TECH TRENDS

WE EXPLORE WE EXHIBIT

MALLA REDDY
COLLEGE OF ENGINEERING

MALLA REDDY
COLLEGE OF ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VOLUME -05 ISSUE NO - 1

EAMCET CODE : MRCE

Permanently affiliated to JNTUH | Approved by AICTE | Accredited by NBA & NAAC

Institution Vision & Mission

Vision

- To emerge as a Centre of Excellence for producing professionals who shall be the leaders in technology innovation, entrepreneurship, management and in turn contribute for advancement of society and human kind.

Mission

- **M1 :** To provide an environment of learning in emerging technologies.
- **M2 :** To nurture a state of art teaching learning process and R&D culture.
- **M3 :** To foster networking with Alumni, Industry, Institutes of repute and other stakeholders for effective interaction.
- **M4 :** To practice and promote high standards of ethical values through societal commitment.

Department of Computer Science and Engineering

Department Vision & Mission

Vision

- To impart futuristic knowledge in Computer Science and to produce highly skilled, imaginative and socially mindful experts who can contribute to industry and architect research fit for working in worldwide condition.

Mission

- To promote strong academic growth by providing fundamental domain knowledge and offering state of art technology for having an excellence in research & development.
- To create an environment for learning analytical skills, advanced programming languages using modern tools and to equip for higher studies.
- To undertake collaborative projects for understanding need of team work in real time environment and to improve communication and inter personnel skills for better employability.
- To promote high standards of ethical values through societal commitment.

Computer Science & Engineering PO's

Engineering Graduates will be able to:

- **PO.1.Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
- **PO.2.Problem analysis:** Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
- **PO.3.Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
- **PO.4.Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
- **PO.5.Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
- **PO.6.The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
- **PO.7.Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
- **PO.8.Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
- **PO.9.Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
- **PO.10.Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

- **PO.11.Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
- **PO.12.Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Computer Science & Engineering PEO's

- **PEO1 –** To make the students understand and implement the engineering concepts in multiple domains.
- **PEO2 –** To provide knowledge based services so as to meet the needs of the society and industry by usage of modern tools.
- **PEO3 –** To understand engineering processes for design and development of software components and products efficiently for improving employability.
- **PEO4 –** To educate students in disseminating the research findings to create interest for higher studies.
- **PEO5 –** To inculcate knowledge with due consideration for ethical and economic issues.

Computer Science & Engineering PSO's

- **PSO1:** Professional Skills: The ability to understand, analyze and develop computer programs in the areas related to algorithms and System Software.
- **PSO2:** Problem Solving Skills: The ability to apply standard practices and strategies in software project development to deliver a quality and defect free product.
- **PSO3:** Employability Skills: The ability to employ modern computer languages and technologies, so as to be industry ready and for better employability and research.

SENTIMENT ANALYSIS USING BERT NEURAL NETWORK

ABSTRACT

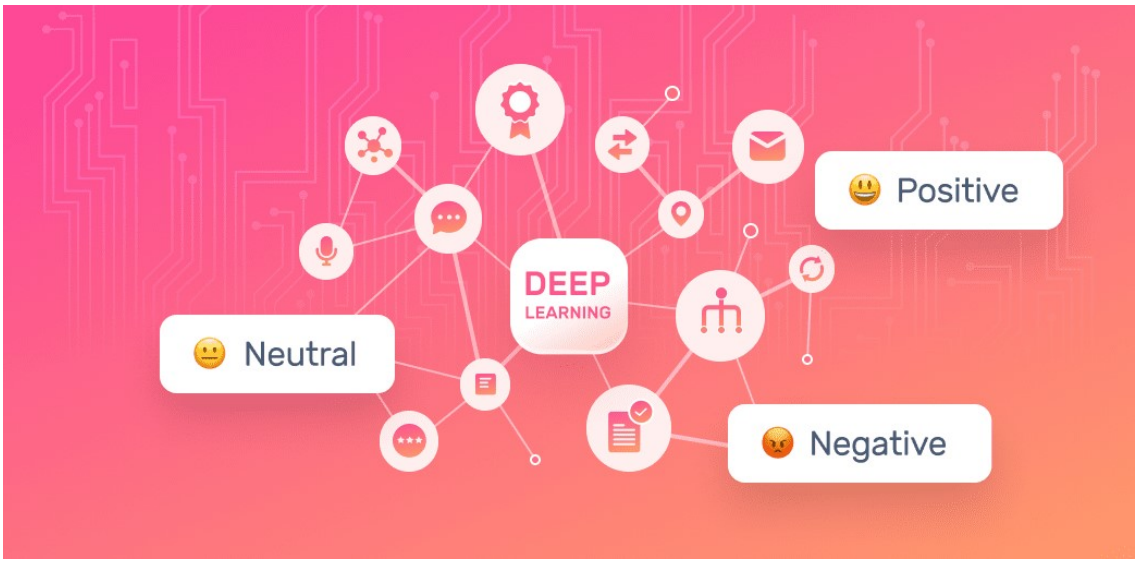
We introduce a new language representation model called BERT, which stands for Bidirectional Encoder Representations from Transformers. Unlike recent language representation models (Peters et al., 2018a; Radford et al., 2018), BERT is designed to pre train deep bidirectional representations fromunlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre- trained BERT model can be fine tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task specific architecture modifications. In this paper, we improve the fine-tuning based approaches by proposing BERT: Bidirectional Encoder Representations from Transformers.

BERT alleviates the previously mentioned unidirectionality constraint by using a "masked language model" (MLM) pre-training objective, inspired by the Cloze task (Taylor, 1953). The masked language model randomly masks some of the tokens from the input, and the objective is to predict the original vocabulary id of the masked word based only on its context. Unlike left-to-right language model pre training, the MLM objective enables the representation to fuse the left and the right context, which allows us to pre train a deep bidirectional Transformer. In addition to the masked language model, we also use a "next sentence prediction" task that jointly pretrains text-pair representations.

