AIES Mini Project Report

Topic

Text-Summarizer

Guided By

Prof. Nitin Pise

Group Members

PRN	NAME	ROLL NO.
1032210790	Vansh Gurnani	25
1032210573	Yuvraj singh Lamba	20
1032210606	Keshav Jha	22
1032210499	Jay Mehta	14

Table of Contents

1) Project Title

Text-Util

2) Team Members

PC14 Jay Mehta
PC20 Yuvraj Singh Lamba
PC22 Keshav Jha
PC25 Vansh Gurnani

3) **What is AI Text Summarizer?**

AI text summarizers are a part of natural language processing (NLP) applications that aim to condense lengthy textual information while retaining its core essence. They're designed to mimic human comprehension and summarization abilities, processing vast amounts of text quickly and efficiently.

There are primarily two types of text summarization techniques:

1. Extractive Summarization: This method identifies and extracts the most important sentences or phrases from the original text. It doesn't create new sentences but selects and rearranges existing ones to form a coherent summary. Algorithms in extractive summarization evaluate the importance of sentences based on factors like word frequency, importance of specific words, sentence position, etc. TextRank and graph-based algorithms are commonly used for extractive summarization.

2. Abstractive Summarization: This approach involves understanding the meaning of the text and generating new sentences to convey the core information in a more concise manner. Abstractive summarization methods leverage advanced NLP techniques like language generation models (such as transformer-based models like GPT, BERT, etc.) to comprehend the text and produce human-like summaries that might not be verbatim but convey the same information in a more condensed form.

AI text summarizers have various applications:

- Information Retrieval: Helping users quickly grasp the main points of lengthy documents, articles, or news stories.

- Content Curation: Creating concise summaries of various articles or pieces of content for newsletters, social media, or websites.

- Document Analysis: Facilitating the processing of large volumes of documents in fields like legal, research, or journalism.

These tools continue to evolve with advancements in machine learning and NLP, aiming to generate more accurate and coherent summaries that effectively capture the essential information from diverse types of texts.

4) Algorithm

1. Imports and Setup:

- The code begins by importing necessary libraries, such as Flask for creating the API endpoints, NLTK for natural language processing tasks like tokenization and frequency distribution, and `heapq` for getting the top sentences.

- It initializes a Flask app and enables CORS (Cross-Origin Resource Sharing) to allow cross-origin requests.

2. Routes:

- There are two routes defined:

  - `'/'` simply returns 'Hello, World!' when accessed.

  - `'/get_summary'` is a POST route that expects JSON data with a 'text' key containing the input text for summarization.

3. Tokenization and Preprocessing:

- The `get_summary` function retrieves the input text from the POST request's JSON payload.

- It tokenizes the text into sentences using `sent_tokenize` from NLTK.

- Tokenizes the words and removes English stopwords (common words that don't carry significant meaning) using NLTK's `stopwords` corpus.

4. Calculating Word Frequencies:

- The code calculates the frequency distribution of the words in the text after filtering out stopwords.

5. Scoring Sentences:

- For each sentence in the input text, it calculates a score based on the sum of frequencies of the words present in that sentence.

6. Generating Summary:

- Selects the top 5 sentences with the highest scores using `heapq.nlargest`.

- Joins these sentences together to create the summary.

7. Calculating Accuracy:

- The `calculate_accuracy` function takes the original text and the generated summary.

- It calculates the accuracy of the summary by comparing the number of shared words between the original text and the summary.

8. Running the App:

- Finally, the app runs in debug mode.

5) Benefits of AI Text Summarizer

1. Time Efficiency: They save time by quickly condensing large volumes of text into shorter, digestible summaries. Users can grasp the main points without reading through entire documents or articles.

2. Content Skimming: For researchers, students, or professionals, these tools help in quickly skimming through numerous documents to find relevant information. This is especially useful when dealing with extensive academic papers or reports.

3. Information Retrieval: When dealing with vast amounts of data or news articles, summarizers help in extracting key details, allowing users to stay updated without delving into every piece of information.

4. Enhanced Productivity: By providing concise summaries, these tools enable professionals to prioritize tasks more effectively and make informed decisions without spending excessive time reading lengthy documents.

5. Language Processing: AI summarizers are capable of processing content in multiple languages, aiding in accessibility and understanding across diverse linguistic barriers.

