

Real Time Air Quality Index from Various Locations



Appasani Geetha, Pasumarthi Sai Ramya, Chenikala Sravani, M.Ramesh

Abstract: The surveys regarding air pollution shows that there has been a hasty growth due to the emission of fuels and exhaust gases from factories. The Air Quality Index (AQI) has been launched to note the contemporary status of the air quality. The intent of AQI is to aid every individual know how the regional air quality will make an impact on them. The Environmental Protection Agency assess the AQI for five major air pollutants namely Nitrogen dioxide (NO₂), ground-level ozone (O₃), particle pollution (PM₁₀, PM_{2.5}), carbon monoxide (CO), and sulphur dioxide (SO₂). The intent of the project is to congregate real-time Air Quality Index from distinct monitoring stations across India, analysing the data and reporting on it. Collect the real-time data using the API key provided by Open Government Data (OGD) platform India. This is done by making use of Microsoft Business Intelligence (MSBI) and Power BI Tools to transform, analyse and visualize the data. This project can be utilized to develop various programs like Ozone today in Europe and in mobile applications which acts as an alert system that can protect people from air pollution.

Keywords: Air Pollution, Air Quality Index, Business Intelligence, Power BI, pollutants, SSIS, SSAS.

I. INTRODUCTION

In India, the environmental conditions have been continuously disintegrated because of the industrial developments, urbanization, lack of awareness and motor vehicles. The emissions from industry and vehicles are the prime purpose behind toxins in India [1]. Various pollutants, namely, sulphur dioxide, nitrogen dioxide, carbon monoxide, ground-level ozone (O₃), PM₁₀ and PM_{2.5} impacts the air quality and consequently leads to the effect of human health. Latterly, these specifications were observed to determine the nature of air for the significant number of states and urban communities in India. A report of all stations, urban communities, and states all over India is generated, using AQI (Air Quality Index). The Environmental Protection Agency (EPA) introduced a tool named AQI to assess the levels of pollution [8] in environment because of the major air pollutants.

The National Air Quality Index (AQI) was pioneered in New Delhi on September 17, 2014, under the Swachh Bharat Abhiyan. The values of AQI and the equivalent concentrations along with the related health impacts for the associated six pollutants are as given below [10]:

Table – I:
AQI Category, Pollutants and Health Break Points

Category	Range
Good	0 – 50
Satisfactory	51 – 100
Moderately Polluted	101 – 200
Poor	201 – 300
Very Poor	301 – 400
Severe	401 – 500

Air pollution has become global atmospheric problem that causes harm to all the living organisms [6]. This can cause a severe threat to the people if it exceeds the satisfactory limit. The status level of air gives rise to the intricate interaction of numerous factors that include the science of the atmosphere and the effusion of various pollutants. Surveys have mentioned that frequent exposure to environmental air pollutants over a prolonged period of time intensify the risk of being vulnerable to air borne sicknesses, for example, cardiovascular disease, respiratory disease, skin diseases and lung cancer. Throughout the world, many cities persistently analyse air quality utilizing monitoring networks drafted to live and record air contamination fixations to personify exposure of the population to those pollutants. Prevailing studies shows that guide-lines of suggested pollution values cannot be regarded as cut-off values below which a zero adverse acknowledgement may be expected [6]. Hence the clean and fresh air is a sole requirement for good human health and wellbeing of the humanity [8]. Several programs have begun to keep track on the quality level of air and to accumulate the vast amount of data on concentration of each air pollutant (e.g., SPM, CO, NO₂, SO₂, etc.) in different parts of the world [12]. The massive data usually don't pass on the air quality status to the scientific community, officialdom, policy makers, and especially to the overall public during a simple and straightforward manner. This issue is approached by deciding the Air Quality Index (AQI) of guaranteed area using Microsoft Business Intelligence, SQL Server Management Studio and Power Bi tools.

Manuscript received on May 25, 2020.
Revised Manuscript received on June 29, 2020.
Manuscript published on July 30, 2020.