INTRODUCTION

Sentiment analysis has emerged as a vital tool for understanding public opinion and gauging sentiment within textual data. In the contemporary era of data-driven decision-making, sentiment analysis holds significant importance across various domains, including marketing, customer feedback analysis, and social media monitoring. The ability to accurately discern sentiment from text data can provide valuable insights for businesses and organizations. In this context, the utilization of advanced deep learning techniques, particularly BERT (Bidirectional Encoder Representations from Transformers) neural network, presents a compelling approach for sentiment analysis. BERT, renowned for its ability to capture contextual information and effectively process bidirectional sequences, has shown remarkable performance in natural language processing tasks. By leveraging the power of BERT, this project endeavours to delve into the domain of sentiment analysis and demonstrate the efficacy of BERT in accurately deciphering sentiment within textual data. The primary objective of this documentation is to outline the methodology, implementation, and outcomes of deploying the BERT neural network for sentiment analysis. By elucidating the process of training and fine-tuning the BERT model, along with the subsequent analysis of results, this documentation aims to provide a comprehensive insight into the potential and applicability of BERT in sentiment analysis tasks.

In this context, the utilization of advanced deep learning techniques, particularly BERT (Bidirectional Encoder Representations from Transformers) neural network, presents a compelling approach for sentiment analysis. BERT, renowned for its ability to capture contextual information and effectively process bidirectional sequences, has shown remarkable performance in natural language processing tasks. By leveraging the power of BERT, this project endeavours to delve into the domain of sentiment analysis and demonstrate the efficacy of BERT in accurately deciphering sentiment within textual data. The primary objective of this documentation is to outline the methodology, implementation, and outcomes of deploying the BERT neural network for sentiment analysis. By elucidating the process of training and fine-tuning the BERT model, along with the subsequent analysis of results, this documentation aims to provide a comprehensive insight into the potential and applicability of BERT in sentiment analysis tasks.



EXISTING SYSTEMS

Historically, sentiment analysis has relied on traditional lexicon-based methods and statistical algorithms to interpret the sentiment of textual data. Lexicon-based techniques involved the usage of sentiment lexicons and dictionaries to assign sentiment scores to words within the text. However, these approaches often encountered difficulties in handling contextual nuances and keeping pace with evolving language usage trends. Early sentiment analysis systems leveraged machine learning algorithms such as Naive Bayes, Support Vector Machines (SVM), and Logistic Regression. These algorithms relied on features extracted from text data, including n-grams and term frequency-inverse document frequency (TF-IDF), to classify sentiment. While effective to a certain extent, statistical approaches had limitations in capturing complex linguistic patterns and discerning intricate contextual dependencies within the text**.**

With the emergence of deep learning, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) gained prominence for sentiment analysis tasks. These neural network architectures were instrumental in capturing sequential data and inferring contextual dependencies within text, marking a significant leap forward in sentiment analysis capabilities.

1. **Naive Bayes:** A probabilistic approach that calculates the likelihood of words occurring in positive or negative contexts to classify text.

2. **Support Vector Machines (SVM):** Leveraging hyperplane separation in high dimensional spaces to divide text into sentiment categories.

3. **Decision Trees:** A decision-making approach that categorizes text based on a series of rules.

4. **CNN's:** Convolutional Neural Networks (CNNs) analyse local and global features in text to capture patterns relevant to sentiment.

5. **RNN's:** Recurrent Neural Networks (RNNs) excel in understanding sequential data, making them suitable for capturing contextual dependencies in sentiment analysis**.**

DRAWBACKS

1. **Limited Contextual Understanding:** Traditional lexicon-based methods and statistical algorithms often struggle to capture the nuanced contextual understanding required for accurate sentiment analysis, leading to potential misinterpretation of sentiment in complex texts.

2. **Difficulty in Handling Sarcasm and Irony:** Existing systems may struggle to discern sentiment in instances of sarcasm and irony, as these rely heavily on contextual and cultural understanding, posing challenges for accurate sentiment classification.

3. **Feature Engineering:** Traditional methods require manual feature engineering, where experts select relevant features for sentiment classification. This approach is

timeconsuming, subjective, and may not capture all the nuances of language that impact sentiment.

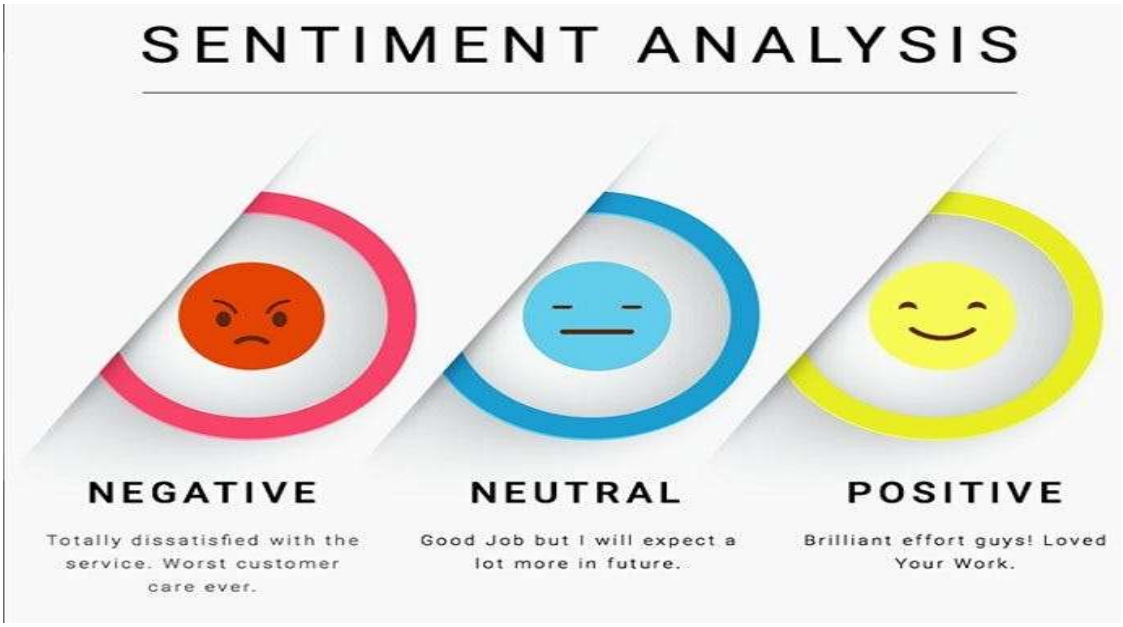
4. **Computational Complexity:** Some approaches, particularly those involving deep learning models, may pose computational challenges due to high computational requirements for training and inference, impacting practical scalability and efficiency.

