

Australian Government
IP Australia
Innovation Patent Australia

Patent Title: **AMD-Data Analysis: ACADEMY DATA ANALYSIS USING MACHINE LEARNING, DEEP LEARNING PROGRAMMING.**

Name and address of patentees(s):

Dr. Ashendra Kumar Saxena (Associate Professor) Address: College of Computing Sciences and IT, Teerthankar Mahaveer University, Moradabad, (UP), India.
MD. Iqbal (Associate Professor) (Computer Science and Engineering) Address: Meerut Institute of Engineering & Technology, Meerut (UP), India-250005.
Mr. Rupesh Kumar Jindal (Assistant Professor) Address: Sharda University, Knowledge Park III, Greater Noida, UP-201310, India.
Dr. Santosh Dilipkumar Parakh (Associate Professor) Address: Arihant Clinic, Aashirwad bungalow, Near Z.P School, Shivaji Nagar, Wadujkar Estate, Tal-Baramati, Dist.- Pune, Maharashtra, India, 413102.
Dr. Rakesh Kumar ER, (Manager Examination) Address: Texila American University, LOT 24-42, Plantation Providence, East Bank Demerara, Georgetown, Guyana, South America.
Dr. Chukka Santhaiah (Associate Professor) Address: Department of CSE, Institute of Aeronautical Engineering, Hyderabad, Telangana, India. E-Mail: chukka.santh@gmail.com[
Prof.(Dr.) S. B. Chordiya (Director-SIMMC-Campus) Address: Suryadatta Institute of Management & Mass Communication (SIMMC) Sr. No: 342, Bavdhan, Pune-411021, MH, India.
Dr. Rambabu Vatti (Principal) Office Address: Principal, Bharat Institute of Engineering and Technology, Hyderabad, Telangana, India-501510. Residential Address: Dr. Rambabu Vatti, Prasanna Lakshmi Nilayam, Lane no.5, Survey no,59, Sai Nagar, Kondhwa(Bk), Pune, Maharashtra-411048, India.

Complete Specification: **Australian Government.**

FIELD OF THE INVENTION

Our invention “**AMD-Data Analysis**” is related to academy data analysis using machine learning, deep learning programming to classifying of data with a deep learning neural network.

BACKGROUND OF THE INVENTION

Rule-based classification systems to classify discrete sets of data are often difficult and expensive to maintain, and often insufficient for tasks involving large, varying, and/or complex data sets. In some cases, these systems may be prone to failure if faced with data that varies or changes over time, or data that contains variations within the classes themselves. In some cases, rule-based algorithms designed in advance may be ineffective at classifying the live data. Also, the manual design of effective rule-based classifiers may become difficult as the classification options become more complex. Also, it may be difficult to identify the features in the source data that may be used for effective automatic classification of data. Thus, it is with respect to these considerations and others that the invention has been made.

Flow cytometers are typically used to analyze the properties of single cells. For example, as a single cell suspension interrupts a laser beam of the flow cytometry system at high velocity, it produces a scattering of light from the beam. Data is generally relayed to a computer for interpretation of the results. These systems are typically designed for the enumeration, identification, and sorting of cells possessing selected properties. Fluorescence-activated cell sorting (FACS) is a specific type of flow cytometry, which utilizes fluorescent markers (e.g., fluorochrome-labeled monoclonal antibodies) to label cells in order to detect and sort the cells as part of multi-parameter analyses.

Flow cytometry fluorescence measurement data is currently displayed using either logarithmic or linear scaling. In most applications linear scaling fails to provide appropriate resolution across the typical data range of up to 10,000:1. Logarithmic displays are unable to deal with negative data values and typically introduce biologically art factual peaks, particularly in data derived through fluorescence compensation. The result is that both the compactness and central tendency of low signal cell populations is severely obscured. Previous attempts to develop improved visualizations (e.g., displaying cytometry data for a human viewer) have not been very successful in that they have involved seriously compromising quantitation and/or introduced their own artifacts into the display (e.g., a simple linear-to-log splice tends to introduce a distinct transition line into the display).

PRIOR ART SEARCH

US20130132331A1 *2010-03-082013-05-23National Ict. Australia Limited
Performance evaluation of a classifier.

US8468385B1 *2010-10-272013-06-18Netapp, Inc. Method and system for
handling error events.

US20150039543A1 *2013-07-312015-02-05Balakrishnan Athmanathan Feature
Based Three Stage Neural Network Intrusion Detection.

US20160305865A1 *2015-04-172016-10-20Hamilton Sundstrand Corporation
Wavelet-based analysis for fouling diagnosis of an aircraft heat exchanger.

US20170017793A1 *2015-07-152017-01-19Cylance Inc. Malware detection.

US20170093977A1 *2015-09-292017-03-30Fujitsu Limited Pattern transmission
method, control method, and system.

CN106709345A *2015-11-172017-05-24 Deep learning method-based method
and system for deducing malicious code rules and equipment.

US9690937B1 *2015-03-302017-06-27EMC IP Holding Company LLC
Recommending a set of malicious activity detection rules in an automated, data-
driven manner.

US9705904B1 *2016-07-212017-07-11Cylance Inc. Neural attention mechanisms
for malware analysis

WO2017132428A1 *2016-01-292017-08-03Yahoo! Inc. Method and system for
distributed deep machine learning.

WO2017165801A1 *2016-03-242017-09-28The Regents Of The University Of
California Deep-learning-based cancer classification using a hierarchical
classification framework.

OBJECTIVES OF THE INVENTION

- 1) The objective of the invention is to an embodiments are directed towards classifying data using machine learning, deep learning that may be incrementally refined based on expert input.
- 2) The other objective of the invention is to the invented technology is also the provided to a deep learning model that may be trained based on a classifiers and sets of mapping, training data and testing data.
- 3) The other objective of the invention is to If the number of classification errors exceeds a defined fixed table data and threshold classifier may be modified based on data corresponding to observed classification errors.
- 4) The other objective of the invention is to the invented technology a three type of learning model (fast, medium, slower) may be trained based on the modified classifiers the data and the data corresponding to the observed

classification errors and another confidence value may be generated and associated with the classification of the data by the all type of learning, mapping model.

- 5) The other objective of the invention is to the invented technology re-information may be generated based on a comparison result of the confidence value associated with the all learning model and the confidence value associated with the machine, deep learning model. The invention also provides methods of analyzing or displaying data.
- 6) The other objective of the invention is to the invention provides methods for visualizing or displaying high dynamic range data obtained from flow cytometry analyses. Related systems and computer programs products are also provided.

