化と解析手法を確立することを目的として実施した．特に，複数の機械学習モデルと多様化した入力を組み合わせて用いることでシステムを高信頼化する手法に着目し，複数の異なる機械学習モデルを用いた場合のシステム全体の信頼性への影響を確率的な解析によって明らかにするとともに，実際に利用可能な機械学習アルゴリズムやデータを用いてその有効性を検証することを目標とした．

３．研究の方法

具体的に次の研究課題に取り組んだ．

(1) 機械学習モデルの多様性の概念定義

複数の機械学習モデルを用いるシステムでは，機械学習モデルの多様性がシステムの高信頼化に寄与すると考えられる．そこで，機械学習モデルの多様性の概念を整理し，システムの出力の信頼性指標と対応付けるための指標を定義する．

(2) 機械学習モデルの多様性のモデル化と情報システムの信頼性への影響分析

多様性の定義を用い，機械学習応用システムの構成や振る舞いを確率モデルで捉えて分析し，信頼性への影響を解析的に示す．

(3) 異なるアルゴリズムや学習モデルを用いた機械学習システムの信頼性向上効果の評価

実際に活用されている機械学習アルゴリズムやデータを用いて多様な機械学習モデルによってシステムの信頼性が向上することを確認する．自動運転における交通標識の認識への応用を想定し，多様な機械学習モデルを用いて画像分類システムを構築した場合に出力の信頼性が向上することを確かめる．

４．研究成果

(1) Ｎバージョン機械学習システムの提案

異なる機械学習モデルと異なる入力データを組み合わせて機械学習システムの出力を高信頼化する構成を N バージョン機械学習システムとして提案し，本研究分野の先駆的研究成果を発表した．Ｎバージョン機械学習システムの信頼性を特徴づけるために機械学習モデルの多様性を測るモデル多様性と入力データの多様性を測る入力データ多様性を定義し，これらの指標が N バージョン機械学習システムの信頼性に与える影響を確率的に解析した．モデルの多様性は推論エラーを引き起こす入力データ集合の共通部分(intersection)を用いて定義し，入力データの多様性は二つの入力が共に推論エラーとなる条件付き確率で定義した．この多様性指標定義により 3 種類の異なる 2 バージョン機械学習システムの信頼性を定式化でき，信頼性の比較評価が行えるようになった．比較評価結果の一例を図１に示す．信頼性を最大化する 2 バージョン機械学習システムの構成は入力データ多様性の度合いとモデル多様性の度合いに依存して変化することが示されている．３バージョン機械学習システムに対しても同様な定式化と比較評価が可能であることも確認できた．本研究成果はディペンダブルコンピューティングの国際会議で発表し，当該分野において研究の新規性と有用性が認められた．この研究成果を土台として現在本研究分野の研究を主導している．