5. **Computational Complexity:** Some approaches, particularly those involving deep learning models, may pose computational challenges due to high computational requirements for training and inference, impacting practical scalability and efficiency.

PROPOSED SYSTEMS

In our project, we introduce a revolutionary shift in sentiment analysis by leveraging the power of BERT. By integrating BERT into sentiment analysis, our project aims to usher in a new era of precision and context-aware sentiment classification, enhancing decision support and insights for individuals and businesses. The landscape of sentiment analysis underwent a paradigm shift with the introduction of BERT (Bidirectional Encoder Representations from Transformers) and transformer-based architectures. BERT, renowned for its bidirectional processing and attention mechanism, has showcased remarkable performance in capturing intricate contextual information and discerning sentiment within textual data. This revolutionized the sentiment analysis landscape, opening doors to more accurate and nuanced sentiment interpretation.The utilization of BERT and transformer-based models for transfer learning has empowered sentiment analysis in tailored domains. By fine-tuning pre-trained models on domain-specific datasets, sentiment analysis systems can now capture and comprehend sentiment nuances distinct to various industries and sectors, leading to more accurate and domain-relevant sentiment insights.

In contemporary sentiment analysis, BERT and its variants, alongside transformer-based models such as Generative Pre-trained Transformer (GPT), have become pivotal in sentiment analysis tasks. Additionally, the prevalence of transfer learning techniques, where pre-trained models are fine-tuned on domain-specific data, has substantially advanced the state-of-theart in sentiment analysis. The integration of BERT and transformer-based models with transfer learning strategies has revolutionized sentiment analysis methodologies, empowered the accurate interpretation and understood of sentiment within textual data.



Advantages

1. **Enhanced Contextual Understanding:** The utilization of BERT facilitates a deeper comprehension of the contextual nuances within textual data, enabling the accurate interpretation of sentiment within diverse and complex language structures.

2. **Improved Accuracy and Consistency:** BERT's bidirectional processing empowers the model to capture intricate dependencies within the text, leading to heightened accuracy and consistency in sentiment analysis compared to traditional methods.

3. **Adaptability to Varied Textual Data:** BERT's pre-trained nature and transfer learning capabilities allow for the seamless adaptation to diverse domains and languages, thereby enhancing the system's versatility in analysing sentiment across a wide spectrum of textual content.

4. **Efficient Handling of Ambiguity:** Through its ability to capture bidirectional contexts, BERT excels in effectively discerning and processing ambiguous language constructs, resulting in more nuanced and precise sentiment analysis outcomes.

5. **Efficient Handling of Ambiguity:** Through its ability to capture bidirectional contexts, BERT excels in effectively discerning and processing ambiguous language constructs, resulting in more nuanced and precise sentiment analysis outcomes.

Algorithm

The BERT (Bidirectional Encoder Representations from Transformers) model is being used as part of the algorithm for sentiment analysis in the project. BERT is a powerful natural language processing model developed by Google, capable of understanding context and meaning in language. It has been widely adopted for various NLP tasks, including sentiment analysis, due to its ability to capture complex language patterns.

This algorithm includes steps for importing packages, loading datasets, and preparing the data for model training.

Algorithm for Sentiment Analysis using BERT:

Step 1: Importing Packages

- Import required packages such as pandas, numpy, matplotlib, seaborn, transformers, torch, and sklearn.

Step 2: Set Key Variables

- Define key variables including random seed, batch size, and the device (GPU or CPU) for running the model.

Step 3: Load Dataset

- Read the dataset for sentiment analysis, for example, a CSV file containing movie reviews or user feedback.

Step 4: Data Preprocessing

- Prepare the data for training, validation, and testing sets.
- Tokenize the text data using the BERT tokenizer and apply padding to ensure consistent input sequence length.

Step 5: Create Data Loaders

- Define functions to create data loaders for the training, validation, and testing sets with tokenized and padded input.

Step 6: Define BERT Model

- Instantiate the BERT model and tokenizer using the transformers library.

Step 7: Model Training

- Define the training loop, including loss function, optimizer, and the number of training epochs.

Step 8: Model Evaluation

- Evaluate the trained model on the validation and testing data using appropriate metrics such as accuracy, precision, recall, and F1 score.

Step 9: Deployment

- Deploy the trained sentiment analysis model for inference on new text data.

This provides an outline of the algorithm flow for integrating the BERT model into sentiment analysis.

FUTURE ENHANCEMENT

1. **Multi-lingual Sentiment Analysis**: Extending the BERT-based sentiment analysis model to support multiple languages, thereby broadening its applicability to diverse linguistic contexts.

2. **Fine-grained Sentiment Classification**: Refining the model to categorize sentiments into more nuanced categories, such as sentiment intensity levels or specific emotions, to provide deeper insights into textual sentiment**.**

3. **Domain-specific Sentiment Analysis:** Adapting the BERT model to specialize in sentiment analysis within specific domains, such as finance, healthcare, entertainment, to cater to industry-specific sentiment interpretation needs**.**

4. **Aspect-based Sentiment Analysis:** Enhancing the model's capability to identify and analyse sentiment towards specific aspects or entities within the text, enabling more targeted sentiment assessment.

5. **Real-time Sentiment Analysis:** Developing a framework for real-time sentiment analysis utilizing BERT, enabling and continuous sentiment monitoring in dynamic textual data sources, such as social media streams and news articles**.** .

