DESIGNS TRADE MARKS PHICAL INDICATIONS	सत्यगव जयत GOVERNMENT OF INDIA	Department of Industrial Polic Ministry of Comme	y and Proi
Design Application [Details		
Application Number:			
321700-001			
Cbr Number:			
17835			
Cbr Date:			
17-09-2019 16:00:38			
Applicant Name:			
1. Nagesh Salimath	2. Sharanabasappa C Gandage	3. Soumya M A	
4. Shruti A G			
Design Application S	Status		
Application Status:			
Application Under Proce	ss(Awating for Technical Examination)	
Back (/designapplicationS	tatus/)		

Disclaimer: Application status is available for the application filed on or after 1st April 2009 with application no 222230. The information under " Design Application Status" is dynamically retrieved and is under testing, therefore the information retrieved by this system is not valid for any legal proceedings under the Design Act 2000. In case of any discrepancy you may contact the appropriate Patent Office or send your comments to following email IDs:

Design Office, Kolkata : controllerdesign.ipo@nic.in

Controller General of Patents, Designs and Trademarks



search by number - 🛛 🗙

💈 Intellectual Property India 🗙 🔇 Search for patents | USPTO 🗙 📳 Intellectual Property India 🗙 🔮 Downloads

atentSearch/PatentSearch/ViewApplicationStatus





Application Details		
APPLICATION NUMBER	201941033341	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	19/08/2019	
APPLICANT NAME	1 . Jayasbree Agarkhed 2 . Vijayalaxmi Kadrolli 3 . Siddarama R. Pati	
TITLE OF INVENTION	A SYSTEM TO MONITOR PIGEON-PEA AGRICULTURE ALONG WITH HELICOVERPA ARMIGERA PEST CONTROL USING	
FIELD OF INVENTION	PHYSICS	
E-MAIL (As Per Record)		
ADDITIONAL-EMAIL (As Per Record)	udachanv@gmail.com	
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE	19/08/2019	
PUBLICATION DATE (U/S 11A)	09/10/2020	

jov.in/PatentSearch/PatentSearch/ViewApplicationStatus

×

Office of the Controller General of Patents. Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India

+



Application Details			
APPLICATION NUMBER	201941037608		
APPLICATION TYPE	ORDINARY APPLICATION		
DATE OF FILING	18/09/2019		
APPLICANT NAME	1 . Dr. Suvarna Nandyal 2 . Mrs. BhavanaR Maale 3 . Mr. Anil Mangalgi 4 . Mrs. Basanti Ghanti 5 . Mr. Anmol Mangalgi 6 . Mr. Anil Pawar 7 . Mr. Ravindra Maale 8 . Dr. Suryakant B Patil		
TITLE OF INVENTION	PROCESS FOR IDENTIFICATION AND CLASSIFICATION OF DANCE TYPES AND SYSTEM THEREOF		
FIELD OF INVENTION	COMPUTER SCIENCE		
E-MAIL (As Per Record)	contact@psp-ipassociates.com		
ADDITIONAL-EMAIL (As Per Record)	contact@psp-ipassociates.com		
E-MAIL (UPDATED Online)			
PRIORITY DATE			
REQUEST FOR EXAMINATION DATE	20/08/2020		
PUBLICATION DATE (U/S 11A)	27/09/2019		
FIRST EXAMINATION REPORT DATE	26/08/2020		
Date Of Certificate Issue	22/03/2022		
POST GRANT JOURNAL DATE	25/03/2022		
REPLY TO FER DATE	23/02/2021		
	Application Status		
APPLICATION STATUS	Granted Application, Patent Number :392544		
E-Register	Order(s)/Decision(s) View Documents		
Filed Published RQ Filed Under Examination Disposed			

Intellectual Property India

/PatentSearch/PatentSearch/ViewApplicationStatus

× +





Application Details	
APPLICATION NUMBER	201941040286
APPLICATION TYPE	ORDINARY APPLICATION
DATE OF FILING	04/10/2019
APPLICANT NAME	1 . JAYASHREE AGARKHED 2 . ASHALATHA R 3 . SIDDARAMA R. PATIL
TITLE OF INVENTION	WORKING OF A SECURED MULTI-TENANT FRAMEWORK IN CLOUD COMPUTING SYSTEMS IN HEALTH CARE MANAGEMENT SYSTEMS
FIELD OF INVENTION	COMPUTER SCIENCE
E-MAIL (As Per Record)	
ADDITIONAL-EMAIL (As Per Record)	koushikdeb009@gmail.com
E-MAIL (UPDATED Online)	
PRIORITY DATE	
REQUEST FOR EXAMINATION DATE	04/10/2019
PUBLICATION DATE (U/S 11A)	09/10/2020
REPLY TO FER DATE	01/09/2021
	Application Status

Intellectual Property India

×

+

n/PatentSearch/PatentSearch/ViewApplicationStatus



Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India



Application Details		
APPLICATION NUMBER	201941042625	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	21/10/2019	
APPLICANT NAME	 Jayashree Agarkhed Patil Yogita Dattatrya Siddarama R. Patil 	
TITLE OF INVENTION	A SYSTEM TO MONITOR SENSOR NETWORK BASED SMART GREENHOUSE FOR THE CULTIVATION OF DENDROBIUM	
FIELD OF INVENTION	MECHANICAL ENGINEERING	
E-MAIL (As Per Record)		
ADDITIONAL-EMAIL (As Per Record)	koushikdeb009@gmail.com	
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE	21/10/2019	
PUBLICATION DATE (U/S 11A)	09/10/2020	
REPLY TO FER DATE	28/08/2021	

ndia

ntellectual Property India

atentSearch/PatentSearch/ViewApplicationStatus

× +





Application Details		
APPLICATION NUMBER	201941044634	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	04/11/2019	
APPLICANT NAME	1 . Jayashree Agarkhed 2 . Siddarama R. Patil 3 . Swapna Kalyan 4 . Vijayalaxmi Patil	
TITLE OF INVENTION	A SYSTEM TO REMOTELY MONITOR THE HEALTH OF THE ELDERY USING SENSOR ASSEMBLY FIMELY MEDICATION	OR
FIELD OF INVENTION	BIO-MEDICAL ENGINEERING	
E-MAIL (As Per Record)		
ADDITIONAL-EMAIL (As Per Record)	koushikdeb009@gmail.com	
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE	04/11/2019	
PUBLICATION DATE (U/S 11A)	09/10/2020	
	Application Status	
o Hi 📀 f	🖹 🖪 🧮 🎐 💽 🖉 🦉	<u>)</u> 19





Application Details			
APPLICATION NUMBER	202021041957		
APPLICATION TYPE	ORDINARY APPLICATION		
DATE OF FILING	28/09/2020		
APPLICANT NAME	1 . Mr. ROHIT KIRAN GAVALI 2 . Mrs. RUPALI KIRAN GAVALI 3 . Prof. SACHIN PRABHAKAR KOMBLE 4 . Dr. SUVARNA NANDYAL		
TITLE OF INVENTION	EXPANDABLE AND RETRACTED VEHICLE		
FIELD OF INVENTION	MECHANICAL ENGINEERING		
E-MAIL (As Per Record)	director@psp-ipassociates.com		
ADDITIONAL-EMAIL (As Per Record)	director@pspipassociates.com		
E-MAIL (UPDATED Online)			
PRIORITY DATE			
REQUEST FOR EXAMINATION DATE	12/03/2021		
PUBLICATION DATE (U/S 11A)	30/10/2020		
REPLY TO FER DATE	05/07/2021		

tellectual Property India

entSearch/PatentSearch/ViewApplicationStatus

× +





Application Details		
APPLICATION NUMBER	202041000764	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	08/01/2020	
APPLICANT NAME	1 . DR. NAGESH SALIMATH 2 . SHARANABASAPPA C GANDAGE 3 . SHRUTI A G 4 . RAJALAXMI S BILGUNDI 5 . CHETANKUMAR KALASKAR 6 . SATISHKUMAR MALLAPPA	
TITLE OF INVENTION	MACHINE LEARNING BASED METHOD FOR CLASSIFICATION OF SPECTRAL IMAGES	
FIELD OF INVENTION	COMPUTER SCIENCE	
E-MAIL (As Per Record)	nageshsalimath@pdaengg.com	
ADDITIONAL-EMAIL (As Per Record)	nageshsalimath@pdaengg.com	
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE		
PUBLICATION DATE (U/S 11A)	17/01/2020	
	Application Status	



.gov.in/PatentSearch/PatentSearch/ViewApplicationStatus

×

Office of the Controller General of Patents. Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India

+



APPLICATION NUMBER APPLICATION TYPE DATE OF FILING APPLICANT NAME	202041039981 ORDINARY APPLICATION 15/09/2020
APPLICATION TYPE	ORDINARY APPLICATION 15/09/2020
DATE OF FILING	15/09/2020
PPLICANT NAME	
	1 . Dr. SUVARNA NANDYAL 2 . Mrs. SUVARNA LAXMIKANT KATTIMANI 3 . Dr. PRAKASH PATTAN 4 . Mr. ANMOL MANGALGI 5 . Mr. SANJEEV KUMAR ANGADI 6 . Dr. BASANTI GHANTI 7 . Mr. LAXMIKANT BHIMARAO KATTIMANI 8 . Mr. SOMASHEKHAR S.DHANYAL 9 . Mrs. VIJAYALAXMI S.PATIL
ITLE OF INVENTION	AUTOMATIC HIGHLIGHTS EXTRACTION FROM CRICKET BASED ON SEMANTIC QUERY PROCESSING
IELD OF INVENTION	COMPUTER SCIENCE
MAIL (As Per Record)	director@psp-ipassociates.com
DDITIONAL-EMAIL (As Per Record)	
MAIL (UPDATED Online)	
RIORITY DATE	
EQUEST FOR EXAMINATION DATE	14/10/2020
UBLICATION DATE (U/S 11A)	16/10/2020
IRST EXAMINATION REPORT DATE	28/10/2020
)ate Of Certificate Issue	08/07/2021
OST GRANT JOURNAL DATE	16/07/2021
EPLY TO FER DATE	27/04/2021
VATE OF Certificate Issue	08/07/2021 16/07/2021 27/04/2021
	Application Status
PPLICATION STATUS	Granted Application, Patent Number :371532
E-Register	Order(s)/Decision(s) View Documents

tellectual Property India

tentSearch/PatentSearch/ViewApplicationStatus

× +





APPLICATION NUMBER	2021/1016356
	202141010550
APPLICATION TYPE	ORDINARY APPLICATION
DATE OF FILING	07/04/2021
APPLICANT NAME	 Dr. ANURADHA T Dr. SHAILAJA Dr. SAYYADA FAHMEEDA SULTANA SUDHA. V. PAREDDY SMITHA.S. PADSHETTY
TITLE OF INVENTION	AN ENVIRONMENT FRIENDLY LIFE SAVING SYSTEM IN VEHICLE FOR PUBLIC SAFETY
FIELD OF INVENTION	ELECTRONICS
-MAIL (As Per Record)	
ADDITIONAL-EMAIL (As Per Record)	anuradhat26@gmail.com
-MAIL (UPDATED Online)	
PRIORITY DATE	
REQUEST FOR EXAMINATION DATE	-
PUBLICATION DATE (U/S 11A)	16/04/2021
	Application Status

ntellectual Property India

atentSearch/PatentSearch/ViewApplicationStatus

+

х



Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India



Application Details		
APPLICATION NUMBER	202141020076	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	02/05/2021	
APPLICANT NAME	1 . Dr. Shubhangi Digamber Chikte (Professor) 2 . Kumari Shilpa (Assistant Professor)	
TITLE OF INVENTION	SYSTEMS AND METHODS FOR PREDICTION OF OSTEOPOROTIC USING BIOMEDICAL IMAGE ONLY.	
FIELD OF INVENTION	COMPUTER SCIENCE	
E-MAIL (As Per Record)	shubhangidc@vtu.ac.in	
ADDITIONAL-EMAIL (As Per Record)	shilpadevappa11@gmail.com	
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE		
PUBLICATION DATE (U/S 11A)	14/05/2021	