SUMMARY OF THE INVENTION

In certain embodiments of the method, using comprises displaying the scaled data for a human viewer. For example, the scaled data is typically displayed on a coordinate grid and the scaling function primarily depends on data in a single data dimension to assure that the coordinate grid is substantially rectilinear. Display values generally increase in size more than corresponding display variables in linear regions of the scaled data as a family-generating variable is adjusted to increase a range of linearity.

The scaling function typically includes at least one generalized hyperbolic sine function. In some embodiments, the generalized hyperbolic sine function is in a form of $V=Z(10^{n/m}-1-G^2(10^{-n/mG}-1))$, where V is a data value to be displayed at channel position n in a plot of said scaled data, m is the asymptotic channels per decade, and G is linearization strength. In certain embodiments, the generalized hyperbolic sine function is a form of $V=a(e^x-p^2e^{-px}+p^2-1)$, where V is a data value to be plotted at display position x in a plot, a is a scaling factor, and p is linearization strength. Optionally, the generalized hyperbolic sine function is a form of $S(x; a, b, c, d, So)=ae^{bx}-ce^{-dx}-So$, for positive x and for negative x , a reflection of the positive x in a form of $Sref(x; a, b, c, d, So)=(x/absx) S(absx; a, b, c, d, So)$, where $absx$ is the absolute value of variable x . In some embodiments, using comprises inputting said scaled data into at least one data analysis algorithm (e.g., automated data analysis software, such as cluster analysis software and the like) to identify the portions of the raw data of interest.

In another aspect, the present invention relates to a computer program product that includes a computer readable medium having one or more logic instructions for receiving raw data in a computer, which raw data comprises data from a plurality of light detectors of a flow cytometry system, and scaling the raw data

using at least one scaling function that provides substantially linear transformations for data values proximal to zero and substantially logarithmic transformations for other data values to generate scaled data. The computer readable medium typically includes one or more of, e.g., a CD-ROM, a floppy disk, a tape, a flash memory device or component, a system memory device or component, a hard drive, a data signal embodied in a carrier wave, or the like.

In still another aspect, the invention provides a system for analyzing flow cytometry data. The system includes (a) at least one flow cytometer, and (b) at least one computer operably connected to the flow cytometer, which computer has system software. The system software includes one or more logic instructions for receiving raw data in the computer, which raw data comprises data from a plurality of light detectors of a flow cytometry system, and scaling the raw data using at least one scaling function that provides substantially linear transformations for data values proximal to zero and substantially logarithmic transformations for other data values to generate scaled data. In preferred embodiments, the system software further includes one or more logic instructions for displaying the scaled data for a human viewer. In some embodiments, the system software further comprises one or more logic instructions for analyzing the scaled data to identify portions of the raw data of interest (e.g., automated data analysis software, such as cluster analysis software or the like).

The analysis according to the invention can be accessed using an information processing system and/or over a communications network. According to specific embodiments of the invention, a client system is provided with a set of interfaces that allow a user to indicate one or more analyses and/or analysis parameters and that may direct a user to input the necessary initial data or option selections. The client system displays information that identifies analysis available and displays an indication of an action that a user is to perform to request an analysis. In response to a user input, the client system sends to a server system the necessary information. The server system uses the request data, and optionally one or more sets of server data, to perform the requested analysis. Subsequently, results data are transmitted to the client system. In specific embodiments, such analysis can be provided over the Internet, optionally using Internet media protocols and formats, such as HTTP, RTTP, XML, HTML, XHTML, VRML, as well as image, audio, or video formats, etc. However, using the teachings provided herein, it will be understood by those of skill in the art that the methods and apparatus of the present invention could be advantageously used in other related situations where users access content over a communication channel, such as modem access systems, institution network systems, wireless systems, etc. Thus, the present invention is involved with a number of unique methods and/or systems that can be used together or

independently to provide analysis related to biologic or other data. In specific embodiments, the present invention can be understood as involving new business methods related to providing such analysis.

The invention and various specific aspects and embodiments will be better understood with reference to the following drawings, appendix, and detailed descriptions. In some of the drawings and detailed descriptions below, the present invention is described in terms of the important independent embodiment of a system operating on a digital data network. This should not be taken to limit the invention, which, using the teachings provided herein, can be applied to other situations, such as cable television networks, wireless networks, etc. For purposes of clarity, this discussion refers to devices, methods, and concepts in terms of specific examples, e.g., flow cytometry. However, the invention and aspects thereof have applications to a variety of types of devices and systems. It is therefore intended that the invention not be limited except as provided in the attached claims.

It is well known in the art that logic systems and methods such as described herein can include a variety of different components and different functions in a modular fashion. Different embodiments of the invention can include different mixtures of elements and functions and may group various functions as parts of various elements. For purposes of clarity, the invention is described in terms of systems and/or methods that include many different innovative components and innovative combinations of innovative components and known components. No inference should be taken to limit the invention to combinations containing all of the innovative components listed in any illustrative embodiment in this specification. The functional aspects of the invention that are implemented on a computer, as will be understood from the teachings herein, may be implemented or accomplished using any appropriate implementation environment or programming language, such as C, C++, Cobol, Pascal, Fortran, Java, Java-script, PLI, LISP, HTML, XML, d-HTML, assembly or machine code programming, etc. All references, publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes. All documents, data, and other written or otherwise available material described or referred to herein, are incorporated by reference

BRIEF DESCRIPTION OF THE DIAGRAM

FIG. 1: is a system environment in which various embodiments may be implemented.

FIG. 2: is a schematic embodiment of a client computer.

FIG. 3: is a schematic embodiment of a network computer.

FIG. 4: is a logical schematic of a portion of a service integration system in accordance with at least one of the various embodiments.

FIG. 5: shows a flowchart for a process for reacting to the discovery of new entities in in a monitored network in accordance with at least one of the various embodiments;

FIG. 6: is a table diagram showing the sample results of a confusion matrix for a test of a system designed to classify documents into one of three classes in accordance with at least one of the various embodiments;

FIG. 7: is a table diagram showing an example of a class-specific confusion matrix in accordance with at least one of the various embodiments;

FIG. 8: is a table diagram showing the sample results of a real-time scoring of website files being analyzed for malicious code.

DESCRIPTION OF THE INVENTION

Various embodiments now will be described more fully hereinafter with reference to the accompanying drawings, which form a part hereof, and which show, by way of illustration, specific exemplary embodiments by which the invention may be practiced. The embodiments may, however, be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the embodiments to those skilled in the art. Among other things, the various embodiments may be methods, systems, media or devices. Accordingly, the various embodiments may take the form of an entirely hardware embodiment, an entirely software embodiment or an embodiment combining software and hardware aspects. The following detailed description is, therefore, not to be taken in a limiting sense.