PRESENTED BY

J . ESHWAR

P . DIVYASRI

S . PRASAD

K . ANAND

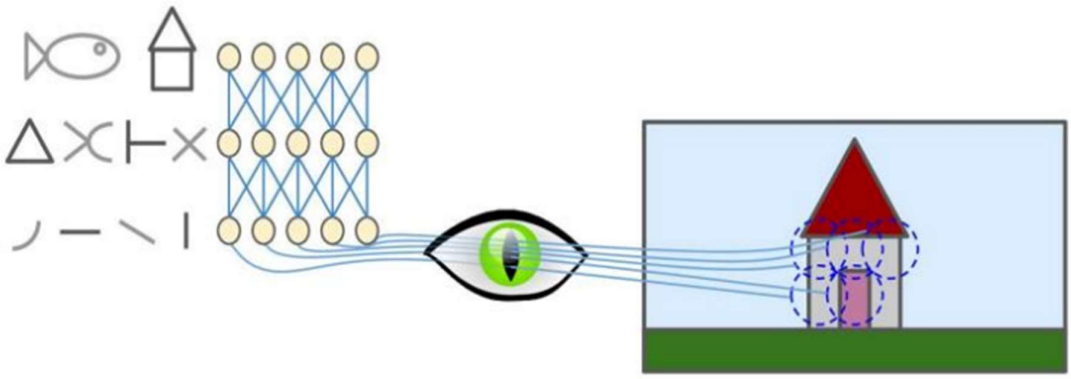
INTELLIGENT TEXT EXPLORATION FOR LANGCHAIN POWERED DOCUMENT ENQUIRY

ABSTRACT

This Project is Query based System represents a sophisticated fusion of cutting-edge Natural anguage processing (NLP) techniques and advanced document retrieval methodologies, tailored to provide users with intuitive access to extensive repositories of textual information. Leveraging state-of-the-art NLP models such as BERT and GPT, the system empowers users to interact effortlessly through natural language queries, eliminating the need for complex search syntax. By intelligently parsing and understanding user queries, the system navigates through a diverse document corpus to retrieve the most relevant information, thereby streamlining the information retrieval process and enhancing user productivity. Furthermore, the Langchain system integrates robust document indexing mechanisms and efficient response generation algorithms, ensuring rapid and accurate retrieval of pertinent documents. Summarization techniques condense retrieved documents into concise summaries or snippets, offering users quick insights without overwhelming them with extensive content. Through continuous refinement driven by user feedback and iterative enhancements, the Langchain Document Querying System stands as a testament to innovation in document retrieval, promising enhanced efficiency, accuracy, and accessibility in navigating it.

INTRODUCTION

Langchain Document Querying System. In an age where access to vast amounts of textual data is a daily reality, our project stands as a beacon of innovation. By seamlessly integrating advanced natural language processing (NLP) capabilities, courtesy of the powerful Langchain model, our system empowers users to interact with a document corpus in a manner that transcends traditional query systems. The primary objective is to enhance the user experience, enabling intuitive and efficient extraction of knowledge from a diverse array of documents. At its core, this project addresses the growing need for a more intelligent and contextaware document querying system. Through meticulous document preprocessing, sophisticated NLP model integration, and an optimized document indexing mechanism, we aim to create a system that not only understands the semantics of user queries but also swiftly retrieves relevant information from the document corpus. As we embark on this journey, our vision is to redefine the way users engage with textual data, providing a robust and user-friendly solution that unlocks the full potential of information retrieval in the digital age. Additionally, the system involves user query processing through tokenization and preprocessing, sophisticated query matching algorithms that consider both NLP insights and document indexing, and a response generation system for delivering coherent and user-friendly answers based on matched documents. Optional features include the exploration of a user interface for enhanced interaction and iterative optimization based on user feedback and system evaluation metrics, ensuring the Langchain Document Querying System evolves into a powerful and user-centric information retrieval solution



EXISTING SYSTEM

Existing systems in the domain of document querying encompass a diverse range of solutions, each tailored to address specific needs and challenges. One prominent example is Elasticsearch, an open-source search engine widely used for full-text search and document retrieval. Elasticsearch offers powerful indexing capabilities, enabling users to store, search, and analyze large volumes of structured and unstructured data efficiently. With its support for complex queries, relevance scoring, and distributed architecture, Elasticsearch serves as a versatile solution for document querying across various industries and applications. Another notable system is Apache Solr, a fast and scalable search platform built on top of Apache Lucene. Similar to Elasticsearch, Solr provides advanced search features, including faceted search, geospatial search, and real-time indexing, making it suitable for diverse use cases such as e-commerce, content management, and enterprise search. Additionally, Solr offers robust administration and monitoring tools, allowing users to manage and optimize their search infrastructure effectively. These existing systems serve as valuable benchmarks and sources of inspiration for the Langchain Document Querying System, offering insights into best practices and proven methodologies for building efficient and scalable document retrieval solutions.