Application Status

Intellectual Property India

Home (http://ipindia.nic.in/index.htm) About Us (http://ipindia.nic.in/about-us.htm) Who's Who (http://ipindia.nic.in/whos-who-page.htm) Policy & Programs (http://ipindia.nic.in/policy-pages.htm) Achievements (http://ipindia.nic.in/achievements-page.htm)

RTI (http://ipindia.nic.in/right-to-information.htm) Feedback (https://ipindiaonline.gov.in/feedback) Sitemap (shttp://ipindia.nic.in/itemap.htm)

Contact Us (http://ipindia.nic.in/contact-us.htm) Help Line (http://ipindia.nic.in/helpline-page.htm)

Skip to Main Content Screen Reader Access (screen-reader-access.htm)







(http://ipindia.nic.in/inc

Patent Search

Inventor		
Classification (IPC)	H04W0016140000, H04W0072020000, H04W0024000000, H04W0088020000, H04W0016220000	
Field Of Invention	COMMUNICATION	
Priority Date		
Priority Country		
Priority Number		
Application Filing Date	04/10/2019	
Application Number	201941040278	
Publication Type	INA	
Publication Date	09/10/2020	
Publication Number	41/2020	
Invention Title	A METHODOLOGY OF SPECTRUM SENSING OF COGNITIVE RADIO SYSTEMS USING CLUSTER-BASED PROCEDURE	

Siddarama R. Patil	Professor And H.O.D., Department of E&Ce, P.D.A. College of Engineeing, Kalaburagi, India.	India	India
Vilaskumar Patil	Research Scholar, Department of ECE, P.D.A College of Engineering, Kalaburagi, India.	India	India
Jayashree Agarkhed	Professor, Department of CSE, P.D.A College of Engineering, Kalaburagi, India,	India	India

Applicant

Name	Address	Country	Nationa
Siddarama R. Patil	Professor And H.O.D., Department of E&Ce, P.D.A. College of Engineeing, Kalaburagi, India.	India	India
Vilaskumar Patil	Research Scholar, Department of ECE, P.D.A College of Engineering, Kalaburagi, India.	India	India
Jayashree Agarkhed	Professor, Department of CSE, P.D.A College of Engineering, Kalaburagi, India,	India	India

Abstract:

ABSTRACT This invention illustrates an optimal cluster based spectrum sensing and resource allocation (OCSR) methodology for cognitive radio networks to overcome the challenge in the spectrum sensing. Spectrum sensing, mobility, sharing, and management are common parameters that are considered in cognitive radio network. The pro system consist of different optimization process. The first contribution of OCSR is for the cluster process which uses equilibrium whale optimization (EWO). The cluster pro includes the division of channels into occupied sub-band set (OSS) and the idle sub-band set (ISS). The second contribution is for spectrum sensing problems and the spec sensing is used to improve the detection performance. The technique proposed is the isolated K-best detectors. The third contribution is for prioritization of the traffic levels is differential evolution algorithm. This methodology is implemented into any wireless network where spectrum sh necessary for effortless and inexpensive communication.

Complete Specification

WE CLAIM THAT

1. A cluster-based cooperative spectrum sensing methodology.for.cognitive radio systems where each cluster head collects energies of a reporting channel measured b cognitive users within the cluster and decides whether a primary user is absent from a given spectrum

2. As claimed in Claim - 1, the methodology conducts a radio scene analysis to ascertain existing signals in the signal space arid to predict where the signals will exist in future and the subject system predicts holes corresponding to White space or Grey space, then the cognitive radio or software-defined radio is configured to transmit sig in the unoccupied part of the spectrum which permits increased use of the spectrum

As claimed in Claim - 1, the the methodologies includes techniques including the 3G:PP"(the 3rd~Generation Partnership Project) LTE (Long Term Evolution) technique
 The methodology configured' is made capable to communicate with other radios the cognitive radio gets the information of the global radio and the differential evolution algorithm divides the cognitive radios in a network into sub-levels which will decrease the scanning time of the cognitive radio.

5. In the system three techniques is used in combination for optimal spectrum sensing consisting of equilibrium whale optimization (EWO) algorithm for clustering proc isolated K-best detector is used for spectrum sensing problems, and the differential evolution algorithm used for prioritizing the traffic levels.

View Application Status



Department of Industrial Policy and Promotion Government of India

Terms & conditions (http://ipindia.gov.in/terms-conditions.htm) Privacy Policy (http://ipindia.gov.in/privacy-policy.htm)

Copyright (http://ipindia.gov.in/copyright.htm) Hyperlinking Policy (http://ipindia.gov.in/hyperlinking-policy.htm)

Accessibility (http://ipindia.gov.in/accessibility.htm) Archive (http://ipindia.gov.in/archive.htm) Contact Us (http://ipindia.gov.in/contact-us.htm) Help (http://ipindia.gov.in/help.htm)

Content Owned, updated and maintained by Intellectual Property India, All Rights Reserved.

Page last updated on: 26/06/2019

Intellectual Property India

Home (http://ipindia.nic.in/index.htm) About Us (http://ipindia.nic.in/about-us.htm) Who's Who (http://ipindia.nic.in/whos-who-page.htm) Policy & Programs (http://ipindia.nic.in/policy-pages.htm) Achievements (http://ipindia.nic.in/achievements-page.htm)

RTI (http://ipindia.nic.in/right-to-information.htm) Feedback (https://ipindiaonline.gov.in/feedback) Sitemap (shttp://ipindia.nic.in/itemap.htm)

Contact Us (http://ipindia.nic.in/contact-us.htm) Help Line (http://ipindia.nic.in/helpline-page.htm)

Skip to Main Content Screen Reader Access (screen-reader-access.htm)

() | | |







(http://ipindia.nic.in/inc

Patent Search

Invention Title	THE DESIGN OF MICROSTRIP PATCH ANTENNA USING THE GEOMETRY OF GOLDEN RATIO
Publication Number	41/2020
Publication Date	09/10/2020
Publication Type	INA
Application Number	201941049444
Application Filing Date	02/12/2019
Priority Number	
Priority Country	
Priority Date	
Field Of Invention	ELECTRONICS
Classification (IPC)	H01Q0009040000, H01Q0009280000, H01Q0001380000, H01Q0001220000, H01Q0003260000
Inventor	

Name	Address	Country	Nat
SIDDARAMA R PATIL	PROFESSOR, ELECTRONICS AND COMMUNICATIONS ENGINEERING E&CE AND DEAN ACADEMICS, P.D.A COLLEGE OF ENGINEERING, KALABURAGI - 585102	India	Indi
RAVI M YADAHALLI	PROFESSOR, DEPT OF E&CE S G BALEKUNDARI INSTITUTE OF TECH BELAGAVI - 585102	India	Indi
REVANASIDDAPPA KINAGI	FACULTY OF ENGINEERING AND TECHNOLOGY SHARNBASVA UNIVERSITY KALABURAGI - 585102	India	Indi
JAYASHREE AGARKHED	PROFESSOR, DEPARTMENT OF CSE P.D.A COLLEGE OF ENGINEERING, KALABURAGI - 585102	India	Indi

Applicant

Name	Address	Country	Nat
SIDDARAMA R PATIL	PROFESSOR, ELECTRONICS AND COMMUNICATIONS ENGINEERING E&CE AND DEAN ACADEMICS, P.D.A COLLEGE OF ENGINEERING, KALABURAGI	India	Indi
RAVI M YADAHALLI	PROFESSOR, DEPT OF E&CE S G BALEKUNDARI INSTITUTE OF TECH BELAGAVI	India	Indi
REVANASIDDAPPA KINAGI	FACULTY OF ENGINEERING AND TECHNOLOGY SHARNBASVA UNIVERSITY KALABURAGI	India	Indi
JAYASHREE AGARKHED	PROFESSOR, DEPARTMENT OF CSE P.D.A COLLEGE OF ENGINEERING, KALABURAGI	India	Indi

Abstract:

This invention consists of a design of microstrip patch antenna using the geometry of Golden Ratio to overcome the disadvantage of narrow impedance bandwidth disadv of patch antenna. The L shaped stub is attached to this antenna which results into a dual band and compact operation of antenna which can be used for wireless communications. The GSA antenna helps to give wider bandwidth with a gain of 2.3dbi and correspondingly GSAs gives dual bandwidth of operation at 5.9 and 2.4GHz respectively which may be used for ISM band applications.

Complete Specification

WE CLAIM THAT

1. A dual-band printed antenna with a modified golden square GSMS A that has efficient operational features namely, gain greater than OdB, almost uniform space distribution of the radiated power, small reflection coefficient at the feeding port.

2. As claimed in Claim — 1, the length of the patch antenna equal to the width and etching out remaining extra copper material from the RMSA and attaching an L shal stub the geometry of the antenna called as Golden square antenna with stub (GSAS), width of the stub being 7mm and length 11.33mm and surface area thickness being 0.5mm.

3. As claimed in Claim - 1, the antenna gives a dual band resonating at 5.9 and 2.4GHz respectively with return loss of -28.15, bandwidth of 224MHz and gain of 2.5dbi 5.9GHz and at 2.4Ghz return loss is -22.19 bandwidth of 45MHz gain of -8.7dbi with size reduction of 61.5%.

FIELD OF THE INVENTION AND USE OF INVENTION

[0001] This invention relates to the field of electronics and communication engineering more particularly to the making of efficient microstrip patch antennas so as to improve the overall performance by introducing a plurality of converging perturbations about the perimeter of the patch so that the flow of electromagnetic current progressing along the perimeter of the patch is perturbed and results in an effective electromagnetic length substantially greater than the actual physical diameter of the patch. Here in the antenna L shaped stub is attached which results into a dual band and compact operation of antenna which can be used for wireless communications.

View Application Status



Terms & conditions (http://ipindia.gov.in/terms-conditions.htm) Privacy Policy (http://ipindia.gov.in/privacy-policy.htm)

Copyright (http://ipindia.gov.in/copyright.htm) Hyperlinking Policy (http://ipindia.gov.in/hyperlinking-policy.htm)

Accessibility (http://ipindia.gov.in/accessibility.htm) Archive (http://ipindia.gov.in/archive.htm) Contact Us (http://ipindia.gov.in/contact-us.htm) Help (http://ipindia.gov.in/help.htm)

Content Owned, updated and maintained by Intellectual Property India, All Rights Reserved.

