

AY- 2018-19

Annual Technical Magazine

EINSTEIN ACADEMY OF TECHNOLOGY AND MANAGEMENTNKAT

**AT:Baniatangi, P.O: Bajpur,Dist:Khurda,
Bhubaneswar-752060,Odisha, India,
PIN: 752060**

ALVARO-CS

Department of Computer Science and Engineering

www.eatm.in

Department of Computer Science and Engineering

VISION

To achieve excellence in the field of computer science and engineering with a goal to create state-of-the-art technocrats to address the future challenges in this fast growing technological world.

MISSION

- Imparting innovative quality education, viz. strong fundamental concepts, analytical capability, programming and problem solving skills, to meet the technological and socio-economic needs of the region and the country as well.
- Facilitating in-house value added courses and professional trainings with an emphasis on basic principles of Computer Science and Engineering.
- Encouraging students for higher studies to inculcate research activities in emerging areas of science and technology.
- Promoting industry-institute interaction for the enhancement of technical and entrepreneurial s

Creative Desk

Prof. Prakash Chandra Dash,
Professor,
(Reviewer, Technical Magazine)

Prof. Riyazuddin Khan,
Associate Professor,
(Reviewer, Technical Magazine)

Khageswar Mohanta,
Student Member,
(Magazine Design Team)

Index

Title	Page No.
Technical Articles	1-18
Non-Technical Articles	19-20
Poetry	21-25
Sketch	26-28

MESSAGES

Chairman's Message

Sri



Mr. Basanta Kumar Bisoi
Chairman, EATM

Alvaro-CS is the manifestation of the desire of Computer Science Faculty and Students to share their innovative ideas on common platform. It gives me great pleasure to know that Alvaro-CS departmental magazine for year 2018-19 is ready for publication. This magazine is a perfect blend of magnificent and groundbreaking articles. It has concentrated in disseminating information to the student community and quenches their thirst for knowledge updates. I am very glad to congratulate the editor for their hard work and bringing out this edition.

Secretary's Message



Sri Krushna Chandra Bisoi
Secretary's EATM

"Alvaro-CS" is the annual magazine of Computer Science & Engineering for year 2018-19. This is a combination of exquisite articles and innovative ideas from faculty members in the computer science and engineering departments and a new generation of students. We are confident that useful articles and innovative ideas in magazines will be appealing and useful to our readers.

Principal's Message



Dr. Suvendu Prasad Sahu
Principal, EATM

Alvaro-CS is the annual magazine released by the Department of Computer Science & Engineering for the year 2018-19. It is a blend of exquisite articles and innovative ideas from the faculty and new – age Students of Computer Science & Engineering Department. I strongly believe that the informative articles & innovative ideas presented in the magazine will be appealing and useful to the readers.

Dean Academics Message



Dr. Biswajit Nayak

Alvaro-CS for the year 2018-19 is the manifestation of the choice of the Computer Science Faculty and Students to percentage their revolutionary thoughts on a not unusual place platform. It offers me brilliant satisfaction to realize that Alvaro departmental mag is prepared for publication. This magazine is a great mixture of magnificent and groundbreaking articles. It has targeting disseminating facts to the pupil network and quenches their thirst for know-how updates. I am very satisfied to congratulate the editor for his or her tough paintings and bringing out this edition.

HOD's Message

Dr.



Anil Kumar Mishra
(Department of CSE)

“Coming together is a beginning, keeping together is progress and working together is success” – This E-magazine “Alvaro”, a flag ship magazine of Computer Science & Engineering Department of Einstein Academy of Technology And Management Engineering College, is the culmination of the never tiring initiative and endeavors taken by the faculty and students of CSE. The Magazine strives to inform, engage, inspire and educate diverse readership on developments in Computer Science field.



Securing the Insecure

Prof. Riyazuddin Khan
Associate Professor

The internet is a war zone. Enter if you dare.

The internet was borne from the need for unfettered scientific communication across different locales. People adopted it into their everyday lives so they could share in that democratization of information. Digital economies emerged and more powerful computing systems allowed for more powerful web services.

Being able to carry out bank transactions on your phone, for example, may be a boon to a traveling salesman, but it's also a recipe for ending up as prey for bad actors looking to make it rich.

Nearly every internet fad has its own ledger of breaches, cybersecurity attacks, or digital ransoms.

It goes back to the dotcom boom. Yahoo dominated the cyberspace in the late 1990s and early 2000s, its email services reigning supreme in the edging digital society. The California-based company was valued at \$54.9 billion – or INR 410,900 crores – in 2006. Following a 2013 data breach that left 3 billion accounts compromised, the poster child of the Y2K era was purchased in 2017 by Verizon for \$4.48 billion – a 92% nosedive.

The early-2010s started with a fad to throw out physical hard drives and toss our personal data into the cloud – as if they were floating, white puffs of ones and zeros. The Cyber security attacks that followed included a 2012 Dropbox hack that exposed 68 million users information, a 2014 phishing campaign that exposed photos of celebrities, and a 2012 LinkedIn breach that involved the theft of countless encrypted passwords. Heck, even Microsoft's cloud services were hacked in 2010.

Moreover, here we are creating vacuum cleaners that connect to the internet.

To no one's surprise, the internet of things has its own record of unglamorous cyber break-ins. Hackers at the 2016 DEF CON security conference – equivalent to an annual county fair for hackers looking to have fun and do some good – found 47 vulnerabilities for 23 devices. Everything from doorlocks to wheelchairs to thermostats walked out of the conference with newly documented weaknesses.

Researchers in 2015 demonstrated the ability to not just kill the engines of a Jeep Cherokee on middle of the highway, but also mess with its air conditioning, radio settings, and windshield wipers – all via its Internet-connected, onboard entertainment system.