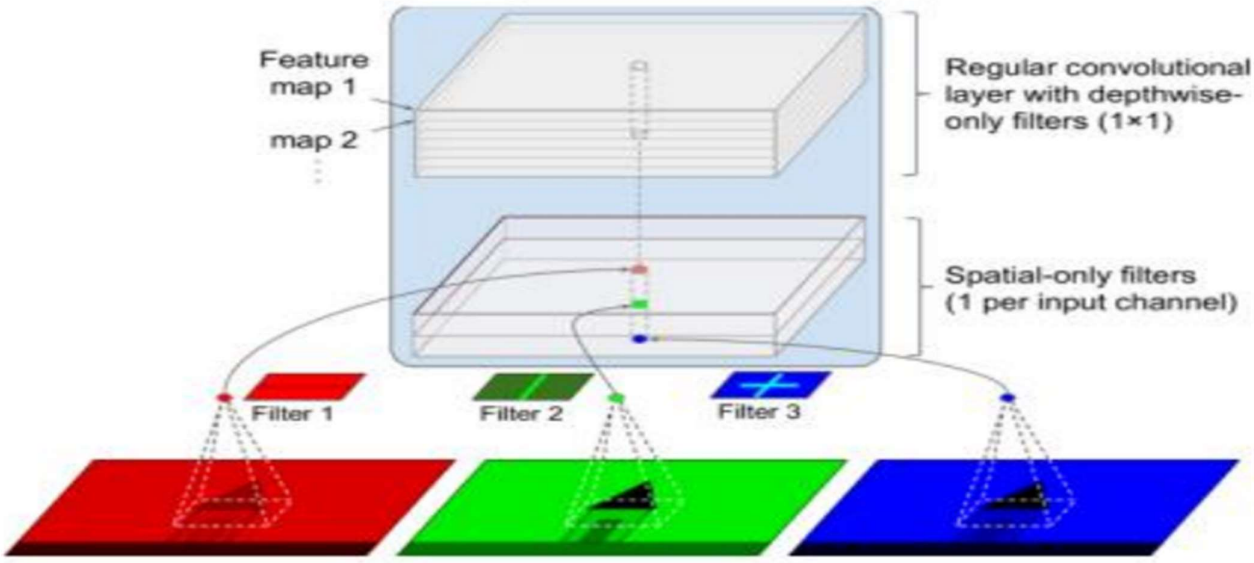
PROPOSED SYSTEM

The Langchain Document Querying System is a revolutionary approach to document querying that combines the most recent developments in document retrieval and natural language processing (NLP) methods into a unified and user-friendly system. Through the utilization of cutting-edge natural language processing (NLP) models such as BERT and GPT, the system goes beyond conventional keyword-based search techniques, allowing users to communicate with the platform through natural language questions. This user-friendly method improves user happiness by making search results more relevant and enhancing accessibility for users. It also makes it easier to comprehend user intent more precisely and nuancedly. Moreover, the system's incorporation with cloud-based services, specifically OpenAI, grants it access to state-of-the-art language models and embeddings, augmenting its comprehension of language and guaranteeing flexibility in response to changing language. The Langchain system has strong language understanding skills in addition to effective document retrieval mechanisms made

possible by sophisticated indexing and vectorization algorithms. The system generates efficient representations of the document corpus by utilizing advanced document indexing algorithms and cloud-based infrastructure. This allows for quick and pertinent document retrieval in response to user queries. In addition to speeding up the search process, this smooth integration of NLP and document retrieval technologies increases the system's scalability and efficiency, establishing it as a trailblazing solution in the field of information retrieval systems. The Langchain Document Querying System is set to revolutionize document querying by combining state-of-the-art technologies with user-centric design concepts. It provides users with an easy-to-use interface for retrieving and exploring large libraries. The Langchain Document Querying System is conceived with the purpose of revolutionizing the landscape of information retrieval. At its core, the project aims to empower users with an intelligent and user-friendly tool that seamlessly navigates through extensive document corpora. By integrating advanced natural language processing capabilities through the Langchain model, the system enables users to pose queries in natural language, bridging the gap between human interaction and machine understanding. The purpose extends beyond traditional document querying, aspiring to deliver an intuitive experience where users can effortlessly extract relevant information from a diverse range of documents. This project seeks to enhance efficiency, accessibility, and user satisfaction in the process of information retrieval, catering to the evolving needs of users grappling with an ever-expanding sea of digital data. 5 | P a g e
The Langchain Document Querying System sets out to achieve several interrelated goals and objectives. Firstly, it aims to create an intelligent and efficient information retrieval platform by seamlessly integrating advanced natural language processing capabilities through the Langchain model. The system seeks to enhance user interaction with a diverse document corpus, allowing users to pose queries in natural language and receive coherent and relevant responses. Additionally, the project endeavors to optimize query speed and accuracy through the implementation of a robust document indexing mechanism, ensuring swift retrieval of pertinent information. Furthermore, user experience is a focal point, with the potential inclusion of a user interface for intuitive interaction. Through iterative optimization and feedback mechanisms, the project aims to deliver a cutting-edge document querying system that not only understands the nuances of user queries but also transforms the way users access and extract knowledge from extensive textual data.

SCOPE OF THE PROJECT

The scope of the Langchain Document Querying System encompasses a comprehensive range of functionalities aimed at providing users with seamless access to vast repositories of textual information. Primarily, the project focuses on developing advanced natural language processing (NLP) algorithms capable of understanding and interpreting user queries in everyday language. These algorithms will facilitate the extraction of key concepts and



keywords from user queries, enabling the system to retrieve relevant documents from a diverse corpus. Moreover, the project aims to integrate efficient document indexing techniques to optimize the retrieval process, ensuring rapid and accurate responses to user queries. Additionally, the system will incorporate advanced response generation algorithms to present retrieved documents in a user-friendly manner, including summarization techniques to provide concise insights. The scope also extends to the implementation of user feedback mechanisms to continuously refine and improve the system's performance. Overall, the Langchain Document Querying System seeks to revolutionize the way users interact with textual information, offering a robust and intuitive platform for efficient document retrieval and analysis

PRESENTED BY:

P SAI KUMAR

M. SRIRAM

P.VILASINI

Traffic Sign Recognition using Deep Learning

ABSTRACT

To ensure a smooth and secure flow of traffic, road signs are essential. A major cause of road accidents is negligence in viewing the Traffic signboards and interpreting them incorrectly. The proposed system helps in recognizing the Traffic sign and sending a voice alert through the speaker to the driver so that he/ she may take necessary decisions. The proposed system is trained using Convolutional Neural Network (CNN) which helps in traffic sign image recognition and classification. A set of classes are defined and trained on a particular dataset to make it more accurate.

In this Project, the German Traffic Sign Benchmarks Dataset was used, which contains approximately 43 categories and 51,900 images of traffic signs. The accuracy of the execution is about 98.52 percent. Following the detection of the sign by the system, a voice alert is sent through the speaker which notifies the driver. The proposed system also contains a section where the vehicle driver is alerted about the traffic signs in the near proximity which helps them to be aware of what rules to follow on the route. The aim of this system is to ensure the safety of the vehicle's driver, passengers, and pedestrians.

