

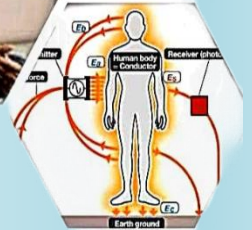
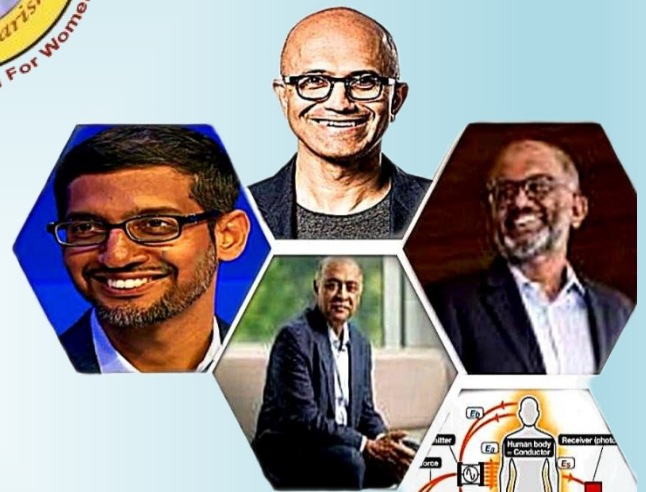
GAYATRI VIDYA PARISHAD COLLEGE OF ENGINEERING FOR WOMEN

[Approved by AICTE NEW DELHI, Affiliated to JNTUK Kakinada]

[Accredited by National Board of Accreditation (NBA) for B.Tech. CSE, ECE & IT – Valid from 2019-20 to 2021-22]

Kommadi, Madhurawada, Visakhapatnam – 530048

Department of Computer Science and Engineering



COSCENGER'S INSPIRE

2020
Volume - 4





GAYATRI VIDYA PARISHAD COLLEGE OF ENGINEERING FOR WOMEN

[Approved by AICTE NEW DELHI, Affiliated to JNTUK Kakinada]

[Accredited by National Board of Accreditation (NBA) for B.Tech. CSE, ECE & IT – Valid from 2019-20 to 2021-22] Kommadi, Madhurawada, Visakhapatnam – 530048

Department of Computer Science and Engineering

Institute Vision

To emerge as an acclaimed centre of learning that provides value-based technical education for the holistic development of students

Institute Mission

- Undertake the activities that provide value-based knowledge in Science, Engineering, and Technology
- Provide opportunities for learning through industry-institute interaction on the state-of-the-art technologies
- Create a collaborative environment for research, innovation, and entrepreneurship
- Promote activities that bring in a sense of social responsibility

Department Vision

To evolve into a Centre of learning that imparts quality education in Computer Science and Engineering to produce highly competent professionals.

Department Mission

- Impart computing and technical skills with an emphasis on professional competency and human values.
- Enrich the learning aptitude to face the dynamic environment of the Computer Industry.
- Enhance the analytical and problem-solving capability through contests and technical seminars.

Program Educational Objectives

After successful completion of the program, the graduates will be able to:

- PEO-1:** Apply both fundamental and advanced knowledge to analyze, design and develop innovative computing products.
- PEO-2:** Design and develop interdisciplinary and innovative software systems for real-world problems.
- PEO-3:** Inculcate soft skills, ethical conduct and an ability to engage in lifelong learning to serve the societal and environmental needs.

Program Specific Outcomes

Engineering Graduates will be able to:

- PSO-1:** Develop real-time applications by applying software engineering principles and implementing with emerging technologies in the field of Computer Science and Engineering.
- PSO-2:** Apply the knowledge of data analytics, soft computing, information security and other domains in Computer Science and Engineering for developing software systems.

From the Principal's DESK



It is pleasure to know that GVPCEW is bringing out the magazine of CSE department" **COSENGER'S INSPIRE - 2020**" for the year 2019-2020.

This institution constantly strives in the all-round development of the students through its endless efforts. COSENGER'S INSPIRE is one such endeavor providing a wide spectrum of engineering and artistic edifice, swaying from serious thinking to playful inventiveness. The inspiring women students at GVPCEW are brimming with zeal for life empowering themselves with skills and creativity.

I am happy that there is a dedicated team of staff and students who have brought out **COSENGER'S INSPIRE - 2020**. They have presented the stupendous achievements of CSE students of GVPCEW in the field of academics, sports and extra- curricular activities.

I extend my heartiest congratulations to the editorial board and all those who have shelved their valuable time to elevate this magazine to unprecedented heights. I wish the readers have a delightful reading. May all our students soar high in uncharted skies and bring glory to the world and their profession with the wings of education.

-Dr.K.V.S.Rao

EDITORIAL

We are happy to bring out an issue of the departmental magazine “**COSENGER'S INSPIRE**” for the year 2020.

This issue has the faculty article by Dr. M. Bhanu Sridhar, Associate Professor, Student articles, snippets on technology, and other regular features. The student's department activities in the preceding semester/year have been listed in brief.

We thank all the department members for continuous help bringing out this issue.

Editorial Team

Faculty Co-Ordinator: Mr. Addanki Udaya Kumar
Assistant Professor,
Department of CSE.

K Hari Chandana : 17JG1A0563

K Sai Himaja : 17JG1A0561

SK Rumana : 18JG1A0587

Y Sai Meghana : 18JG1A05A9

A Tejaswini sai : 18JG1A0507

E MahaLakshmi : 18JG1A0524

For Any Suggestions, Mail to: csemagazine@gvpcew.ac.in

Table of Contents

Sl. No	Contents
1	A Study of Data Analytics in Cloud Computing with an Emphasis on Security
2	Text Classification Using Natural Language Processing Techniques
3	Human Area Network Technology – Red Tacton
4	Robotic Process Automation
5	Know Your CEO's
6	Puzzles
7	Story
8	Crossword
9	Picking Up the Pieces from The Art Studio
10	NPTEL & Hacker rank
11	Contributions & Achievements
12	Placements Information

A Study of Data Analytics in Cloud Computing with an Emphasis on Security

Dr. M. Bhanu Sridhar

Associate Professor

Department of CSE

Abstract— Data Analytics is a sub area of Data Science with massive scope for research. With the advent of social media, the data has crossed the limits of a system, server and data center. Cloud Computing is another area in the IT field where different services like Software, Infrastructure, storage etc. are offered as services online. The main advantage here is that small business firms need not accumulate data storage as a physical existence. It has to be realized that data in Cloud is not only for storage but also for analyzing. This paper concentrates upon the recent trends in data storage and analyzing, in the clouds, and points out the security short comings. The future scope of this concept is concentrated on where Internet of Things (IoT) beckons the researchers through Analytics.

Keywords- Big Data, Cloud Computing, Security, SaaS, IaaS, IoT, Analyzing of Big Data

1. INTRODUCTION

Storage of data is a common procedure followed by anyone. The data starts to grow as years pass and is accumulated. It is up to the concerned company/individual to utilize the data so as to find out exactly what has happened and what might happen in the future.

Since the advent of Internet and much more after social media introduction, the data has grown by leaps and bounds every day. The word 'BIG DATA' was first mentioned by Michael Cox and David Ellsworth [1]. Apache Software Foundation took up the challenge of storing and processing the 'big data' and out came Apache Hadoop, Apache Pig, Apache Hive and Apache Mahout. Most recently, Apache Spark has been creating waves in the corporate IT field.

Data [2] refers to managing, analyzing and capturing different data sets where size, complexity and rate of growth varies for each of them. The benefits that can be accomplished with Data analytics are products can be redeveloped effectively, maintenance costs can be reduced, offering deeper insight from enterprise perspective, customizing websites in real time, creating new revenue streams and analysis of risks can be performed effectively.

Cloud computing [3], [4] is also an emerging research area where different services can be provided to the users on demand. The different services that can be provided are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS).

In SaaS, software is deployed over Internet is delivered as on demand service and its main characteristics are easy access to commercial software, no handling of software upgrades and patches and provides an API for integration between different pieces of software. In PaaS, platform for creation of software over the web is delivered as on demand service and its main characteristics are integration with web services and databases via the common standards, providing web-based user interface (UI) tools that help to create, modify, test and deploy different UI scenarios, providing support for multiple concurrent users utilizing same development application and support for development team collaboration. In IaaS servers, network, storage and operating systems are delivered as on- demand service and its main characteristics are including multiple users on single piece of hardware, resources are distributed as a service and providing support for dynamic scaling.