You get the point: The more devices we hook up to the World Wide Web, the more we hook up our lifelines to this network. Add in enough internet-connected utilities and you cook up a cyber-storm that can wipe out large swaths of our daily lives.



That is what happened in the fall 2016 Mirai botnet attack. The attack exploited one of the essential facets of how the internet works. Communication devices connected to the internet communicate via multi-byte units of information known as packets. They come with sender and recipient addresses, which are known as Internet Protocol Addresses – IP for short – and hop along network hubs stationed across the world via fiber optic cables deep under the sea or satellites up in space – a digital form of hot potato, if you will.

Every time you are up your web browser and type in google.com, you are sending packets along this distributed network, eventually to one of Google's many high-powered web servers, which then sends packets of data hopping back to you. Every time a new page loads, the more the two of you exchange data.

Now imagine that back-and-forth happening simultaneously, trillions of times over. That's what happened in 2016.

The Mirai botnet was quite simple: An unknowing user would download the malware, which would then infect internet-connected devices – webcams and routers, for example – by guessing every password from its in-built dictionary of common passwords.

Didn't know your webcam and router had passwords? Neither did many people.

Many users kept and still keep the default passwords on their internet-connected devices, and the internet of things was no exception. The malware webbed together hundreds, if not thousands of infected devices – from smart toasters to printers to smart washers to smart ovens to DVRs to even baby monitors – and had them do one thing: repeatedly send web requests to internet services.

That's all it took to break the internet. Fittingly so, “mirai” is Japanese for “future”. It's a dark one, alright. Consequently, we shouldn't be asking ourselves how to sign out of this digital nightmare. The better question is whether the internet of things can be tamed.

If the boom and bust of businesses like Yahoo have taught us anything, it's that rushing to embrace something just because it's a fad – and because it's connected to the internet – is a bad idea.

In other words, when Nike's light up, self-lacing shoes hit stores near you, wait before rushing to the cash register. A security researcher might just show you why an app-controlled shoe that needs an nightly charge is asking for trouble.

Zuckerberg and Pichai can change their tunes all they want, but their message is surprisingly consistent: The internet is inevitable – and so is the loss of your privacy.

Green Tomorrow (GT) - An Energy Efficient Novel Approach Using Crowd Computing

**Dr.Sanjaya Kumar Sen
Professor**

'Save Energy – Save Earth' is a novel approach for GT using crowd computing. Crowd computing is one of the upcoming research areas allowing the people around the world to work together in assistance with the machines. A decade after the machines will completely take over the control with the heavy boom in industrial revolution (IR). The exponential increase in data and its advance techniques, if not managed properly will result in the increase of high-energy consumption (EC) and e-wastes.

Currently, all of us are living and experiencing the era of 'Advance Computer/Network Age', where the computer software handles almost everything. Today, software alone might oversee the world. But, there will always be a margin between what humans understand and what computers can perform. Certain things are difficult to model with computers but can be performed by the human beings with their extraordinary intelligence. Consider an example, where the medical prescription written by your family doctor, which is hard to recognize is given to the computer/robot to read for the medicines. Here, the various available algorithms, which involves huge computations and requires large amount of power consumption (PC); fails to decipher the scrawls.

Rather the same data when given to a group of people, and asked to recognize, with each other's help, they're able to fairly accurately glean words out of the seemingly unreadable. A human perceps and classifies the samples (or patterns) in the original space (pattern space) whereas a computer does not understand the pattern space. Thus, crowd computing allows the humans to proactively use their own natural energy rather than artificial computations, requiring a large power consumption. We all have learnt the law of energy conservation in science, which says; in a closed system, the amount of energy is fixed. Energy inside the system can neither be created nor be destroyed. The energy can be converted from one form to another (and sometimes back again). Today, the distributed computing infrastructure makes use of millions of devices, which are the parts of IR and results in the increase of carbon emissions, which is a serious threat for energy and power. The environmental cost function $C(x)$ considered in many existing studies is as follows:

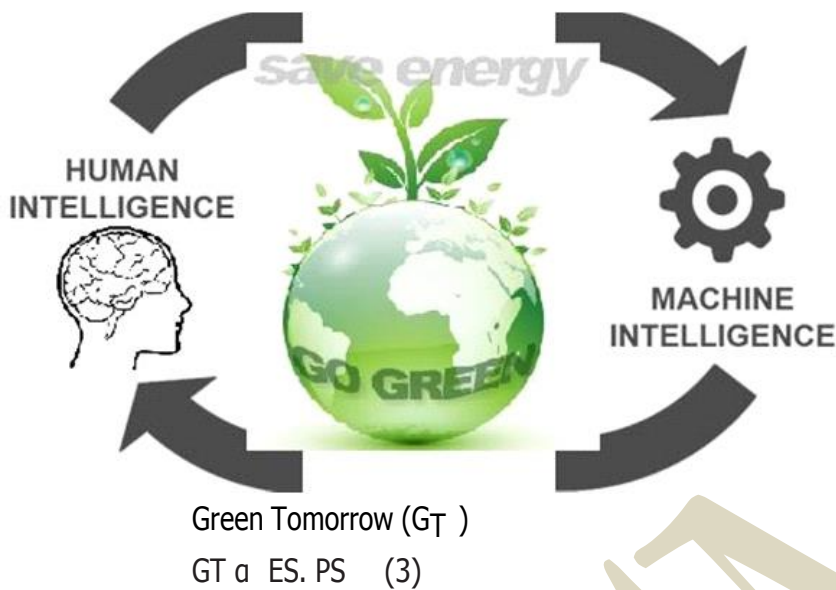
$$C(x) = c.r.PUE.P(x) \quad \dots\dots(1)$$

c – carbon footprint cost (\$/gm)

r – avg. carbon emission rate (gm/KWh) PUE – Power Usage Effectiveness

P(x) – server power = server utilization

$$P(x) = \text{Requirement/Capacity} = P_{\text{busy}}/(P_{\text{busy}}+P_{\text{idle}}) \quad (2)$$