INTRODUCTION

Human factor remains the most common cause of road mortality. Indeed, the potentially dangerous choices made by the driver might be intentional (speed driving, for example) as they might be the result of physical tiredness, drowsiness or a poor perception and interpretation of seen scenes. The introduction of autonomous vehicles will certainly reduce these causes or even make them disappear.

As part of the development of these autonomous vehicles, particularly driving assistance systems, several manufacturers and laboratories have oriented their works towards the exploitation of visual information because of its usefulness for the detection of road, vehicles, pedestrians and traffic signs. The principle of driving assistance systems aiming at road signs recognition is to detect signs, interpret their meaning, then transmit the information to the driver (by a projection on a windshield, a screen or a smartphone) or even better, transmit the information to the vehicle that carries out the execution without needing a human decision. However, given that the classical approach has been bounded by well-structured models of traffic signs (undistorted and completely visible models) only, it became necessary to consider real characteristics of the road environment. For this, the current researches are moving towards the development of recognition systems which are more adapted to real images of road signs that do not generally look like their models.

Motivated by the success of classification and recognition methods, in different domains, based on Deep learning, we are interested in the use of these new advances in Machine learning for traffic signs recognition.

EXISTING SYSTEM

- Burgoon et al. used 16 linguistic features categorized in four classes, which achieved an accuracy of 60.72% using a DT algorithm with 15-fold cross-validation.
- Vicario et al. used different features like text (e.g., number of characters, words, sentences, question marks, and negations), user-specific, and message specific (e.g., number of replies, likes) to identify hoaxes and fake news on social media using linear

regression, logistic regression, support vector machine (SVM), K-nearest neighbor (KNN), and NNs. The validation on an Italian Facebook data set with new features achieved an accuracy of 91% on the linear regression classification algorithm.

- Pérez-Rosas et al. used major linguistic features (e.g., n-grams, punctuation, psycho-linguistic, readability, and syntax) and achieved an accuracy of 76% on two novel data sets covering seven domains.

DRAWBACKS

- The fast scattering of data at a high rate with minimal effort enables the widespread of false information, such as fake news, which are harmful to society and people.
- One major challenge that is associated with OSNs is verification of messages exchanged as well as the authenticity of users. Some messages that are spread through these social networks may create horrible situations regarding peace and harmony in society. Such messages, currently coined as fake news, can also be life-threatening.
- calculations are inefficient to run on top.
- Larger the vocabulary size, larger the matrix size (not scalable to large vocabulary).
- Not all word associations can be understood using this technique.

PROPOSED SYSTEM

- In our proposed system, we develop the Traffic Sign Board Recognition and Voice Alert System using Convolutional Neural Network. Our system will able to detect, recognize and infer the road traffic signs would be a prodigious help to the driver.
- The objective of an automatic road signs recognition system is to detect and classify one or more road signs from within live color images.
- In this base paper we provide alertness to the driver about the sign using voice of the detected sign board. The system provides the driver with real time information from road signs, which consist the most important and challenging tasks. Next generate an acoustic

warning to the driver in advance of any danger. This warning then allows the driver to take appropriate corrective decisions in order to mitigate or completely avoid the event.

ADVANTAGES

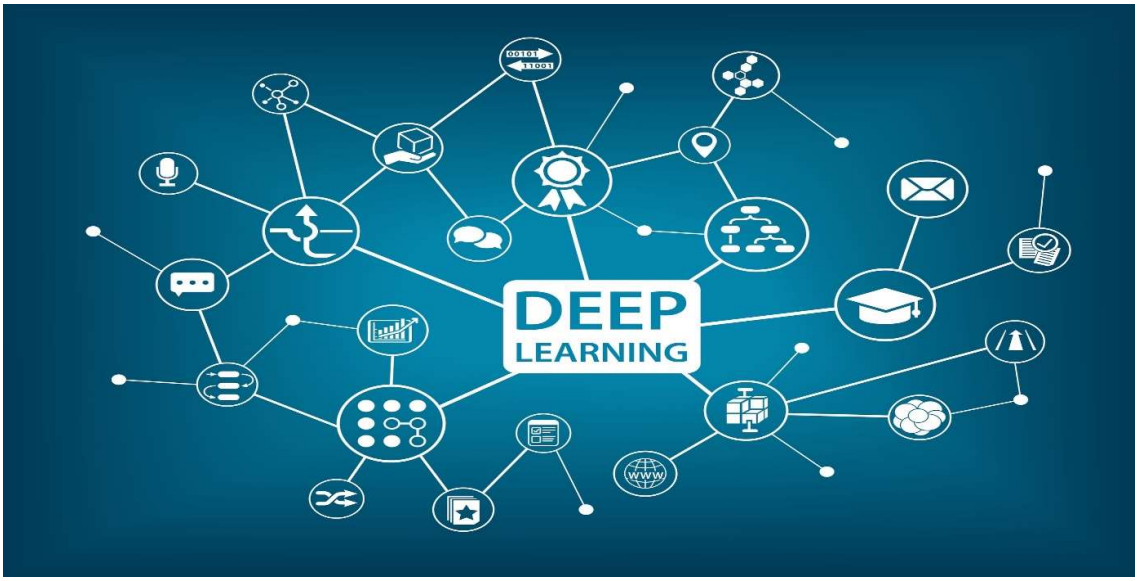
- The accuracy of the proposed system is 97% and this model turned out to give the best accuracy as compared to the other models that we analysed in the existing.
- If a certain image is not containing a traffic sign, then the user gets a prompt of “No Sign Detected” is also implemented in our model.
- Despite some missed detections, the detector still preforms extremely well even for several difficult cases.
- A good performance is also presented, where all traffic-sign detections are displayed for a couple of full-resolution images.
- The system provides an efficient deep network for learning a large number of categories with efficient and fast detection.

IMPLEMENTATION

DEEP LEARNING

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the back propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

Machine-learning technology powers many aspects of modern society: from web searches to content filtering on social networks to recommendations on e-commerce websites, and it is increasingly present in consumer products such as cameras and smartphones. Machine-learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant results of search.