* Correspondence Author

Appasani Geetha, Student-IT, Velagapudi Ramakrishna Siddhartha Engineering College, India,

Pasumarthi Sai Ramya, Student-IT, Velagapudi Ramakrishna Siddhartha Engineering College, India,

Chenikala Sravani, Student-IT, Velagapudi Ramakrishna Siddhartha Engineering College, India,

M. Ramesh, Guide, Assistant Professor-IT, Velagapudi Ramakrishna Siddhartha Engineering College, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

II. LITERATURE SURVEY

This chapter presents a study of analysing the Air Quality Index using individual methods and tools and to interpret the significance of AQI in atmosphere. This empowers us to know the perception of various creators and gives a clear view of data interpretation and comparison between the various tools.

Analysis of Air Quality Index – In this work the authors have proposed a detailed research on Air Quality Index. They analysed the Air Quality Index Values for three cities Delhi, Bangalore and Chennai. They found Air quality Index can be calculated by two methods, one as taking the greatest value among the pollutants and the other is finding the average of the considerable number of poisons. Also this paper predicted the AQI for the future years using Regression analysis method. All the pollutants SO₂, NO₂, RSPM, and PM are analysed with analysis of Variance test. The classification for the calculated Air Quality Index values are as Good (0-50), Moderate (51-100), Unhealthy for sensitive groups (101-150), Unhealthy (151-200), Very unhealthy (210-300), and Hazardous (301-500). The Critical difference, Standard Error and Mean are calculated and a Post hoc test is used to find the difference between the mean and to find which mean is highest. Utilizing these qualities the work is additionally reached out to fit an Auto Regressive Model and predict the values for proceeding days [1]. A Visualization approach of Air quality Index using R – Martand Pratap Singh and Dileep Kumar proposed an efficient visualization approach to certainly view the spatial and temporal details involved in the data of air pollution. The tool R has been utilized to examine the air pollution information which comprises of an Open air package developed usually for the atmospheric sciences. They considered the outgrowth of package with the motive of demonstrating how intuitively the data of air pollution can be analysed. One of the focal themes of open air is the utilization of conditioning plots and analyses, which hugely enrich inference possibilities. The author added specific tools that can be used for air quality monitoring are using wireless sensor network, IAQ tool kit, Wireless monitoring of environmental parameters through Zigbee and AQI estimation as per EPA standard. They summarized that the vehicular outflows are of the specific concern of the fact that these are beneath sources and therefore have the utmost effect on widespread community [4].

Common Air Quality Index (CAQI) - The CAQI was evolved in the course of CITEAIR project and has been around since 2006. The author found that numerous urban communities present their air quality in different, hard to compare ways. Sef vend den elshout then created this index for the purpose of observing the air quality of Europe in real-time easily. The Common Air Quality Index (CAQI) is developed to view and contrast air quality on hourly or daily basis. The CAQI has a scale from 1 to 100 with the smaller rankings producing better Air Quality and the color ranges from light green to dark red. There are core pollutants such as NO₂, PM₁₀ and other pollutants (CO, PM_{2.5}), without those core pollutants index cannot be calculated. A sub-index is computed for every pollutant corresponding to a grid that converts concentration measurements into a

ranking on a scale. The city index is the utmost of overall index. They have this procedure separately for traffic and background indices. The CQAI is determined in accordance with the grid system by linear interpolation between the class borders. They collated all the existing methods for the identification of suitable alternative [11].

III. ARCHITECTURE DIAGRAM

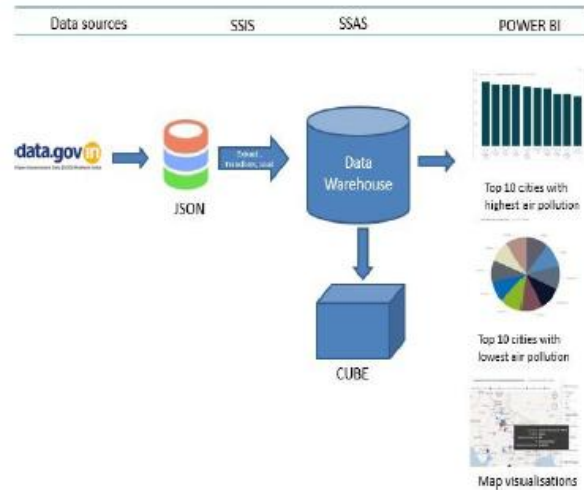


Fig 1. Architecture diagram

The information is gathered from the Open Government Data (OGD) India – data.gov.in. It gives open access to Government possessed shareable information and the information is updated periodically. The information is accessible in the format of JSON, XML. Here, the data taken is of JSON format.