Throughout the specification and claims, the following terms take the meanings explicitly associated herein, unless the context clearly dictates otherwise. The phrase "in one embodiment" as used herein does not necessarily refer to the same embodiment, though it may. Furthermore, the phrase "in another embodiment" as used herein does not necessarily refer to a different embodiment, although it may. Thus, as described below, various embodiments may be readily combined, without departing from the scope or spirit of the invention.

In addition, as used herein, the term “or” is an inclusive “or” operator, and is equivalent to the term “and/or,” unless the context clearly dictates otherwise. The term “based on” is not exclusive and allows for being based on additional factors not described, unless the context clearly dictates otherwise. In addition, throughout the specification, the meaning of “a,” “an,” and “the” include plural references. The meaning of “in” includes “in” and “on.”

For example, embodiments, the following terms are also used herein according to the corresponding meaning, unless the context clearly dictates otherwise.

As used herein the term “deep learning model” refers to classification models that may require longer training times in exchange for more accurate classifications. In some embodiments, deep learning neural network models, as described in detail below, may be considered a deep learning model. However, other machine learning and/or classification techniques may be employed to generate deep learning model.

As used herein the term “fast learning model” refers to classification model that may sacrifice accuracy in exchange for reduced training times as compared to a deep learning model. In some cases, the deep learning model and the fast learning model may be the same kind of classification model. In such cases, the deep learning model may be configured to have a longer training time to improve accuracy as compared with the fast learning model version.

As used herein the term “high speed memory cache” refers to high speed persistent storage that is persistent and large, such as a solid-state drive (SSD) as well fast, size-constrained, non-persistent memory. In some embodiment these caches may comprise high speed volatile memory, persistent, somewhat slower storage like SSD, or the like, or combination thereof. In some embodiments, a sensor computer may include one or more high speed memory caches that enable the real-time capture of an information stream, such as, network information, network traffic, or the like, or combination thereof. In some embodiments, cache may be arranged to purge some or all of its contents as needed. Further, in at least one of the various embodiments, purging may include offloading the contents of a cache to another data store.

Briefly stated, embodiments are directed towards classifying data using machine learning that may be incrementally refined based on expert input. In at least one of the various embodiments, data may be provided to a deep learning model that has been trained using a plurality of classifiers and one or more sets of training data and/or testing data.

In at least one of the various embodiments, the data provided for classification may be real-time network information, captured/buffered network information, or the like. Also, in at least one of the various embodiments, a sensor computer may be employed to monitor and buffer some or all of the data, such as, network information in real-time.

In at least one of the various embodiments, the data may be classified using the deep learning model and the one or more classifiers. In at least one of the various embodiments, a confidence value may be generated and associated with the classification of the data depending on how close the data matches the classifier.

In at least one of the various embodiments, if the number of classification errors exceeds one or more defined thresholds, additional actions may be performed. In at least one of the various embodiments, one or more of the classifiers may be tuned and/or modified based on data corresponding to one or more observed classification errors. In at least one of the various embodiments, a fast learning model may be trained based on the one or more modified classifiers, the data, and the data corresponding to the one or more observed classification errors. In at least one of the various embodiments, the data may be classified based on the fast learning model and the one or more modified classifiers. And, another confidence value may be generated and associated with the classification of the data by the fast learning model.

In at least one of the various embodiments, exceeding a defined threshold may include exceeding one or more different thresholds that are defined for different types of classification errors, such that the classification errors related to dangerous events may have a lower defined threshold than classification errors related to safe events.

In at least one of the various embodiments, report information may be generated based on a comparison result of the confidence value associated with the fast learning model and the confidence value associated with the deep learning model. If the confidence value associated with the classification made by the deep learning model is greater the confidence value associated with the fast learning model, the classification information generated by the deep learning model may be used; otherwise, the classification information generated by the fast learning model may be used in as report information. In at least one of the various embodiments, the report information may be employed to generate on or more reports for storage and/or display to a user.

In at least one of the various embodiments, if the deep learning model may be retrained based on the one or more modified classifiers the trained fast learning

model may be discarded. Also, in at least one of the various embodiments, the deep learning model may be retrained at other times based on a defined schedule. Also, in at least one of the various embodiments, if the data is classified as being associated with a new network entity, historical network information may be associated with the new network entity based on a type of the new network entity. And, in at least one of the various embodiments, real-time network information associated with the new network entity may be buffered. In at least one of the various embodiments, if the data is classified as being associated with anomalous activity, one or more notifications may be generated depending on a type of the anomalous activity.

FIG. 1: shows components of one embodiment of an environment in which embodiments of the invention may be practiced. Not all of the components may be required to practice the invention, and variations in the arrangement and type of the components may be made without departing from the spirit or scope of the invention. As shown, system 100 of FIG. 1 includes local area networks (LANs)/wide area networks (WANs)—(network) 110, wireless network 108, client computers 102-105, Classification Server Computer 116, one or more services provided by servers, such as, Data Sensor Computer 118, Enterprise Server Computer 120, or the like.

At least one embodiment of client computers 102-105 is described in more detail below in conjunction with FIG. 2. In one embodiment, at least some of client computers 102-105 may operate over one or more wired and/or wireless networks, such as networks 108, and/or 110. Generally, client computers 102-105 may include virtually any computer capable of communicating over a network to send and receive information, perform various online activities, offline actions, or the like. In one embodiment, one or more of client computers 102-105 may be configured to operate within a business or other entity to perform a variety of services for the business or other entity. For example, client computers 102-105 may be configured to operate as a web server, firewall, client application, media player, mobile telephone, game console, desktop computer, or the like. However, client computers 102-105 are not constrained to these services and may also be employed, for example, as for end-user computing in other embodiments. It should be recognized that more or less client computers (as shown in FIG. 1) may be included within a system such as described herein, and embodiments are therefore not constrained by the number or type of client computers employed.

Computers that may operate as client computer 102 may include computers that typically connect using a wired or wireless communications medium such as personal computers, multiprocessor systems, microprocessor-based or

programmable electronic devices, network PCs, or the like. In some embodiments, client computers 102-105 may include virtually any portable computer capable of connecting to another computer and receiving information such as, laptop computer 103, mobile computer 104, tablet computers 105, or the like. However, portable computers are not so limited and may also include other portable computers such as cellular telephones, display pagers, radio frequency (RF) devices, infrared (IR) devices, Personal Digital Assistants (PDAs), handheld computers, wearable computers, integrated devices combining one or more of the preceding computers, or the like. As such, client computers 102-105 typically range widely in terms of capabilities and features. Moreover, client computers 102-105 may access various computing applications, including a browser, or other web-based application.