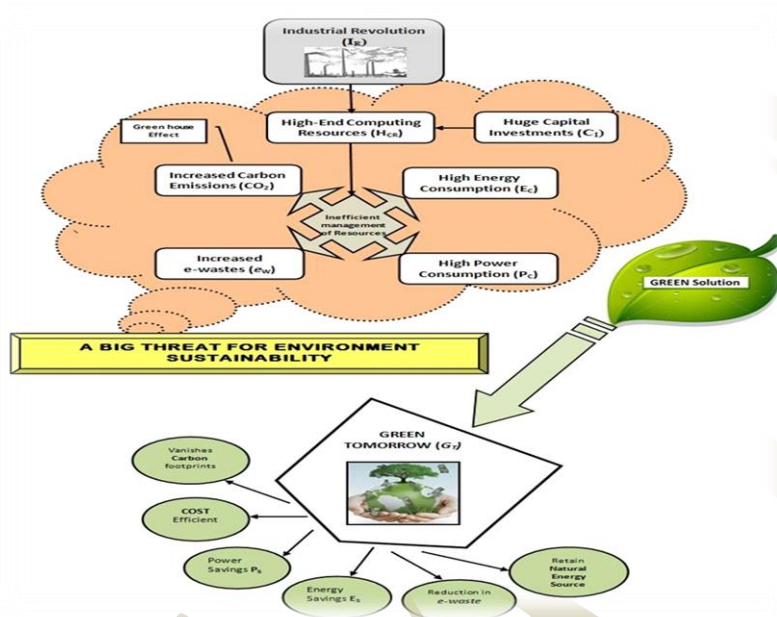


Green tomorrow (GT) as shown in above Figure proposes an energy efficient approach where the Humans and Machines make a great team to save 'Mother Earth'. GT is directly proportional to the integration of energysavings (ES) and power savings (PS). The reduction in energy consumption and reduction in power consumption will collectively contribute to the vast amount of energy and power savings. These valuable savings will drive us all to 2080 rather than having hazardous energy crises in 2050, which might lay the critical scenarios for the living beings to survive.

Rob Miller, a professor of computer science at User Interface Design Group, Massachusetts Institute of Technology (MIT) - Computer Science and Artificial Intelligence Laboratory (CSAIL), highlights the '*crowd computing*' revolution as a challenge that makes the workers/employers and machines/software as colleagues rather than competitors. The dependence of human beings on the various advanced gadgets and their accompanying software for the daily use is increasing day by day, thus superseding the human intelligence. The industrial revolution has given an inferiority complex to the humans as they may become secondary to the technologies they create, which requires huge amount of power and energy consumption. The day-to-day technological innovations & advancements are demanding for more and more amount of energy resources. Thus, energy conservation is one of the challenging tasks for the humans to survive in future.

There are various definitions and views for crowd computing, which takes a different approach for solving different problems. Srini Devadas, professor of electrical engineering and computer science at the MIT CSAIL, states crowd computing as a symbiotic relationship between software and humans. Crowd computing analyses and recognizes the strong and weak points of both, and efficiently utilizes those qualities for improved energy savings. The human knowledge is combined with the varying technological advances to yield efficiency and profitability. It leverages the human intelligence through its experience and subjectivity, against that of artificial intelligence through its speed and objectivity. Brain-guided computations can perform certain tasks at a great speed, that computers alone cannot, such as transcription, video moderation, etc. and result in the increase of techno power savings. The power of crowd computing has benefitted the digital universe with increase in the

energy savings and reduction in harmful emissions, e-wastes. According to Murray et al. in crowd computing, opportunistic networks can be used to spread computation and collect results. The mobile phones having large bandwidth (BW) is used as nodes and crowd computing allows the distributed human interaction tasks with optimized utilization of various resources. Cooke and Gillam describes crowd computing as the group of people who are offering their intellect and computers to solve problems which are at present unsuitable for computational approaches. According to Schneider crowd, computing is a myriad of human interaction tools that allow the exchange of ideas, nonhierarchical decision making, and full use of the world's mind space.



The emergence of Ubiquitous computing is rapidly changing the scenario of industries from ownership-based approach to subscription-oriented approach, where the access to scalable infrastructure and services is on-demand: anytime-anywhere. As shown in Figure, this massive computational operations, resulting in low resource utilization and wastage of energy.

There is an increasing demand for heavy power consumption (PC), for example a typical datacenter with 2000 racks needs nearly 25 Megawatt of power to operate, which results in higher operational cost along with the additional cost incurred on cooling process.

The continuous increase in the level of Carbon emissions (CO₂) by the industry is a dangerous pollutant that leads to adverse effects on the environment and all the living organisms. Gartner et al. [2007] estimated that the Information and Communication Technologies (ICT) industry generates about 2% of the total global CO₂ emissions. The proposed approach suggest that a decrease in emission volume of 15% – 30% is essential before year 2025 to manage the global warming situation and keep the global temperature increase below 20C. It is predicted that if the global temperature rises by 3.60 C, the polar ice caps and glaciers would melt, which would increase the water level of oceans by about 100 m and hence lead to the flooding of low-lying coastal areas of the earth.

The increase in e-waste due to the sophisticated advancements in electronic gadgets with large amount of energy and power requirements.