MSBI contains of tools which offers better responses for Business Intelligence and Data Mining Queries. It utilises Visual studio along with SQL server. It offers various instruments for various frameworks, which are required in Business Intelligence solutions.

SQL Server Integration Services (SSIS): SSIS could be a segment of the Microsoft SQL Server database programming that can be utilized to execute a wide scope of information movement tasks. SSIS gives a quick and adaptable information warehousing instrument utilized for information extraction, stacking and change like cleaning, gathering, like cleaning, gathering, combining data. SSIS in like way fuses graphical instruments and wizards for accomplishing workflow functions like sending email messages, FTP tasks, objectives and data sources. SSIS ETL tools assists with building enterprise level information change and information integration arrangements. ETL means

- E - Extricating the information from different sources
- T - Changing the information to meet the prerequisites
- L - Loads the information into target information distribution center.

SQL Server Analysis Services (SSAS): SSAS could be a multi-dimensional OLAP server just as an analytics engine that permits you to cut up huge volumes of information. It is a part of Microsoft SQL Server and performs analysis utilizing different measurements. SSAS is employed to make cubes using the data from data warehouse for deeper and faster data analysis.

The important terminology in SSAS are:

Data Source:

Data Source is a sort of connection string and it sets up an association between the analysis data base.

Data Source View:

Data source view is a consistent model of Database.

Cube:

A cube is a basic unit of storage. It is an assortment of information which has been aggregated to permit queries to return information rapidly.

Dimension Table:

A dimension table contains dimensions of a fact. They are joined to fact table utilizing a foreign key.

Measure:

For every fact table there ought to be at least one table which should be analysed.

Power BI:

It is information perception and Business insight that converts data from different data sources to BI reports. Power BI suite gives numerous software, connector, and services - Power BI desktop, Power BI service dependent on Saas, and portable Power BI applications accessible for various platforms. These arrangement of services are utilized by business clients to consume information and manufacture BI reports.

IV. PROPOSED METHOD

Business intelligence (BI) is the arrangement of procedures and tools for the change of raw information into important and valuable data for business analysis purposes. Microsoft gives a few tools to change your business information into data. We can utilize these tools with the interface of Visual Studio.

- **SSIS:** It is a combination tool and represents SQL Server Integration Services.
- **SSAS:** It is an analysis tool and represents SQL Server Analysis Services.
- **SSRS:** It is a reporting tool and represents SQL Server Reporting Services.

A. SQL Server Integration Service (SSIS):

SSIS is a data warehousing tool and is one of the services provided by the Microsoft. The JSON Source component of SSIS will extract the information from HTTP URL or any file that is locally saved in your system and can be uploaded manually which provides the output in the form of columns and rows. The component Conditional Split is utilized to check for a specific condition and then transfers the data to required destination based on given condition. The data conversion component can translate one data type to another and it gives the new output column. Now the packages for the required columns can be created and

prevent the duplicate values entering into the fields, finally the output of each package is to be stored in particular destination component. The Look Up component here is generally used to fetch the values of id in the place of string values. You can also use sql queries using OLEDB command components to procure the desired results. Construct Sql Schema in order to provide the relationships between the fact and dimension tables. All the above batch of rules and functions are implemented on raw data to convert it into a standard format. By executing all the packages of the project the data is loaded into the SQL server. In the SQL Server Management Studio (SSMS) by making use of SQL server Agent you can automatically process the project for a specific time by creating a job in it. It incorporates graphical tools and wizards for establishing and troubleshooting the packages. SSIS ETL tool assists with building enterprise level information change and information integration solutions. This SSIS ETL tool is helpful in loading information from transactional information into Data Warehouse. So you can make reports utilizing that information with the assistance of Power BI. SSIS additionally incorporates graphical tools, wizards for performing tasks like sending email messages and information sources.