The different criteria that should be considered when analyzing Data in clouds are how to incorporate scalable, efficient and low-cost data storage platform and support for application development which involves modeling, mining, exploration and analysis of parallel execution of massive amount of data sets [3]. This paper deals with deeper analysis of Data in clouds, frameworks used for analysis and its pros and cons, security issues and challenges to be considered when analyzing Data in cloud and Internet of Things(IoT) and paves way for future research and development.

2. DATA IN CLOUD

Due to massive amount of information available worldwide Data is becoming tremendous challenge for today's rapidly changing traditional markets when performing in-depth analysis. The different areas where Data is being available are social media, mobile phone details, transactional data, documentations such as financial statements, insurance forms, medical records and customer correspondence, RFID tags, weather information, Internet of things (IoT), traffic patterns, communication events etc. Data is generally defined by five V's (Variety, Velocity, Volume, Veracity, and Visibility) [2] [3].

A. Data Business Drivers

The Data goals for many different types of organizations fall into major categories [2][5] such as

- **Revenue**
Design and execute Data analytics use cases that increase revenue, lower costs and improve efficiency in business operations
- **Customer services**
Improve customer understanding, obtain behavioral insight into client transactions and attract variety of customers.
- **Business development**
Introduce new products and services, decide what to outsource without affecting customer experience, gaining new competitive insight into markets.
- **Business agility and governance**
Plan with greater confidence, make better decisions, ensure regulatory compliance and lower costs.
- **IT and operational optimization**
Develop a strategy that uses existing enterprise areas for optimizing the applications

B. Cloud Computing

It refers to providing different types of services on demand to various categories of users. The different types of cloud for deployment of Data in cloud are as follows [4]

- **Public Cloud**
This allows different services to be accessible to the public on commercial basis by Cloud service provider. The major benefits are cost effectiveness, reliability, flexibility, scalability and its disadvantages are low security, less customizable.
- **Private Cloud**
It allows different services to be accessible within an organization. Its main benefits are cost efficiency, more control, high security and privacy and its disadvantages are inflexible pricing, limited scalability and additional skills are required to maintain cloud deployment.
- **Hybrid Cloud**
It is combination of public and private clouds which allows data and applications to move from one cloud to another. Its main benefits are scalability, flexibility and security and its disadvantages are network issues, security compliance and infrastructural dependency.
As it can be viewed in Figure 1 [16], a profound analysis is essential for the classification of clouds and the capabilities of each type. In comparison, each style will have its own pros and cons but after taking a specific view, it is for the customer to choose which

type can be relied upon. Public clouds are generally offered by cloud computing firms for enterprise organizations and can provide more services. Here, the need of the service and the security provided by the firm play a key role in the final decision of the user. Private clouds are used for small or middle-level business companies or even for personal usage. It becomes transparent for a customer that depending on the confidentiality of the project, and the level of resources or services provided in public and private clouds, he might use the ones that are best suited for his firm. Thus, comes out the hybrid cloud where some resources from the public cloud and some others from private cloud can be used to provide utmost satisfaction to the customers.

The services provided by the Public, Private and Hybrid clouds are depicted in an able manner in Figure 1 and a succinct look on the diagram provides the all needed information about the taxonomy of cloud computing.

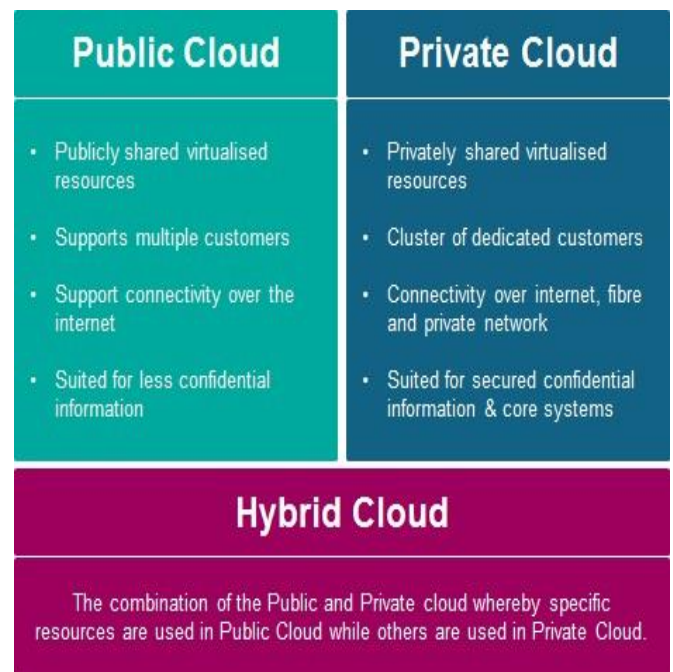


Fig 1: Classification of Clouds

C. Architectural Decisions for deploying Data in cloud

Data requires large amounts of storage and processing but traditional platforms for analysis such as data warehouses can't handle these gigantic data. So cloud computing is a means for accommodating these databases by using divide and conquer approach [3] [5]. Several architectural decisions are to be finalized before handling Big Data:

- Performance
- Scalability
- Reliability

- Availability
- Location and placement
- Sensitive data
- Disaster recovery

The key considerations to be taken for successful justification, management and deployment of data in cloud are build business case, assess data application workloads for deployment of data into cloud, develop technical approach for deploying and managing data and operationalize cloud based data infrastructure.

3. CLOUD BASED DATAANALYTICS FRAMEWORKS

There are several frameworks available for storing and processing of data Figure 2 elaborates the usage of Cloud Computing in Data Analytics [6]:

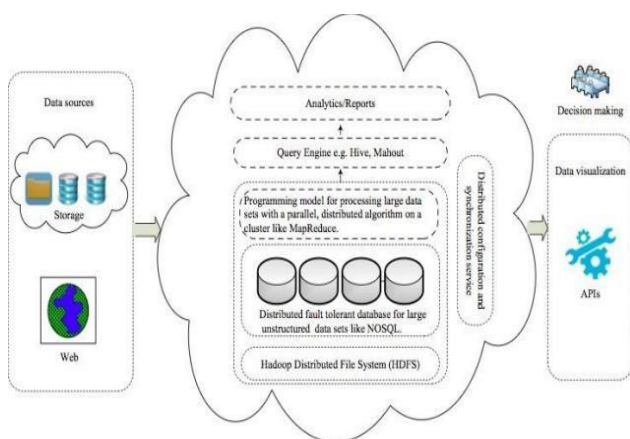


Fig 2: Usage of Cloud Computing in Data Analytics

Data Analytics require high performance processors to produce effective results for computation of data mining algorithms. There are several data mining techniques and tools that are available for extracting useful knowledge from large data sets but effective response in shorter time is the major criterion. A combination of Data Analytics and knowledge discovery with cloud computing systems offers an effective solution producing useful insights in shorter time. The several models for implementing Data analytics services are as follows [17]:

- **Data Analytics Software as a Service:**
It provides well defined data mining algorithms as service to end users.
- **Data Analytics Platform as a Service:**
It provides suitable platform for developers to build their own applications without concern about underlying infrastructure.
- **Data Analytics Infrastructure as a Service:**
It provides set of resources to run data mining applications

4. SECURITY ISSUES AND CHALLENGES FOR DATA ANALYTICS IN CLOUD

Security is becoming major issue for data storage in cloud-based networks. Cloud computing technology comes with security issues which include networks, databases, operating systems, virtualization, resource scheduling and allocation, transaction management, load balancing and memory management [15]. The security issues associated with cloud computing environment can be categorized into several levels such as [13][14]:

- **Network level**

The issues and challenges associated with the network level include network protocols and security in networks such as distributed nodes, distributed data etc.

- **User Authentication level**

The issues and challenges associated with this level include encryption/decryption techniques, authentication methods which includes authentication of distributed applications, access rights for nodes, logging etc.

- **Data level**

The issues and challenges associated with this level include integrity of data and availability issues with data such as protection of data and distributed data.

- **Generic level**

The issues and challenges associated with this level includes different usage of security tools and usage of different technologies

Cloud security alliance in 2013 identified top ten challenges for Data security such as

- Secure computation in distributed programming frameworks
- Security best practices for non-relational data bases
- Secure data storage and transactions logs
- End-point input validation/filtering
- Real-time security monitoring
- Scalable and compassable privacy-preserving data mining and analytics
- Cryptographically enforced data centric security
- Granular access control
- Granular audits
- Data Provenance

5. INTERNET OF THINGS (IOT)

It is becoming the next technological revolution where the revenue and data that IoT products generate will force the entire organizations to upgrade their tools and processes, technologies to accommodate this new data volume. The different ways in which IoT will impact data are data storage, data security, data tools and technologies.