A web-enabled client computer may include a browser application that is configured to receive and to send web pages, web-based messages, and the like. The browser application may be configured to receive and display graphics, text, multimedia, and the like, employing virtually any web-based language, including a wireless application protocol messages (WAP), and the like. In one embodiment, the browser application is enabled to employ Handheld Device Markup Language (HDML), Wireless Markup Language (WML), WMLScript, JavaScript, Standard Generalized Markup Language (SGML), Hypertext Markup Language (HTML), eXtensible Markup Language (XML), JavaScript Object Notation (JSON), or the like, to display and send a message. In one embodiment, a user of the client computer may employ the browser application to perform various activities over a network (online). However, another application may also be used to perform various online activities.

Client computers 102-105 also may include at least one other client application that is configured to receive and/or send content between another computer. The client application may include a capability to send and/or receive content, or the like. The client application may further provide information that identifies itself, including a type, capability, name, and the like. In one embodiment, client computers 102-105 may uniquely identify themselves through any of a variety of mechanisms, including an Internet Protocol (IP) address, a phone number, Mobile Identification Number (MIN), an electronic serial number (ESN), or other device identifier. Such information may be provided in a network packet, or the like, sent between other client computers, classification server computer 116, data sensor computer 118 and enterprise server computer 120, or other computers.

Client computers 102-105 may further be configured to include a client application that enables an end-user to log into an end-user account that may be managed by

another computer, such as classification server computer 116, data sensor computer 118, enterprise server computer 120, or the like. Such an end-user account, in one non-limiting example, may be configured to enable the end-user to manage one or more online activities, including in one non-limiting example, project management, software development, system administration, configuration management, search activities, social networking activities, browse various websites, communicate with other users, or the like. Further, client computers may be arranged to enable users to provide configuration information, or the like, to classification server computer 116. Also, client computers may be arranged to enable users to display reports, interactive user-interfaces, and/or results provided by classification server computer 116.

Wireless network 108 is configured to couple client computers 103-105 and its components with network 110. Wireless network 108 may include any of a variety of wireless sub-networks that may further overlay stand-alone ad-hoc networks, and the like, to provide an infrastructure-oriented connection for client computers 103-105. Such sub-networks may include mesh networks, Wireless LAN (WLAN) networks, cellular networks, and the like. In one embodiment, the system may include more than one wireless network. Wireless network 108 may further include an autonomous system of terminals, gateways, routers, and the like connected by wireless radio links, and the like. These connectors may be configured to move freely and randomly and organize themselves arbitrarily, such that the topology of wireless network 108 may change rapidly.

Wireless network 108 may further employ a plurality of access technologies including 2nd (2G), 3rd (3G), 4th (4G) 5th (5G) generation radio access for cellular systems, WLAN, Wireless Router (WR) mesh, and the like. Access technologies such as 2G, 3G, 4G, 5G, and future access networks may enable wide area coverage for mobile computers, such as client computers 103-105 with various degrees of mobility. In one non-limiting example, wireless network 108 may enable a radio connection through a radio network access such as Global System for Mobile communication (GSM), General Packet Radio Services (GPRS), Enhanced Data GSM Environment (EDGE), code division multiple access (CDMA), time division multiple access (TDMA), Wideband Code Division Multiple Access (WCDMA), High Speed Downlink Packet Access (HSDPA), Long Term Evolution (LTE), and the like. In essence, wireless network 108 may include virtually any wireless communication mechanism by which information may travel between client computers 103-105 and another computer, network, a cloud-based network, a cloud instance, or the like.

Network 110 is configured to couple network computers with other computers, including, classification server computer 116, data sensor computer 118, enterprise server computer 120, client computers 102-105 through wireless network 108, or the like. Network 110 is enabled to employ any form of computer readable media for communicating information from one electronic device to another. Also, network 110 can include the Internet in addition to local area networks (LANs), wide area networks (WANs), direct connections, such as through a universal serial bus (USB) port, other forms of computer-readable media, or any combination thereof. On an interconnected set of LANs, including those based on differing architectures and protocols, a router acts as a link between LANs, enabling messages to be sent from one to another.

In addition, communication links within LANs typically include twisted wire pair or coaxial cable, while communication links between networks may utilize analog telephone lines, full or fractional dedicated digital lines including T1, T2, T3, and T4, and/or other carrier mechanisms including, for example, E-carriers, Integrated Services Digital Networks (ISDNs), Digital Subscriber Lines (DSLs), wireless links including satellite links, or other communications links known to those skilled in the art. Moreover, communication links may further employ any of a variety of digital signaling technologies, including without limit, for example, DS-0, DS-1, DS-2, DS-3, DS-4, OC-3, OC-12, OC-48, or the like. Furthermore, remote computers and other related electronic devices could be remotely connected to either LANs or WANs via a modem and temporary telephone link. In one embodiment, network 110 may be configured to transport information of an Internet Protocol (IP).

Additionally, communication media typically embodies computer readable instructions, data structures, program modules, or other transport mechanism and includes any information non-transitory delivery media or transitory delivery media. By way of example, communication media includes wired media such as twisted pair, coaxial cable, fiber optics, wave guides, and other wired media and wireless media such as acoustic, RF, infrared, and other wireless media. One embodiment of classification server computer 116 is described in more detail below in conjunction with FIG. 3. Briefly, however, classification server computer 116 includes virtually any network computer capable of service integration in network environment.

Although FIG. 1 illustrates classification server computer 116, data sensor computer 118, and enterprise server computer 120, each as a single computer, the innovations and/or embodiments are not so limited. For example, one or more functions of classification server computer 116, and data sensor computer 118,

and enterprise server computer 120, or the like, may be distributed across one or more distinct network computers. Moreover, classification server computer 116, and data sensor computer 118, and enterprise server computer 120, are not limited to a particular configuration such as the one shown in FIG. 1. Thus, in one embodiment, classification server computer 116, and data sensor computer 118, and enterprise server computer 120 may be implemented using a plurality of network computers. In other embodiments, server computers may be implemented using a plurality of network computers in a cluster architecture, a peer-to-peer architecture, or the like. Further, in at least one of the various embodiments, classification server computer 116, and data sensor computer 118, and enterprise server computer 120 may be implemented using one or more cloud instances in one or more cloud networks. Accordingly, these innovations and embodiments are not to be construed as being limited to a single environment, and other configurations, and architectures are also envisaged.

FIG. 2: shows one embodiment of client computer 200 that may be included in a system in accordance with at least one of the various embodiments. Client computer 200 may include many more or less components than those shown in FIG. 2. However, the components shown are sufficient to disclose an illustrative embodiment for practicing the present invention. Client computer 200 may represent, for example, one embodiment of at least one of client computers 102-105 of FIG. 1. As shown in the figure, client computer 200 includes a processor device, such as processor 202 in communication with a mass memory 226 via a bus 234. In some embodiments, processor 202 may include one or more central processing units (CPU) and/or one or more processing cores. Client computer 200 also includes a power supply 228, one or more network interfaces 236, an audio interface 238, a display 240, a keypad 242, an illuminator 244, a video interface 246, an input/output interface 248, a haptic interface 250, and a global positioning system (GPS) receiver 232.

