(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 March 2005 (17.03.2005)

PCT

(10) International Publication Number
WO 2005/024648 A1

(51) International Patent Classification[7]: G06F 15/18, 19/00

(21) International Application Number:
PCT/AU2004/001199

(22) International Filing Date:
6 September 2004 (06.09.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2003904855 5 September 2003 (05.09.2003) AU

(71) Applicant *(for all designated States except US)*: **AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH** [SG/SG]; 30 Biopolis Street, #09-02 Matrix, Singapore 138671 (SG).

(71) Applicant *(for MG only)*: **IP ORGANISERS PTY LTD** [AU/AU]; Level 21, 367 Collins Street, Melbourne, Victoria 3000 (AU).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **LI, Jiriyan** [CN/SG]; 30 Biopolis Street, #09-02 Matrix, Singapore 138671 (SG).

(74) Agent: **PHILLIPS ORMONDE & FITZPATRICK**; 367 Collins Street, Melbourne, Victoria 3000 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2005/024648 A1

(54) Title: METHODS OF PROCESSING BIOLOGICAL DATA

(57) Abstract: The present invention relates to methods useful for processing large amounts of high-dimensional biological data, such as that provided by microarray analysis of gene expression. The methods are useful for providing rules applicable to the classification, diagnosis and prognosis of diseases such as cancer. The inventive methods implement iterative decision trees to process the training data and generate the rules. However, unlike the prior art methods the methods described avoid the use of bootstrap data and considers substantially the entire training data set at each iteration of the decision tree generation process.