6. CONCLUSION

Data is becoming tremendous challenge for today's rapidly changing traditional markets when performing in depth analysis. The different areas where Data is being available are social media, Mobile phone details, transactional data, documentation such as financial statements, insurance forms, medical records and customer correspondence, RFID tags, Weather information, Internet of Things (IoT), traffic patterns, communication events etc. Cloud Computing is another area in the IT field where different services like Software, Infrastructure, storage etc. are offered as services online.

The different frameworks like Hadoop, Spark, and Twister are used for analyzing and processing Data in cloud computing. In spite of its benefits, there are many security issues and challenges to be faced for storing of data in cloud-based networks at different levels such as Network level, User Authentication level, Data level and Generic level. Lastly Internet of Things (IoT) is also starting to have an enormous impact on Data analytics. IoT and cloud computing are inter-related and the security of sensor messages surely presents a topic for research. Data Analytics, which is being loaded and utilized from the Clouds these day's poses a serious challenge to the corporate business companies in the area of security.

The paper, after discussing Big Data, Analytics and Cloud Computing, proposes the budding idea of using the Data Analytics from the Cloud – like MS Azure Solutions, SAP HANA and other services. The paper proposes that security of these services must be more beefed up to avoid irreparable damages to the mined data in the Clouds. The area of Data Analytics in Cloud Computing ultimately is the tower to climb for the researchers to bring out satisfactory solutions to the ever-emerging problems in the field.

References:

- [1] Cox, Michael, and David Ellsworth. "Application-controlled demand paging for out-of-core visualization." Proceedings of the 8th conference on Visualization'97. IEEE Computer Society Press, 1997.
- [2] Rohit Chandrashekar, Maya Kala, Dashrath Mane, "Integration of Data in Cloud computing environments for enhanced data processing capabilities". International Journal of Engineering Research and General Science Volume 3, Issue 3, Part-2, May-June, 2015.
- [3] Charlotte Castelino, Dhaval Gandhi, Harish G. Narula, Nirav H.

- [4] Chokshi, "Integration of Data and Cloud Computing", International Journal of Engineering Trends and Technology (IJETT) – Volume 16 Number 2 – Oct 2014.
- [5] Ali Gholami, ErwinLaure."Security and Privacy of Sensitive data in Cloud Computing: A Survey of Recent Developments", NETCOM, NCS, WiMoNe, CSEIT, SPM – 2015,pp. 131–150, 2015.
- [6] Zhigao Zheng, Ping Wang, Jing Liu, Shengli Sun, "Real-Time Big
- [7] Data Processing Framework: Challenges and Solutions", Appl. Math. Inf. Sci. 9, No. 6, 3169-3190 (2015).
- [8] Samiya Khan, Kashish AraShakil, MansafAlam, "Cloud Based
- [9] DataAnalytics: A Survey of Current Research and Future Directions", Journal of Contemporary Psychotherapy,2015.
- [10] Ritu Aggrawal, Gautam Kumar, "Business Intelligence by Cloud Computing Using Hadoop Single Node Cluster", International Journal of Electrical Electronics & Computer Science Engineering, Special Issue - TELMISR 2015, ISSN: 2348-2273.
- [11] SanthoshVoruganti, "Map Reduce a Programming Model for Cloud Computing Based On Hadoop Ecosystem", International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014, 3794-3799.
- [12] Nikhil Gupta, Komal Saxena, "Cloud Computing Techniques for Dataand Hadoop Implementation", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 3 Issue 4, April – 2014.
- [13] Samira Daneshyar, Majid Razmjoo, "Large-scale data processing using MapReduce in Cloud computing Environment", International
- [14] Journal on Web Service Computing (IJWSC), Vol.3, No.4, December 2012.
- [15] J. Boehm, K. Liu, C. Alis, "Sideloaded - Ingestion of Large Point Clouds Into the Apache Spark DataEngine", ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B2, 2016, pp.343-348.
- [16] Damien Graux, Louis Jachiet, Pierre Genevs,NabilLaya, "SPARQLGX: Efficient Distributed Evaluation of SPARQL with Apache Spark", International Semantic Web Conference,2016. [13] R.Saranya, V.P.MuthuKumar, "Security issues associated with Datain cloud computing", International Journal of Multidisciplinary Research and Development, Volume 2, Issue 4, 580-585, April 2015.
- [17] Elmustafa Sayed Ali Ahmed, Rashid A. Saeed, "A Survey of DataCloud Computing Security", International Journal of Computer Science and Software Engineering (IJCSSE), Volume 3, Issue 1, December 2014,ISSN (Online): 2409-4285.
- [18] Shantanu Kalbhor, Hiteshkumar Jain, Kaushiki Upadhyay, "Providing classification and security of Datain Cloud computing", International Journal of Technical Research and Applications e-ISSN: 2320-8163, Volume 4, Issue 2(March-April, 2016), PP. 302-304.
- [19] <http://cloudacademy.com/blog/cloud-migration-benefits-risks/>
- [20] Domenico Talia ,"Clouds for Scalable Data Anaytics", IEEE Computer Society, Volume 46, Issue 5, May 2013, Pages 98-101.

TEXT CLASSIFICATION USING NATURAL LANGUAGE PROCESSING TECHNIQUES

SARAGADA DIVYA

(18JG1A05A9)

Department of CSE

Abstract— Data has been the New Oil in the 21st century. As a result of the rapid growth of globalization and digitalization, the size of data has increased where text classification has become a vital task. The task of text classification using Machine Learning can be developed using various Natural Language Processing techniques such as stemming and lemmatization, but the use of Term Frequency-Inverse Document Frequency (TF-IDF) helps refine text data even further, as TF-IDF produces feature values for training a classifier. Here the steps that should be followed for TF-IDF implementation are elaborated.

1. INTRODUCTION

NLP is a part of Computer Science, Human Language, and Artificial Intelligence. It is the technology that is used by machines to understand, analyse, manipulate, and interpret human's languages.

Applications

- Machine Translation
- chatbot
- Smart assistants
- Speech Recognition
- Data and Text analysis
- Sentiment Analysis

Machine Learning algorithms and almost all Deep Learning Architectures are not capable of processing strings or plain text in their raw form. They require numerical numbers as inputs to perform any sort of task, such as classification, regression, clustering. Also, from the huge amount

of data that is present in the text format, it is imperative to extract some knowledge out of it and build any useful applications.

The dataset consists of articles related to business, entertainment, politics, sport, technology. The model segregates the data into different classes based on the text content and nuances involved in the data. The plain text need to be processed by the model.

To convert the text data into numerical data, we need some smart ways which are known as **vectorization**, or in the NLP world, it is known as **Word embeddings**. There are other word embedding Techniques like One-hot Encoding (OHE), Bag-of-Words (BOW).

2. LITERATURE SURVEY

Shahzad et. al [1] proposed working principle and steps which should be followed for implementation of TF-IDF and built a model that combined TF-IDF with other techniques such as Naïve Bayes to get better results. Presented the strengths and weaknesses of TD-IDF algorithm.

Juan Ramos et. Al [2] concluded that Classification of huge data has associated problems like sparseness and higher feature dimensions in the extraction method this reduces the models generalization ability. To solve this problem Deep learning based Bi-LSTM-CNN using TF-IDF is employed to get higher accuracy in predicting the class.

3. BLOCK DIAGRAM

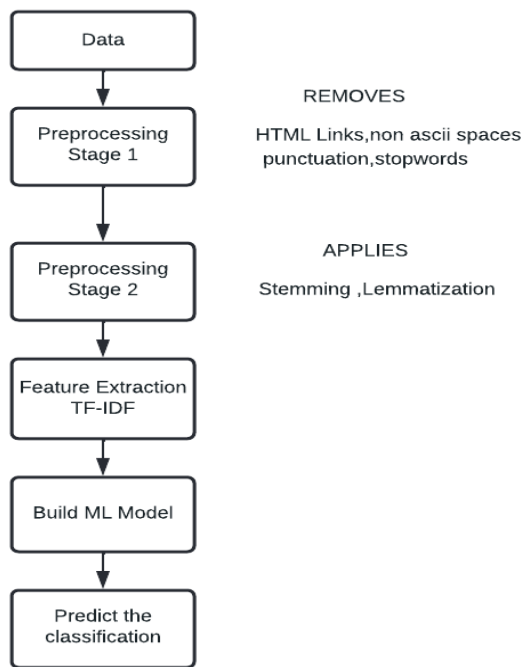


Fig. Flow chart of the proposed model

4. METHODOLOGY

Preprocessing of original data makes the text more predictable and analyzable for classifiers to work effectively therefore the input data is first channeled through various stages.

