

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年3月31日現在

機関番号：12608

研究種目：若手研究(A)

研究期間：2008～2011

課題番号：20680007

研究課題名（和文）

動的に変化する環境下での機械学習

研究課題名（英文）

Machine Learning under Changing Environments

研究代表者

杉山 将 (SUGIYAMA MASASHI)

東京工業大学・大学院情報理工学研究科・准教授

研究者番号：90334515

研究成果の概要（和文）：

従来の機械学習の研究では、データが生成される環境が時間とともに変化しないという大前提のもとで理論やアルゴリズムが構築されてきた。しかしながら、近年の応用場面ではそのような定常性が成り立たない場合が多い。そこで本研究プロジェクトでは、環境の変化に対応するための基礎理論、および、実用的なアルゴリズム開発を行なった。そして、その成果をロボティクス、画像認識、脳波解析、音声識別、自然言語処理など様々な分野に応用した。

研究成果の概要（英文）：

Existing machine learning research has developed theories and algorithms based on the assumption that data generating environments do not change over time. However, recent applications do not satisfy such a stationarity assumption. In this project, we therefore developed fundamental theories and practical algorithms for coping with changing environments. We further applied the developed algorithms to various real-world applications including robotics, image recognition, brain signal analysis, speech recognition, and natural language processing.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	7,000,000	2,100,000	9,100,000
2009年度	4,700,000	1,410,000	6,110,000
2010年度	4,700,000	1,410,000	6,110,000
2011年度	3,200,000	960,000	4,160,000
総計	19,600,000	5,880,000	25,480,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：学習と知識獲得，知能情報処理，機械学習，データマイニング

## 1. 研究開始当初の背景

機械学習とは、データの背後に潜む規則をコンピュータに自動的に獲得させる技術の総称である。これにより、人間がタスクに合わせてコンピュータを制御するプログラムを作る必要がなくなり、効率よく大量のデータを処理することが可能となる。

機械学習における最も基本的な前提は、コンピュータが規則を獲得するときに用いる訓練データが、将来予測を行ないたいテストデータと同じ確率分布に従うという仮定である。この仮定に基づいて、訓練データからテストデータに関する情報を予測する。しかし、近年の応用場面ではこの大前提が成り立たないことが多い。

## 2. 研究の目的

本研究の目的は、このような非定常な環境下でも精度が理論的に保証される統計的機械学習手法を開発することである。具体的には、非定常環境適応のための基礎理論の構築、実用的な機械学習アルゴリズムの開発、そして、実世界応用を行う。

## 3. 研究の方法

基礎理論研究として、環境が変化した場合に既存の学習法がどのような影響を受けるかを数理的に解明し、環境が変化した場合でもアルゴリズムが正しい答えに収束することを理論的に保証するための一般論を構築する。そして、理論的な知見に基づき、環境の変化の影響を吸収できる機械学習アルゴリズムを開発する。また、脳やロボット学習などの実システムの非定常性を数理的にモデル化する。

そして、応用研究として、ブレインコンピュータインターフェース、ロボット制御、信号画像処理などに開発した非定常環境適応アルゴリズムを適用する。

また、開発したアルゴリズムの産業応用の可能性を探るとともに、ソフトウェアをウェブなどで無償公開し、一連の研究成果を専門書として出版する。

## 4. 研究成果

理論的には、統計学の分野で知られている重点サンプリングの技法を機械学習における損失関数に適用することにより、訓練データとテストデータの確率分布の違いを漸近的に吸収することができる。この考えに基づき、まずは重点サンプリングに必要となる重要度重みの推定方法を考案した。重要度重みとはテストデータと訓練データの確率密度関数の比であり、我々が提案した手法では、分母と分子の確率密度関数を個別に推定するのではなく、それらの比を直接推定する。これにより、重要度の推定精度を大幅に向上

させることができた。更に、この手法を大規模・高次元データに適用できるよう様々な改良を行うとともに、推定量の振る舞いを数学的に明らかにした。

応用研究としては、ブレイン・コンピュータインターフェース、自然言語処理、話者識別、顔画像からの年齢予測、ロボット制御の研究を行い、提案した非定常環境適応学習手法の有効性を実証した。