6. Content Creation: In content marketing or social media management, summarizers assist in creating short, engaging summaries for sharing articles or blog posts, attracting audience attention.

7. Learning Aid: For students or learners, summarizers can distill complex information into simpler formats, aiding comprehension and facilitating better retention of key concepts.

8. Consistency and Objectivity: They provide consistent and objective summaries, eliminating biases that might be present in manually written summaries.

9. Resource Optimization: By minimizing the time spent on reading lengthy texts, these tools optimize resources, allowing individuals to focus on higher-value tasks that require human input and creativity.

10. Decision Support: In fields like business analysis or market research, summaries generated by AI summarizers can provide crucial insights for decision-making processes.

6) Use case and application

1. Time Efficiency: They save time by quickly condensing large volumes of text into shorter, digestible summaries. Users can grasp the main points without reading through entire documents or articles.

2. Content Skimming: For researchers, students, or professionals, these tools help in quickly skimming through numerous documents to find relevant information. This is especially useful when dealing with extensive academic papers or reports.

3. Information Retrieval: When dealing with vast amounts of data or news articles, summarizers help in extracting key details, allowing users to stay updated without delving into every piece of information.

4. Enhanced Productivity: By providing concise summaries, these tools enable professionals to prioritize tasks more effectively and make informed decisions without spending excessive time reading lengthy documents.

5. Language Processing: AI summarizers are capable of processing content in multiple languages, aiding in accessibility and understanding across diverse linguistic barriers.

6. Content Creation: In content marketing or social media management, summarizers assist in creating short, engaging summaries for sharing articles or blog posts, attracting audience attention.

7. Learning Aid: For students or learners, summarizers can distill complex information into simpler formats, aiding comprehension and facilitating better retention of key concepts.

8. Consistency and Objectivity: They provide consistent and objective summaries, eliminating biases that might be present in manually written summaries.

9. Resource Optimization: By minimizing the time spent on reading lengthy texts, these tools optimize resources, allowing individuals to focus on higher-value tasks that require human input and creativity.

10. Decision Support: In fields like business analysis or market research, summaries generated by AI summarizers can provide crucial insights for decision-making processes.

7) Future prospects and achievements

1. Enhanced Accuracy: Ongoing advancements in machine learning models, particularly in language understanding and generation, will likely lead to more accurate and contextually relevant summaries. This involves better comprehension of nuances, context, and domain-specific information.

2. Multimodal Summarization: Integrating text with other media formats like images, videos, and audio to create comprehensive summaries. This could involve summarizing video content or combining information from multiple sources into cohesive summaries.

3. Customization and Personalization: Tailoring summaries to individual preferences or specific domains. Customizable summarization could adapt to different reading levels, language proficiency, or prioritize information based on user interests.

4. Real-time Summarization: Advancements might enable real-time summarization of live events, news updates, or rapidly changing data streams, providing immediate and concise updates.

5. Improved Abstractive Techniques: Further developments in abstractive summarization could lead to summaries that not only condense information but also generate more human-like, coherent, and creative text.

6. Domain-specific Summarization: Specialized summarization for industries like healthcare, law, finance, or technical fields, catering to specific jargon, terminology, and requirements of these domains.

7. Ethical and Bias Considerations: Efforts will continue to address ethical concerns, ensuring AI text summarizers avoid perpetuating biases or misinformation in the generated summaries.

8. Interactive Summarization Interfaces: Creation of interfaces that allow users to interact with summaries, seek more details on specific points, or provide feedback for better refinement of summaries.

9. Cross-lingual Summarization: Advancements in creating summaries that can easily translate and summarize information across multiple languages, fostering global accessibility to information.

10. Explainable AI Summarizers: Developing AI models that can explain how they arrived at a particular summary, providing transparency and trust in the summarization process.