Cleaning the HTML links and removing non-ASCII characters, stop words and punctuation smoothens the text data and two key preprocessing techniques Stemming, and Lemmatization are also performed to attain the base forms of text and reduce size of corpus and improve text quality to make it more analyzable.

After the preprocessing stage, Feature extraction from the corpus has to be performed using TFIDF this vectorization method helps in dimensionality reduction as it takes both the Term Frequency and Inverse Document Frequency into account. Now Data is fit for preparing the machine learning algorithms.

4.1 Stemming

It is normalization of words where the base form of the word is achieved by shedding the prefixes and suffixes. Stemming and lemmatization are methods used by search engines and chatbots to analyze the meaning behind a word.

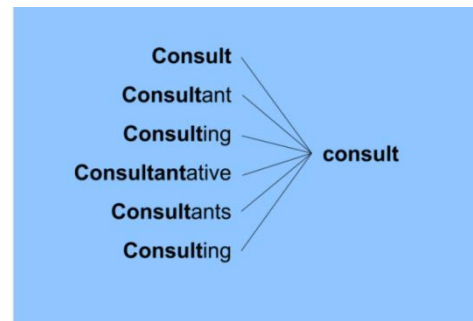


Fig. Stemming example

4.2 Lemmatization

It takes morphological analysis of the words into consideration. It converts the word into its base form and makes it a meaningful word for understanding the context in which it's used.

The difference between stemming and lemmatization is that a stemmer operates on a single word without knowledge of the context, and therefore cannot discriminate between words which have different meanings depending on part of speech.

4.3 TF-IDF VECTORIZATION

TF-IDF is a numerical statistic that shows the relevance of keywords to some specific documents or it can be said that, it provides those keywords, using which some specific documents can be identified or categorized. TF-IDF helps in further refinement of text data as TF-IDF[3] produces feature values for training a classifier.

In TF-IDF method each cell contains a weight value that signifies how important a word is for an individual text message or document [4].

TF(Term Frequency)

$$TF(\text{word}) = \frac{\text{Frequency of word in document}}{\text{Total no of words in document}}$$

Example:

Document: cat loves to play with ball

Tokenization: ['cat', 'loves', 'to', 'play', 'ball', 'with']

For the above sentence, the term frequency value for word cat will be:

$$tf('cat') = 1 / 6$$

When the term frequency of a document is calculated, it can be observed that the algorithm treats all keywords equally, doesn't matter if it is a stop word like "of", which is incorrect. All keywords have different importance. Let's say, the stop word "of" is present in a document 2000 times but it is of no use or has a very less significance, that is exactly what IDF is for.

IDF(Inverse Document Frequency)

The inverse document frequency assigns lower weight to frequent words and assigns greater weight for the words that are infrequent.

$$IDF(\text{term}) = \log\left(\frac{\text{No of documents}}{\text{No of documents containing term in it}}\right)$$

Here the main idea is to find the common or rare words.

For example, we have 10 documents and the term "technology" is present in 5 of those documents, few words like 'is' or 'and' are very common, and most likely, they will be present in almost all 10 documents IDF can be calculated as

$$IDF('is') = \log(10/10) = 0$$

$$IDF('technology') = \log(10/5) = 0.3010$$

TF-IDF Score

The greater or higher occurrence of a word in documents will give higher term frequency and the less occurrence of word in documents will yield higher importance (IDF) for that keyword .So multiply both to get TF-IDF.

$$TFID(\text{term}) = TF(\text{term}) * IDF(\text{term})$$

It calculates a weight value that signifies how important a word in a document

5. OUTPUT

TF-IDF produces feature values for training a classifier. If the input data contains the articles related to different classes, then the TF-IDF scores for the words belonging to a particular class will be as shown below

The words belonging to same class will have similar TF-IDF score as shown below

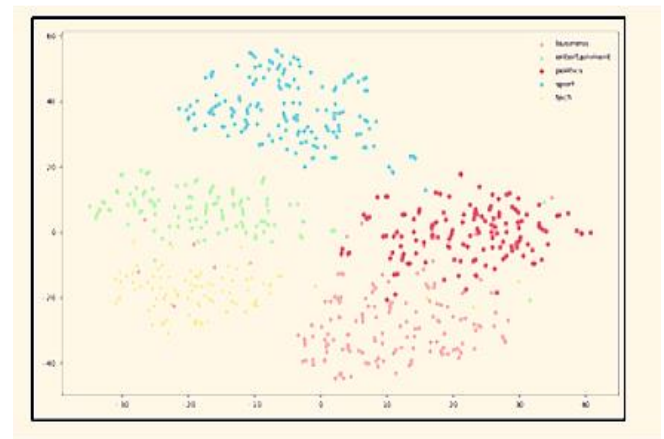


Fig. Classification based on TF-IDF Vectorization

6. CONCLUSION

TF-IDF vectoriser categorizes the words based on TF-IDF Score. Applying stemming,

lemmatization recognizes the tenses of words which improves the model also a further improvement is needed for capturing the semantic meaning of the words accurately. It can be combined with machine learning and deep learning model to predict the classification.

Search engine giants like Google has adapted latest algorithms such as PageRank for bringing out the most relevant results when a user place a query. In future research, world is going to witness some new techniques that can overcome the limitations of TF-IDF like check the semantics of the text in documents.

7. REFERENCES

- [1] Qaiser, Shahzad & Ali, Ramsha. (2018) "Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents". International Journal of Computer Applications. 181. 10.5120/ijca2018917395.
- [2] Juan Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries", pp. 1-4
- [3] Kowsari et al. —Text Classification Algorithms: A Survey, 201
- [4] Sujatha, M.M et al. | Metrics for assessing quality of a web site, International Journal of Innovative Technology and Exploring Engineering, 2019.

HUMAN AREA NETWORK TECHNOLOGY – RED TACTON

Chelluri Sai Kusuma Keerthi
18JG1A0516, Department of CSE

I. INTRODUCTION

At the end of 2002, the Japanese telecommunications group NTT announced that it would develop a new data transmission technology that uses the conductive properties of the human body to exchange information between electronic devices. With Red Tacton the company has now scarcely two and a half years later presented its first prototype of a Human Area Network (HAN).



Fig.1 RED TACTON Device

II. WHAT IS RED TACTON?

Red Tacton is a breakthrough technology that, for the first time, enables reliable high-speed HAN. In the past, Bluetooth, infrared communications (IrDA), radio frequency ID systems (RFID), and other technologies have been proposed to solve the "last meter" connectivity problem.

However, they each have various fundamental technical limitations that constrain their usage, such as the precipitous fall-off in transmission speed in multi-user environments producing network congestion

1. Red Tacton uses the minute electric field emitted on the surface of the human body. Technically, it is completely distinct from wireless and infrared.
2. A transmission path is formed at the moment when a part of the human body comes in contact with a Red Tacton transceiver. Physically

separating ends the contact and thus ends the communication.

3. Using Red Tacton, communication starts when terminals carried by the user or embedded in devices are linked in various combinations according to the user's natural, physical movements.
4. Communication is possible using any body surface, such as the hands, fingers, arms, feet, face, and legs. Red Tacton works through shoes and clothing as well.

III. MECHANISM OF RED TACTON

Data is received using a photonic electric field sensor that combines an electro-optic crystal and a laser light to detect fluctuations in the minute electric field. The naturally occurring electric field induced on the surface of the human body dissipates into the earth. Therefore, this electric field is exceptionally faint and unstable. The photonic electric field sensor developed by NTT enables weak electric fields to be measured by detecting changes in the optical properties of an electro-optic crystal with a laser beam.

IV. RED TACTON TRANSCEIVER

The figure below shows the block diagram of a RED TACTON transceiver. The signal from the interface is sent to the data sense circuit and the transmitter circuit. The data sense circuit senses the signal and if the data is present, it sends a control signal to the transmitter which activates the transmitter circuit. The transmitter circuit varies the electric field on the surface of our body. This change in the electric field is detected by the electro-optic sensor. The output of the electro-optic sensor is given to the detector circuit, which in turn is given to the interface of the receiving RED TACTON device.

