Fung, Glenn M.; **Mangasarian, Olvi L.**

Privacy-preserving linear and nonlinear approximation via linear programming. (English)

Zbl 1270.90029
Optim. Methods Softw. 28, No. 1, 207-216 (2013).

Summary: We propose a novel privacy-preserving random kernel approximation based on a data matrix $A \in \mathbb{R}^{m \times n}$ whose rows are divided into privately owned blocks. Each block of rows belongs to a different entity that is unwilling to share its rows or make them public. We wish to obtain an accurate function approximation for a given $y \in \mathbb{R}^m$ corresponding to each of the m rows of A. Our approximation of y is a real function on $\mathbb{R}^n$ evaluated at each row of 4 and is based on the concept of a reduced kernel $K(A, B')$, where B' is the transpose of a completely random matrix B. The proposed linear-programming-based approximation, which is public but does not reveal the privately held data matrix A, has an accuracy comparable to that of an ordinary kernel approximation based on a publicly disclosed data matrix A.

MSC:

90C05 Linear programming
90C90 Applications of mathematical programming

Keywords:

privacy-preserving approximation; random kernels; support vector machines; linear programming

Software:

RSVM

Full Text: DOI

References:

[1] Bednarz, A., Bean, N. and Roughan, M. Hiccups on the road to privacy-preserving linear programming. Proceedings of the 8th ACM Workshop on Privacy in the Electronic Society. pp.117–120.

[2] DOI: 10.1162/neco.1995.7.1.108 · doi:10.1162/neco.1995.7.1.108

[3] Chen, K. and Liu, L. Privacy preserving data classification with rotation perturbation. Proceedings of the Fifth International Conference of Data Mining (ICDM'05). November27–30. pp.589–592. Houston, TX, USA: IEEE.

[4] Cherkassky V., Learning from Data – Concepts, Theory and Methods (1998) · Zbl 0960.62002

[5] Cristianini N., An Introduction to Support Vector Machines (2000)

[6] DOI: 10.1016/j.amc.2006.07.076 · Zbl 1162.15015 · doi:10.1016/j.amc.2006.07.076

[7] Laur, S., Lipmaa, H. and Mielikäinen, T. Cryptographically private support vector machines. KDD '06: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp.618–624. New York: ACM. · Zbl 1122.68494

[8] Laur, S., Lipmaa, H. and Mielikäinen, T. Cryptographically private support vector machines. KDD '06: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp.618–624. New York: ACM. · Zbl 1122.68494

[9] DOI: 10.1109/TNN.2006.883722 · doi:10.1109/TNN.2006.883722

[10] Lee, Y.J. and Mangasarian, O. L. RSVM: Reduced support vector machines. Proceedings First SIAM International Conference on Data Mining. April5–7. Chicago CD-ROM; available atftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/00-07.pdf

[11] L. Liu, J. Wang, Z. Lin, and J. Zhang,Wavelet-based data distortion for privacy-preserving collaborative analysis, Tech. Rep. 482-07, Department of Computer Science, University of Kentucky, Lexington, KY, 2007, available athttp://www.cs.uky.edu/jzhang/pub/MINING/

[12] Mangasarian O. L., Advances in Large Margin Classifiers pp 135– (2000)

[13] O.L. Mangasarian,Privacy-preserving horizontally partitioned linear programs, Tech. Rep. 10-02, Data Mining Institute, Computer Sciences Department, University of Wisconsin, Madison, WI, April 2010; available atftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/10-02.pdf, Optimization Letters, 6(3) (2012), pp. 431–436.

[14] O.L. Mangasarian,Privacy-preserving linear programming, Tech. Rep. 10-01, Data Mining Institute, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, March 2010; available atftp://ftp.cs.wisc.edu/pub/dmi/tech-

reports/10-01.pdf, Optim. Lett. 5(01) (2011), pp. 165–172.

[15] DOI: 10.1007/s10957-006-9157-x · Zbl 1139.90017 · doi:10.1007/s10957-006-9157-x

[16] O.L. Mangasarian and E.W. Wild,Privacy-preserving classification of horizontally partitioned data via random kernels, Tech. Rep. 07-03, Data Mining Institute, Computer Sciences Department, University of Wisconsin, Madison, WI, November 2007. Proceedings of the 2008 International Conference on Data Mining, R. Stahlbock, S.V. Crone, and S. Lessman, eds., Vol. II, DMIN08, Las Vegas, July 2008, pp. 473–479.

[17] Schölkopf B., Learning with Kernels (2002)

[18] Tikhonov A. N., Solutions of Ill–Posed Problems (1977) · Zbl 0354.65028

[19] Vapnik V. N., The Nature of Statistical Learning Theory, 2. ed. (2000) · Zbl 0934.62009 · doi:10.1007/978-1-4757-3264-1

[20] Xiao, M.J., Huang, L.S., Shen, H. and Luo, Y.L. Privacy preserving id3 algorithm over horizontally partitioned data. Sixth International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT'05). December5–8. pp.239–243. Dalian, China: IEEE Computer Society.

[21] Yu, H., Jiang, X. and Vaidya, J. Privacy-preserving SVM using nonlinear kernels on horizontally partitioned data. SAC '06: Proceedings of the 2006 ACM Symposium on Applied Computing. pp.603–610. New York: ACM Press. · doi:10.1145/1141277.1141415