8) Source Code :-

https://github.com/vanshgurnani/Text-Util

https://github.com/vanshgurnani/flask-test

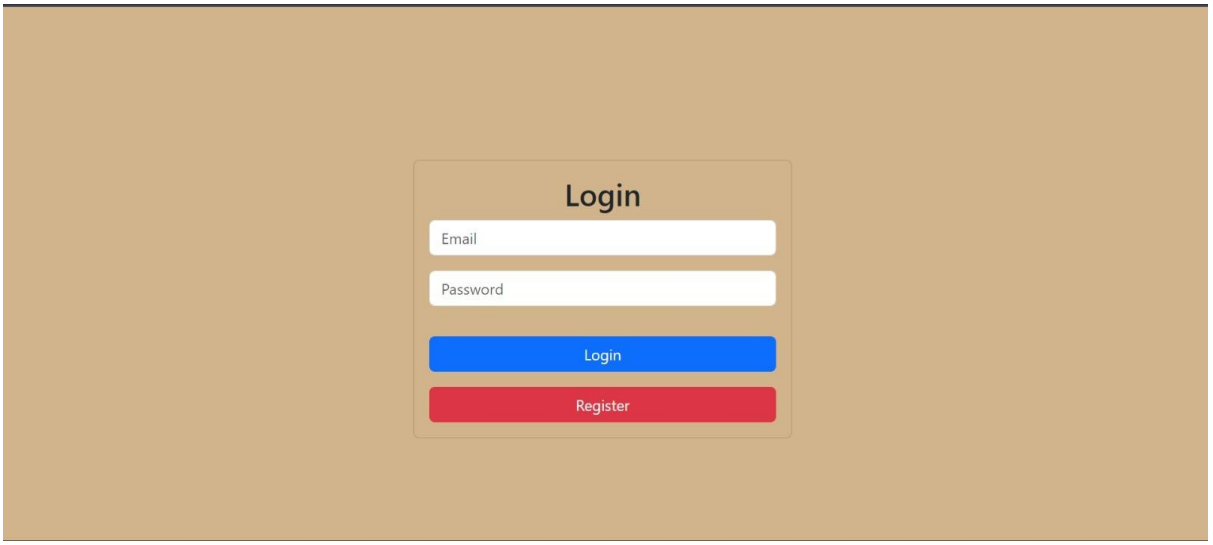
https://text-util-dj8q.vercel.app/

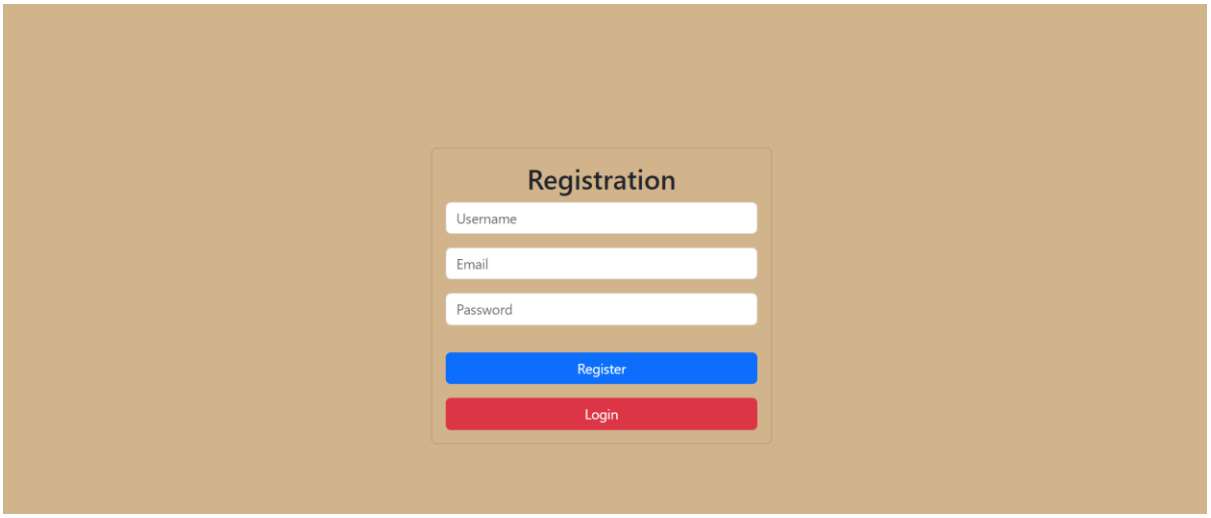
9) Conclusion :-

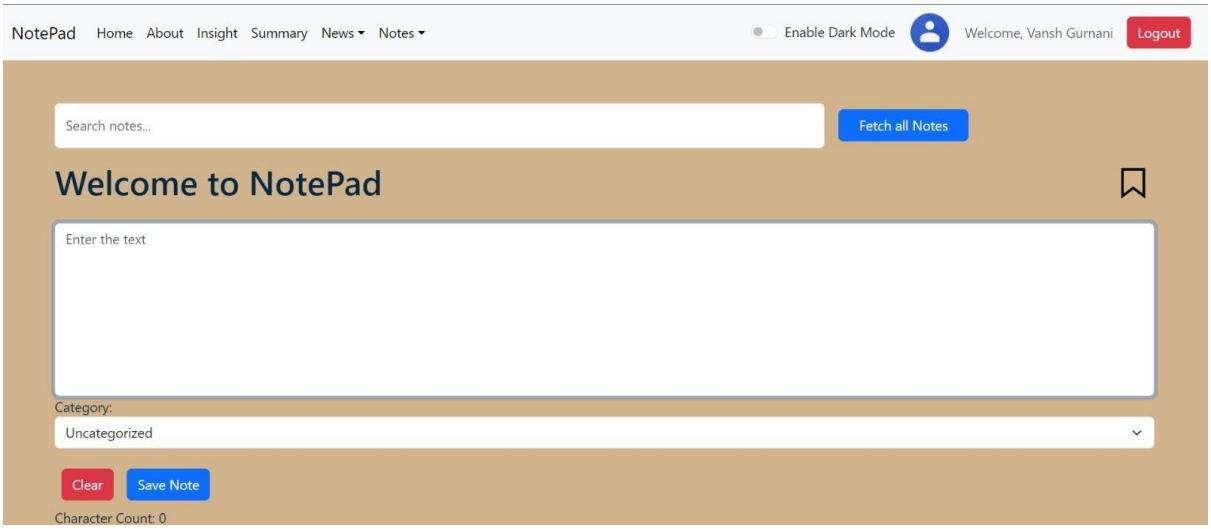
AI text summarizers represent an innovative solution harnessing advanced algorithms to swiftly condense extensive text into succinct, meaningful summaries. Their emergence signifies a transformative tool, streamlining productivity, accessibility, and time efficiency. These summarizers mitigate information overload, enabling quicker, informed decision-making. As AI technology evolves, the landscape for further enhancements and novel applications appears promising, positioning AI text summarizers as pivotal in navigating the deluge of information in the future.