Power supply 228 provides power to client computer 200. A rechargeable or non-rechargeable battery may be used to provide power. The power may also be provided by an external power source, such as an alternating current (AC) adapter or a powered docking cradle that supplements and/or recharges a battery.

Client computer 200 may optionally communicate with a base station (not shown), or directly with another computer. Network interface 236 includes circuitry for coupling client computer 200 to one or more networks, and is constructed for use with one or more communication protocols and technologies including, but not limited to, GSM, CDMA, TDMA, GPRS, EDGE, WCDMA, HSDPA, LTE, user datagram protocol (UDP), transmission control protocol/Internet protocol (TCP/IP), short

message service (SMS), WAP, ultra wide band (UWB), IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMax), session initiated protocol/real-time transport protocol (SIP/RTP), or any of a variety of other wireless communication protocols. Network interface 236 is sometimes known as a transceiver, transcribing device, or network interface card (NIC).

Audio interface 238 is arranged to produce and receive audio signals such as the sound of a human voice. For example, audio interface 238 may be coupled to a speaker and microphone (not shown) to enable telecommunication with others and/or generate an audio acknowledgement for some action. Display 240 may be a liquid crystal display (LCD), gas plasma, light emitting diode (LED), organic LED, or any other type of display used with a computer. Display 240 may also include a touch sensitive screen arranged to receive input from an object such as a stylus or a digit from a human hand. Keypad 242 may comprise any input device arranged to receive input from a user. For example, keypad 242 may include a push button numeric dial, or a keyboard. Keypad 242 may also include command buttons that are associated with selecting and sending images.

Is 244 may provide a status indication and/or provide light. Illuminator 244 may remain active for specific periods of time or in response to events. For example, when illuminator 244 is active, it may backlight the buttons on keypad 242 and stay on while the client computer is powered. Also, illuminator 244 may backlight these buttons in various patterns when particular actions are performed, such as dialing another client computer. Illuminator 244 may also cause light sources positioned within a transparent or translucent case of the client computer to illuminate in response to actions. Video interface 246 is arranged to capture video images, such as a still photo, a video segment, an infrared video, or the like. For example, video interface 246 may be coupled to a digital video camera, a web-camera, or the like. Video interface 246 may comprise a lens, an image sensor, and other electronics. Image sensors may include a complementary metal-oxide-semiconductor (CMOS) integrated circuit, charge-coupled device (CCD), or any other integrated circuit for sensing light.

Client computer 200 also comprises input/output interface 248 for communicating with external devices, such as a headset, or other input or output devices not shown in FIG. 2. Input/output interface 248 can utilize one or more communication technologies, such as USB, infrared, Bluetooth™, or the like. Haptic interface 250 is arranged to provide tactile feedback to a user of the client computer. For example, the haptic interface 250 may be employed to vibrate client computer 200 in a particular way when another user of a computer is calling. In some embodiments, haptic interface 250 may be optional.

Client computer 200 may also include GPS transceiver 232 to determine the physical coordinates of client computer 200 on the surface of the Earth. GPS transceiver 232, in some embodiments, may be optional. GPS transceiver 232 typically outputs a location as latitude and longitude values. However, GPS transceiver 232 can also employ other geo-positioning mechanisms, including, but not limited to, triangulation, assisted GPS (AGPS), Enhanced Observed Time Difference (E-OTD), Cell Identifier (CI), Service Area Identifier (SAI), Enhanced Timing Advance (ETA), Base Station Subsystem (BSS), or the like, to further determine the physical location of client computer 200 on the surface of the Earth. It is understood that under different conditions, GPS transceiver 232 can determine a physical location within millimeters for client computer 200; and in other cases, the determined physical location may be less precise, such as within a meter or significantly greater distances. In one embodiment, however, client computer 200 may through other components, provide other information that may be employed to determine a physical location of the computer, including for example, a Media Access Control (MAC) address, IP address, or the like.

Mass memory 226 includes a Random Access Memory (RAM) 204, a Read-only Memory (ROM) 222, and other storage means. Mass memory 226 illustrates an example of computer readable storage media (devices) for storage of information such as computer readable instructions, data structures, program modules or other data. Mass memory 226 stores a basic input/output system (BIOS) 224, or the like, for controlling low-level operation of client computer 200. The mass memory also stores an operating system 206 for controlling the operation of client computer 200. It will be appreciated that this component may include a general-purpose operating system such as a version of UNIX, or LINUX™, or a specialized client communication operating system such as Microsoft Corporation's Windows Mobile™, Apple Corporation's iOS™, Google Corporation's Android™, or the like. The operating system may include, or interface with a Java virtual machine module that enables control of hardware components and/or operating system operations via Java application programs.

Mass memory 226 further includes one or more data storage 208, which can be utilized by client computer 200 to store, among other things, applications 214 and/or other data. For example, data storage 208 may also be employed to store information that describes various capabilities of client computer 200. The information may then be provided to another computer based on any of a variety of events, including being sent as part of a header during a communication, sent upon request, or the like. Data storage 208 may also be employed to store social networking information including address books, buddy lists, aliases, user profile information, user credentials, or the like. Further, data

storage 208 may also store messages, web page content, or any of a variety of user generated content.

At least a portion of the information stored in data storage 208 may also be stored on another component of client computer 200, including, but not limited to processor readable storage media 230, a disk drive or other computer readable storage devices (not shown) within client computer 200. Processor readable storage media 230 may include volatile, non-transitory, nonvolatile, removable, and non-removable media implemented in any method or technology for storage of information, such as computer- or processor-readable instructions, data structures, program modules, or other data. Examples of computer readable storage media include RAM, ROM, Electrically Erasable Programmable Read-only Memory (EEPROM), flash memory or other memory technology, Compact Disc Read-only Memory (CD-ROM), digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other physical medium which can be used to store the desired information and which can be accessed by a computer. Processor readable storage media 230 may also be referred to herein as computer readable storage media and/or computer readable storage device.

Applications 214 may include computer executable instructions which, when executed by client computer 200, transmit, receive, and/or otherwise process network data. Network data may include, but is not limited to, messages (e.g. SMS, Multimedia Message Service (MMS), instant message (IM), email, and/or other messages), audio, video, and enable telecommunication with another user of another computer. Applications 214 may include, for example, a browser 218, and other applications 220.