Page last updated on: 26/06/2019

Welcome Saurabh Kumar Jain Sig

Sign out

Controller General of Patents, Designs & Trade Marks G.S.T. Road, Guindy, Chennai-600032 Tel No. (091)(044) 22502081-84 Fax No. 044 22502066 E-mail: chennai-patent@nic.in Web Site: www.ipindia.gov.in



सत्यमेव जयते G.A.R.6 [See Rule 22(1)] RECEIPT



Date/Time 2022/01/19 17:51:22

UserId: rsaurabh26

Docket No 5177

To Saurabh Kumar Jain

F-440, Delta-1, Greater Noida

CBR Detail:

Sr. No.	Ref. No./Application No.	App. Number	Amount Paid	C.B.R. No.	Form Name	Remarks
1	E- 12/453/2022/CHE	202241003064	2500	2262	FORM 9	
2	202241003064	TEMP/E- 1/3472/2022- CHE	1600	2262	FORM 1	METHOD FOR MAXIMIZING SEAMLESS VERTICAL MEDIA INDEPENDENT HANDOVER USING MODIFIED INVASIVE WEED OPTI

TransactionID	Payment Mode	Challan Identification Number	Amount Paid	Head of A/C No
N-0000910550	Online Bank Transfer	1901220019469	4100.00	1475001020000001

Total Amount : ₹ 4100

Amount in Words: Rupees Four Thousand One Hundred Only

Received from Saurabh Kumar Jain the sum of ₹ 4100 on account of Payment of fee for above mentioned Application/Forms.

* This is a computer generated receipt, hecnce no signature required.

Print





Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India

(http://ipindia.nic.in/index.htm)



(http://ipindia.nic.in/index.htm)

Application Details		
APPLICATION NUMBER	202241007315	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	11/02/2022	
APPLICANT NAME	1 . Dr Vishwanath Patil 2 . <mark>Mr. Ratnakar Karjo</mark> l 3 . <mark>Mr. Omkar Yatgal</mark> 4 . Mrs.Smita C Chetti	
TITLE OF INVENTION	MPPT in solar photo voltaic system under fluctuating irradiance condition technique	
FIELD OF INVENTION	ELECTRICAL	
E-MAIL (As Per Record)	vishwanath.patil09@gmail.com	
ADDITIONAL-EMAIL (As Per Record)		
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE		
PUBLICATION DATE (U/S 11A)	18/02/2022	
Application Status		

APPLICATION STATUS

Awaiting Request for Examination

office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India

ttp://ipindia.nic.in/index.htm)



(http://ipindia.nic.in/index.htm)

	Application Details
ILICATION NUMBER	202241025889
ILCATION TYPE	ORDINARY APPLICATION
TE OF FILING	04/05/2022
LICANT NAME	1 . Dr. GEETA HANJI 2 . Dr. RAKESH 3 . RATNAMALA S PATIL
LE OF INVENTION	DIAGNOSING IMAGE CLASSIFICATION USING DEEP CONVOLUTION NEURAL NETWORK WITH BIG DATA APPLICATIONS
LD OF INVENTION	COMPUTER SCIENCE
IAIL (As Per Record)	9 BROWNER STR
DITIONAL-EMAIL (As Per Record)	sharangandage@pdaengg.com
AIL (UPDATED Online)	
ORITY DATE	A A MARCH C
UEST FOR EXAMINATION DATE	
BLICATION DATE (U/S 11A)	13/05/2022

Application Status

PLICATION STATUS

2:53 AM

Awaiting Request for Examination

fipindiaservices.gov.in/PatentSearch/PatentSearch/ViewApplicationStatus

(12) INNOVATION PATENT (19) AUSTRALIAN PATENT OFFICE

(11) Application No. AU 2020102129 A4

(54) Title IML- Data Cleaning: INTELLIGENT DATA CLEANING USING MACHINE LEARNING PROGRAMMING

(51) International Patent Classification(s) **G06F 16/35** (2019.01) G06N 5/02 (2006.01) G06K 9/00 (2006.01) G06N 20/00 (2019.01)

(21) Application No: 2020102129 (22) Date of Filing: 2020.09.03

(45)	Publication Date:	2020.10.15

- (45) Publication Journal Date: 2020.10.15
- Granted Journal Date: 2020.10.15 (45)

(71) Applicant(s)

Manohar Madgi;Ramesh Shahabadkar;Nagaraj B. Patil;Bharati Harsoor;Ashok Patil;Padmapriya Patil;Basavaraj Chunchure;R. Balakrishna;Rajesh K. S.;Sai Madhavi D.

(72) Inventor(s)

Madgi, Manohar;Shahabadkar, Ramesh;Patil, Nagaraj B.;Harsoor, Bharati;Patil, Ashok;Patil, Padmapriya;Chunchure, Basavaraj;Balakrishna, R.;K. S., Rajesh;D., Sai Madhavi

(74) Agent / Attorney Manohar Madgi, F-10, Site-iv Mt Roskill, Auckland, 1041, NZ Patent Title: **IML- Data Cleaning**: INTELLIGENT DATA CLEANING USING MACHINE LEARNING PROGRAMMING.

ABSTRACT

Our invention" IML- Data Cleaning "is a system and article of manufacture enabling adapting to a shift in document content and also the instructions for receiving at least one labeled mapped seed document receiving unlabeled mapped documents receiving at least one predetermined cost factor training data a transductive classifier using the at least one predetermined cost factor calculated data and at least one seed document and the unlabeled documents. The invention also classifying the unlabeled documents having a confidence level above a predefined threshold into a plurality of indexing and categories using the classifier reclassifying at least some of the categorized documents into the categories using the classifier and outputting identifiers of the categorized documents to at least one of a user another system and another process. The invented systems and articles of also manufacture for separating documents are also presented and the systems and articles of manufacture for document searching are also presented a business and other information service provides data cleansing to correct and update both domestic and global addresses. The invented system a integrate and combination of processes generates cleansed data for input into a matching and mapping process and the matching, the mapping process matches information about a business. The invented technology data cleaning process includes the steps of validating data loaded from at least two source systems appending the validated data to a normalized data cleaning repository selecting the priority of the source system creating a clean database; loading the consistent, normalized, and cleansed data from the clean database into a format required by data systems and software tools using the data. The invented technology also creating reports and updating the clean database by a user without updating the source systems. The data cleaning process distributed and collecting, analyzing data from available sources for optimization models enabling consistent analysis. The invented technology the data cleaning process further provides complete auditability to the inputs and outputs of data systems and software tools that use a dynamic data set.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)

Data Cleaning Process Parsing Correction Standardizing Consolidation



FIG.2: DATA CLEANING PROCESS.

Australian Government IP Australia Innovation Patent Australia

ning INTELLIGENT DATA CLEANING USING

Patent Title: IML- Data Cleaning: INTELLIGENT DATA CLEANING USING MACHINE LEARNING PROGRAMMING. Name and address of patentees(s): Manohar Madgi (Associate Professor) (Department of Computer Science and Engineering) OFFICE ADDRESS: KLE INSTITUTE OF TECHNOLOGY, HUBBALLI 580027, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: C/O VIJAYALAXMI GURMITAKAL, PLOT NO. 115/116, KANAKA NAGAR, BRAHMAPUR, KALABURAGI 585103, KARNATAKA, INDIA. Dr. Ramesh Shahabadkar (Professor) (Department of Computer Science and Engineering) OFFICE ADDRESS: AMBO UNIVERSITY WOLISO CAMPUS, ETHIOPIA RESIDENTIAL ADDRESS: AMITY HARMONY D-329, SHASHIDHAR LAYOUT, DWARAKANAGAR, CHANNASANDRA, BSK 6TH STAGE, BENGALURU 560098, KARNATAKA, INDIA. Dr. Nagarai B. Patil (Principal) OFFICE ADDRESS: GOVT. ENGINEERING COLLEGE, ANEGUNDI ROAD, GANGAVATHI 583227, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: H. No. 1-11-38/26, BASAVESHWAR COLONY, LINGASUGUR ROAD, RAICHUR 584101, KARNATAKA, INDIA. Dr. Bharati Harsoor (Professor & Head) (Department of Information Science and Engineering) OFFICE ADDRESS: PDA COLLEGE OF ENGINEERING, KALABURAGI 585102, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: W/o ARUNKUMAR JIRAGI, H. NO. 1-9/28A, KHUBA PLOTS, KALABURAGI 585102, KARNATAKA, INDIA. Ashok Patil (Assistant Professor) (Department of Information Science and Engineering) OFFICE ADDRESS: PDA COLLEGE OF ENGINEERING, KALABURAGI 585102, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: SAISHANT, PLOT NO. 66, NEAR VANI VILAS PUBLIC SCHOOL, YAMUNA NAGAR, KUSNOOR ROAD, KALABURAGI 585105, KARNATAKA, INDIA. Padmapriva Patil (Assistant Professor) (Department of Electronics and Communication Engineering) OFFICE ADDRESS: PDA COLLEGE OF ENGINEERING, KALABURAGI 585102, KARNATAKA, INDIA RESIDENTIAL ADDRESS: SAISHANT, PLOT NO. 66, NEAR VANI VILAS PUBLIC SCHOOL, YAMUNA NAGAR, KUSNOOR ROAD, KALABURAGI 585105, KARNATAKA, INDIA. Dr. Basavaraj Chunchure (Associate Professor) (Department of Computer Science and Engineering) OFFICE ADDRESS: VIGNAN'S INSTITUTE OF MANAGEMENT AND TECHNOLOGY FOR WOMENS, KONDAPUR (V), GHATKESAR (M), MEDCHAL-MALKAJGIRI (D), HYDERABAD- 501 301, INDIA. RESIDENTIAL ADDRESS: PO: BATGERA, TQ: BASAVAKALYAN, DIST.: BIDAR-585419, KARNATAKA, INDIA. Dr. R. Balakrishna (Professor & Dean) (Department of Computer Science and Engineering) OFFICE ADDRESS: RAJARAJESWARI COLLEGE OF ENGINEERING, No. 14, RAMOHALLI CROSS, KUMBALGODU, MYSORE ROAD, BENGALURU 560074, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: No. 118, SRI SAI NILAYAM, 6TH STAGE, 11TH BLOCK, BANSHANKRI, BANGALORE 560074, KARNATAKA, INDIA. Dr. Rajesh K. S. (Associate Professor) (Department of Computer Science and Engineering) OFFICE ADDRESS: RAJARAJESWARI COLLEGE OF ENGINEERING, No. 14, RAMOHALLI CROSS, KUMBALGODU, MYSORE ROAD, BENGALURU 560074, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: No. 118/2, 3RD MAIN, 8TH CROSS, CHAMRAJPET, BENGALURU 560018, KARNATAKA, INDIA. Dr. Sai Madhavi D. (Associate Professor) (Department of Computer Science and Engineering) OFFICE ADDRESS: RAO BAHADUR Y MAHABALESWARAPPA ENGINEERING COLLEGE, VIJAYA NAGAR, CANTONMENT, BALLARI 583275, KARNATAKA, INDIA. RESIDENTIAL ADDRESS: No. 165, 25TH WARD, VIDYA NAGAR, NORTH BLOCK, NEAR OPD, BATRI ROAD, BALLARI 583104, KARNATAKA, INDIA. Complete Specification: Australian Government.

FIELD OF THE INVENTION

Our invention "IML- **Data Cleaning** "is related to intelligent data cleaning using machine learning programming and also relates to methods and apparatus for data classification and novel applications using machine learning techniques. The t invention also provides a business information service, and more particularly, to cleansing data associated with customer lists.

BACKGROUND OF THE INVENTION

The invention generally relates to data processing and management processes and, more particularly, to an adaptive data cleaning process and system. [003] The quality of a large real world data set depends on a number of issues, but the source of the data is the crucial factor. Data entry and acquisition is inherently prone to errors both simple and complex. Much effort is often given to this front-end process, with respect to reduction in entry error, but the fact often remains that errors in a large data set are common. The field error rate for a large data set is typically around 5% or more.

