

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 December 2007 (21.12.2007)

PCT

(10) International Publication Number
WO 2007/145789 A3

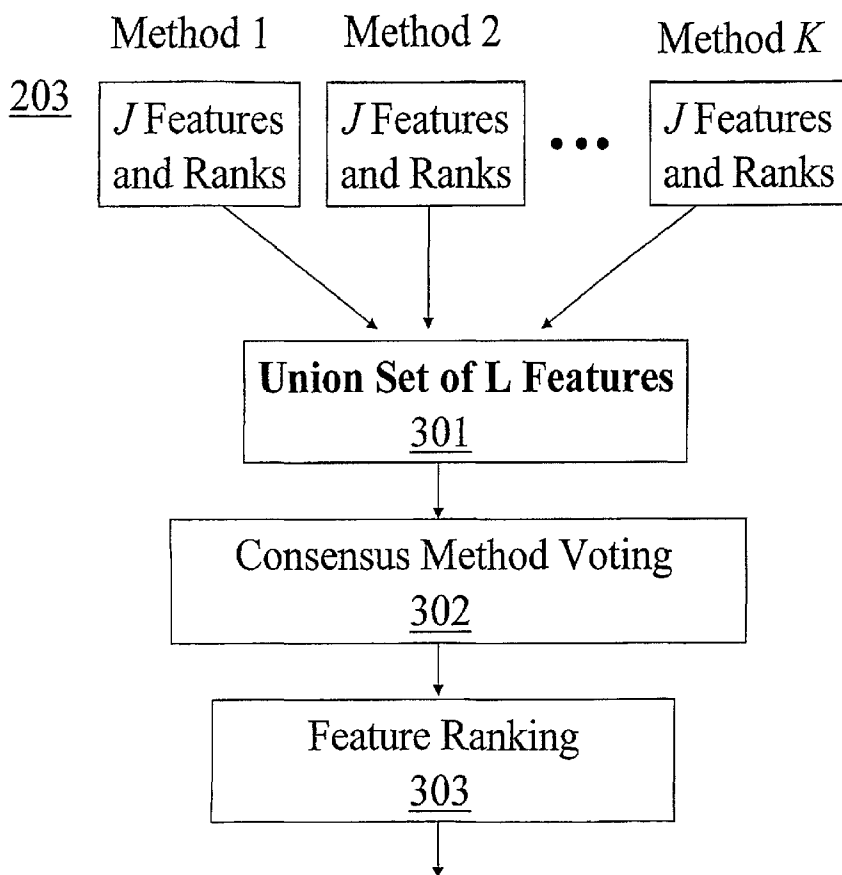
- (51) International Patent Classification:
G06N 5/00 (2006.01)
- (21) International Application Number:
PCT/US2007/012231
- (22) International Filing Date: 17 May 2007 (17.05.2007)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/801,348 18 May 2006 (18.05.2006) US
- (71) Applicant and
- (72) Inventor: ZHANG, John [US/US]; 57 Brentwood Circle, Needham, MA 02492 (US).
- (72) Inventors: LUO, Jun; 100 Broadway, Apt.6, Arlington, MA 02474 (US). CARLSON, An, C.; 6 Hawks Ridge Road, Billerica, MA 01862 (US). WANG, Eric; 294 Bedford Street, Lexington, MA 02420 (US).
- (74) Agents: CHOW, Stephen, Y. et al.; Burns & Levinson, Llp, 125 Summer Street, Boston, MA 02110 (US).

- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, 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, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, IS, IT, LT, LU, LV, MC, MT, 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

[Continued on next page]

(54) Title: METHOD AND IMPLEMENTATION OF RELIABLE CONSENSUS FEATURE SELECTION IN BIOMEDICAL DISCOVERY



(57) Abstract: Process and apparatus for combining multiple processes for choosing features such as biomarkers in statistical data using consensus voting among the multiple processes and their chosen features.

WO 2007/145789 A3



— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
28 August 2008

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US07/12231

A. CLASSIFICATION OF SUBJECT MATTER

IPC: G06N 5/00(2006.01)
G06N 5/00(2006.01)

USPC: 706/45

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 706/45

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST, google and scholar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Bing Liu, Qinghua Cui, Tianzi Jiang, Songde Ma "A combinational feature selection and ensemble neural network method for classification of gene expression data" BMC Bioinformatics 2004	1-3, 14-31, 34-36
A	Alexey Tsymbal, Seppo Puuronen and Iryna Skrypnik "Ensemble Feature Selection with Dynamic Integration of Classifiers" 2001	1-36
A	Han Yu Chuang, Hongfan Liu, Stuart Brown, Cameron McMunn Coffran, Cheng Yan Kao, D Frank, Hus et al "Identifying Significant Genes from Microarray Data" BIBE'04	1-36
A	Kee Jong, Jeremie Mary, Antoine Cornuejols, Elena Marchiori, and Michele Sebag, "Ensemble Feature Ranking", 2004	1-36

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search: 17 June 2008 (17.06.2008) Date of mailing of the international search report: 09 JUL 2008

Name and mailing address of the ISA/US: Mail Stop PCT, Attn: ISA/US, Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450, Facsimile No. (571) 273-3201
Authorized officer: LUT WONG, Telephone No. (571) 270-1123