Browser 218 may include virtually any application configured to receive and display graphics, text, multimedia, messages, and the like, employing virtually any web based language. In one embodiment, the browser application is enabled to employ HDML, WML, WMLScript, JavaScript, SGML, HTML, HTML5, XML, and the like, to display and send a message. However, any of a variety of other web-based programming languages may be employed. In one embodiment, browser 218 may enable a user of client computer 200 to communicate with another network computer, such as classification server computer 116, and data sensor computer 118, and enterprise server computer 120, or the like, as shown in FIG. 1. Other applications 220 may include, but are not limited to, calendars, search programs, email clients, IM applications, SMS applications, voice over Internet Protocol (VOIP) applications, contact managers, task managers, transcoders, database programs, word processing programs, software development tools,

security applications, spreadsheet programs, games, search programs, and so forth.

FIG. 3: shows one embodiment of a network computer 300, according to one embodiment of the invention. Network computer 300 may include many more or less components than those shown. The components shown, however, are sufficient to disclose an illustrative embodiment for practicing the invention. Network computer 300 may be configured to operate as a server, client, peer, a host, cloud instance, or any other computer. Network computer 300 may represent, for example classification server computer 116, and/or other network computers, such as, data sensor computer 118 and enterprise server computer 120, or the like. Network computer 300 includes one or more processor devices, such as, processor 302. Also, network computer 300 includes processor readable storage media 328, network interface unit 330, an input/output interface 332, hard disk drive 334, video display adapter 336, and memory 326, all in communication with each other via bus 338.

As illustrated in FIG. 3, network computer 300 also can communicate with the Internet, or other communication networks, via network interface unit 330, which is constructed for use with various communication protocols including the TCP/IP protocol. Network interface unit 330 is sometimes known as a transceiver, transcribing device, or network interface card (NIC). Network computer 300 also comprises input/output interface 332 for communicating with external devices, such as a keyboard, or other input or output devices not shown in FIG. 3. Input/output interface 332 can utilize one or more communication technologies, such as USB, infrared, NFC, Bluetooth™, or the like.

Memory 326 generally includes RAM 304, ROM 322 and one or more permanent mass storage devices, such as hard disk drive 334, tape drive, optical drive, and/or floppy disk drive. Memory 326 stores operating system 306 for controlling the operation of network computer 300. Any general-purpose operating system may be employed. Basic input/output system (BIOS) 324 is also provided for controlling the low-level operation of network computer 300.

Although illustrated separately, memory 326 may include processor readable storage media 328. Processor readable storage media 328 may be referred to and/or include computer readable media, computer readable storage media, and/or processor readable storage device. Processor readable storage media 328 may include volatile, nonvolatile, non-transitory, non-transitive, removable, and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules, or other data. Examples of processor readable storage media

include RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, solid state storage devices, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other media which can be used to store the desired information and which can be accessed by a computer.

Memory 326 further includes one or more data storage 308, which can be utilized by network computer 300 to store, among other things, applications 314 and/or other data. For example, data storage 308 may also be employed to store information that describes various capabilities of network computer 300. The information may then be provided to another computer based on any of a variety of events, including being sent as part of a header during a communication, sent upon request, or the like. Data storage 308 may also be employed to store messages, web page content, or the like. At least a portion of the information may also be stored on another component of network computer 300, including, but not limited to processor readable storage media 328, hard disk drive 334, or other computer readable storage medias (not shown) within network computer 300.

Data storage 308 may include a database, text, spreadsheet, folder, file, or the like, that may be configured to maintain and store user account identifiers, user profiles, email addresses, IM addresses, and/or other network addresses; or the like. Data storage 308 may further include program code, data, algorithms, and the like, for use by a processor device, such as processor 302 to execute and perform actions. In one embodiment, at least some of data store 308 might also be stored on another component of network computer 300, including, but not limited to processor-readable storage media 328, hard disk drive 334, or the like. Data storage 308 may include source data 310. In at least one of the various embodiments, source data 310 may include data that is collected and/or ingested for processing but a machine learning engine, or the like. Also, in at least one of the various embodiments, data storage 308 may include classified data objects 312 representing training data, test data, and/or classified source data.

Applications 314 may include computer executable instructions, which may be loaded into mass memory and run on operating system 306. Examples of application programs may include transcoders, schedulers, calendars, database programs, word processing programs, Hypertext Transfer Protocol (HTTP) programs, customizable user-interface programs, IPsec applications, encryption programs, security programs, SMS message servers, IM message servers, email servers, account managers, and so forth. Applications 314 may also include, web server 316, machine learning engine 318, interactive tuning application 321, or the like. Web server 316 may represent any of a variety of information and services

that are configured to provide content, including messages, over a network to another computer. Thus, website server 316 can include, for example, a web server, a File Transfer Protocol (FTP) server, a database server, a content server, email server, or the like. Website server 316 may provide the content including messages over the network using any of a variety of formats including, but not limited to WAP, HDML, WML, SGML, HTML5, XML, Compact HTML (cHTML), Extensible HTML (xHTML), or the like.

FIG. 4: shows a logical representation of system 400 to classify data using machine learning that may be incrementally refined based on expert input in accordance with at least one of the various embodiments. In at least one of the various embodiments, one or more computers, such as, client computer 402, laptop computer 404, mobile computer 406, tablet computer 408, enterprise server computer 410, or the like, may be coupled using one or more networks, such as, local network 412. In at least one of the various embodiments, local network 412 may be a portion or instance of network 110, and/or network 108 as shown in FIG. 1.

In at least one of the various embodiments, data sensor computer 414 may be disposed between one or more portions of network 412 and network 416. In at least one of the various embodiments, network 416 may represent one or more wide-area networks (WANs), including the Internet and may be described similarly to network 110. In at least one of the various embodiments, one or more data sensor computers, such as, data sensor computer 414 may be positioned to monitor some or all of the network traffic on network 412. In at least one of the various embodiments, monitored traffic may include data sent between computer on network 412 as well as communication with endpoints outside of network 412.

In at least one of the various embodiments, sensor computer 414 may be one or more network computers, such as, network computer 300. Also, in at least one of the various embodiments, sensor computer 414 may be one or more client computers, such as, client computer 200. In at least one of the various embodiments, sensor computer 414 may include one or more high speed memory data caches for real-time capturing and/or buffer the network information that occurs on network 412. In at least one of the various embodiments, sensor computer 414 may be arranged to execute applications, such as, machine learning engine 318, classifier application 319, interactive tuning application 321, or the like. In at least one of the various embodiments, classifier applications, such as, classifier application 319, may be arranged to employ one or more trained models to classify the observed network information that occurs on the network. Also, in at least one of the various embodiments, the network information buffered in

sensor computers, such as, sensor computer 414 may be employed as training data and/or test data for re-training the one or more classification models using a machine learning application, such as, machine learning application 318.