Up to half of the time needed for a data analysis is typically spent for cleaning the data. Generally, data cleaning is applied to large data sets. Data cleaning is the process of scrubbing data to improve accuracy of a large data set. Ideally, data cleaning should be able to eliminate obvious transcription errors, to correct erroneous entries, such as erroneous part numbers or invalid codes, to update missing data, such as pricing or lead times, and to recognize that there may exist multiple sources and definitions of data. Effective data cleaning should incorporate electronic notes to explain the rational for rule based or manual selections, should provide an audit trail, and should be easy to operate.

Data cleaning is often done using a manual process, which is laborious, time consuming, and prone to errors. Consequently, methods that enable automated detection of errors in large data sets or that assist in detecting errors are of great interest. The process of automated data cleaning is typically multifaceted and a number of problems must be addressed to solve any particular data cleaning problem. Generally, possible error types need to be defined and determined, a search for errors needs to be conducted and the errors need to be identified, and the uncovered errors need to be corrected. [005] For example, current supply chain software solution vendors, such as i2 Technologies, IBM, Manugistics, MCA Solutions, Systems Exchange, or Xelus have well developed and thought out internal data structures.

These structures must be mapped to a customer's source system and must be updated on a periodic basis. The mapping is "hardwired" during implementation, requiring recoding when sources or business rules change. Furthermore, the development of an intermediate database that stores customer data prior to loading into the supply chain software is often needed. Also, current supply chain software solutions do not support archiving results, archiving the inputs that lead to the results, or versioning data over time. This prevents a customer from auditing the decision process which leads, for example, to the stocking recommendations for a piece of heavy equipment, such as aircraft, trucks, ships or machinery. With service part stock levels for repairable items, such as heavy equipment having a long life, running into the tens to hundreds of millions of dollars, auditability is an important requirement for many customers.

Extract, Transform, and Load (ETL) tools are typically used to bridge the gap between source systems and an intermediate database. ETL tools are used to convert data from one operating system and brand of database software to another. ETL tools apply limited business rules to transform and filter data. ETL tools are not designed to handle multiple sources of the same data. Furthermore, when business rules are applied to multiple sources of data, they are applied during the data collection process, which precludes later visibility of changes to more than one source of data. ETL tools also do not support versioning of data, which includes tracking changes in data over time.

In 2000, Ventana Systems, Inc, Harvard, Massachusetts, U.S.A., developed a data cleaning solution for The Boeing Company, Long Beach, California, U.S.A. for the supply software solution for the C-17 airlift program. This prior art cleaning solution is written in Oracle and C⁺⁺, with an Excel-like user interface. The data cleaning solution advances the prior art by allowing users to change data in a database and color-coding the data that was changed, by developing a way to allow changes to data to persist over time using simple decision tree logic, and by allowing users to select the data elements, which they wish to clean. Still, this prior art data cleaning solution incorporates several limitations. For example, the supply chain software solution uses global variables that can be changed by any routine versus using data encapsulation, the data cleaning solution uses a complex internal data structure that makes it difficult to maintain, and the loading of the data by the application must adhere to a strict procedure or the data may become corrupted.

As can be seen, there is a need for a method for data cleaning that is automated and enables selection of data from multiple sources. Furthermore, there is a need for a data cleaning process that allows support for archiving results, archiving the inputs that lead to the results, or versioning data over time. Still further, there is a need for a data cleaning process that can be easily implemented into existing data management systems.

There has, therefore, arisen a need to provide a process for data cleaning that offers standardized procedures, that complements corporate common data warehouse projects, and that selects data from multiple sources. There has further arisen a need to provide a process for data cleaning that recognizes that different customers may need to see different sources of ostensibly the same data element, and that there may exist multiple versions of what should theoretically be the same data. There has still further arisen a need to provide a process for adaptive data cleaning that enables archiving both the data used for an analysis and the results of the analysis.

How to handle data has gained in importance in the information age and more recently with the explosion of electronic data in all walks of life including, among others, scanned documents, web material, search engine data, text data, images, audio data files, etc. One area just starting to be explored is the non-manual classification of data. In many classification methods the machine or computer must learn based upon manually input and created rule sets and/or manually created training examples. In machine learning where training examples are used, the number of learning examples is typically small compared to the number of parameters that have to be estimated, i.e. the number of solutions that satisfy the constraints given by the training examples is large. A challenge of machine learning is to find a solution that generalizes well despite the lack of constraints. There is thus a need for overcoming these and/or other issues associated with the prior art.

PRIOR ART SEARCH

US5506897A *1993-02-221996-04-09Murex Securities, Ltd. Automatic routing system for telephonic services.

WO2001029780A11999-10-192001-04-26Stamps.ComAddress matching system and method.

US6272495B1 *1997-04-222001-08-07Greg Hetherington Method and apparatus for processing free-format data.

US6311178B1 *1997-09-292001-10-30Webplus, Ltd. Multi-element confidence matching system and the method therefor.

US20020059142A1 *2000-04-212002-05-16Robert Krause Systems and methods for providing change of address services over a network.

US6457012B1 *1997-06-102002-09-24Pitney Bowes Inc. Method and system of updating address records utilizing a clientserver interface.

US20020136381A1 *1996-06-102002-09-26Shaffer James D. One number, intelligent call processing system.

WO2002077769A22001-03-232002-10-03United States Postal Service Residential delivery indicator.

US3195107A *1961-01-241965-07-13Siemens Ag Secured transmission of coded binary symbols.

US5287363A *1991-07-011994-02-15Disk Technician Corporation System for locating and anticipating data storage media failures.

US5574898A *1993-01-081996-11-12Atria Software, Inc. Dynamic software version auditor which monitors a process to provide a list of objects that are accessed.

US5491818A *1993-08-131996-02-13Peoplesoft, Inc. System for migrating application data definition catalog changes to the system level data definition catalog in a database.

US5745753A *1995-01-241998-04-28Tandem Computers, Inc. Remote duplicate database facility with database replication support for online DDL operations.

SE509645C2 *1996-02-081999-02-15Ericsson Telefon Ab L MA method for simultaneously with protocol-based functional changes in a database to perform verification of the converted data.

US20080086432A1 *2006-07-122008-04-10Schmidtler Mauritius A RData classification methods using machine learning techniques.

US20100169250A1 *2006-07-122010-07-01Schmidtler Mauritius A RMethods and systems for transductive data classification.

US20110196870A1 *2006-07-122011-08-11Kofax, Inc. Data classification using machine learning techniques.

US20120237116A1 *2011-03-182012-09-20Jing Xiao Identifying Text Pixels in Scanned Images.

US20130073514A1 *2011-09-202013-03-21Microsoft Corporation Flexible and scalable structured web data extraction.

US8855375B22012-01-122014-10-07Kofax, Inc. Systems and methods for mobile image capture and processing.

US8885229B12013-05-032014-11-11Kofax, Inc. Systems and methods for detecting and classifying objects in video captured using mobile devices.

OBJECTIVES OF THE INVENTION

- 1. The objective of the invention is to a system and article of manufacture enabling adapting to a shift in document content and also the instructions for receiving at least one labeled mapped seed document receiving unlabeled mapped documents receiving at least one predetermined cost factor training data a transductive classifier using the at least one predetermined cost factor calculated data and at least one seed document and the unlabeled documents.
- 2. The other objective of the invention is to the invention also classifying the unlabeled documents having a confidence level above a predefined threshold into a plurality of indexing and categories using the classifier reclassifying at least some of the categorized documents into the categories using the classifier and outputting identifiers of the categorized documents to at least one of a user another system and another process.
- 3. The other objective of the invention is to the invented systems and articles of also manufacture for separating documents are also presented and the systems and articles of manufacture for document searching are also presented a business and other information service provides data cleansing to correct and update both domestic and global addresses.
- 4. The other objective of the invention is to the invented system a integrate and combination of processes generates cleansed data for input into a matching and

mapping process and the matching, the mapping process matches information about a business.

- 5. The other objective of the invention is to the invented technology data cleaning process includes the steps of validating data loaded from at least two source systems appending the validated data to a normalized data cleaning repository selecting the priority of the source system creating a clean database loading the consistent, normalized, and cleansed data from the clean database into a format required by data systems and software tools using the data.
- 6. The other objective of the invention is to the invented technology also creating reports and updating the clean database by a user without updating the source systems. The data cleaning process distributed and collecting, analyzing data from available sources for optimization models enabling consistent analysis.
- 7. The other objective of the invention is to the invented technology the data cleaning process further provides complete auditability to the inputs and outputs of data systems and software tools that use a dynamic data set.

SUMMARY OF THE INVENTION

A method for data cleansing at least one input address is received the input address is compared to at least one standard and a single best address corresponding to the input address is provided based on the comparison. In some embodiments, the single best address is matched to a database having unique business identifiers associated with addresses to find a matching address, which is provided. In some embodiments, the database is an advanced office system (AOS). In some embodiments, a match project analysis report is provided.

The input address is converted to a predetermined record layout, before comparing it to the standard. In some embodiments, the input address is associated with at least one code that is used to determine the single best address. In some embodiments, the input address is associated with at least one score that is used to determine the single best address. In some embodiments, the standard is at least one of the following: ZIP+4 coding, coding accuracy support system (CASS), Locatable Address Conversion System (LACS), delivery sequence file (DSF), and National Change of Address (NCOA). In some embodiments, a report is provided. In some embodiments, the report is a postal summary report or a pre-audit report. In some embodiments, at least one status notification is sent to the user, who supplied the input address.

A system for data cleansing comprising a pre-auditor, a verifier, a vendor interface, and a user interface. The pre-auditor is for generating a report having a number of views of an input address file, which contains a plurality of addresses. The verifier is for finding and removing any invalid records from the input address file. The vendor interface is for sending the input address file and an order to at least one vendor and for receiving an output file from the vendor(s). The user interface is for providing a single best address for each address in the input address file.

The system includes a matcher for attempting to match any address in the output file or the invalid records to a matching address in a database that contains unique business identifiers associated with addresses. The system includes an investigator for investigating any address not matched, upon request. In some embodiments, the preauditor calculates a plurality of counts associated with the input address file. In some embodiments, the input address file includes a plurality of records and each record includes a plurality of fields. In some embodiments, the counts are at least one of the following: a number of distinct values by field, a missing field count, a total number of records, or a percent of distinct values. In some embodiments, the views are one of the following: alphabetical, most frequent content, and alpha characters only. In some embodiments, the vendor standardizes addresses using one of the following: Locatable Address (NCOA).

A machine readable medium having instructions stored thereon to perform a method for data cleansing. A machine readable medium is any storage medium, such as a compact disk (CD). At least one input address is received. The input address is compared to at least one standard and a single best address corresponding to the input address is provided based on that comparison. In some embodiments, the single best address is matched to a database having unique business identifiers associated with addresses to find a matching address and a matching address is provided.

The invention, a data cleaning process comprises the steps of: validating data loaded from at least two source systems using data formatting utilities and data cleaning utilities; appending the validated data to a normalized data cleaning repository; selecting the priority of the source systems; creating a clean database; creating and maintaining a cross-reference between the unique data identifiers; loading consistent, normalized, and cleansed data from the clean database into a format required by data systems and software tools using the data; creating standardized data cleaning and management reports using the consistent, normalized, and cleansed data; and updating the consistent, normalized, and cleansed data by a user without updating the source systems. The clean database contains unique data identifiers for each data element from the at least two source systems.