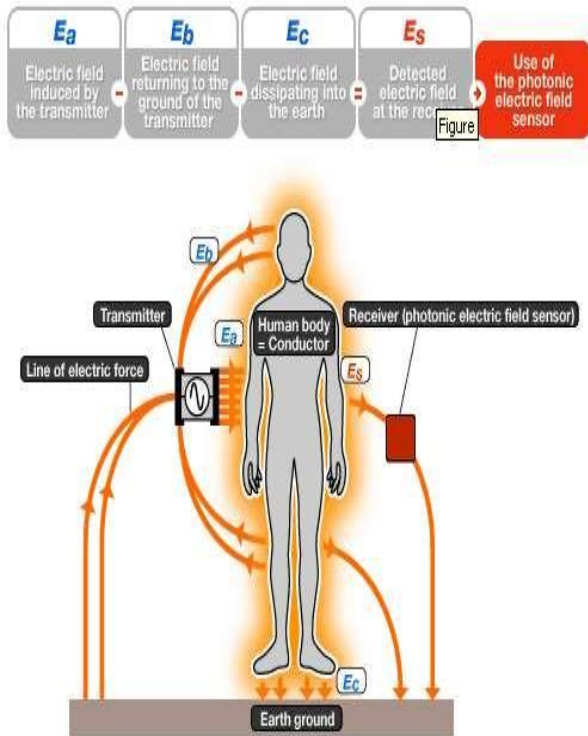


Fig.5 various electric fields on the surface of our body induced by the RED TACTON device. Only a portion of the induced electric field is sensed by the receiving RED TACTON device. The remaining electric fields are dissipated to the ground.

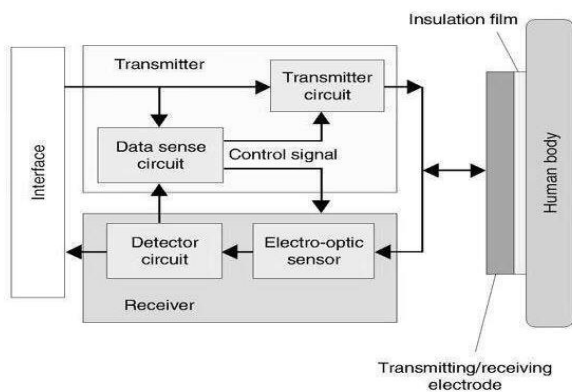


Fig.6 Block diagram of Red TACTON transceiver.

V. HUMAN AREA NETWORKS

In addition to the WANs (Internet) and LANs, there are applications best served by Human Area Networks (HANs) that connect the last meter.

Human society is entering an era of ubiquitous computing, where everything is networked. By making Human Area Networks feasible, Red Tacton will enable ubiquitous services based on human-centered interactions and therefore more intimate and easier for people to use.

VI. APPLICATIONS

1. One to one services
2. Intuitive operation of personal information
3. Device personalization
4. 4.new behavior patterns
5. Security applications

VII. CONCLUSION

Red Tacton technology could put the use of cables to an end. The problem faced by the Red Tacton technology is the cost of development. This technology brings a new dimension to communication which effectively links the user to anyone he wants to communicate with. Since it provides high-speed communication, it can provide seamless service wherever, whenever, and whoever uses it.

VIII. REFERENCES

- [1] Sunitha, C.. (2017). An Overview: RedTacton- Human Area Networking (HAN). International Journal of Trend in Research and Development., 4. 539-541.
- [2] Bobby, Murugesan & Devi, Ajantha & Madhubala, C. (2019). Red TACTON-Innovation of Human Area Networking. 5. 497-499.
- [3] Bobby, Murugesan. (2017). Red TACTON-Innovation of Human Area Networking.
- [4] Z. Wu, D. Xiao, X. Peng, H. Xu and X. Zhuang, "Human Body Network: Network in the Future?," 2008 IEEE International Symposium on Knowledge Acquisition and Modeling Workshop, 2008, pp. 578-581, doi: 10.1109/KAMW.2008.4810554.
- [5] Y. Kado et al., "Human-area networking technology based on near-field coupling transceiver," 2012 IEEE Radio and Wireless Symposium, 2012, pp. 119-122, doi: 10.1109/RWS.2012.6175318.

Robotic Process Automation

Abstract

RPA, or Robotic Process Automation, is a software that mimics the steps a human takes to complete rules-based, repetitive tasks. The robot carries out work with speed and precision, utilizing the same applications your employees use every day. In traditional automation, all the actions are primarily based on the programming/scripting, APIs or other ways of integration methods to the backend systems or internal applications. In distinction, RPA automates software that can migrate the work from the human to the computer which can stop paying humans to do work ripe for automation, faster front and back-office transaction processing, near “Instant On” integration at the lowest cost, optimization of User Interface to drive long call/transaction times down, accelerate digital transformation objectives, eliminate errors thereby improving productivity by making workers smarter.

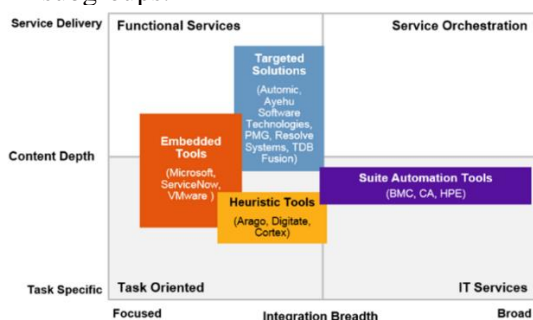
Introduction

Robotic Process Automation (RPA) is being leveraged by many businesses around the world in a wide variety of sectors to provide a vital competitive advantage in business process automation.

RPA is an application of technology which is programmed as a “robot” to capture and interface with existing applications for processing a transaction, manipulating data, triggering responses and communicating with other digital systems. RPA intends to assist or remove a human activity, using software to carry out tasks. RPA tool can be triggered manually or automatically, move or populate data between prescribed locations, document audit trails, conduct calculations, perform actions and trigger downstream activities.

Robotic Automation is a sum of two parts:

- **Robotic Desktop Automation (RDA)** – Used for personal robots for every employee, front office (call center, retail, branches) and back office, Toolbars, Wizards, UI enhancements and Task automation.
- **Robotic Process Automation (RPA)** – Unattended robots replicating 100% of algorithmic work, back-office operations, repetitive in nature, offload front office work, 100% improvement across smaller subgroups.



Note: Cortex was previously Inovivox
 Source: Gartner (March 2016)

RPA Evolution

Gartner predicts that by 2020, company spending on Robotic Process Automation software will reach \$1 billion, at a rate of 41% from 2015 through 2020. By 2020, RPA tools will have evolved to include functionalities such as Artificial Intelligence and Machine Learning capabilities.

Challenges in Traditional RPA Adoption

- **Pace of change** – Most companies are experiencing a typical two-year journey to make any real progress with automation.
- **Skilled People** – Limited availability of skilled resources – domain + technology against the market demand.
- **Governance** – Poor/Inefficient operating model of governance in managing and mitigating risks.
- **Data intake** – many companies are inundated by the challenges of getting started that they overlook data architecture, intake and coding.
- **Industry expertise** – As automation program evolves, the path for what should be automated becomes less clear and more industry-specific.
- **Digital Workforce** – Many companies don’t plan for the maintenance of resources required when automation breaks or underlying systems changes.
- **Technology overhead** – one especially large obstacle is the number of automation vendors and the larger-than-life claims some intake 93% of organizations say they are not fully prepared to handle the issues arising out of their automation journey.

RPA Effectiveness Test

The vital role of Robotic Process Automation is to offer improved customer experience and operation excellence by increasing performance, efficiencies, and agility in the process. However, the integration of RPA requires proper planning beforehand to ensure processes are fit for robotic automation.