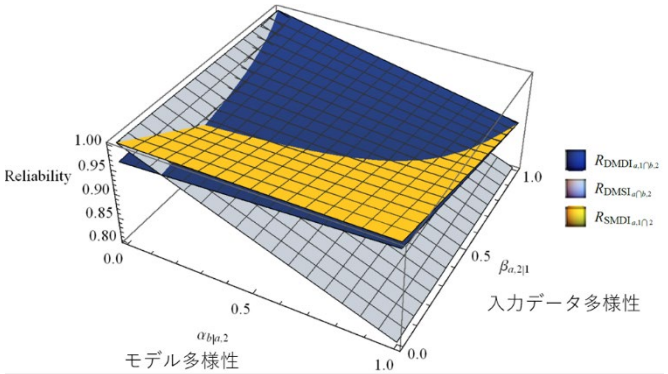

図１２バージョン機械学習システムの信頼性比較評価結果

(2) 入力データ多様化による高信頼化手法の有効性評価

研究当初の計画では機械学習モデルの多様性を生かした高信頼化手法に着目していたが，一つの機械学習モデルであっても異なる入力データを与えることで複数の出力結果を得て最終的な出力を高信頼化できる場合がある．このことを確かめるため，ニューラルネットワークで画像分類を行うシステムを用い，入力画像を加工することで入力データを多様化し，分類結果の信頼性向上効果を調べた．図 2 は手書き数字のデータセット MNIST を畳み込みニューラルネットワーク（CNN）を用いて分類した際に正しく分類されなかった 0 の画像を示している．図には無加工の画像を用いた場合と，ノイズ加工した画像を用いた場合，画像を 20° 回転させた場合の結果が含まれている．同じ CNN を使った場合でもデータ加工方法が異なれば誤分類を起こすサンプルも変化することがわかる．この性質を利用すれば，モデルの多様性を生かした高信頼化と同様の効果を，入力画像データの加工によって達成できる．ベルギーの交通標識のデータセットでも同様の結果が得られることを確認した．本研究成果もディペンダブルコンピューティングの国際会議で発表した．この実験を通じて得た知見から，機械学習システムの信頼性を向上させるデータ多様化手法についての研究の着想を得て現在継続研究を進めている．

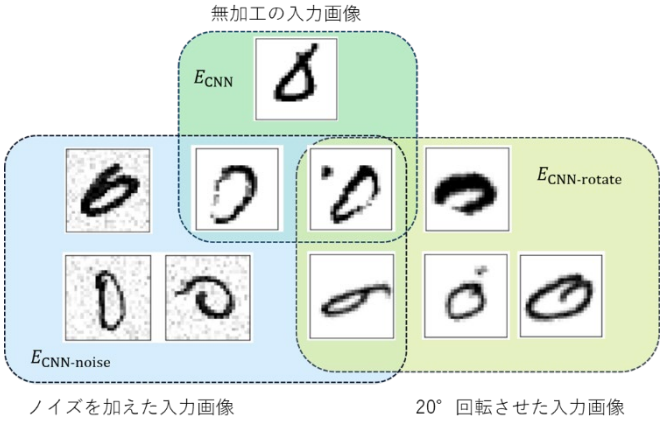

図 2　CNN で誤分類された手書き数字 0 のサンプル

(3) システム性能を考慮した機械学習システム高信頼化手法

自動運転など実際のシステムで高信頼化手法を導入する場合は，性能オーバーヘッドやコストへの配慮が必要となる．学会発表で有識者から指摘をいただき，N バージョン機械学習システムの異なる構成形態についてシステムのスループットや応答時間，消費電力を評価する手法を検討した．最も基本的な構成要素となる 2 つの入力データソースを持つシステムを想定し，機械学習モジュールを並列して動作させる並列型と一つの機械学習モジュールを共有する共有型のアーキテクチャを比較評価した（図 3 参照）．性能評価では待ち行列モデルによる解析評価をまず行い，さらに実機を用いた性能測定により評価した．この結果，並列型がスループットに優れるのに対し，共有型は応答時間や消費電力の面で利点があることが示された．並列型の応答時間分布は裾が重いのに対し，共有型は比較器で比較する二つの推論結果が揃わない場合に処理を棄却するため待ち時間が全体的に短くなっている（図 4 参照）．本研究に関連する成果は国内外のワークショップおよび研究会で発表した．今後は，より多くの入力を持つシステムや異なる機械学習モデルを持つシステムの性能やコストの評価，および信頼性とのトレードオフの解析などについて研究を進める．

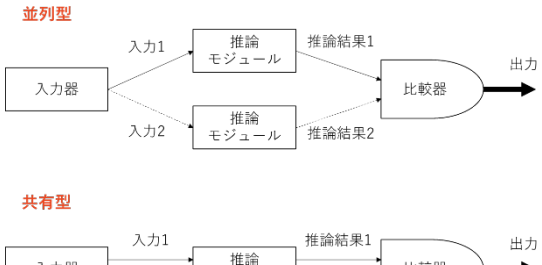

図 3　並列型と共有型の機械学習システム構成

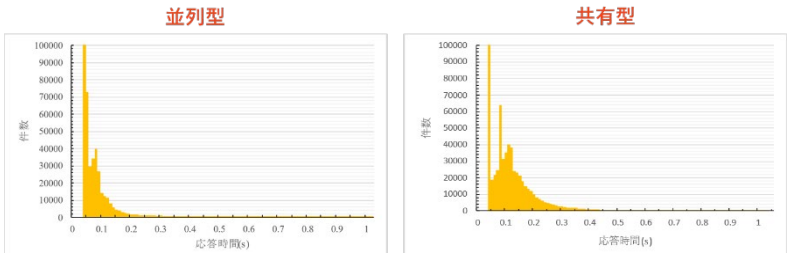

図 4　並列型および共有型機械学習システムの応答時間分布

（参考文献）

[1] F. Machida, N-version machine learning models for safety critical systems, DSN Workshop on Dependable and Secure Machine Learning, pp. 48-51, 2019.