The invention, a data cleaning process for a supply chain comprises the steps of: loading data from multiple source systems to a master table of data elements and sources; selecting precedence of the source systems; reviewing high driver and error reports; cleaning logistics data contained in the master table of data elements and sources; approving consistent, normalized, and cleansed data of the master table of data elements

and sources and providing the cleansed data to data systems and software tools using the data; initiating inventory optimization of stock level and reorder points using a strategic inventory optimization model using the cleansed data; providing spares analysis including stock level and reorder point recommendations; archiving supporting data for customer audit trail; creating reports; and purchasing spares to cover shortfalls according to the reports.

The invention, a data cleaning system includes data formatting utilities, data cleaning utilities, a normalized data cleaning repository, source prioritization utilities, a clean database, cross-reference utilities, and a data cleaning user interface. The data formatting utilities are used to validate data downloaded from at least two source systems. The data cleaning utilities are used to clean the data. The source prioritization utilities are used to select the priority of the at least two source systems. The normalized data cleaning repository receives the formatted and cleansed data. The clean database combines the cleansed and prioritized data. The clean database is a single source of item data containing the best value and unique data identifiers for each data element. The cross-reference utilities are used to create and maintain a cross-reference between the unique data identifiers. The data cleaning user interface enables a user to update the clean database.

These and other features, aspects and advantages of the present invention will become better understood with reference to the following drawings, description and claims. A system and article of manufacture for adapting to a shift in document content according to one embodiment of the present invention includes instructions for: receiving at least one labeled seed document; receiving unlabeled documents; receiving at least one predetermined cost factor; training a transductive classifier using the at least one predetermined cost factor, the at least one seed document, and the unlabeled documents; classifying the unlabeled documents having a confidence level above a predefined threshold into a plurality of categories using the classifier; reclassifying at least some of the categorized documents into the categories using the classifier; and outputting identifiers of the categorized documents to at least one of a user, another system, and another process.

A system and article of manufacture for separating documents according to another embodiment of the present invention includes instructions for: receiving labeled data; receiving a sequence of unlabeled documents; adapting probabilistic classification rules using transduction based on the labeled data and the unlabeled documents; updating weights used for document separation according to the probabilistic classification rules; determining locations of separations in the sequence of documents; outputting indicators of the determined locations of the separations in the sequence to at least one of a user, another system, and another process; and flagging the documents with codes, the codes correlating to the indicators. A system and article of manufacture for document searching according to another embodiment of the present invention includes instructions for: receiving a search query; retrieving documents based on the search query; outputting the documents; receiving user-entered labels for at least some of the documents, the labels being indicative of a relevance of the document to the search query; training a classifier based on the search query and the user-entered labels; performing a document classification technique on the documents using the classifier for reclassifying the documents; and outputting identifiers of at least some of the documents based on the classification thereof.

BRIEF DESCRIPTION OF THE DIAGRAM

FIG.1: Data Cleaning Block Diagram.

Fig.2: Data cleaning Process.

FIG.3: Data Cleaning importance.

FIG. 4: is a control flow diagram for the classification of unlabeled data in accordance with one of the invention using a scaled cost factor.

FIG. 5: is a control flow diagram for the classification of unlabeled data in accordance with one of the invention using user defined prior probability information.

FIG. 6: is a detailed control flow diagram for the classification of unlabeled data in accordance with one embodiment of the invention using Maximum Entropy Discrimination with scaled cost factors and prior probability information.

FIG. 7: is a network diagram illustrating a network architecture in which the various described herein may be implemented.

FIG. 8: is a system diagram of a representative hardware environment associated with a user device.

FIG. 9: is a block diagram representation of the apparatus of one embodiment of the present invention.

FIG. 10: is a flowchart, a classification process.

FIG. 11: is a flowchart, a classification process performed by in accordance with one embodiment.

DESCRIPTION OF THE INVENTION

The following detailed description is of the best currently contemplated modes of carrying out the invention. The description is not to be taken in a limiting sense, but is made merely for the purpose of illustrating the general principles of the invention, since the scope of the invention is best defined by the appended claims.

Broadly, the invention provides an adaptive data cleaning process and system that standardizes the process of collecting and analyzing data from disparate sources for optimization models. The present invention further generally provides a data cleaning process that provides complete auditability to the inputs and outputs of optimization models or other tools or models that are run periodically using a dynamic data set, which changes over time. The adaptive data cleaning process and system as in one embodiment of the present invention enables consistent analysis, eliminates one-time database coding, and reduces the time required to adjust to changing data sources, and may be used, for example, for inventory optimization models or during the development of supply chain proposals. One embodiment of the present invention provides a data cleaning process that is suitable for, but not limited to, applications in aircraft industry, both military and commercial, for example for supply chain management.

The invention provides a data cleaning process that is further suitable for, but not limited to, applications in industries that utilize heavy equipment having a long life. The data cleaning process as in one embodiment of the present invention may be used where a large database needs to be managed, where the database receives data from multiple sources, for example, large corporations that need to combine data from several sub organizations, and where the data to be managed relate to high value goods, such as heavy equipment in transportation industries. The data cleaning process as in one embodiment of the present invention may further be used, for example, for inventory management, order management, consumer data management, or in connection with industrial maintenance.

The invention provides a data cleaning process that selects data from multiple sources and uses heuristics based on precedence to select the best source from the multiple sources and to select the best value for forecasting. Existing ETL (Extract, Transform, and Load) tools are not designed to handle multiple sources of the same data. Current ETL tools may load data from multiple sources but require a software developer or user to create custom logic to select one source over another.

Furthermore, sources may not be added or deleted after initial implementation of a typical ETL tool without manual intervention of a software developer or user. Contrary to the prior art, the data cleaning process, as in one embodiment of the present invention, allows unlimited numbers of data elements and sources to be added or dropped at any time. Contrary to prior art data cleaning processes, the data cleaning process as in one embodiment of the present users, such as

customers, may need to see different sources of ostensibly the same data element, such as a unit price, which may have an internal value for buying a part and an external value for selling the part. For this example, both values of the price are valid and which one is used depends upon the application.

The data cleaning process as in one embodiment of the present invention may have the ability to display multiple values for selected data elements from different sources. The user may override the original selection with information that may be more accurate than the information in the source system. Unlike traditional databases, where only one value for each data element is visible, the data cleaning process as in one embodiment of the present invention may provide versioning to previous values and traceability to all versions of each data element available from different source systems.

The invention provides a data cleaning process that has the ability to capture and identify all changes being made to data elements in the data repository area, and redisplay the changes back to the user. Information about changes to the data element, regardless if the changes are screen changes or mass updates, may be captured by tracking the user changing the data, the date of the change, and comments including why changes were done. This is an advantage over prior art data cleaning processes, which generally allow only flagging the suspected data and which generally require the change to be made to the system of record. In many cases, the system of record is a customer database, or a departmental database, that the data cleaner does not have update authority for. As a result, prior art data cleaning solutions which force the user to update the system of record are often impractical. Contrary to the prior art, the data cleaning process as in one embodiment of the present invention provides dated versioning to both input and outputs to computer models, tracking changes to data over time.

Existing ETL tools do not support versioning data over time. The data cleaning process, as in one embodiment of the present invention, allows auditability of both results and the data and data sources upon which the results were based. The data cleaning process, as in one embodiment of the present invention, further ensures data integrity by screening the data against user definable business rules. Furthermore, the data cleaning process, as in one embodiment of the present invention, allows user additions and deletions, for example, to part numbers from source systems, maintaining traceability to what was added and flagging deleted data for traceability, rather than physically deleting the data. Consequently, data is electronically tagged as deleted, but not physically removed from the data repository. Still further, the data cleaning process, as in one embodiment of the present invention, adds automated notes, and allows for manual notes, which may be attached to each data element and provide information on automated processing, format conversions, and other data quality information. This provides auditability when data must be converted for an analysis, for example, when normalizing currency from Great Britain Pounds to United States Dollars. The invention provides a data cleaning process that may be used, for example in connection with supply chain software tools and that may allow archiving and sharing the results of such supply chain software tools. Currently existing data repositories will store current input data required to perform an analysis. The data cleaning process, as in one embodiment of the present invention, will allow archiving both the data used at the time the analysis was performed, and the results of the analysis. This provides complete auditability to the source of data and the model results based upon that data. This is important, for example, for government supply chain contracts and commercial contracts, where auditability to the rationale behind the purchase of costly maintenance spares is required. There are no known supply chain tools which support archiving of data and results.

The data cleaning process, as in one embodiment of the present invention allows thresholds and triggers to be established at the data element level providing alerts, which notify, for example, asset managers and data owners that specific data elements are suspect and should be reviewed. These thresholds are particularly important when large amounts of data are being updated, as it may be physically impossible as well as error prone to scan each and every data element for errors. Furthermore, the data cleaning process, as in one embodiment of the present invention provides defaults to fill in critical missing data, while flagging the missing data for manual review. This makes it more likely that all parts will be included in an analysis, compared with traditional solutions of deleting an entire item if any data element for that item is missing or invalid. The data cleaning process, as in one embodiment of the present invention provides traceability to all data elements for which defaults have been used.

The following description is the best mode presently contemplated for carrying out the present invention. This description is made for the purpose of illustrating the general principles of the present invention and is not meant to limit the inventive concepts claimed herein. Further, particular features described herein can be used in combination with other described features in each of the various possible combinations and permutations. Unless otherwise specifically defined herein, all terms are to be given their broadest possible interpretation including meanings implied from the specification as well as meanings understood by those skilled in the art and as defined in dictionaries, treatises, etc.

The interest and need for classification of textual data has been particularly strong, and several methods of classification have been employed. A discussion of classification methods for textual data is described in U.S. Pat. No. 6,192,360 to Dumais, the content and substance of which is incorporated herein by reference. The current state of the art in commercially used automatic classification systems is either rule based or utilizes inductive machine learning, i.e. using manually labeled training examples. Both methods typically entail a large manual setup effort compared to transductive methods. The

solutions provided by rule based systems or inductive methods are static solutions that cannot adapt to drifting classification concepts without manual effort.

Inductive machine learning is used to ascribe properties or relations to types based on tokens (i.e., on one or a small number of observations or experiences); or to formulate laws based on limited observations of recurring patterns. Inductive machine learning involves reasoning from observed training cases to create general rules, which are then applied to the test cases. Particularly, preferred embodiments use transductive machine learning techniques. Transductive machine learning is a powerful method that does not suffer from these disadvantages. Transductive machine techniques may be capable of learning from a very small set of labeled training examples, automatically adapting to drifting classification concepts, and automatically correcting the labeled training examples. These advantages make transductive machine learning an interesting and valuable method for a large variety of commercial applications.

Transduction learns patterns in data. It extends the concept of inductive learning by learning not only from labeled data but also from unlabeled data. This enables transduction to learn patterns that are not or only partly captured in the labeled data. As a result, transduction can, in contrast to rule based systems or systems based on inductive learning, adapt to dynamically changing environments. This capability enables transduction to be utilized for document discovery, data cleanup, and addressing drifting classification concepts, among other things.

FIG. 4: a control flow diagram is illustrated showing the method of classification of unlabeled data of one embodiment of the present invention. The method 100 begins at step 102 and at step 104 accesses stored data 106. The data is stored at a memory location and includes labeled data, unlabeled data and at least one predetermined cost factor. The data 106 includes data points having assigned labels. The assigned labels identify whether a labeled data point is intended to be included within a particular category or excluded from a particular category.