The total value of automation goes well beyond financial impact. Consider the below factors when determining ROI:

- **Optimized operation costs**
- **Reduced cycle time**
- **Increased quality: rework, errors and variance**
- **Total flexibility: timeliness, scalability, seasonality**
- **Reduced penalties: payment interest, government**
- **Greater compliance: onshore, extremely detailed audit logs**

- Better insights: real-time visibility into work and outcomes
- Improved customer experience

RPA Implementation Approaches

Robotic Process Automation (RPA) tools can help organizations optimize business efficiency and effectiveness of their operations much faster, improves the accuracy of execution by reducing human interventions and lowers the overall cost. With the use of RPA tools, Organization can use the saved A key to the successful RPA implementation is in finding ways to overcome these obstacles by following a proven approaches like stated below:

- Project identification and prioritization
- Use Case identification and scoring
- Use Case development
- RPA development and operations



Conclusion

For successful implementation of RPA it cannot be a single functional ownership, instead it must be a more collaborative approach between IT and business. This all stands on the shoulders on fundamentals that should be reminded of in a big complex transformational change, need of a blueprint, need of a plan, change management and education internally, need to come up with a holistic plan of roll-out journey. Organize all the initiatives around true enterprise strategy around optimization, efficiency, audit, compliance for better visibility to create a better customer experience.

References:

- [1] Capgemini (2017). Robotic Process Automation: Gearing up for greater integration. White paper Available at: <https://www.capgemini.com/wp-content/uploads/2017/08/robotic-process-automation-paper.pdf>.
- [2] A Study of Robotic Process Automation Among Artificial Intelligence — Available at: <https://0x9.me/23EJU> Robotic Process Automation. Available at: <https://www.intellidyne-llc.com/sites/default/files/2019-02/ITD%20RPA%20Initiative.pdf>.

- [3] Maximizing Your RPA ROI. Available at: <https://www.helpsystems.com/resources/guides/maximizing-roi-robotic-process-automation>.
- [4] Robotic Process Automation (2017). A Pragmatic Approach to Digital Transformation. White paper Available at: https://contextor.eu/wp-content/uploads/2017/06/Livre_Blanc_Contextor_EN.pdf.
- [5] Dr. John Hindle, Dr. Mary Lacity, Dr. Leslie Willcocks, Dr. Shaji Khan (2018). ROBOTIC PROCESS AUTOMATION: Benchmarking the Client Experience. White paper Available at: <https://0x9.me/GSzhj>.
- [6] Mark Davison and James Manos (2018). WHAT IS YOUR BOT VISION? when and How to Scale Robotic Process and Cognitive Automation. White paper Available at: <https://0x9.me/CH6X1>.
- [7] David Chappell (2017). Introducing Blue Prism. Robotic Process Automation for the Enterprise. White paper Available at: http://www.davidchappell.com/writing/white_papers/Introducing_Blue_Prism_v2--Chappell.pdf.

KNOW YOUR CEO'S

S.NIHARIKA-18JG1A0586
T.SYAMALA-18JG1A0595

1.Sundar Pichai:



Pichai Sundararajan was born on June 10, 1972, better known as Sundar Pichai, is an Indian-American business executive.[6] He is the chief executive officer (CEO) of Alphabet Inc. and its subsidiary Google.

Pichai began his career as a materials engineer. Following a short stint at the management consulting firm McKinsey & Co., Pichai joined Google in 2004, where he led the product management and innovation efforts for a suite of Google's client software products, including Google Chrome and Chrome OS, as well as being largely responsible for Google Drive. Later Pichai was selected to become the next CEO of Google on August 10, 2015, after previously being appointed Product Chief by CEO Larry Page. On October 24, 2015, he stepped into the new position at the completion of the formation of Alphabet Inc., the new holding company for the Google company family. He was appointed to the Alphabet Board of Directors in 2017.



2.Satya Nadella:



Satya Narayana Nadella was born on 19 August 1967 is an Indian American business executive. He is the executive chairman and CEO of Microsoft, succeeding Steve Ballmer in 2014 as CEO and John W. Thompson in 2021 as chairman. Before becoming CEO, he was the executive vice president of Microsoft's cloud and enterprise group, responsible for building and running the company's computing platforms.

Nadella worked at Sun Microsystems as a member of its technology staff before joining Microsoft in 1992. At Microsoft, Nadella has led major projects that included the company's move to cloud computing and the development of one of the largest cloud infrastructures in the world. Nadella worked as the senior vice-president of research and development (R&D) for the Online Services Division and vice-president of the Microsoft Business Division. On 4 February 2014, Nadella was announced as the new CEO of Microsoft, the third CEO in the company's history, following Bill Gates and Steve Ballmer.



3.Arvind Krishna:



Krishna was born into a Telugu-speaking family in West Godavari District in the Coastal Andhra region of Andhra Pradesh, India. His father, Major General Vinod Krishna, was an army officer who worked for the Indian Army and his mother, Aarathi Krishna, worked for the welfare of Army widows.

Krishna joined IBM in 1990, at IBM's Thomas J. Watson Research Center,[5] and continued in Watson Research for 18 years till 2009. Next he held General Manager role in Information management software and systems and technology group of IBM. In 2015, he was promoted to senior vice president of IBM Research.[16] He later became senior vice president of IBM's cloud, and cognitive software division.

He also led the building and expansion of new markets for IBM in artificial intelligence, cloud, quantum computing, and blockchain technology.[17][18] He was a driving force behind IBM's \$34 billion acquisition of Red Hat, which closed in July 2019.

He was appointed IBM's CEO in January 2020, effective April 6, 2020, succeeding Ginni Rometty, who had served as CEO since 2012. In 2021, he was named by CRN as the year's "Most Influential Executive".

4. Shantanu Narayen:



Shantanu Narayen was born on May 27,

1963. Narayen grew up in Hyderabad, India, in a Telugu-speaking family, and was the second son of a mother who taught American literature and a father who ran a plastics company.

Narayen joined Adobe in 1998 as senior vice-president of worldwide product development, a position he held through 2001. From 2001 to 2005 he was executive vice-president of worldwide products.

In 2005 he was appointed president and chief operating officer. In November 2007, Adobe announced that Bruce Chizen would step down as CEO effective December 1, 2007, to be replaced by Narayen.

As CEO, Narayen led the transformation of the company, moving its creative and digital document software franchises – which include flagship programs such as Photoshop, Premiere Pro, and Acrobat/PDF – from the desktop to the cloud. In addition, during his tenure as CEO, Adobe has entered the digital experiences category, an expansion which began with the company's acquisition of Omniture in 2009.

In 2018 Adobe exceeded \$100 billion in market cap and joined the Fortune 400 for the first time. In 2018 it also ranked No. 13 on Forbes' Most Innovative Companies list.

PUZZLES

- 1) What mathematical symbol can be placed between 5 and 9 to get a number greater than 5 and smaller than 9.
- 2) There is only one ten letter word in the English language which can be typed using only the top row of the keys on a keyboard. What is it?
- 3) I am an odd number. Take away a letter and I become even. What number am I?

Did you know?

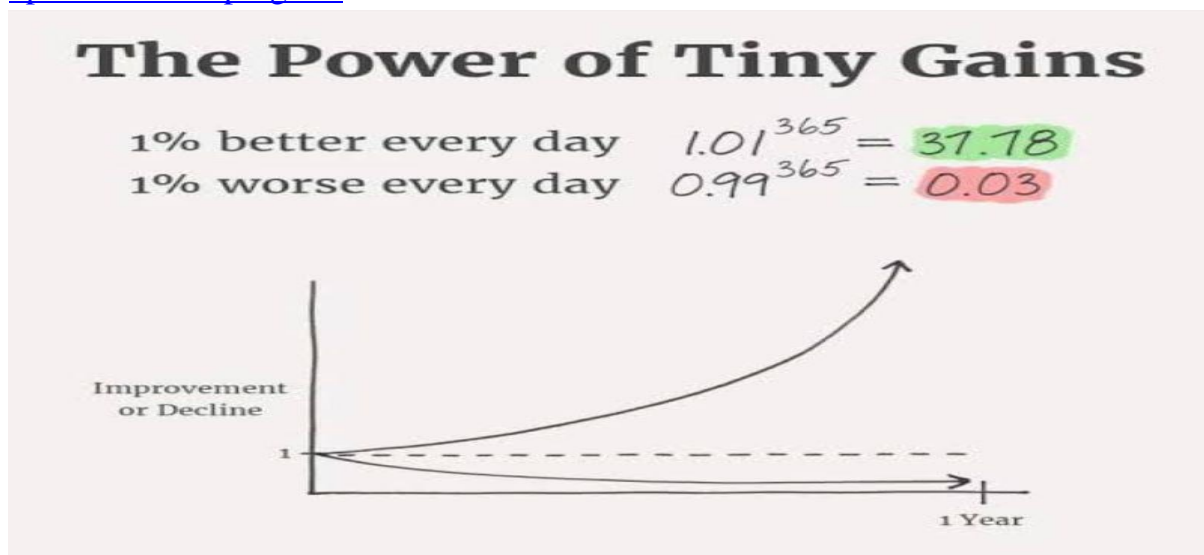
- Windows doesn't allow you to create a folder whose name is CON, PRN, AUX, or NUL. These are keywords reserved by DOS.
- CAPTCHA means... “Completely Automated Public Turing Test to tell Computers and Humans Apart”.