B. SQL Server Analysis Service (SSAS):

In SSAS, another Data Source is made dependent on existing or new connection and then new Data Source View is created by selecting all available objects into included objects. Add all the dimension tables to obtain dimensions from sql server and process them individually and also create a new time hierarchy. Cube is formed by using existing tables. A new KPI is created such that Air Quality Index range should be displayed by the pollution value. Process the cube and visit the browser for getting the results and prepare the required queries. Now, open SSMS and connect to Analysis Services Engine. We found all the projects processed in the Analysis Services Project. Now, the SSAS project is loaded into SSMS. Run the whole project in SSMS.

C. Power BI:

Microsoft Power BI is used to intuitive BI reports and dashboards. The connection in power BI is made live so that all the live information from analysis services is associated consequently. Select the required project, so that all the data from SSAS Project will be loaded into power bi. From Visualizations, select table chart and select required items for showing the top 10 cities with highest population. Similarly, select another table chart and select required objects for displaying top 10 cities with lowest population. Select Maps from Visualizations and select required objects in order to display all the cities in the map. Apply all the required filters in order to view the map and understand easily. Similarly display the top 10 cities with highest health risks and lowest health risks.



V. RESULTS

A. SSIS Package:

The Figure2 shows one of the SSIS package which loads the data through JSON Source tool. All the errors and null values have been sent to one table and the structured data have been sent to another tables.

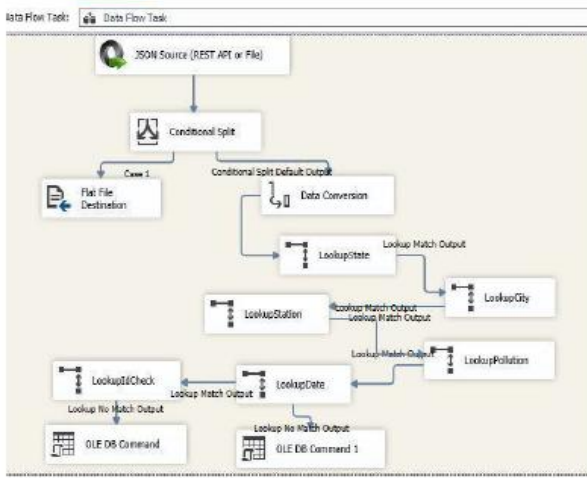


Fig 2. SSIS Package

B. SQL Schema:

The Figure3 shows the SQL Schema which shows all the primary key relations from Fact table to all Dimension tables.

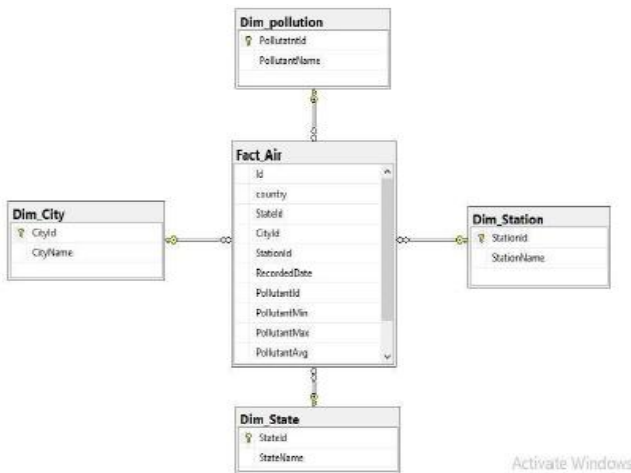


Fig 3. SQL Schema

C. SSAS Cube:

The Figure4 shows the SSAS Cube which shows the Connections between the Fact table and the dimension tables.

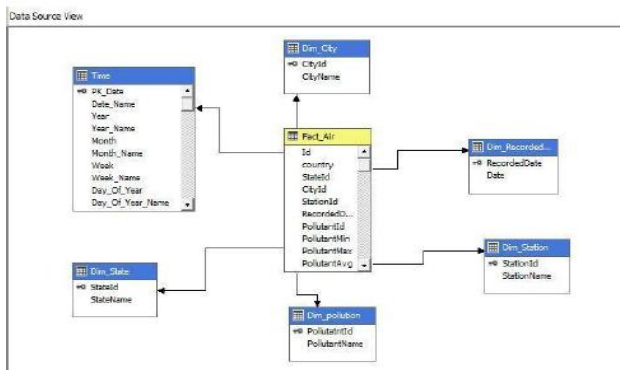


Fig 4. SSAS Cube

D. Top 10 Cities with Highest Health Risk:

The Figure5 shows the topmost 10 cities with highest health risks plotted as Donut Chart in Power Bi. This is done by checking the maximum pollutant and the KPI values. The city with Maximum Pollution value and Low KPI has the highest health risks.