FIG. 5: shows a flowchart for process 900 for reacting to the discovery of new network entities in a monitored network in accordance with at least one of the various embodiments. After a start block, at decision block 902, in at least one of the various embodiments, if the system detects a new entity on the network, control may flow to block 904; otherwise, control may flow block 908. In at least one of the various embodiments, a system may include sensor computer, such as, data sensor computer 414, may be arranged to monitor and classify network information, such as, network traffic that is communicated over a monitored network.

In at least one of the various embodiments, a detected entity may be a client computer, network computer, a mobile computer, user, user group, file-type, hostname, source host computer, destination host computer, router, hub, network interfaces, or the like. Also, in at least one of the various embodiments, a detected entity may include the detection of an previously unknown/unseen instance of an application, such as, a web server, database, domain name server, user applications (e.g., games, office applications, and so on), file sharing applications, or the like.

In at least one of the various embodiments, new computers, such as mobile computers, may be detected if they a wireless portion of the network. For example, if an employ enters her place of work with a new mobile computer it may be configured to automatically join the network. In this example, the presence of the new mobile computer may trigger a new entity detection corresponding to the new mobile computer.

At block 904, in at least one of the various embodiments, previously collected historical network information may be associated with the detected entity based on the class of the detected entity. In at least one of the various embodiments, the detected entity may be a previously unknown instance of a known class. For example, if the detected entity is a new user that has been added to the network, the user may be considered a new instance of the class user. Likewise, for example, if the detected entity is a previously unseen employee's personal mobile computer, it may be a new instance of a known class. Accordingly, in at least one of the various embodiments, process 900 may associate historical network information for another previously detected instance of the same class as the detected entity. In at least one of the various embodiments, the historical network information may provide a baseline history that may be used to classify the network information that may be associated with the new entity.

At block 906, in at least one of the various embodiments, real-time network information for the detected entity may be captured and buffered. In at least one of the various embodiments, if a new entity is detected the system may begin capturing network information that is associated with that particular entity. At block 908, in at least one of the various embodiments, incoming network information associated with the detected entity may be classified using the trained model. However, since the historical information used to train the model may not have included information generated by the new detected entity, the classification of new entities activity may be based on historical information associated with a previously known entity having the same class.

At decision block 910, in at least one of the various embodiments, if the network information for the detected entity is buffered, control may flow to block 912; otherwise, control may flow to decision block 914. In at least one of the various embodiments, newly detected entities may initially be marked and/or tagged as new entities. If a sufficient amount of network information for the detected entity may have been buffered, the detected entity may be considered normal rather than new, and marked/tagged as such. In at least one of the various embodiments, the amount of buffered information to cause a new entity to be considered a normal entity may vary depending on the class of entity and the type of classifiers it may be associated with. In at least one of the various embodiments, threshold values may be defined in configuration to indicate the amount of network information that must be captured for a given class and/or entity.

At block 912, in at least one of the various embodiments, anomalies and/or classifications associated with the detected entity may now be included in the report information. In at least one of the various embodiments, since the detected entity is no longer considered a new entity, its classification information, if any, may be included in the report information for the network. At decision block 914, in at least one of the various embodiments, if the process remains active, control may loop back to decision block 902; otherwise, control may be returned to a calling process.

It will be understood that each block of the flowchart illustration, and combinations of blocks in the flowchart illustration, can be implemented by computer program instructions. These program instructions may be provided to a processor to produce a machine, such that the instructions, which execute on the processor, create means for implementing the actions specified in the flowchart block or blocks. The computer program instructions may be executed by a processor to cause a series of operational steps to be performed by the processor to produce a computer-implemented process such that the instructions, which

execute on the processor to provide steps for implementing the actions specified in the flowchart block or blocks. The computer program instructions may also cause at least some of the operational steps shown in the blocks of the flowchart to be performed in parallel. These program instructions may be stored on some type of machine readable storage media, such as processor readable non-transitive storage media, or the like. Moreover, some of the steps may also be performed across more than one processor, such as might arise in a multi-processor computer system. In addition, one or more blocks or combinations of blocks in the flowchart illustration may also be performed concurrently with other blocks or combinations of blocks, or even in a different sequence than illustrated without departing from the scope or spirit of the invention.

Accordingly, blocks of the flowchart illustration support combinations of means for performing the specified actions, combinations of steps for performing the specified actions and program instruction means for performing the specified actions. It will also be understood that each block of the flowchart illustration, and combinations of blocks in the flowchart illustration, can be implemented by special purpose hardware-based systems, which perform the specified actions or steps, or combinations of special purpose hardware and computer instructions. The foregoing example should not be construed as limiting and/or exhaustive, but rather, an illustrative use case to show an implementation of at least one of the various embodiments of the invention.

FIG.6: is a table diagram showing the sample results of a confusion matrix for a test of a system designed to classify documents into one of three classes in accordance with at least one of the various embodiments. Confusion matrix 1002 may be made up of rows 1004-1008, containing values for the actual counts of those classes; and columns 1010-1014 containing the values of the predicted classes by the system. For example, the testing corpus for this system contained 8 documents in Class A, but only 5 of these documents were predicted to be in Class A. The other three were predicted into Class B.

In at least one of the various embodiments, using this method of evaluating performance, once the confusion matrix has been computed, the system generates counts of True and False Positive and Negatives results for individual classes.

FIG. 7: is a table diagram showing an example of a class-specific confusion matrix in accordance with at least one of the various embodiments. Confusion matrix 1102 may be made up of table cells 1104-1110. Table cell 1104 contains the True Positive results of Class A. The true positives are the number of Class A documents correctly classified as Class A. Table cell 1106 contains the False Positive results of Class A. False positives are the number of non-Class A

documents incorrectly classified into Class A. Table cell 1108 contains the False Negative results of Class A. The false negatives may be the number of Class A documents incorrectly classified as non-Class A. Table cell 1110 contains the True Negative results for Class A. The true negatives are the number of non-Class A documents correctly classified into non-Class A classes.

In at least one of the various embodiments, using the confusion matrix of a specific class, the model's ability at classifying documents into or outside of that class may be computed and represented with Accuracy, Specificity, Sensitivity (or Recall), and Precision. In at least one of the various embodiments, the Accuracy of a Model is the proportion of total samples identified correctly. Accuracy may be calculated with the following equation: In at least one of the various embodiments, a higher value of Precision represents a class containing a high ratio of correctly identified positives. In Model Performance Visualization 528, these performance metrics may be presented to the Domain Expert or other user and then used to determine whether a re-training is appropriate. In some embodiments, the decision is automated to use pre-configured heuristics. Those skilled in the art will appreciate that there are other methods of measuring performance, which vary depending on the model used. Some performance measures include receiver operating characteristic curves ("ROC Curves").