QUICK LESSONS

Practice does not make perfect!
Practice makes progress, here's how!

Suppose you want to learn a new skill, instead of trying to master that skill, make small improvements everyday. Don't practice to be perfect, practice to proceed.

[#practicemakesprogress](#)



The Magic Computer

I had a difficult geography project to finish by the end of the **semester**. My teacher wanted it to be **typewritten**, so I went to the school computer room. But when I got there, all the computers were turned off. Apparently there was a **recent** problem, and technicians were fixing it.

I knew of some **private** study rooms downstairs. They were small and dark, and the computers were very old, but I had no choice. At least the computers were **operating** correctly. I typed and **highlighted** the assignment's title: "**Evaluate** the Government's Response to **Global** Warming." But I didn't know what to write in my **essay**. Finally I decided to find a book to help me. I went to the library, checked the book **indexes** and eventually found a useful book. Then I returned to the computer.

When I looked at the screen, I saw something so **weird** that I nearly **fainted**! The essay was complete! Had somebody in **cyberspace** written it? I didn't know, but I was very happy. I printed it out and handed it in. I got an "A."

After that, I used the computer for all my assignments. I'd type the title, wait **awhile**, and the computer would do it. Every assignment was perfect; I never had to **edit** anything. I stopped paying attention to my teacher's **lectures** and spent my extra time in the **gymnasium**. And my grades got better and better.

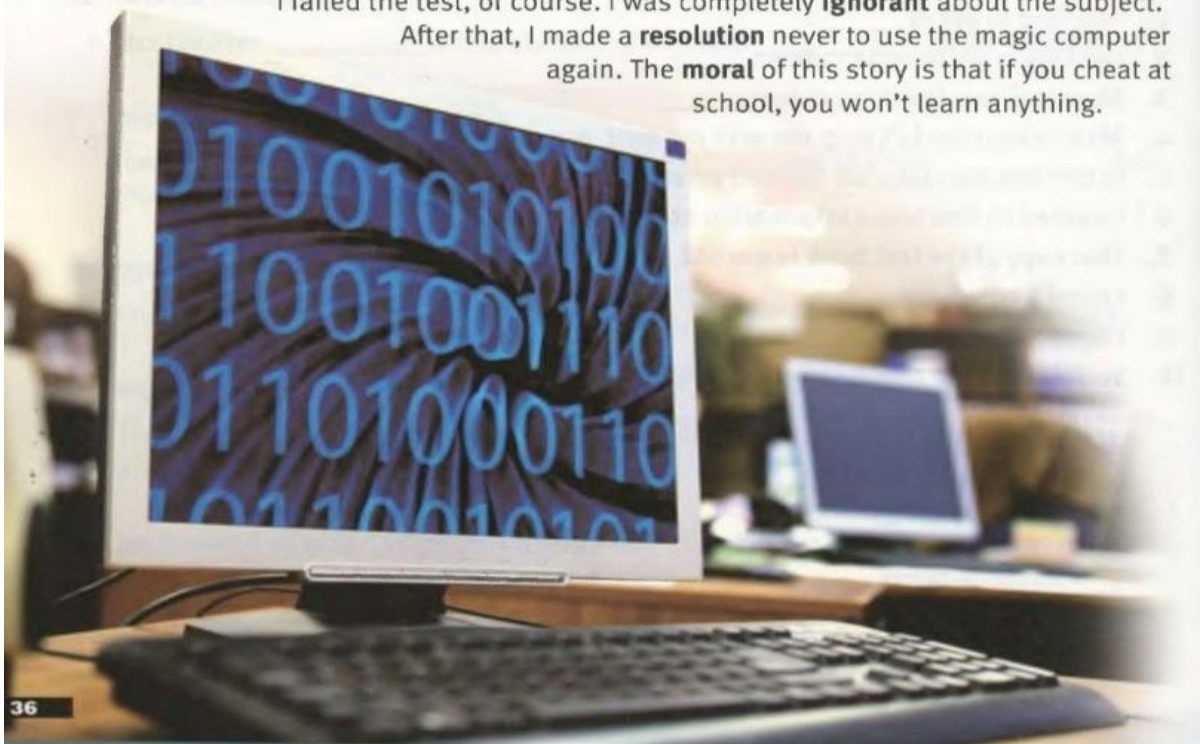
A month later, I was walking into class when my friend said, "Are you prepared for the test?"

"What test?" I asked.

"The geography test!" he replied. "I hope you studied. It's worth seventy percent of our final grade!"

I failed the test, of course. I was completely **ignorant** about the subject.

After that, I made a **resolution** never to use the magic computer again. The **moral** of this story is that if you cheat at school, you won't learn anything.



Here's a quick lesson -

Suppose you have Rs. 86,400 and you lost Rs.50, would you be upset and spend the rest of your money to find it.

No right!

Similarly we also have 86,400 seconds .

And we shouldn't let a negative 50 seconds ruin your entire day.

[#lifelessons](#) [#adviceoftheday](#)

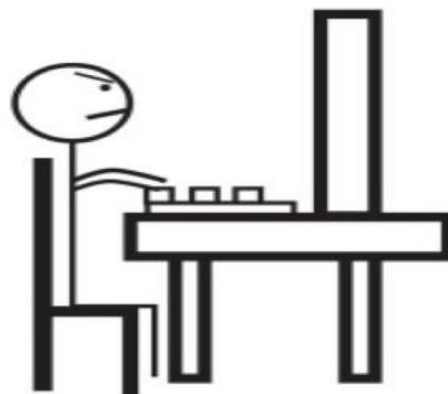
Bob is on the internet.

Bob sees something on the internet that offends him.

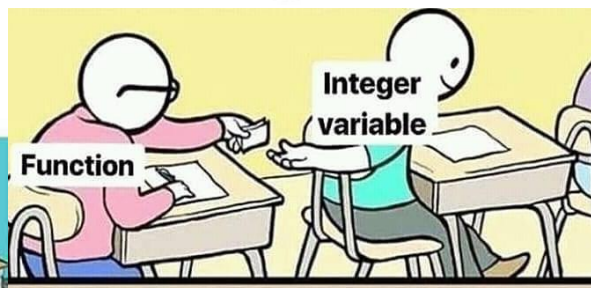
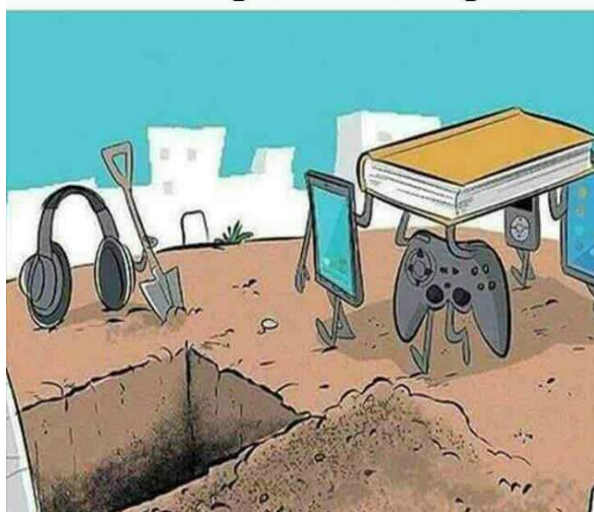
Bob stops what he was doing and posts comments about how he is offended.

Bob is stupid.

Don't be like Bob.



Today's Reality.