ところで、機械学習や統計学の分野で重要な研究課題の一つである能動学習は、数理的には非定常環境適応と密接な関係がある。能動学習とは、訓練データの入力点の配置を最適化する問題であり、能動学習を行うことにより、訓練データとテストデータは自然と異なる分布を持つことになる。従って、能動学習の文脈においても、重点サンプリングの技法が重要な働きをすると考えられる。この考え方に基づき、重点サンプリングを応用した新しい能動学習の基礎理論を構築すると共に、実用的なアルゴリズムを開発した。そして、そのアルゴリズムをロボット制御と半導体露光装置のウェファー位置合わせに応用し、有効性を実証した。

上記の一連の研究成果を英語の単行本としてまとめ、MIT Press 社より出版した。また、ソフトウェアをウェブで無償公開した。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 14 件)

(1) Ueki, K., Sugiyama, M., & Ihara, Y. Lighting condition adaptation for perceived age estimation. (査読有) IEICE Transactions on Information and Systems, vol. E94-D, no. 2, pp. 392-395, 2011.

(2) Hachiya, H., Peters, J., & Sugiyama, M. Reward weighted regression with sample reuse for direct policy search in reinforcement learning. (査読有) Neural Computation, vol. 23, no. 11, pp. 2798-2832, 2011.

(3) Sugiyama, M., Kawanabe, M., & Chui, P. L. Dimensionality reduction for density ratio estimation in high-dimensional spaces. (査読有) Neural Networks, vol. 23, no. 1, pp. 44-59, 2010.

(4) Yamada, M., Sugiyama, M., & Matsui, T. Semi-supervised speaker identification under covariate shift. (査読有) Signal Processing, vol. 90, no. 8, pp. 2353-2361, 2010.

(5) Akiyama, T., Hachiya, H., & Sugiyama, M. Efficient exploration through active learning for value function approximation in reinforcement learning. (査読有) Neural Networks, vol.23, no.5, pp.639-648, 2010.

(6) Li, Y., Kambara, H., Koike, Y., & Sugiyama, M. Application of covariate shift adaptation techniques in brain computer interfaces. (査読有) IEEE Transactions on Biomedical Engineering, vol.57, no.6, pp.1318-1324, 2010.

(7) Tsuboi, Y., Kashima, H., Hido, S., Bickel, S., & Sugiyama, M. Direct density ratio estimation for large-scale covariate shift adaptation. (査読有) Journal of Information Processing, vol.17, pp.138-155, 2009.

(8) Sugiyama, M. & Nakajima, S. Pool-based active learning in approximate linear regression. (査読有) Machine Learning, vol.75, no.3, pp.249-274, 2009.

(9) Kanamori, T., Hido, S., & Sugiyama, M. A least-squares approach to direct importance estimation. (査読有) Journal of Machine Learning Research, vol.10 (Jul.), pp.1391-1445, 2009.

(10) Sugiyama, M., Kanamori, T., Suzuki, T., Hido, S., Sese, J., Takeuchi, I., & Wang, L. A density-ratio framework for statistical data processing. (査読有) IPSJ Transactions on Computer Vision and Applications, vol.1, pp.183-208, 2009.

(11) Yamada, M. & Sugiyama, M. Direct importance estimation with Gaussian mixture models. (査読有) IEICE Transactions on Information and Systems, vol.E92-D, no.10, pp.2159-2162, 2009.

(12) Hachiya, H., Akiyama, T., Sugiyama, M., & Peters, J. Adaptive importance sampling for value function approximation in off-policy reinforcement learning. (査読有) Neural Networks, vol.22, no.10, pp.1399-1410, 2009.

(13) Sugiyama, M., Suzuki, T., Nakajima, S., Kashima, H., von Büna, P., & Kawanabe, M. Direct importance estimation for covariate shift adaptation. (査読有) Annals of the Institute of Statistical

Mathematics, vol.60, no.4, pp.699-746, 2008.

(14) Sugiyama, M. & Rubens, N. A batch ensemble approach to active learning with model selection. (査読有) Neural Networks, vol.21, no.9, pp.1278-1286, 2008.