In at least one of the various embodiments, the Domain Expert Analysis phase 530 may be the process through which the system collects input from and displays scoring results of new data to a Domain Expert. Analysis phase 530 includes components which provide a method of examining the data and scores (a User Interface 532, Alert System 534, or Decision System 536) and for adjusting the classification applied by the system. In at least one of the various embodiments, returning again to the example of e-discovery, as the system processes data it may identify particular documents that are of interest to have a human specifically review. The system may then communicate that information to various users. In some instances, simply a list of relevant files may be provided to a Domain Expert such as a paralegal. In other instances, the system might alert more than one Domain Expert, such as a paralegal and a group of attorneys working on the case. In other instances, that alert might go to paralegals, attorneys, and IT personnel working with training the system. In at least one of the various embodiments, different types of notices may be sent to each type of Domain Expert using a specific User Interface 532. A Decision System 536 may be set up for specific data. In an e-discovery system for example, the system might identify possible attorney-client privileged materials which may be handled in a different manner from other data. A Decision System 536 may be set up to remove such

documents for any further review or production until a Domain Expert specifically reviews the identified document.

FIG.8: of visualizations that a system might provide to Domain Expert in a system that is looking for malicious code in accordance with at least one of the various embodiments. FIG. 12 is a table diagram showing the sample results of a real-time scoring of website files being analyzed for malicious code. In this example, the system predicts the probability that a given file contains malicious code and should be removed from the file system. The new data table 1200 is made up of rows 1202-1210, each representing a file that was processed through the system. Each row may be divided into the following columns: an identifier column 1212 containing an identifier for the file; a path column 1214 containing the location of the file; a prediction score column 1216 containing the probability that a file contains malicious code; a review status column 1218 containing a summary of the manual review the file has undergone; and a reason column 1220 indicating the detected threat. For example, row 1202 indicates that file number 1 at location “/user1/public_html/products/submit.php” contains code which the system has predicted yields an 0.85 probability that the document is malicious, the file has been reviewed by a Domain Expert as being malicious, and a Base64 type of exploit was detected. While the contents of new data table 1200 were included as an illustrative example, those skilled in the art will appreciate that the system can use a new data table having columns corresponding to different and/or a larger number of attributes, as well as a larger number of rows.

While FIG. 8: shows a table whose contents and organization are designed to make them more comprehensible by a human reader, those skilled in the art will appreciate that actual data structures used by the system to store this information may differ from the table shown, in that they, for example, may be organized in a different manner; may contain more or less information than shown; may be compressed, encrypted, or the like. the classification task may require more than the display of groupings of data into various classes as described above. The Domain Expert Analysis phase 530 may also include a User Interface 532, Alert System 534, or Decision System 536 for analyzing the results. In one embodiment, the system includes a User Interface 532, which presents the Domain Expert with a detailed view of data elements, and the opportunity to provide input to the system through adjustments. Adjustments may be necessary when the system produces result which are less accurate than required by the user or task being completed. This may be caused by a number of reasons but may include applying a Training Corpus 508 with too few representative samples of each class, or using samples

not accurately characterizing each class. In such a situation, the adjustment component is used to add documents to the Training Corpus 508.

WE CLAIMS

1. Our Invention “**AMD-Data Analysis**” is an embodiments are directed towards classifying data using machine learning, deep learning that may be incrementally refined based on expert input. The invented technology is also the provided to a deep learning model that may be trained based on a classifiers and sets of mapping, training data and testing data. If the number of classification errors exceeds a defined fixed table data and threshold classifier may be modified based on data corresponding to observed classification errors. The invented technology a three type of learning model (fast, medium, slower) may be trained based on the modified classifiers the data and the data corresponding to the observed classification errors and another confidence value may be generated and associated with the classification of the data by the all type of learning, mapping model. The invented technology re-information may be generated based on a comparison result of the confidence value associated with the all learning model and the confidence value associated with the machine, deep learning model. The invention alos provides methods of analyzing or displaying data. The invention provides methods for visualizing or displaying high dynamic range data obtained from flow cytometry analyses. Related systems and computer programs products are also provided.
2. According to claim1# the invention is to an embodiments are directed towards classifying data using machine learning, deep learning that may be incrementally refined based on expert input.
3. According to claim1,2# the invention is to the invented technology is also the provided to a deep learning model that may be trained based on a classifiers and sets of mapping, training data and testing data.
4. According to claim1,2,3# the invention is to If the number of classification errors exceeds a defined fixed table data and threshold classifier may be modified based on data corresponding to observed classification errors.
5. According to claim1,2,4# the invention is to the invented technology a three type of learning model (fast, medium, slower) may be trained based on the modified classifiers the data and the data corresponding to the observed classification errors and another confidence value may be generated and associated with the classification of the data by the all type of learning, mapping model.
6. According to claim1,2,5# the invention is to the invented technology re-information may be generated based on a comparison result of the confidence value associated with the all learning model and the confidence

value associated with the machine, deep learning model. The invention also provides methods of analyzing or displaying data.

7. According to claim 1, 2, 4, 6# the invention is to the invention provides methods for visualizing or displaying high dynamic range data obtained from flow cytometry analyses. Related systems and computer programs products are also provided.

Date: 22/9/2020

Dr. Ashendra Kumar Saxena (Associate Professor)
MD. Iqbal (Associate Professor)
Mr. Rupesh Kumar Jindal (Assistant Professor)
Dr. Santosh Dilipkumar Parakh (Associate Professor)
Dr. Rakesh Kumar ER, (Manager Examination)
Dr. Chukka Santhaiah (Associate Professor)
Prof.(Dr.) S. B. Chordiya (Director-SIMMC-Campus)
Dr. Rambabu Vatti (Principal)

Patent Title: **AMD-Data Analysis:** ACADEMY DATA ANALYSIS USING MACHINE LEARNING, DEEP LEARNING PROGRAMMING.

ABSTRACT

Our Invention “**AMD-Data Analysis**” is an embodiments are directed towards classifying data using machine learning, deep learning that may be incrementally refined based on expert input. The invented technology is also the provided to a deep learning model that may be trained based on a classifiers and sets of mapping, training data and testing data. If the number of classification errors exceeds a defined fixed table data and threshold classifier may be modified based on data corresponding to observed classification errors. The invented technology a three type of learning model (fast, medium, slower) may be trained based on the modified classifiers the data and the data corresponding to the observed classification errors and another confidence value may be generated and associated with the classification of the data by the all type of learning, mapping model. The invented technology re-information may be generated based on a comparison result of the confidence value associated with the all learning model and the confidence value associated with the machine, deep learning model. The invention alos provides methods of analyzing or displaying data. The invention provides methods for visualizing or displaying high dynamic range data obtained from flow cytometry analyses. Related systems and computer programs products are also provided.