Once data is accessed at step 104, the method of one embodiment of the present invention at step 108 then determines the label prior probabilities of the data point using the label information of data point. Then, at step 110 the expected labels of the data point are determined according to the label prior probability. With the expected labels calculated in step 110, along with the labeled data, unlabeled data and cost factors, step 112 includes iterative training of the transductive MED classifier by the scaling of the cost factor unlabeled data points. In each iteration of the calculation the unlabeled data points' cost factors are scaled. As such, the MED classifier learns through repeated iterations of calculations. The trained classifier then accessed input data 114 at step 116. The trained classifier can then complete the step of classifying input data at step 118 and terminates at step 120. It is to be understood that the unlabeled data of 106 and the input data 114 may be derived from a single source. As such, the input data/unlabeled data can be used in the iterative process of 112 which is then used to classify at 118. Furthermore, one embodiment of the present invention contemplates that the input data 114 maybe include a feedback mechanism to supply the input data to the stored data at 106 such that the MED classifier of 112 can dynamically learn from new data that is input.

FIG. 5: a control flow diagram is illustrated showing another method of classification of unlabeled data of one embodiment of the present invention including user defined prior probability information. The method 200 begins at step 202 and at step 204 accesses stored data 206. The data 206 includes labeled data, unlabeled data, a predetermined cost factor, and prior probability information provided by a user. The labeled data of 206 includes data points having assigned labels. The assigned labels identify whether the labeled data point is intended to be included within a particular category or excluded from a particular category.

At step 208, expected labels are calculated from the data of 206. The expected labels then used in step 210 along with labeled data, unlabeled data and cost factors to conduct iterative training of a transductive MED classifier. The iterative calculations of 210 scale the cost factors of the unlabeled data at each calculation. The calculations continue until the classifier is properly trained. The trained classifier then accessed input data at 214 from input data 212. The trained classifier can then complete the step of classifying input data at step 216. As with the process and method described in FIG. 4, the input data and the unlabeled data may derive from a single source and may be put into the system at both 206 and 212. As such, the input data 212 can influence the training at 210 such that the process my dynamically change over time with continuing input data.

In both methods as shown in FIGS. 4 and 5 a monitor may determine whether or not the system has reached convergence. Convergence may be determined when the change of the hyperplane between each iteration of the MED calculation falls below a predetermined threshold value. In an alternative embodiment of the present invention, the threshold value can be determined when the change of the determined expected label falls below a predetermined threshold value. If convergence is reached, then the iterative training process may cease.

Referring particularly to FIG. 6, illustrated is a more detailed control flow diagram of the iterative training process of at least one embodiment of the method of the present invention. The process 300 commences at step 302 and at step 304 data is accessed from data 306 and may include labeled data, unlabeled data, at least one predetermined cost factor, and prior probability information. The labeled data points of 306 include a label identifying whether the data point is a training example for data points to be included in the designated category or a training example for data points to be excluded form a

designated category. The prior probability information of 306 includes the probability information of labeled data sets and unlabeled data sets.

In step 308, expected labels are determined from the data from the prior probability information of 306. In step 310, the cost factor is scaled for each unlabeled data set proportional to the absolute value of the expected label of a data point. An MED classifier is then trained in step 312 by determining the decision function that maximizes the margin between the included training and excluded training examples utilizing the labeled as well as the unlabeled data as training examples according to their expected labels. In step 314 classification scores are determined using the trained classifier of 312. In 316 classification scores are calibrated to class membership probability. In step 318, label prior probability information is updated according to the class membership probability. An MED calculation is performed in step 320 to determine label and margin probability distributions, wherein the previously determined classification scores are used in the MED calculation.

As a result, new expected labels are computed at step 322 and the expected labels are updated in step 324 using the computations from step 322. At step 326 the method determines whether convergence has been achieved. If so, the method terminates at step 328. If convergence is not reached, another iteration of the method is completed starting with step 310. Iterations are repeated until convergence is reached thus resulting in an iterative training of the MED classifier. Convergence may be reached when change of the decision function between each iteration of the MED calculation falls below a predetermined value. In an alternative embodiment of the present invention, convergence may be reached when the change of the determined expected label value falls below a predetermined threshold value.

FIG. 7: is a network architecture 700, in accordance with one embodiment. As shown, a plurality of remote networks 702 are provided including a first remote network 704 and a second remote network 706. A gateway 707 may be coupled between the remote networks 702 and a proximate network 708. In the context of the present network architecture 700, the networks 704, 706 may each take any form including, but not limited to a LAN, a WAN such as the Internet, PSTN, internal telephone network, etc.

In use, the gateway 707 serves as an entrance point from the remote networks 702 to the proximate network 708. As such, the gateway 707 may function as a router, which is capable of directing a given packet of data that arrives at the gateway 707, and a switch, which furnishes the actual path in and out of the gateway 707 for a given packet.

Further included is at least one data server 714 coupled to the proximate network 708, and which is accessible from the remote networks 702 via the gateway 707. It should be noted that the data server(s) 714 may include any type of computing device/groupware.

Coupled to each data server 714 is a plurality of user devices 716. Such user devices 716 may include a desktop computer, lap-top computer, hand-held computer, printer or any other type of logic. It should be noted that a user device 717 may also be directly coupled to any of the networks, in one embodiment.

A facsimile machine 720 or series of facsimile machines 720 may be coupled to one or more of the networks 704, 706, 708. It should be noted that databases and/or additional components may be utilized with, or integrated into, any type of network element coupled to the networks 704, 706, 708. In the context of the present description, a network element may refer to any component of a network.

FIG. 8: shows a representative hardware environment associated with a user device 716 of FIG. 7, in accordance with one embodiment. Such FIG. illustrates a typical hardware configuration of a workstation having a central processing unit 810, such as a microprocessor, and a number of other units interconnected via a system bus 812.

The workstation shown in FIG. 8 includes a Random Access Memory (RAM) 814, Read Only Memory (ROM) 816, an I/O adapter 818 for connecting peripheral devices such as disk storage units 820 to the bus 812, a user interface adapter 822 for connecting a keyboard 824, a mouse 826, a speaker 828, a microphone 832, and/or other user interface devices such as a touch screen and a digital camera (not shown) to the bus 812, communication adapter 834 for connecting the workstation to a communication network 835 (e.g., a data processing network) and a display adapter 836 for connecting the bus 812 to a display device 838.

Referring particularly to FIG. 9 there is shown the apparatus 414 of one embodiment of the present invention. One embodiment of the present invention comprises in memory device 404 for storing labeled data 416. The labeled data points 416 each include a label indicating whether the data point is a training example for data points being included in the designated category or a training example for data points being excluded from a designated category. Memory 404 also stores unlabeled data 418, prior probability data 420 and the cost factor data 422.

The processor 402 accesses the data from the memory 404 and using transductive MED calculations trains a binary classifier enable it to classify unlabeled data. The processor 402 uses iterative transductive calculation by using the cost factor and training examples from labeled and unlabeled data and scaling that cost factor as a function of expected label value thus effecting the data of the cost factor data 422 which is then reinput into processor 402. Thus the cost factor 422 changes with each iteration of the MED classification by the processor 402. Once the processor 402 adequately trains an MED classifier, the processor can then construct the classifier to classify the unlabeled data into classified data 424.

FIG. 10: labeled data points are received at step 1002, where each of the labeled data points has at least one label which indicates whether the data point is a training example for data points for being included in a designated category or a training example for data points being excluded from a designated category. In addition, unlabeled data points are received at step 1004, as well as at least one predetermined cost factor of the labeled data points and unlabeled data points. The data points may contain any medium, e.g. words, images, sounds, etc. Prior probability information of labeled and unlabeled data points may also be received. Also, the label of the included training example may be mapped to a first numeric value, e.g. +1, etc., and the label of the excluded training example may be mapped to a second numeric value, e.g. -1, etc. In addition, the labeled data points, unlabeled data points, and at least one predetermined cost factor of the labeled data points.

Further, at step 1006 a transductive MED classifier is trained through iterative calculation using said at least one cost factor and the labeled data points and the unlabeled data points as training examples. For each iteration of the calculations, the unlabeled data point cost factor is adjusted as a function of an expected label value, e.g. the absolute value of the expected label of a data point, etc., and a data point label prior probability is adjusted according to an estimate of a data point class membership probability, thereby ensuring stability. Also, the transductive classifier may learn using prior probability information of the labeled and unlabeled data, which further improves stability. The iterative step of training a transductive classifier may be repeated until the convergence of data values is reached, e.g. when the change of the decision function of the transductive classifier falls below a predetermined threshold value, when the change of the determined expected label value falls below a predetermined threshold value, etc.

Additionally, in step 1008 the trained classifier is applied to classify at least one of the unlabeled data points, the labeled data points, and input data points. Input data points may be received before or after the classifier is trained, or may not be received at all. Also, the decision function that minimizes the KL divergence to the prior probability distribution of the decision function parameters given the included and excluded training examples may be determined utilizing the labeled as well as the unlabeled data points as learning examples according to their expected label. Alternatively, the decision function may be determined with minimal KL divergence using a multinomial distribution for the decision function parameters.

In step 1010 a classification of the classified data points, or a derivative thereof, is output to at least one of a user, another system, and another process. The system may be remote or local. Examples of the derivative of the classification may be, but are not limited to, the classified data points themselves, a representation or identifier of the classified data points or host file/document, etc.

The computer executable program code is deployed to and executed on a computer system. This program code comprises instructions for accessing stored labeled data points in a memory of a computer, where each of said labeled data points has at least one label indicating whether the data point is a training example for data points for being included in a designated category or a training example for data points being excluded from a designated category. In addition, the computer code comprises instructions for accessing unlabeled data points from a memory of a computer as well as accessing at least one predetermined cost factor of the labeled data points and unlabeled data points from a memory of a computer. Prior probability information of labeled and unlabeled data points stored in a memory of a computer may also be accessed. Also, the label of the included training example may be mapped to a first numeric value, e.g. +1, etc., and the label of the excluded training example may be mapped to a second numeric value, e.g. -1, etc.

Further, the program code comprises instructions for training a transductive classifier through iterative calculation, using the at least one stored cost factor and stored labeled data points and stored unlabeled data points as training examples. Also, for each iteration of the calculation, the unlabeled data point cost factor is adjusted as a function of the expected label value of the data point, e.g. the absolute value of the expected label of a data point. Also, for each iteration, the prior probability information may be adjusted according to an estimate of a data point class membership probability. The iterative step of training a transductive classifier may be repeated until the convergence of data values is reached, e.g. when the change of the decision function of the transductive classifier falls below a predetermined threshold value, when the change of the determined expected label value falls below a predetermined threshold value, etc.

Additionally, the program code comprises instructions for applying the trained classifier to classify at least one of the unlabeled data points, the labeled data points, and input data points, as well as instructions for outputting a classification of the classified data points, or derivative thereof, to at least one of a user, another system, and another process. Also, the decision function that minimizes the KL divergence to the prior probability distribution of the decision function parameters given the included and excluded training examples may be determined utilizing the labeled as well as the unlabeled data as learning examples according to their expected label.

The processing apparatus comprises at least one memory for storing: (i) labeled data points, wherein each of said labeled data points have at least one label indicating whether the data point is a training example for data points being included in a designated category or a training example for data points being excluded from a designated category; (ii) unlabeled data points; and (iii) at least one predetermined cost factor of the labeled data points and unlabeled data points. The memory may also store prior probability information of labeled and unlabeled data points. Also, the label of the included training

example may be mapped to a first numeric value, e.g. +1 , etc., and the label of the excluded training example may be mapped to a second numeric value, e.g. -1, etc.