[学会発表] (計11件)

(1) Ueki, K., Sugiyama, M., & Ihara, Y. Perceived age estimation under lighting condition change by covariate shift adaptation. In Proceedings of 20th International Conference on Pattern Recognition (ICPR2010), pp.3400-3403, Istanbul, Turkey, Aug. 23-26, 2010.

(2) Wichern, G., Yamada, M., Thornburg, H., Sugiyama, M., & Spanias, A. Automatic audio tagging using covariate shift adaptation. In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2010), pp.253-256, Dallas, Texas, USA, Mar. 14-19, 2010.

(3) Li, Y., Koike, Y., & Sugiyama, M. A framework of adaptive brain computer interfaces. In Proceedings of the 2nd International Conference on BioMedical Engineering and Informatics (BMEI09), pp.473-477, Tianjin, China, Oct. 17-19, 2009.

(4) Hachiya, H., Peters, J., & Sugiyama, M. Efficient sample reuse in EM-based policy search. In W. Buntine, M. Grobelnik, D. Mladenic, and J. Shawe-Taylor (Eds.), Machine Learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science, vol.5781, pp.469-484, Berlin, Springer, 2009. (Presented at the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD2009), Bled, Slovenia, Sep. 7-11, 2009)

(5) Akiyama, T., Hachiya, H., & Sugiyama, M. Active policy iteration: Efficient exploration through active learning for value function approximation in reinforcement learning. In Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI2009), pp.980-985, Pasadena, California, USA, Jul. 11-17, 2009.

(6) Yamada, M., Sugiyama, M., & Matsui, T. Covariate shift adaptation for semi-supervised speaker identification. In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2009), pp.1661-1664, Taipei, Taiwan, Apr. 19-24, 2009.

(7) Kanamori, T., Hido, S., & Sugiyama, M. Efficient direct density ratio estimation for non-stationarity adaptation and outlier detection. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou (Eds.), Advances in Neural Information Processing Systems 21, pp.809-816, Cambridge, MA, MIT Press, 2009. (Presented at Neural Information Processing Systems (NIPS2008), Vancouver, British Columbia, Canada, Dec. 8-13, 2008)

(8) Sugiyama, M. & Nakajima, S. Pool-based agnostic experiment design in linear regression. In W. Daelemans, B. Goethals, and K. Morik (Eds.), Machine Learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science, vol.5212, pp.406-422, Berlin, Springer, 2008. (Presented at the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD2008), Antwerp, Belgium, Sep. 15-19, 2008)

(9) Hachiya, H., Akiyama, T., Sugiyama, M., & Peters, J. Adaptive importance sampling with automatic model selection in value function approximation. In Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI2008), pp.1351-1356, Chicago, Illinois, USA, Jul. 13-17, 2008.

(10) Sugiyama, M. & Rubens, N. Active learning with model selection in linear regression. In M. J. Zaki, K. Wang, C. Apte, and H. Park (Eds.), Proceedings of the Eighth SIAM International Conference on Data Mining (SDM2008), pp.518-529, Atlanta, Georgia, USA, Apr. 24-26, 2008.

(11) Tsuboi, Y., Kashima, H., Hido, S., Bickel, S., & Sugiyama, M. Direct density ratio estimation for large-scale covariate shift adaptation. In M. J. Zaki, K. Wang, C. Apte, and H. Park (Eds.), Proceedings of the Eighth SIAM

International Conference on Data Mining (SDM2008), pp.443-454, Atlanta, Georgia, USA, Apr. 24-26, 2008.

[図書] (計2件)

(1) Sugiyama, M. & Kawanabe, M. Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation, MIT Press, Cambridge, MA, USA, 2012. 308 ページ

(2) Quiñero-Candela, J., Sugiyama, M., Schwaighofer, A., & Lawrence, N. D. (Eds.), Dataset Shift in Machine Learning, MIT Press, Cambridge, MA, USA, 2009. 248 ページ

[その他]

ホームページ等

<http://sugiyama-www.cs.titech.ac.jp/~sugi/>

## 6. 研究組織

### (1) 研究代表者

杉山 将 (SUGIYAMA MASASHI)

東京工業大学・大学院情報理工学研究所・  
准教授

研究者番号：90334515