[2] F. Machida, On the diversity of machine learning models for system reliability, IEEE Pacific Rim Int'l Symp. on Dependable Computing (PRDC), pp. 276-285, 2019.

[3] 脇上和也,町田文雄, 2 入力機械学習システムの信頼性と性能評価, 第 20 回ディペンダブルシステムワークショップ, 2022.

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　0件／うち国際共著　0件／うちオープンアクセス　1件）

１．著者名 Fumio Machida	４．巻 -
２．論文標題 Using Diversities to Model the Reliability of N-version Machine Learning System	５．発行年 2021年
３．雑誌名 TechArchive	６．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.36227/techrxiv.16435656.v1	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計11件（うち招待講演　2件／うち国際学会　5件）

１．発表者名 Qiang Wen, Fumio Machida
２．発表標題 Reliability Models and Analysis for Triple-model with Triple-input Machine Learning Systems
３．学会等名 IEEE Conference on Dependable and Secure Computing (DSC)（国際学会）
４．発表年 2022年

１．発表者名 高橋満帆, 町田文雄
２．発表標題 Nバージョン機械学習モデルによるシステム高信頼化のための入力データ多様化
３．学会等名 第19回ディペンダブルシステムワークショップ（DSW 2021）
４．発表年 2021年

１．発表者名 Yuta Makino, Tuan Phung-Duc, Fumio Machida
２．発表標題 A Queueing Analysis of Multi-model Multi-input Machine Learning Systems
３．学会等名 Dependable and Secure Machine Learning（国際学会）
４．発表年 2021年

１．発表者名
巻野 侑大, Phung-Duc Tuan, 町田 文雄

２．発表標題
マルコフ連鎖を用いた多モデル多入力型機械学習システムの性能評価

３．学会等名
2020年度待ち行列シンポジウム

４．発表年
2021年

１．発表者名
Fumio Machida

２．発表標題
On the diversity of machine learning models for system reliability

３．学会等名
IEEE Pacific Rim Int'l Symp. on Dependable Computing (PRDC)（国際学会）

４．発表年
2019年

１．発表者名
Fumio Machida

２．発表標題
N-version machine learning models for safety critical systems

３．学会等名
DSN Workshop on Dependable and Secure Machine Learning（国際学会）

４．発表年
2019年

１．発表者名
町田文雄

２．発表標題
多様な分類器を用いた機械学習応用システムの信頼性

３．学会等名
日本OR学会「4部合同研究会 ～確率モデルの新展開～」（招待講演）

４．発表年
2019年

１．発表者名
町田文雄

２．発表標題
Nバージョンモデルによる機械学習応用システムの高信頼化

３．学会等名
日本ソフトウェア科学会第17回 ディペンダブルシステムワークショップ（招待講演）

４．発表年
2019年

１．発表者名
Mitsuho Takahashi, Fumio Machida, Qiang Wen

２．発表標題
How Data Diversification Benefits the Reliability of Three-version Image Classification Systems

３．学会等名
IEEE 27th Pacific Rim International Symposium on Dependable Computing (PRDC)（国際学会）

４．発表年
2022年

１．発表者名
町田文雄

２．発表標題
Nバージョン機械学習分類システムによる分類結果の正確性と安全性評価

３．学会等名
日本信頼性学会 第30回春季信頼性シンポジウム

４．発表年
2022年

１．発表者名
脇上和也，町田文雄

２．発表標題
2入力機械学習システムの信頼性と性能評価

３．学会等名
第20回ディペンダブルシステムワークショップ（DSW 2022）

４．発表年
2022年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

N-version Machine Learning System https://www.sd.cs.tsukuba.ac.jp/en/n-version_mls.html

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関