Thus, energy consumption, power consumption, carbon emission and e-wastes by industrial revolution/infrastructures have become a key environmental concern. All these factors together pose a Big

Threat for Environment Sustainability – A state in which the demands placed on the environment can be met without reducing its capacity to allow all living beings to live well, today and in Future. Hence, GT is considered to be the function of IR.

$$GT = f(IR) \quad \dots\dots(4)$$

IR \propto HCR . CI

IR directly depends on the heavy demand of HCR with huge CI. Therefore to have the green solution (GS) we need to control IR with appropriate utilization of various resources.

$$GT \propto 1 / IR$$

$$IR = (EC.PC.CO2.ew. CI)^n \quad \dots\dots(5)$$

$n > 0$ and $n \in \{ HCR \mid \text{As per the Application requirements} \} \Rightarrow \text{Max. (GT) = Min. (IR)} \approx (\text{Appropriate Resource utilization})$

growth of industrial revolution demands for high-end computing resources (HCR) with huge capital investments (CI). The inefficient management and handling of these resources result in critical penalty. The intensive increase in IR has following impact:

The proposed framework of GT gives the GREEN Solution for the Big Threat. GT is the result of Crowd Computing, which is a new energy efficient research paradigm. Here, the human intelligence and machine intelligence work hand-in-hand to create a miracle called 'GT'. EENACC is a great boon for green future: Save Energy – Save Earth. Crowd computing allows improvements by consuming less amount of energy. It's a fantastic solution to get a massive amount of computing power at fairly cheap amount of energy. The energy consumption can be minimized by actually measuring the amount of energy being used by monitoring the machines of the people who participate or contribute for the solution of a problem as the part of Crowd Computing process. There should be efficient resource management for proper industrial work with less power consumption. One of the important objective as well as economic incentive for an organization is the cutback in the energy budget of a datacenter. Crowd sourcing makes the traditional datacenters more energy efficient by using technologies such as resource virtualization and efficient workload allocation. Server consolidation reduces the energy consumption by allowing different workloads to share the same physical host using virtualization and also switching off the unused servers. Thus, energy optimization can be achieved by combining resources as per the current utilization, efficient virtual network topologies and thermal status of computing hardware and nodes. The EC and PC can be further reduced by performing the execution of massive computations on slow speed with an additional advantage of doing some other work simultaneously using all the computing resources without disturbing the know-how of the machines. Thus, bring together the unused computational power (PS) and save the additional amount of energy (ES) which would have been required in Future. Today we can have lots of computational power at cheap amount of energy with efficient use of resources and heavy financial gains-profit.

A recent research survey of various energy efficient solutions and strategies show that shifting the business applications to crowd computing can reduce and vanish the carbon footprints of organizations in the upcoming days. The use of crowd sourcing practices has shown following good results in reduction of carbon emissions:



Small Scale industries – upto 80 - 85 percent

Large Scale industries – upto 40 - 65 percent

Small Scale industries – upto 65 - 80 percent

Electricity is a significant source of energy, which is used to power homes, business, and industries. The combustion of fossil fuels to generate electricity is the largest single source of CO₂ emissions in the world. The lifetime of CO₂ is hard to define as the gas is not destroyed over time, but instead moves among different parts of the ocean–atmosphere–land system and has a negative effect on all living organisms. Hence by accepting and using the green solution there will definitely be a significant reduction in the carbon wastes. Saving the electricity will directly minimize the corresponding CO₂ emissions. Also, there must be efficient management and handling of the advanced electronic gadgets to reduce the e-wastes. There should be proper dispersal of the e-wastes or appropriate recycling methods must be used for the e-wastes to have reduction in energy and power consumption.



Breast Cancer Analysis **Prof. Rekhanjali Sahoo and Rajalaxmi Samal**

Introduction:

There has been a growing increase in the incidence of breast cancer, which is still the most significant cancer related cause of female mortality. In spite of significant progress in the management of breast cancer, the search for a curative treatment is still ongoing. Although a number of crucial studies and clinical trials have significantly contributed to the improvement of breast cancer care, many often remain unknown to the majority of clinicians, suggesting a need to identify at least the top 100 most cited studies in the field. Breast cancer is most frequently discovered as an asymptomatic nodule on a mammogram. A new breast symptom should be taken seriously by both patients and their doctors by the possibility of an underlying breast cancer at almost any age.

Machine learning a sub-field of Artificial Intelligence is used to achieve thorough understanding of the learning process and to implant learning capabilities in computer system. It has various applications in the areas of science, engineering and the society. Machine learning approaches can provide generalized solutions for a wide range of problems effectively and efficiently. The machine learning approaches make computers more intelligent. Machine learning helps in solving prognostic and diagnostic problems in a variety of medical domains. It is mainly used for prediction of disease progression, for therapy planning, support and for overall patient management. Hypothesis from the patient data can be drawn from expert systems mechanisms that use medical diagnostic reasoning. As mentioned earlier breast cancer is dreadful, so there is a need for computerized systems that emulate the doctors expertise in detecting the disease and help in accurate diagnosis. Machine learning has various approaches for building such systems. There is no single approach for all the problems and each approach perform differently for different problems. Therefore, there is a need for finding the approaches that perform well for a particular problem. In this thesis, various approaches are used for breast cancer diagnosis and they are compared to find the best performing ones.

I. Problem Statement

In the current system, the tumor images and the screenings take a lot of time to be analyzed by the radiologists and give a mammogram report. The mammogram report consists of certain characteristics of the tumor such as its radius, shape and texture. All these characteristics are later analyzed by oncologists and lets them decide which factors contribute to the malignant tumor. The entire process takes a few weeks and also puts a lot of pressure on the patient. To reduce the stress and cost a new system is required which generates instant results and also gives the patient some relief.

II. Proposed System

In the presence of tumor, this project will make predictions more accurately about the presence of tumor in a patient based upon the test report. The project eliminates the presence of a Doctor for the consultation in order to find out about the presence of benign or malignant tumor. As it is a well-trained machine-learning model, the accuracy with which it gives us results that are very high. The accuracy of the model is calculated to be 90%, which is very high, and an effective way to solve the existing problem.

III. Domain Information

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Some machine learning methods:

Machine learning algorithms are often categorized as supervised or unsupervised.