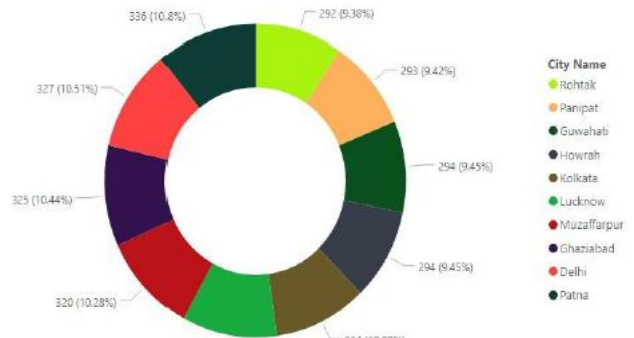


Fig 5. Top 10 Cities with Highest Health Risk

E. Top 10 Cities with Lowest Health Risk:

The Figure6 shows the top 10 cities with lowest health risks plotted as Donut Chart in Power Bi. This is done by checking the maximum pollutant and the KPI values. The city with Minimum Pollution value and Highest KPI has the lowest health risks.

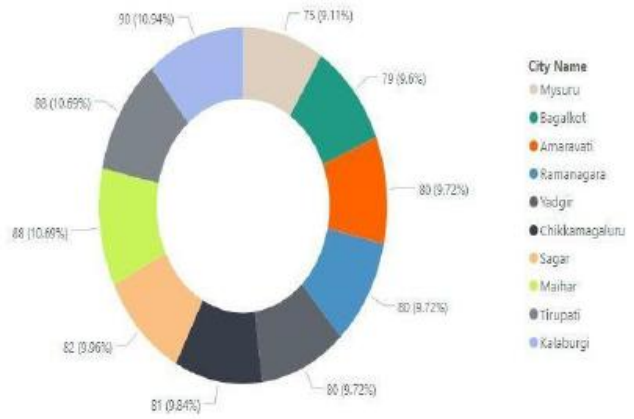


Fig 6. Top 10 cities with Lowest Health Risk

F. Map Visualization:

The Figure7 shows the Map Visualization of all the Stations, Maximum Pollutant Avg, KPI and Maximum Recorded Date. The bubble colors for all stations in the map are based on the maximum pollution value. The pollution value ranges from 0 to 100 with KPI good and satisfactory have a “light pink” color. And the pollution values ranges from 101 to 200 with KPI “Moderately Polluted have light “blue” color. And the pollution values ranges from 201 to 300 with KPI Poor have “green” color. And the pollution values ranges from 301 to 400 with KPI Very Poor have “Purple” color. And the pollution values ranges from 401 to 500 with KPI Severe have “Red” color.