In addition, the data processing apparatus comprises a transductive classifier trainer to iteratively teach the transductive classifier using transductive Maximum Entropy Discrimination (MED) using the at least one stored cost factor and stored labeled data points and stored unlabeled data points as training examples. Further, at each iteration of the MED calculation the cost factor of the unlabeled data point is adjusted as a function of the expected label value of the data point, e.g. the absolute value of the expected label of a data point, etc. Also, at each iteration of the MED calculation, the prior probability information may be adjusted according to an estimate of a data point class membership probability. The apparatus may further comprise a means for determining the convergence of data values, e.g. when the change of the decision function of the transductive classifier calculation falls below a predetermined threshold value, when the change of the determined expected label values falls below a predetermined threshold value, when the change of the determined convergence.

In addition, a trained classifier is used to classify at least one of the unlabeled data points, the labeled data points, and input data points. Further, the decision function that minimizes the KL divergence to the prior probability distribution of the decision function parameters given the included and excluded training examples may be determined by a processor utilizing the labeled as well as the unlabeled data as learning examples according to their expected label. Also, a classification of the classified data points, or derivative thereof, is output to at least one of a user, another system, and another process.

A article of manufacture comprises a program storage medium readable by a computer, where the medium tangibly embodies one or more programs of instructions executable by a computer to perform a method of data classification. In use, labeled data points are received, where each of the labeled data points has at least one label which indicates whether the data point is a training example for data points for being included in a designated category or a training example for data points being excluded from a designated category. In addition, unlabeled data points are received, as well as at least one predetermined cost factor of the labeled data points and unlabeled data points. Prior probability information of labeled and unlabeled data points may also be stored in a memory of a computer. Also, the label of the included training example may be mapped to a second numeric value, e.g. -1, etc.

Further, a transductive classifier is trained with iterative Maximum Entropy Discrimination (MED) calculation using the at least one stored cost factor and the stored labeled data points and the unlabeled data points as training examples. At each iteration of the MED calculation, the unlabeled data point cost factor is adjusted as a function of an

expected label value of the data point, e.g. the absolute value of the expected label of a data point, etc. Also, at each iteration of the MED calculation, the prior probability information may be adjusted according to an estimate of a data point class membership probability. The iterative step of training a transductive classifier may be repeated until the convergence of data values is reached, e.g. when the change of the decision function of the transductive classifier falls below a predetermined threshold value, when the change of the determined expected label value falls below a predetermined threshold value, etc.

Additionally, input data points are accessed from the memory of a computer, and the trained classifier is applied to classify at least one of the unlabeled data points, the labeled data points, and input data points. Also, the decision function that minimizes the KL, divergence to the prior probability distribution of the decision function parameters given the included and excluded training examples may be determined utilizing the labeled as well as the unlabeled data as learning examples according to their expected label. Further, a classification of the classified data points, or a derivative thereof, is output to at least one of a user, another system, and another process.

A method for classification of unlabeled data in a computer-based system is presented. In use, labeled data points are received, each of said labeled data points having at least one label indicating whether the data point is a training example for data points for being included in a designated category or a training example for data points being excluded from a designated category. Additionally, labeled and unlabeled data points are received, as are prior label probability information of labeled data points and unlabeled data points. Further, at least one predetermined cost factor of the labeled data points and unlabeled data points is received.

Further, the expected labels for each labeled and unlabeled data point are determined according to the label prior probability of the data point. The following sub steps are repeated until substantial convergence of data values:

- 1. generating a scaled cost value for each unlabeled data point proportional to the absolute value of the data point's expected label;
- 2. training a Maximum Entropy Discrimination (MED) classifier by determining the decision function that minimizes the KL divergence to the prior probability distribution of the decision function parameters given the included training and excluded training examples utilizing the labeled as well as the unlabeled data as training examples according to their expected label;
- 3. determining the classification scores of the labeled and unlabeled data points using the trained classifier;
- 4. calibrating the output of the trained classifier to class membership probability;

- 5. updating the label prior probabilities of the unlabeled data points according to the determined class membership probabilities;
- 6. determining the label and margin probability distributions using Maximum Entropy Discrimination (MED) using the updated label prior probabilities and the previously determined classification scores;
- 7. computing new expected labels using the previously determined label probability distribution; and
- 8. updating expected labels for each data point by interpolating the new expected labels with the expected label of previous iteration.

Also, a classification of the input data points, or derivative thereof, is output to at least one of a user, another system, and another process.

Convergence may be reached when the change of the decision function falls below a predetermined threshold value. Additionally, convergence may also be reached when the change of the determined expected label value falls below a predetermined threshold value. Further, the label of the included training example may have any value, for example, a value of +1, and the label of the excluded training example may have any value, for example, a value of -1.

The invention, a method for classifying documents is presented in FIG. 11. In use, at least one seed document having a known confidence level is received in step 1100, as well as unlabeled documents and at least one predetermined cost factor. The seed document and other items may be received from a memory of a computer, from a user, from a network connection, etc., and may be received after a request from the system performing the method. The at least one seed document may have a label indicative of whether the document is included in a designated category, may contain a list of keywords, or have any other attribute that may assist in classifying documents. Further, in step 1102 a transductive classifier is trained through iterative calculation using the at least one predetermined cost factor, the at least one seed document, and the unlabeled documents, wherein for each iteration of the calculations the cost factor is adjusted as a function of an expected label value. A data point label prior probability for the labeled and unlabeled documents may also be received, wherein for each iteration of the calculations the data point label prior probability may be adjusted according to an estimate of a data point class membership probability.

Additionally, after at least some of the iterations, in step 1104 confidence scores are stored for the unlabeled documents, and identifiers of the unlabeled documents having the highest confidence scores are output in step 1106 to at least one of a user, another system, and another process. The identifiers may be electronic copies of the document themselves, portions thereof, titles thereof, names thereof, file names thereof, pointers to the documents, etc. Also, confidence scores may be stored after each of the iterations,

wherein an identifier of the unlabeled document having the highest confidence score after each iteration is output.

The invention is capable of discovering patterns that link the initial document to the remaining documents. The task of discovery is one area where this pattern discovery proves particularly valuable. For instance, in pre-trial legal discovery, a large amount of documents has to be researched with regard to possible connections to the lawsuit at hand. The ultimate goal is to find the "smoking gun." In another example, a common task for inventors, patent examiners, as well as patent lawyers is to evaluate the novelty of a technology through prior art search. In particular, the task is to search all published patents and other publications and find documents within this set that might be related to the specific technology that is examined with regard to its novelty.

The task of discovery involves finding a document or a set of documents within a set of data. Given an initial document or concept, a user may want to discover documents that are related to the initial document or concept. However, the notion of relationship between the initial document or concept and the target documents, i.e. the documents that are to be discovered, is only well understood after the discovery has taken place. By learning from labeled and unlabeled documents, concepts, etc., the invention can learn patterns and relationships between the initial document or documents and the target documents.

WE CLAIMS

- 1. Our invention" IML- Data Cleaning "is a system and article of manufacture enabling adapting to a shift in document content and also the instructions for receiving at least one labeled mapped seed document receiving unlabeled mapped documents receiving at least one predetermined cost factor training data a transductive classifier using the at least one predetermined cost factor calculated data and at least one seed document and the unlabeled documents. The invention also classifying the unlabeled documents having a confidence level above a predefined threshold into a plurality of indexing and categories using the classifier reclassifying at least some of the categorized documents into the categories using the classifier and outputting identifiers of the categorized documents to at least one of a user another system and another process. The invented systems and articles of also manufacture for separating documents are also presented and the systems and articles of manufacture for document searching are also presented a business and other information service provides data cleansing to correct and update both domestic and global addresses. The invented system a integrate and combination of processes generates cleansed data for input into a matching and mapping process and the matching, the mapping process matches information about a business. The invented technology data cleaning process includes the steps of validating data loaded from at least two source systems appending the validated data to a normalized data cleaning repository selecting the priority of the source system creating a clean database loading the consistent, normalized, and cleansed data from the clean database into a format required by data systems and software tools using the data. The invented technology also creating reports and updating the clean database by a user without updating the source systems. The data cleaning process distributed and collecting, analyzing data from available sources for optimization models enabling consistent analysis. The invented technology the data cleaning process further provides complete auditability to the inputs and outputs of data systems and software tools that use a dynamic data set.
- 2. According to claim1# the invention is to a system and article of manufacture enabling adapting to a shift in document content and also the instructions for receiving at least one labeled mapped seed document receiving unlabeled mapped documents receiving at least one predetermined cost factor training data a transductive classifier using the at least one predetermined cost factor calculated data and at least one seed document and the unlabeled documents.
- 3. According to claim1,2# the invention is to the invention also classifying the unlabeled documents having a confidence level above a predefined threshold into a plurality of indexing and categories using the classifier reclassifying at least some of the categorized documents into the categories using the classifier and

outputting identifiers of the categorized documents to at least one of a user another system and another process.

- 4. According to claim1,2,3# the invention is to the invented systems and articles of also manufacture for separating documents are also presented and the systems and articles of manufacture for document searching are also presented a business and other information service provides data cleansing to correct and update both domestic and global addresses.
- 5. According to claim1,2,4# the invention is to the invented system a integrate and combination of processes generates cleansed data for input into a matching and mapping process and the matching, the mapping process matches information about a business.
- 6. According to claim1,2,4,5# the invention is to the invented technology data cleaning process includes the steps of validating data loaded from at least two source systems appending the validated data to a normalized data cleaning repository selecting the priority of the source system creating a clean database loading the consistent, normalized, and cleansed data from the clean database into a format required by data systems and software tools using the data.
- 7. According to claim1,2,5# the invention is to the invented technology also creating reports and updating the clean database by a user without updating the source systems. The data cleaning process distributed and collecting, analyzing data from available sources for optimization models enabling consistent analysis.
- 8. According to claim1,2,4,6,7# the invention is to the invented technology the data cleaning process further provides complete auditability to the inputs and outputs of data systems and software tools that use a dynamic data set.

Data: 1/9/2020

Manohar Madgi (Associate Professor) Dr. Ramesh Shahabadkar (Professor)

Dr. Nagaraj B. Patil (Principal)

Dr. Bharati Harsoor (Professor & Head)

Ashok Patil (Assistant Professor)

Padmapriya Patil (Assistant Professor)

Dr. Basavaraj Chunchure (Associate Professor)

Dr. R. Balakrishna (Professor & Dean)

Dr. Rajesh K. S. (Associate Professor)

Dr. Sai Madhavi D. (Associate Professor)

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG.1: DATA CLEANING BLOCK DIAGRAM.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)

Data Cleaning Process Parsing Correction Standardizing Consolidation



FIG.2: DATA CLEANING PROCESS.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG.3: DATA CLEANING IMPORTANCE.



FIG. 4: IS A CONTROL FLOW DIAGRAM FOR THE CLASSIFICATION OF UNLABELED DATA IN WITH ONE OF THE INVENTION USING A SCALED COST FACTOR.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG. 5: IS A CONTROL FLOW DIAGRAM FOR THE CLASSIFICATION OF UNLABELED DATA IN ACCORDANCE WITH ONE OF THE INVENTION USING USER DEFINED PRIOR PROBABILITY INFORMATION.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG. 6: IS A DETAILED CONTROL FLOW DIAGRAM FOR THE CLASSIFICATION OF UNLABELED DATA IN ACCORDANCE WITH ONE EMBODIMENT OF THE INVENTION USING MAXIMUM ENTROPY DISCRIMINATION WITH SCALED COST FACTORS AND PRIOR PROBABILITY INFORMATION.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG. 7: IS A NETWORK DIAGRAM ILLUSTRATING A NETWORK ARCHITECTURE IN WHICH THE VARIOUS DESCRIBED HEREIN MAY BE IMPLEMENTED.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG. 8: IS A SYSTEM DIAGRAM OF A REPRESENTATIVE HARDWARE ENVIRONMENT ASSOCIATED WITH A USER DEVICE.