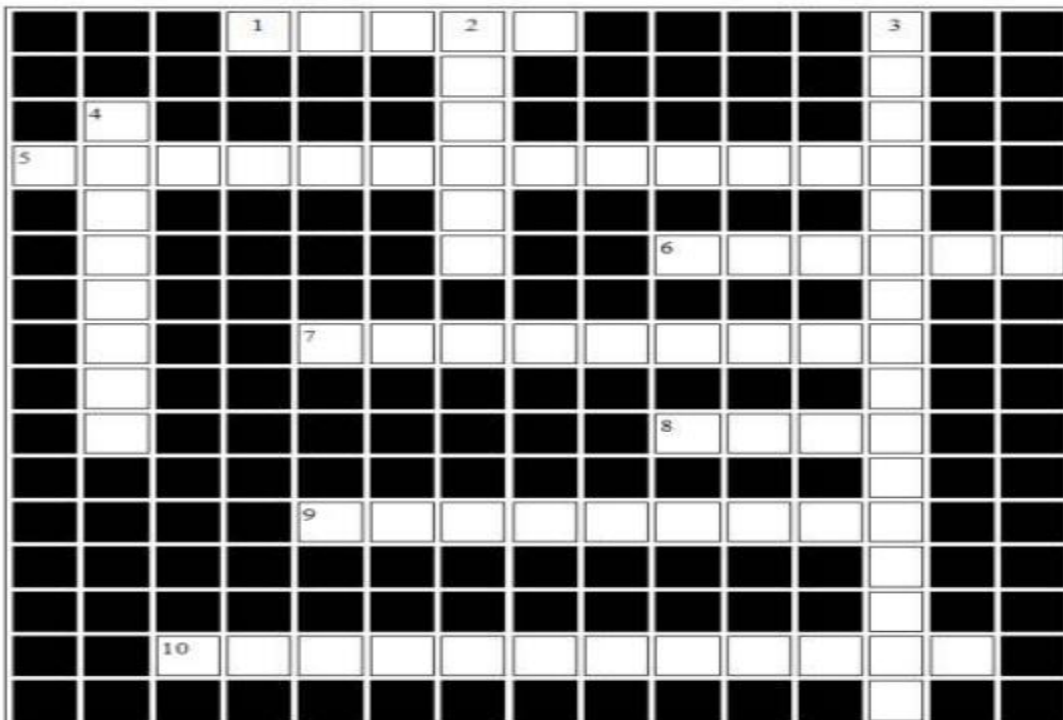
CROSSWORD

ACROSS

1. A ___ statement must be placed in between cases
5. Execute a loop a specific amount of times determined in advance
6. Classes require a main ___ in order to be an executable program
7. Variables that cannot change
8. Program used to code
9. Variable names must start with a ___ letter
10. Used when a task is needed to be done if the outcome of the comparison is false, you need to use the else part of this to be true

DOWN

2. Stores information in contiguous memory blocks
3. Used to evaluate an expression and tries to match the result to several possible cases
4. Arrays are used to a counted loop



PICKING UP THE PIECES FROM

THE ART STUDIO

Y ALEKHYA
18JG1A05A5

Y PAVANI
18JG1A05A4



TOPPERS IN NPTEL 2019 JAN-APRIL SEMESTER

Elite		Elite	
 NPTEL Online Certification (Funded by the Ministry of HRD, Govt. of India)		 NPTEL Online Certification (Funded by the Ministry of HRD, Govt. of India)	
This certificate is awarded to S DIVYA for successfully completing the course Introduction to R Software		This certificate is awarded to TEKI YASASWINI for successfully completing the course Introduction to R Software	
with a consolidated score of 95 %		with a consolidated score of 91 %	
Online Assignments 23.46/25	Proctored Exam 72/75	Online Assignments 25.00/25	Proctored Exam 66/75
Total number of candidates certified in this course: 1097		Total number of candidates certified in this course: 1097	
 Prof. Rajesh M. Hegde Chairman, Centre for Continuing Education IIT Kanpur	 Prof. Satyaki Roy NPTEL Coordinator IIT Kanpur	 Prof. Rajesh M. Hegde Chairman, Centre for Continuing Education IIT Kanpur	 Prof. Satyaki Roy NPTEL Coordinator IIT Kanpur
Aug-Oct 2019 (8 week course)		Aug-Oct 2019 (8 week course)	
 Indian Institute of Technology Kanpur		 Indian Institute of Technology Kanpur	
			
Roll No: NPTEL19MA33562340096		Roll No: NPTEL19MA33552340659	
To validate and check scores: https://npTEL.ac.in/hoc		To validate and check scores: https://npTEL.ac.in/hoc	

The National Programme on Technology Enhanced Learning (NPTEL) is a programme of the Ministry of Human Resources Development (MHRD) that was started in 2003 by seven Indian Institutes of Technology and the Indian Institute of Science, Bangalore, to give excellent education to anyone interested in learning from the IITs. The main purpose was to develop web and video courses at the undergraduate and postgraduate levels in all major fields of engineering and physical sciences, as well as postgraduate management courses.

Since 2013, 4-week, 8-week and 12-week online courses have been available through an online portal. These classes are free to enrol in and learn from. An optional in-person, proctored certification exam will be conducted at a cost of Rs. 1000/- per course, with a certificate issued by the participating institutions and industry.

TOP HACKER RANK HACKOS FROM CSE DEPARTMENT

S. No.	NAME	ROLLNO.	HACKOS
1.	Y.Nikitha	18JG1A05A8	5614
2.	Nagalakshmi Sowmya Mantha	18JG1A0566	4455

Hacker Rank is a technology company that focuses on competitive programming challenges for consumers and corporations, in which developers compete by attempting to programme according to given specifications. It is a gathering place for programmers from all over the world to solve challenges in a variety of computer science domains. Students from our college exhibit a strong interest in the topic and actively participate in it.

CONTRIBUTIONS & ACHIEVEMENTS

List of Department Activities conducted in the Year: 2019-20

Sl. No.	Date	ACTIVITY	Organizing Body
1	10-06-19 – 04-10-19	Value Added Course on “Programming approaches for Problem Solving “ by Dr. NB Venkateswarlu	Department
2	20-06-19	Seminar on “IEEE Student Branch and Activities” by Dr.N Deepika Rani, Vice Chair, IEEE, Vizag Bay Sub-Section for all IEEE student members of GVPCEW.	IEEE Chapter
3	22-06-19, 29-06-19, 20-07-19, 31-08-19	A 4-Day workshop on “ Use of Chatbots: A Web Technologies Latest Trend” by Dr. PVSL Jagadamba, Mr. V L Rao, Mr. K P Naidu, Mr., Mrs.B V Lakshmi.	Department
4	01-07-19 - 6/7/2019	INFYTO Training Program, Mr. K Nagaraju, Ms. BV Lakshmi, Ms. D Indu.	Department
5	07-06-2019	“A Coding Quiz Competition “by COSCENGERS Team	COSCENGERS
6	27-07-19	One day seminar on “The Role of IT in New Age Broadcasting” by Managing & Technical Team, Broadcasting Engineering Society (INDIA), Hyderabad	Department of ECE
7	13-08-19	Expert talk on “Importance of Higher Education to surge”, Sravani, USIES, Hyderabad.	CSI Chapter
8	30-08-19	A Guest Lecture on “Statistics Concepts on Data Science “ by Mr. Kamal Behera, Data Scientist, northdoor, UK.	Department
9	09-09-2019	Seminar on “Reporting an IEEE event and filing of L-31” by Dr.S.Lakshminarayana, Executive member, IEEE, Vizag Bay Sub-Section for all IEEE student members of GVPCEW.	IEEE Chapter
10	09-11-2019	A Guest Lecture on “IBM Cognos Analytics” by Sandeep Kaur, Software Developer, IBM, Vizag.	CSI Chapter
11	10-01-2019	Seminar on “Machine Learning Applications using Python Programming” by B.L.V.Vinay Kumar, Assistant Professor, GVPCEW as a part of IEEE- Day Celebrations.	IEEE Chapter
12	28-09-19	A Guest Lecture on “ Latest Trends in Data Analytics: An Introduction” by Mr. T Suresh Kumar, WIPRO, Visakhapatnam	COSENGERS
13	27-09-19, 7/11/2019	2-day training program on “Google Explore on Machine Learning”, Dr. M Swapna	Department
14	18-11-19 - 23-11-19	A workshop on “Salesforce student program ADX 201” by Mr. Kranthi Kumar, ICT Academy	Department
15	17-12-19	One day Student Convention os CSI	CSI Chapter

A COUPLE OF PHOTOS FROM THE EVENTS



Speech by
Dr. P. V. S. L. Jagadamba,
HOD, DEPARTMENT OF CSE

On the day of the inauguration,
lamps are lit.
Siva Jyothi, secretary
secretary's introduction speech.



Gathering at the venue.

Culturals of the event



Placements Information

For the batch of 2016-2020:

- Highest Package: **11CTC**
- No. of students placed: **63**
- No. of offers received: **90**

*Following companies have taken more than 10 students in their organization
Companies that have offered more than or equal to 4.5 Lakhs package.*