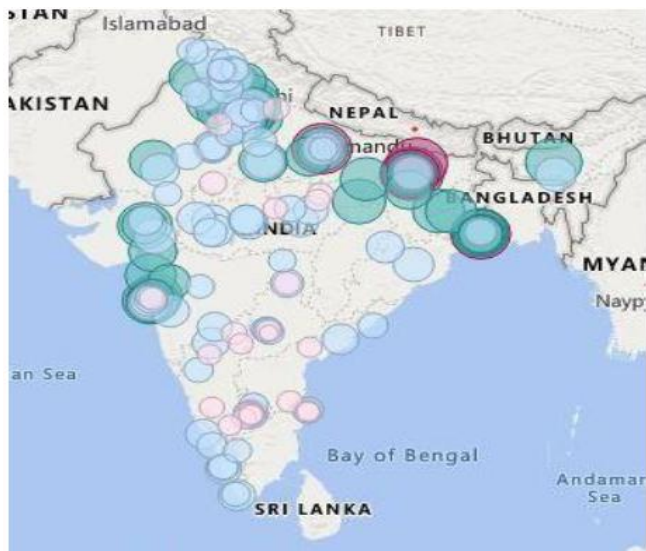


Fig 7. Map Visualization

VI. CONCLUSION

Previously there are several methodologies for collecting Real-time Air Quality Index from various locations in India for analyzing the data and generating reports but in this paper the work has been implemented using Microsoft Business Intelligence tools, which have easy data exploration and efficient visualization techniques. SSIS gives a quick and adaptable information utilized for information extraction, cleaning, gathering, and combining data. The cubes have been created for the faster analysis of data. The live data is connected to Power BI generated from SSAS which in turn generate reports showing the highest health risk cities and a Map Visualisation of overall states stating its condition of air quality. All the figures above describes how the data is processed and results acquired. Power BI tool, which has high level security in publishing reports compared to other technologies and tools.

FUTURE SCOPE

Until now, much is being done to control, monitor and rectify damage done by pollutants. In the future the historical data will be collected and the AQI across multiple years will be compared. The problems are diverse and a few are only being recognized but it's important to keep a close control over pollutants so that we can maintain the environment in an acceptable condition for future generations.

REFERENCES

1. Nikila Varshini, Sreeha.MR, Lhavanya Roobini. VN, Vijayarangam.J, Sujithra.M Coimbatore Institute of Technology, "Analysis of Air Quality Index"
2. Kanchan, Amit Kumar Gorai, and Pramila Goyal, "A Review on Air Quality Indexing System", Asian Journal of Atmospheric Environment, Vol. 92, pp. 101-113, June 2015
3. Bezuglaya, E.Y., Shchutskaya, A.B., Smirnova, "I.V. Air Pollution Index and Interpretation of Measurements of Toxic Pollutant Concentrations. Atmospheric Environment" Vol. 27, pp. 773-779
4. Martand Pratap Singh, R. Dileep Kumar, "A Visualization Approach of Air Quality Index using R Review Paper", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 7 Issue V, ISSN: 2321-9653 May 2019

5. Biswanath Bishoi, Amit Prakash, V.K. Jain, Bishoi et al., "A Comparative Study of Air Quality Index Based on Factor Analysis and US-EPA Methods for an Urban Environment", Aerosol and Air Quality Research, Vol. 9, No. 1, pp. 1-17, 2009
6. Amit Kumar Gorai, Pramila Goyal, "A Review on Air Quality Indexing System", Asian Journal of Atmospheric Environment, 9(2):101-113 • June 2015,
7. Dalila Taieb and Ammar Ben Brahim, Int. J. Renewable Energy Technology, "Methodology for developing an air quality index (AQI) for Tunisia", Vol. 4, No. 1, 2013,
8. Gufran Beig Sachin D. Ghude and Aparna Deshpande, "Scientific Evaluation of Air Quality Standards and Defining Air Quality Index for India Indian Institute of Tropical Meteorology, Pune, India Ministry of Earth Sciences, Govt. of India", August 2010
9. Yuzhe Yang, Zijie Zheng, Student Member, IEEE, Kaigui Bian, Member, IEEE, Lingyang Song, Senior Member, "Realtime Profiling of Fine- Grained Air Quality Index Distribution using UAV Sensing", IEEE, 8 Nov 2017
10. Gowtham. Sarella, Mrs. Dr. Anjali. K. Khambete, "AMbient Air Quality Analysis using Air Quality Index – A Case Study of Vapi", IJRST – International Journal for Innovative Research in Science & Technology| Volume 1 | Issue 10 | March 2015,
11. Sef van den Elshout, Hans Bartelds, Hermann Heich, Karine Leger, citeair_ji_caqi_update.doc, May 23, 2012
12. US Environmental Protection Agency (USEPA) (2008), "Air quality index: a guide to air quality and your health. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park"

AUTHORS PROFILE



Chenikala Sravani studying IV/IV B.tech in Information Technology in Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, affiliated to JNTUK.



Appasani Geetha pursuing IV/IV B.tech in Information Technology in Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, affiliated to JNTUK.



Pasumarthi Sai Ramya pursuing IV/IV B.tech in Information Technology in Velagapudi Ramakrishna Siddhartha Engineering College,



Vijayawada, affiliated to JNTUK.
M.Ramesh, Assistant Professor, Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College.