FIG. 9: IS A BLOCK DIAGRAM REPRESENTATION OF THE APPARATUS OF ONE EMBODIMENT OF THE PRESENT INVENTION.

Dr. Ramesh Shahabadkar (Professor) Dr. Bharati Harsoor (Professor & Head) Padmapriya Patil (Assistant Professor) Dr. R. Balakrishna (Professor & Dean) Dr. Sai Madhavi D. (Associate Professor)



FIG. 10: IS A FLOWCHART, A CLASSIFICATION PROCESS.



FIG. 11: IS A FLOWCHART, A CLASSIFICATION PROCESS PERFORMED BY IN ACCORDANCE WITH ONE EMBODIMENT.

(12) PATENT APPLICATION PUBLICATION

(19) INDIA

(22) Date of filing of Application :28/12/2019

(21) Application No.201941054348

2

(43) Publication Date : 03/01/2020

(54) Title of the invention : ICM-PERFORMANCE: INCREASE THE CACHE MEMORY PERFORMANCE USING NEURAL

(51)		-
Internationa	:G06F0012121000,G06N0003080000 C0 C0	(71)Name of Applicant :
classification	1 George Contraction Contracti	DMANOHAR COSUL
(31) Priority		Address of Applicant
Document	:NA	DEPARTMENT OF COMPUTER
No		SCIENCE & ENGINEERING. H.NO
(22) Priority		40/1, MANJULA NILAYA,
(32) I month	:NA	PAVANGANGA COLONY,
(22) Name		RAJAPUR, NANDUR ROAD,
(33) Name	·NA	GULBARGA-585105, KARNATAKA,
of priority	1 V F	INDIA. e-mail:
country		manohargosul5@gmail.com Andhra
(86)		Pradesh India
International		2)MANOHAR MADGI
Application	PC1//	3)DR .SHANTALA GIRADDI
No	-01/01/1900	4)DR. VINEET TIRTH
Filing	·	5)DR. NEERAJ KUMAR
Date		SHUKLA (ASSOCIATE
(87)		PROFESSOR ELECTRICAL
International	. NA	ENGINEERING DEPARTMENT)
Publication		6)MR. UDAY BALGAR
No		(72)Name of Inventor :
(61) Patent		1)MANOHAR GOSUL
61) I dition		2)MANOHAR MADGI
of Authon		3)DR .SHANTALA GIRADDI
to	NA	4)DR. VINEET TIRTH
Application	NA	5)DR. NEERAJ KUMAR
Number		SHUKLA (ASSOCIATE
Filing		PROFESSOR ELECTRICAL
Date		ENGINEERING DEPARTMENT)
(62)	3	OMR. UDAY BALGAR
Divisional to	N	
Annlication	NA	•
Number	NA	
Filing		
rung		

Date

ICM-PERFORMANCE: INCREASE THE CACHE MEMORY PERFORMANCE USING NEURAL NETWORKS, DEEP LEARNING. ABSTRACT My Invention • ICM- PERFORMANCE • is a Systems and methods for selecting an appropriate caching algorithm to be used when temporarily storing data accessed by an executing application using a neural network may dynamically and/or iteratively replace an initial caching algorithm being used for the application. An input layer of the neural network may gather values of performance related parameters, such as cache hit rates, data throughput rates, or memory access request response times. The neural network may detect a pattern or change in a pattern of accesses, or a change in a workload, a hardware component, or an operating system parameter. Dependent on these and/or other inputs, the neural network may select and apply a caching algorithm likely to improve performance of the application. Other inputs to the neural network may include values of hardware configuration parameters and/or operating system parameters. The neural network may perform a training exercise or may be self-training,

e.g., using reinforcement learning.

No. of Pages : 27 No. of Claims : 9

The Patent Office Journal No. 01/2020 Dated 03/01/2020

245

|\u101#inbox/FMfcazGaPzHKHCArpIDFJnxmwhBFlkTq?projector=1&messagePartId=0.5



Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India



Application Details			
APPLICATION NUMBER	292041009754		
APPLICATION TYPE	ORDINARY APPLICATION		
DATE OF FILING	08/01/2020		
APPLICANT NAME	1 . DR. NAGESH SALIMATH 2 . SHARANABASAPPA C GANDAGE 3 . SHRUTI A G 4 . RAJALAXNI S BILGUNDI 5 . CHETANKUMAR KALASKAR 6 . SATISHKUMAR MALLAPPA		
TITLE OF INVENTION	MACHINE LEARNING BASED METHOD FOR CLASSIFICATION OF SPECTRAL IMAGES		
FIELD OF INVENTION	COMPUTER SCIENCE		
E-MAIL (As Per Record)	nageshsallmath@pdaengg.com		
ADDITIONAL-EMAIL (As Per Record)	nageshsallmath@pdaengg.com		
E-MAIL (UPDATED Online)			
PRIGRITY DATE			
REQUEST FOR EXAMINATION DATE			
PUBLICATION DATE (UVS 11A)	17/01/2020		

Application Status



PATENT APPLICATION PUBLICATION

NDIA

mate of filing of Application :23/06/2020

(21) Application No.202041026402 A

(43) Publication Date : 17/07/2020

Diffe of the invention : COVID19 PROTECTED ROOM: DISINFECTING ROOM AIR USING MACHINE LEARNING STEM

(57) Abstract

Patent Title: Covid19 Protected Room: DISINFECTING ROOM AIR USING MACHINE LEARNING SYSTEM. ABSTRACT My ivention Covid19 Protected Room• is a system for disinfecting a room having first, second and third air inlets and an air intake control using machine learning assembly to prefer control air flow into the enclosure through the first, second and third air inlets. The Air flows between the exterior and interior of the enclosure through the second, third air inlet passes through a Covid19 killer filter assembly. The invention also includes an air dispersion outlet having a fan that draws air into the invention through the first, second and third air inlets and forces air out of the invention. The chemical dispersion assembly generates a disinfecting clean fog relative to the fan. The controller system (Machine learning Programming) controls the air intake control assembly to disperse the disinfecting clean fog into the room, and subsequently draw the disinfecting clean fog from the room and through the Covid19 killer filter assembly. Also UVC radiation onto complete room, disinfection controlled by Machine learning programming. The invention mprises a disinfection kit comprising a housing comprising a like a blower or fan for blowing air, a transparent tubular member for eiving air blown by the blower or fan and for directing it in a fixed direction. The invented kit/ devices is shown for providing ubstantially simultaneous disinfection of covid19, bacteria or other pathogens in air passing through the system using ultraviolet radiation from an ultraviolet light source and that is also capable of substantially simultaneous and efficient disinfecting room air and surfaces outside the device using the same ultraviolet light source.

No. of Pages : 23 No. of Claims : 6

The Patent Office Journal No. 29/2020 Dated 17/07/2020

6
G S D O

iffice of the Controller General of Patents, Designs & Trade Marks repartment of Industrial Policy & Promotion. Anistry of Commerce & Industry, Sovernment of India



Reprint and	
APPLICATION NUMBER	202141036593
APPLICATION TYPE	ORDINARY APPLICATION
DATE OF FILING	12/08/2021
APPLICANT NAME	1. Ingeniouz
is a	 Dr.N.Raghavendra Sai, Koneru Lakshmaiah Education Foundation Poornima Hulipalled, Vijayanagara Sri Krishnadevaraya University Drakshaveni G. BMSITM
	 Anuradha Fatil, sharnbasva university kalaburagi Vishwanath Burkpalli, PDA Engineering college Shweta Madiwal, Sharn basva University, Kalaburagi Prof. fDr. Walini Kanta Sahar Sawa Source (new set to be University)
	 Mr. Bhavani Sankar Panda, SOET, GIET University Dr. S. Hariharan, Shadan Women's College of Engineering and Technology Mr. D.Naveen Raju, Sri Sairam Institute of Technology Dr. Urmila R. Ppl, Shivaji University Vinit Gupta, Medi-Caps University
TATLE OF-INVENTION	Image processing & Machine Learning Based Detection of Malignant Skin Cancer Using Support Vector Machine (SVM) classifier
FIELD OF INVENTION	BIO-MEDICAL ENGINEERING
E-MAIL (As Per Record)	Ingeniouz1@gmail.com
ADDITIONAL-EMAIL (As Per Record)	
E-MAIL (UPDATED Online)	

PUBLICATION DATE (U/S 11A) 24/09/2021

1.1

REQUEST FOR EXAMINATION DATE

;

PRIORITY DATE

_
5
Ð
lectua
Prop
Perty
Indi
5

~



Government of India Ministry of Commerce & Industry, Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion,

http://ipindia.nic.in/index.htm)



(F	
ttp://i	
pindi	
a.nic.	
in/inc	
dex.h	
itm)	

Application Status

Awaiting Request for Examination

APPLICATION STATUS

View Documents



APPLICATION STATUS

ę

Awaiting Request for Examination

Application Status

(http://ipindia.r	nic.in/index.htm)
惫	
9 <u>1</u>	INTELLECTUAL (http://ipindia.nic.in/index.i PROPERTY INDIA
	Application Details
APPLICATION NUMBER	202141044110
APPLICATION TYPE	ORDINARY APPLICATION
DATE OF FILING	29/09/2021
APPLICANT NAME	Dr. Anand R
TITLE OF INVENTION	FARMER WITH STREAMLINED AGRICULTURE PROCESS
FIELD OF INVENTION	COMMUNICATION
E-MAIL (As Per Record)	anand.rajendran@reva.edu.in
ADDITIONAL-EMAIL (As P Record)	
E-MAIL (UPDATED Online)	
PRIORITY DATE	
REQUEST FOR EXAMINATION DATE	
PUBLICATION DATE (U/S 11A)	10/06/2022



T

010124, 12.00 PM

Intellectual Property India

Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion,



Office of the Controller General of Patents, Designs & Trade Marks Department of Industrial Policy & Promotion, Ministry of Commerce & Industry, Government of India

(http://ipindia.nic.in/index.htm)



(http://ipindia.nic.in/index.htm)

Application Details		
APPLICATION NUMBER	202241035711	
APPLICATION TYPE	ORDINARY APPLICATION	
DATE OF FILING	22/06/2022	
APPLICANT NAME	 GURAPPA KALYANI DR.VISWANATHAN RAMASAMY DR.VISHAL RATANSING PATIL DR.GARIMA SRIVASTAVA DR.R.THIAGARAJAN DR.R.RAMKUMAR MS.J.OMANA MR.SUBBARAYADU 	
TITLE OF INVENTION	A SAFE GUARD SYSTEM FOR MINE WORKERS USING WIRELESS SENOR NETWORKS	
FIELD OF INVENTION	ELECTRONICS	
E-MAIL (As Per Record)		
ADDITIONAL-EMAIL (As Per Record)	rthiyagarajantpt@gmail.com	
E-MAIL (UPDATED Online)		
PRIORITY DATE		
REQUEST FOR EXAMINATION DATE		
PUBLICATION DATE (U/S 11A)	01/07/2022	

Application Status

APPLICATION STATUS

Awaiting Request for Examination

Intellectual Property India