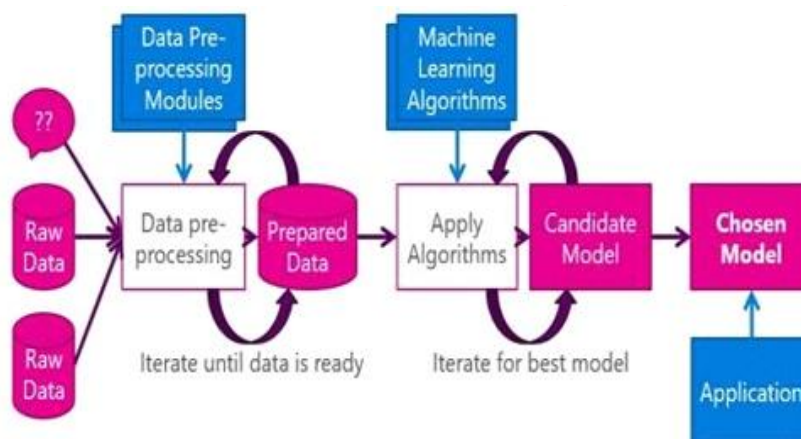
- Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.
- In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

IV. Experimentation Analysis

The dataset containing the report details is initially loaded into the program and the attributes which have high correlation are chosen. The `getdummies()` method is used to get the dummy values such as 'M' and 'B'. The test cases are split into test and train cases and analyzed. The train data is fit into the model and trained. Later the model predicts the values for the test data. The accuracy of the model is calculated by seeing how many test cases have been predicted accurately and how many test cases are wrongly predicted.

V. Architecture of Proposed System

Architecture diagram is a diagram of a system, in which the principal parts or functions are represented by blocks connected by lines that show the relationships of the blocks. The block diagram is typically used for a higher level, less detailed description aimed more at understanding the overall concepts and less at understanding the details of implementation.



VI. Algorithms

Logistic Regression:

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression). Mathematically, a binary logistic model has a dependent variable with two possible values, such as pass/fail, which is represented by an indicator variable, where the two values are labelled "0" and "1". In the logistic model, the log-odds (the logarithm of the odds) for the value labelled "1" is a linear combination of one or more independent variables ("predictors"); the independent variables can each be a binary variable (two classes, coded by an indicator variable) or a continuous variable (any real value). The corresponding probability of the value labelled "1" can vary between 0 (certainly the value "0") and 1 (certainly the value "1"), hence the labelling; the function that converts log-odds to probability is the logistic function, hence the name. The unit of measurement for the log-odds scale is called a logit, from logistic unit, hence the alternative names. Analogous models with a different sigmoid function, instead of the logistic function, can also be used such as the probit model; the defining characteristic of the logistic model is that increasing one of the independent variables multiplicatively scales the odds of the given outcome at a constant rate, with each independent variable having its own parameter; for a binary dependent variable this generalizes the odds ratio.

Logistic regression is a statistical method for predicting binary classes. The outcome or target variable is dichotomous in nature. Dichotomous means there are only two possible classes. For example, it can be used for cancer detection problems. It computes the probability of an event occurrence.

It is a special case of linear regression where the target variable is categorical in nature. It uses a log of odds as the dependent variable. Logistic Regression predicts the probability of occurrence of a binary event utilizing a logistic function.

Linear Regression Equation:

Where, y is dependent variable and $x_1, x_2 \dots$ and X_n are explanatory variables. Sigmoid Function: Logistic regression is named for the function used at the core of the method, the logistic function.

The logistic function, also called the sigmoid function was developed by statisticians to describe properties of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment. It's an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1, but never exactly at those limits.

Apply

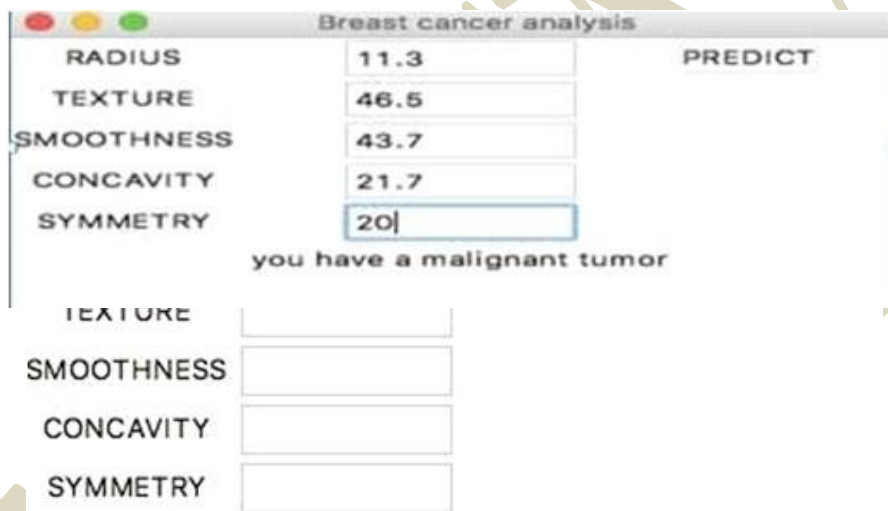
$$p = 1 / (1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)})$$

sigmoid function on linear regression:

Properties of Logistic Regression:

- ★ The dependent variable in logistic regression follows Bernoulli distribution.
- ★ Estimation is done through maximum likelihood.

VII. Results And Discussions



Feature	Value
RADIUS	11.3
TEXTURE	46.5
SMOOTHNESS	43.7
CONCAVITY	21.7
SYMMETRY	20

PREDICT

you have a malignant tumor

TEXTURE

SMOOTHNESS

CONCAVITY

SYMMETRY

...