Increasingly, these applications make use of a class of techniques called deep learning. Conventional machine- learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input. Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification.

PYTHON

Python is an object-oriented, high level language, interpreted, dynamic and multipurpose programming language. Python is easy to learn yet powerful and versatile scripting language which makes it attractive for Application Development. Python's syntax and dynamic typing with its interpreted nature, make it an ideal language for scripting and rapid application development in many areas.

Python supports multiple programming pattern, including object oriented programming, imperative and functional programming or procedural styles. Python is not intended to work on special area such as web programming. That is why it is known as multipurpose because it can be used with web, enterprise, 3D CAD etc. We don't need to use data types to declare variable because it is dynamically typed so we can write a=10 to declare an integer value in a variable. Python makes the development and debugging fast because there is no compilation step included in python development and edit-test-debug cycle is very fast.

FUTURE ENHANCEMENT

The prototype can be expanded to include an inbuilt alert system with a camera in the vehicle's center. Also, the feature of getting the estimated time for reaching that particular traffic sign can be added. This system can also be expanded for identification of traffic signals and hence prompt the user about the time to reach that particular signal and its status as well. The user can accordingly plan their trip start time and hence cross all signals without having to wait. Also the driver verification will be done with the help of an API providing the information about the license holder and the license number.

PRESENTED BY

R. KAVYA SRI
SARITHA DANU
R. SAI SHESHU
J. SATHVIK

An Efficient Key Management and Multi-Layered Security Framework for SCADA Systems

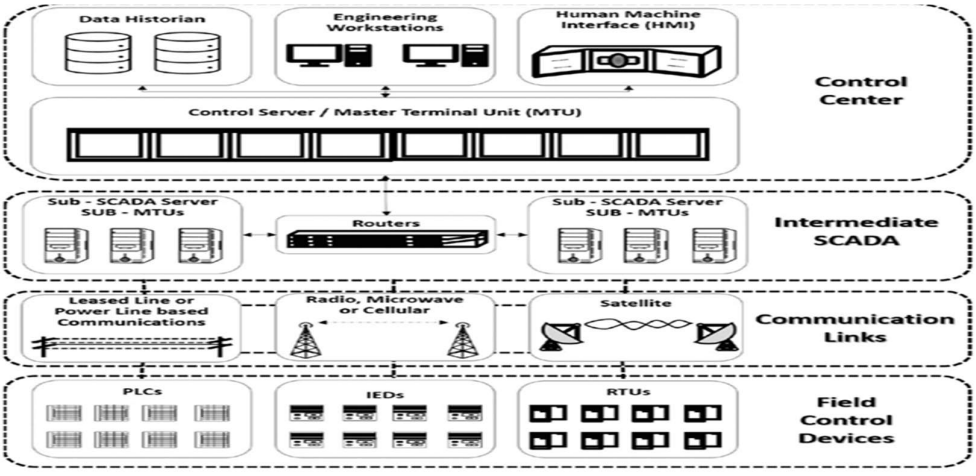
ABSTRACT

Supervisory Control and Data Acquisition (SCADA) networks play a vital role in industrial control systems. Industrial organizations perform operations remotely through SCADA systems to accelerate their processes. However, this enhancement in network capabilities comes at the cost of exposing the systems to cyber-attacks. Consequently, effective solutions are required to secure industrial infrastructure as cyber-attacks on SCADA systems can have severe financial and/or safety implications. Moreover, SCADA field devices are equipped with microcontrollers for processing information and have limited computational power and resources. This makes the deployment of sophisticated security features challenging. As a result, effective lightweight cryptography solutions are needed to strengthen the security of industrial plants against cyber threats. In this paper, we have proposed a multi-layered framework by combining both symmetric and asymmetric key cryptographic techniques to ensure high availability, integrity, confidentiality, authentication and scalability. Further, an efficient session key management mechanism is proposed by merging random number generation with a hashed message authentication code. Moreover, for each session, we have introduced three symmetric key cryptography techniques based on the concept of Vernam cipher and a pre-shared session key, namely, random prime number generator, prime counter, and hash chaining. The proposed scheme satisfies the SCADA requirements of real-time request response mechanism by supporting broadcast, multicast, and point to point communication.

INTRODUCTION

There has been a surge in the deployment of Supervisory Control and Data Acquisition (SCADA) systems to control and monitor industrial infrastructure over the Internet. Organizations such as oil and natural gas, power stations, water & sewage systems, chemical plants, manufacturing units, railway, and other transportation use SCADA systems to monitor and control their infrastructure such as oil pipelines, solar panels, water pipelines, boilers, railway tracks, and plant floor components across open access networks. A SCADA system typically includes a control server (also known as Master Terminal Unit (MTU)), SUB-MTUs, communication links (e.g., satellite, radio or microwave links, cellular network, switched or lease lines and powerlines), and geographically dispersed field control devices, namely, Programmable Logic Controllers (PLCs), Remote Terminal Units (RTUs), and Intelligent Electronic Devices (IEDs). The block diagram of a typical SCADA system is depicted.

For continuous monitoring and control of plant floor devices, sensors and actuators are used to measure different attributes of machinery and transmit that information to field devices. Further, the field control devices, namely, PLCs, RTUs, and IEDs supply digital status information to the MTU (typically placed at the remote location) to determine the acceptable ranges according to parameters set in the server. This information will then be transmitted back to the field control device(s) where actions may be taken to optimize the performance of the system. Moreover, the status information is stored in a database and is displayed on a Human Machine Interface (HMI) at the control centre, where operators can interact with the plant floor



machinery for centralized monitoring and system control those on a power plant require hundreds of field devices and dedicated subsystems to reduce the load on the centralized server.