VIII. Conclusion

In this project, we have outlined and technique for the problem of breast cancer detection. Although different classification techniques have been developed for cancer classification, there are still many drawbacks in their classification capability. In order to enhance breast cancer classification, in this project we proposed a new framework for breast cancer classification by combining mammogram wavelet transformation and neural network. According to results, classification based on locations of any abnormalities that may be present, character of background tissue, class of abnormality present, does not always shows the desired result. Finally, the evaluation and performance analysis of the proposed approach clearly shows that the preliminary 'results are promising in breast cancer discovery at early stage.



A Basic Analysis of Machine Learning and Its algorithm Types

Prof. Sasmita Pradhan

Abstract

Over the past few decades, Machine Learning (ML) has evolved from the Endeavour of few computer enthusiasts exploiting the possibility of computers learning to play games and a part of Mathematics (Statistics) that seldom considered computational approaches, to an independent research discipline that has not only provided the necessary base for statistical-computational principles of learning procedures, but also has developed various algorithms that are regularly used for text interpretation, pattern recognition, and a many other commercial purposes and has led to a separate research interest in data mining to identify hidden regularities or irregularities in social data that growing by second. This paper focuses on explaining the concept and evolution of Machine Learning, some of the popular Machine Learning algorithms.

Conclusion

The foremost target of ML researchers is to design more efficient (in terms of both time and space) and practical general purpose learning methods that can perform better over a widespread domain. In the context of ML, the efficiency with which a method utilizes data resources that is also an important performance paradigm along with time and space complexity. Higher accuracy of prediction and humanly interpretable prediction rules are of high importance. Being completely data-driven and having the ability to examine a large amount of data in smaller intervals of time, ML algorithms have an edge over manual or direct programming. In addition, they are often more accurate and not prone to human bias.

Consider the following scenarios: Development of software to solve perception tasks using sensors, like speech recognition, computer vision etc. It is easy for anyone to label an image of a letter by the alphabet it denotes, but designing an algorithm to perform this task is difficult. Customization of software according to the environment it is deployed to. Consider speech recognition software's that has to be customized according to the needs of the customer. Like ecommerce sites that customize the products displayed according to customers or email reader that enables spam detection as per user preferences. Direct programming lacks the ability to adapt when exposed to different environment.



Abstract

Planning a trip not only depends on the travelling cost, time and path, but also on the socio-economic status of the traveler. Data mining techniques support numerous applications of intelligent transportation systems (ITSs). This paper critically reviews various data mining techniques for achieving trip planning in ITSs. The literature review starts with the discussion on the contributions of descriptive and predictive mining techniques in ITSs, and later continues on the contributions of the clustering techniques. Being the largely used approach; the use of cluster analysis in ITSs is assessed. The relevance of the socio-economic constraints is defined using correlations, whereas the frequent as well as the feasible attributes are mined through the sequential pattern mining approach. The proposed model maintained a substantial trade-off between multiple performance metrics, though the trip mean model performed statistically.

Conclusion

This paper has introduced a new trip-planning model using data mining approaches. Real-time travel information has been acquired from the Indian city of Hyderabad, and the experimentation has been carried out to demonstrate the performance of the proposed planning model. The proposed planning model was able to produce the socio-economic constraints, which are highly relevant to the trip, rather than its frequency. Three levels of performance investigation have revealed that the proposed model has maintained an adequate trade-off between all these performance metrics.



Abstract

The healthcare industry collects huge amounts of health care data, which unfortunately are not mined and analyzed in a proper manner to discover hidden information, to take decisions effectively, to discover the relations that connect patterns. The aim of this paper is to develop a decision support in Heart Disease Prediction System (HDPS) using machine learning's effective algorithms. Using medical profile of the patient (age, gender, blood pressure, blood sugar, cholesterol, chest pain, ECG graph etc.), it can predict the likelihood of patient getting a heart disease. The likelihood (class label) may be of 5 stages: no, low, medium, high and very high. If an unknown sample comes, then the system will predict the class label of the sample. Hence two basic functions namely classification and prediction will be performed. Initially binary classification is performed to find whether there is a likelihood of disease. If yes, then multiclass classification is used to classify the disease among the remaining four stages. It is implemented in python as an application which takes medical test's parameter as an input. Comparison between the algorithm's performances is also depicted. It can be used as a training tool to train nurses and medical students to diagnose patients with heart disease.

Conclusion

Heart Disease Estimation System is an enhancement that makes estimation systems more effective by having more accuracy and confidence level on the data at present, based on its knowledge and behavior. The aim of Heart Disease Estimation System is to provide the right patient with the right information at the right time. A heart disease estimation system is based on two principal classification models: a binary model (which is all about classifying the test data into either of the 2 classes available), a multiclass classification (which classifies the test data into either of the classes available). An important aspect of this model is its high capability of showing up good confidence level. It is highly accurate at determining the accuracy of a Yes case. The proposed system can recommend and prescribe a suitable medication to the patient and immediately let the patient know the information needed at the right time.



Abstract

Human Activity Recognition (HAR) is a well-researched area that aims to recognize the activity performed by a person. However, the practical applications often encounter complications such as “The curse of dimensionality” and “Redundant features” which results in their poor performance. Hence, the need for feature selection is very imperative in such cases. This paper aims at identifying the subsets of HAR dataset that consists of most important and relevant features using Boruta Feature Selection Algorithm. Because the dataset with less number of features that are more relevant requires less computational time to train the classifier and it also improves the accuracy rate of the classification model. Upon identification, this paper also implements Support Vector Machines (SVMs) classification algorithm on identified subsets as well as HAR dataset. Furthermore, comparing the accuracy rates attained by the classifier on different subsets as well as their computational time.