EXISTING SYSTEM

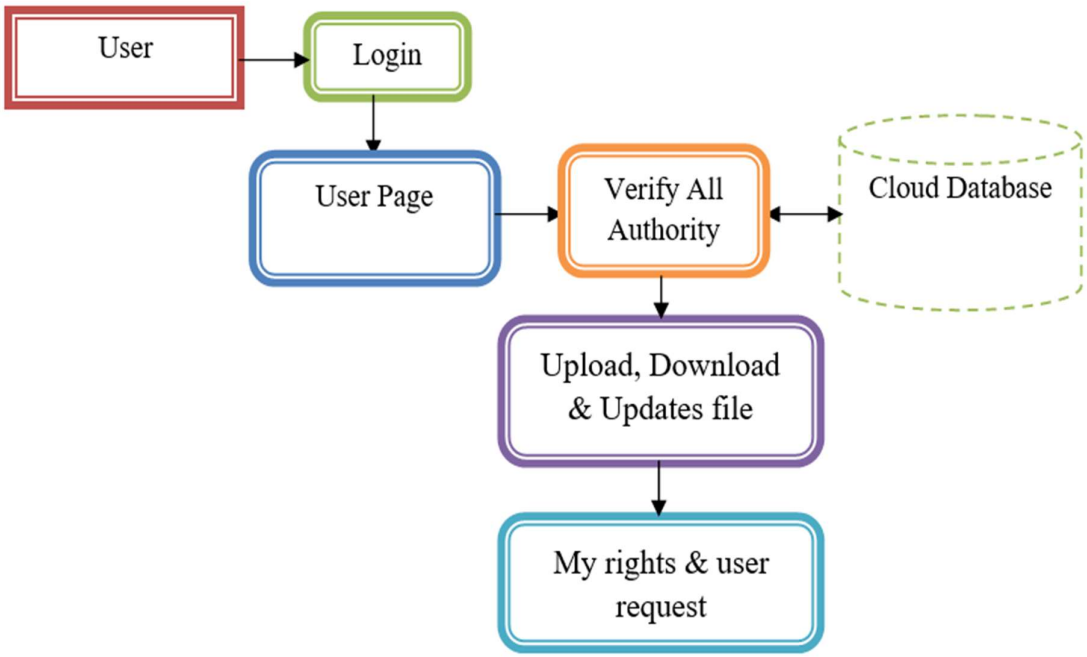
1.These techniques fall under two main categories, namely, centralized key management and decentralized key management schemes.

2.Moreover, each of these categories uses three approaches to generate and extract the session key, namely, symmetric, asymmetric, and hybrid.

3.The drawback of the centralized scheme is that if the key distribution center (KDC) is down, the communication is cut off, which is not acceptable in SCADA systems.

4.In a decentralized approach, the keys are created using keying material and may only affect the single communication link in case of a breakdown.

5.Symmetric key-based approach is efficient in terms of message integrity and high availability, but does not provide authentication and confidentiality.



DRAWBACKS

1.In a decentralized approach, the keys are created using keying material and may only affect the single communication link in case of a breakdown.

2.Less Security.

3.Approach is efficient in terms of message integrity and high availability, but does not provide authentication and confidentiality.

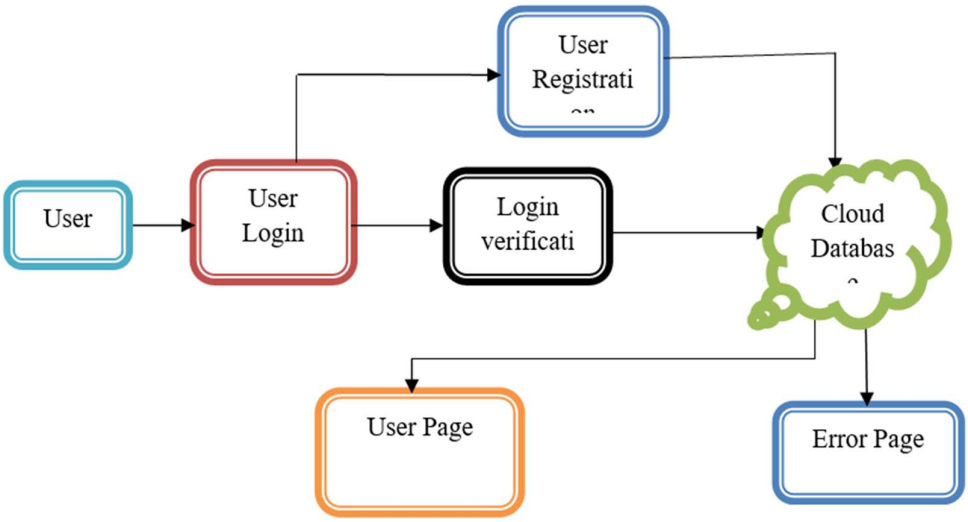
PROPOSED SYSTEM

1.The proposed work aims to offer a multi-layered security framework for industrial infrastructures by combining both symmetric and asymmetric key cryptography techniques.

2.This novel approach follows a layered architecture, where the MTU and sub-MTU can communicate using a hybrid technique for an entire session while the sub-MTU and RTU can communicate using symmetric key cryptography once the session key is securely exchanged.

3.Also, we have proposed a novel approach to generate symmetric keys using vernam cipher rather than using existing methods such as 3DES, AES, etc.

4.We have proposed a multi-layered framework by combining both symmetric and asymmetric key cryptographic techniques to ensure high availability, integrity, confidentiality, authentication and scalability.



ADVANTAGES

1..A multi-layered framework by combining both symmetric and asymmetric key cryptographic techniques.

2.High security.

3.It’s provided authentication and confidentiality.

FUTURE ENHANCEMENT

Further, an efficient session key management mechanism is proposed by merging random number generation with a hashed message authentication code. Moreover, for each session, we have introduced three symmetric key cryptography techniques based on the concept of Vernam cipher and a pre-shared session key, namely, random prime number generator, prime counter, and hash chaining. The proposed scheme satisfies the SCADA requirements of real-time request response mechanism by supporting broadcast, multicast, and point to point communication.

PRESENTED BY:

P BHAVYA

S MOUNIKA

SANJEETH KUMAR

G RAMU