Conclusion

On implementing the Boruta feature selection algorithm, the computational time taken to train the classifier has decreased, and the accuracy rate of the trained classification model has improved with the right subsets of HAR dataset. From all the results obtained, we can say that HAR dataset is more linearly separable as the SVM linear kernel has obtained highest accuracy for all test cases. The scope of this paper is to enhance the accuracy rate of a classification model by implementing feature selection on HAR dataset, which identifies subsets that consist of only most important and relevant features. This research has assisted us in understanding the HAR dataset and the mechanism of algorithms.



Feature Selection on Deep CNN features used for Image Classification

Abhishek DAS
(3rd yr CSE), B.Tech

Abstract

Feature Extraction forms the core of vision-based applications such as image classification, recognition, retrieval etc. Due to success of deep learning in several domains, feature learning has gained importance as compared to conventional image feature extraction. It has been observed that activation values extracted from pre trained convolution neural nets such as AlexNet, VGGNet give efficient results as compared to generic feature extractors in matching, retrieval etc. However, features obtained from fully connected layers of such deep nets are of considerable size as compared to conventional features of an image. Hence, it is important to select the best features, which can represent the image distinctly without reducing the efficiency of operation being performed. In this paper, we have experimented with features extracted from fully connected layers of VGG19 for classification of images. We have experimentally shown that feature selection when applied on fc1 or fc2 features greatly reduces time to build a classification model without affecting accuracy across all kinds of classifiers.

Conclusion

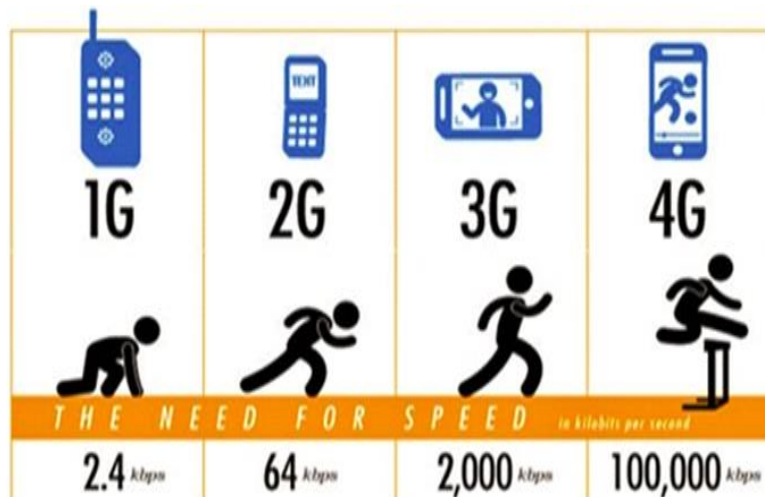
The features extracted from VGG16 are compared with conventional global features in terms of classification accuracy.

Advancements In The Field Of Wireless Network Communication

Sagar Ray (2nd yr CSE)
B.Tech



Telecommunication and networking has been and will be one of the core technologies in helping the evolution of humankind and technology itself. If it was not for these channels of communications and data transmission, we would probably still be in an era where technology isn't as advanced as today. We are currently in 4th generation trying to move to 5th Generation.



But where did it all start ?

If you remember, we all started with 2G. It was only about to be able to make a call to someone else far away from us (CDMA). We then moved to Global Systems for Mobile Communication (GSM) that enabled data transfer on top of voice communications. When we started off with data transfers the max we had achieved was (30-35 kbps).

Then with GPRS General Packet Radio Service had technology similar to GSM but data speeds improved to 110 kbps. Then with EDGE enhanced Data rates for GSM Evolution in 2003 we were able to achieve 135 kbps of data speed. It's still used in many parts of the world and by many operators as it satisfies basic needs of both carriers and users.

Then when we entered in 3rd Generation where there was a drastic development, we achieved the speeds of 2 mbps. This is when we could easily send emails and messages from our smartphones. Then in 4G Data rates of 100 mbps were achieved. However, the main drawback was the cost. 1GB of data cost us more than 150 rupees.

The game changer technology that came into the market was 4G LTE (long term Evolution). It's a complete redesign of previous architecture.

Why do you think the very famous telecomm JIO offers only 4G LTE services ?



Technologies until 4G and that from 4G LTE have different architectures. It's going to cost the company a lot to maintain 2 different architectures. Even companies like Airtel are planning to withdraw services below 4G.

This LTE technology has drastically got down the charges. 1 GB data on average costs us around 5 rupees now.

We are now in the era of the Internet of Things. We have so many smart devices, smart vehicles, smart homes etc. The need for a reliable network is greater than ever. That is what drives us to transmit to the 5th generation that aims at meeting the increasing demands and is more reliable. There are so many features that 5G offers. one of them is 5G New radio. 5G NR is specifically meant for vehicular technologies.

There has been a lot of advancements in the field of wireless network communication over the years in terms of overall development and change in core functionality, which has been crucial to put us in a era that is driven by technology all around us and with 5G a couple years away, technologies such as IoT, Cloud computing and AI will completely redefine our world by 2025.

Reference : <https://mse238blog.stanford.edu/2017/07/ssound/1g-2g-5g-the-evolution-of-the-gs/>

Jokes

1. A programmer gets stopped at an airport and is asked, "Do you have anything to declare?"

He answers, yes, three variables and a constant.

2. Why didn't the integer and string fall in love?

It was a type miss-match.

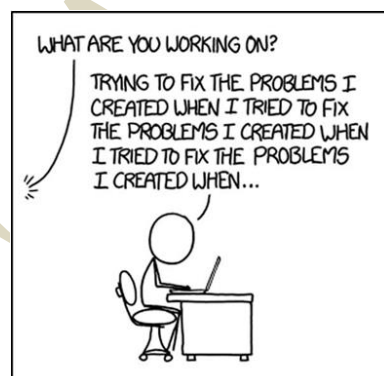
3. Why do Java programmers tend to wear glasses?

Because they can't C#.

4. Why computers are like women :

- No one but the Creator understands their internal logic.
- The native language they use to communicate with other computers is incomprehensible to everyone else.
- Even your smallest mistakes are stored in long-term memory for later retrieval.
- As soon as you make a commitment to one, you find yourself spending half your paycheck on accessories for it.

5.

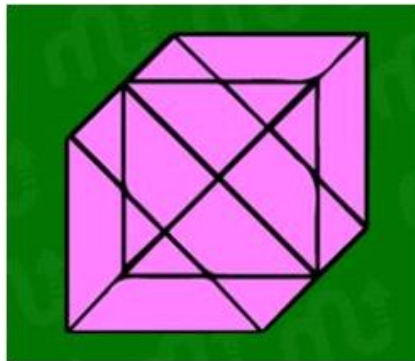


Jharana Patra

3rd year, B.TECH

Tricky Riddles

- 1) WHICH WORD IS WRITTEN INCORRECTLY IN A DICTONARY?
- 2) PEOPLE BUY ME TO EAT, BUT NEVER EAT ME. WHAT AM I?
- 3) WHO MAKES MOVES WHILE BEING SEATED?
- 4) EVA'S MOTHER HAD THREE CHILDREN. THE FIRST WAS CALLED APRIL, THESECOND WAS CALLED MAY. WHAT WAS THE NAME OF THE THIRD?
- 5) CAN YOU FIND THE NUMBER OF TRIANGLES IN THE GIVEN FIGURE?



Answers:

- 1) THE WORD IS WRITTEN "**INCORRECTLY**".
- 2) A PLATE
- 3) A CHESS PLAYER
- 4) ITS EVA!
- 5) THE NUMBER OF TRIANGLES IN THE GIVEN FIGURE IS '**24**'

Pravudatta Kar

3rd year, B.TECH



POETRY

EATM



May Day

A delicate fabric of bird song
Floats in the air,
The smell of wet wild earth
Is everywhere.
Red small leaves of the maple
Are clenched like a hand,
Like girls at their first communion
The pear trees stand.
Oh I must pass nothing by
Without loving it much,
The raindrop try with my lips,
The grass with my touch;
For how can I be sure
I shall see again
The world on the first of May
Shining after the rain?

Soumitree Jena (1901322107)



Someday

Someday you will cry for me

Like I cried for you.

Someday you'll miss me

Like I missed you.

Someday you'll need me

Like I needed you.

Someday you'll love me

But I won't love you.

Bibek Samal (1801322029)



Childhood

**Going back to the time, doing all that I want,
Just living up my life, getting all that I can't!**

**Great was the childhood, innocent were the smiles.
Walking through the woods was better than today's long drive of miles!**

**Getting nostalgic at the sight of old places,
Feeling the magic of love, affection and kisses.**

**Wish I could be selfless and joyous once again
And forget the mess that ends up only in pain!**

Luv Kumar (2001322078)



Mother Doesn't Want a Dog

Mother doesn't want a dog.

Mother says they smell,

And never sit when you say sit,

Or even when you yell.

And when you come home late at night

And there is ice and snow,

You have to go back out because

The dumb dog has to go.

Mother doesn't want a dog.

Mother says they shed,

And always let the strangers in

And bark at friends instead,

And do disgraceful things on rugs,

And track mud on the floor,

And flop upon your bed at night

And snore their doggy snore.

Mother doesn't want a dog.

She's making a mistake.

Because, more than a dog, I think

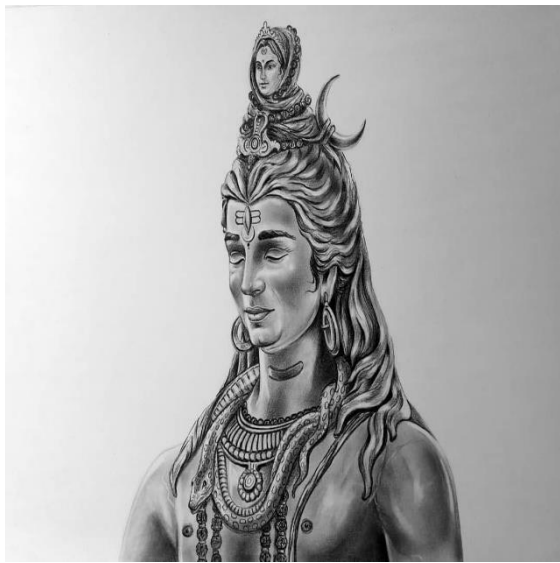
She will not want this snake.

Soubhagya Paikaray (2001322106)



SKETCH

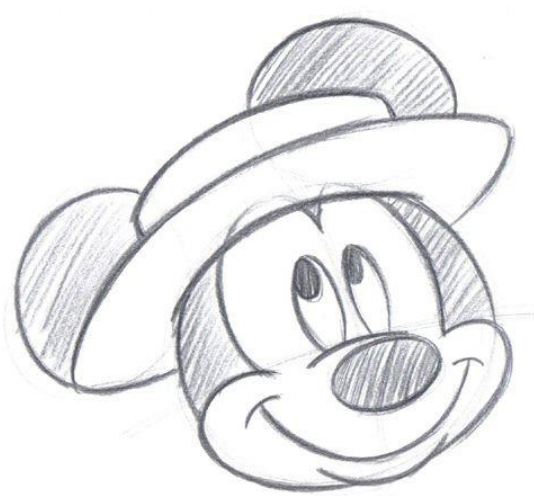
EATM



Laki Naika (1901322078)



Shubhalaxmi Nayak (2001322049)



Deepa Patra (1921322089)



SUDAMA RAY (1801322189)



EINSTEIN ACADEMY OF TECHNOLOGY AND MANAGEMENT KAT

AT: Baniatangi, P.O: Bajpur, Dist: Khurda,
Bhubaneswar-752060, Odisha, India,
PIN: 752060