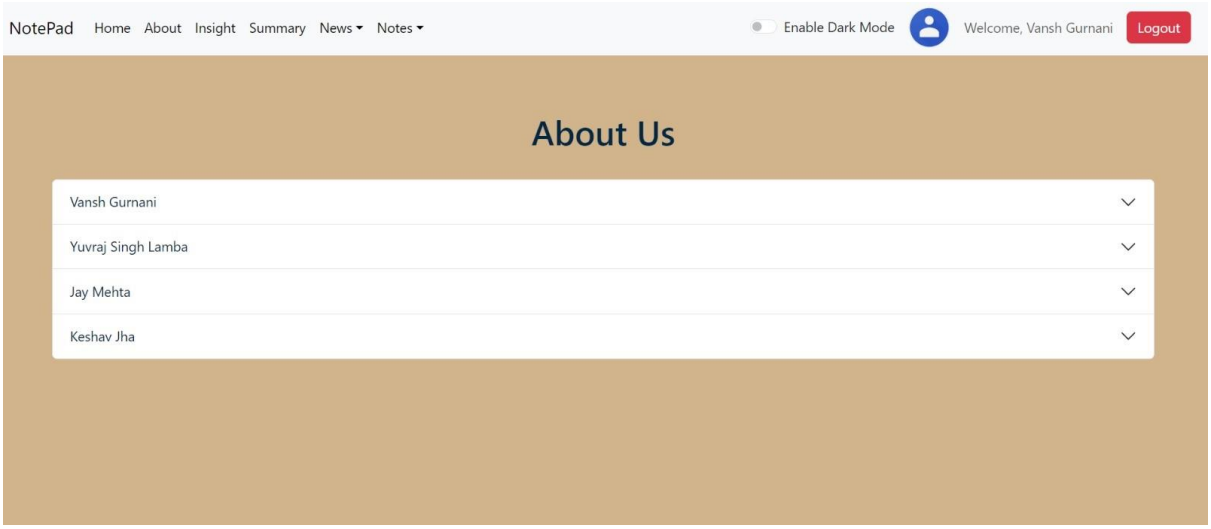
Screenshots of the Project:-

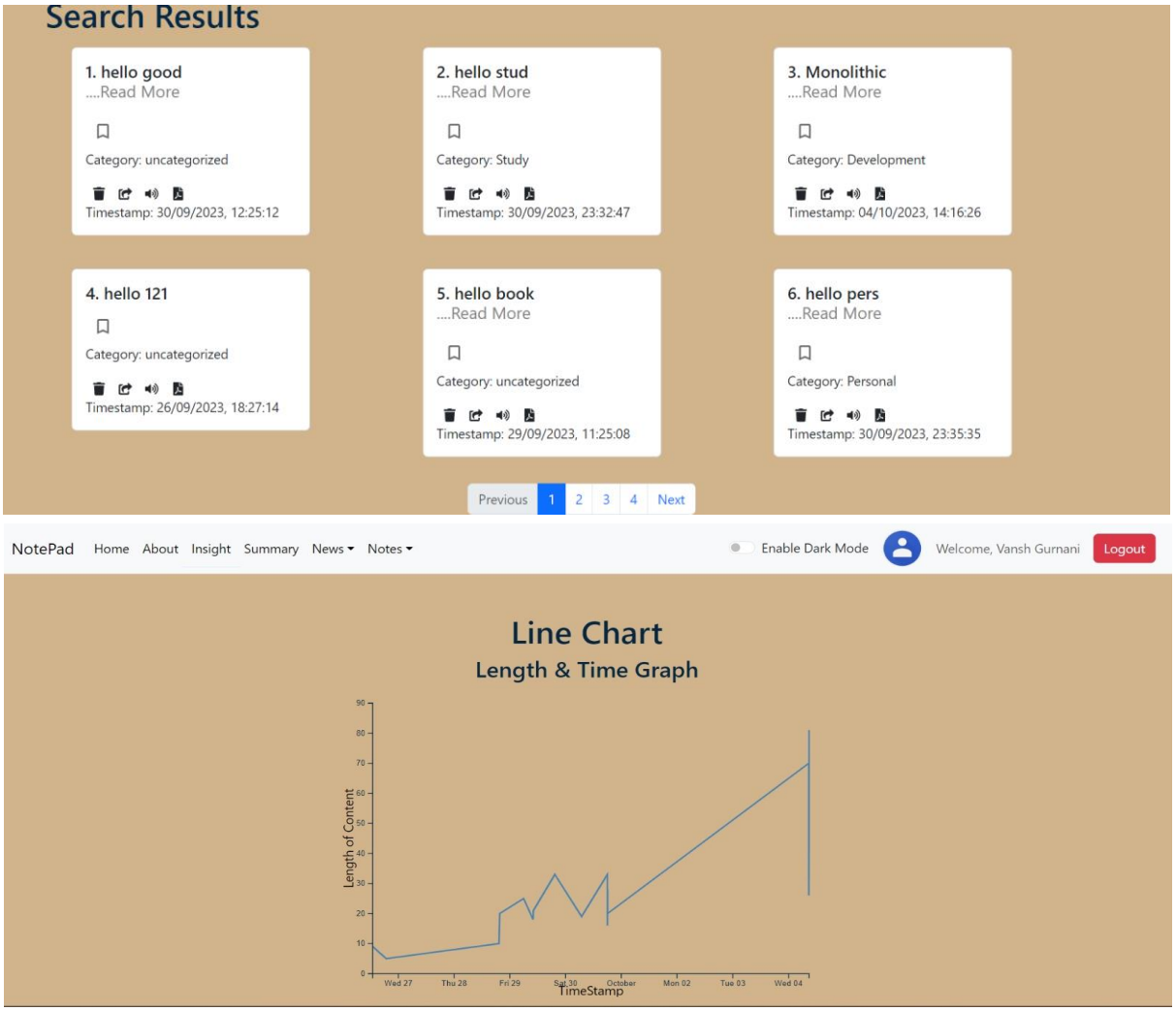
Light Mode :-




Registration
Username
Email
Password
Register
Login


NotePad
Home
About
Insight
Summary
News
Notes
Enable Dark Mode
Welcome, Vansh Gurnani
Logout
Search notes...
Fetch all Notes
Welcome to NotePad
Enter the text
Category:
Uncategorized
Clear
Save Note
Character Count: 0

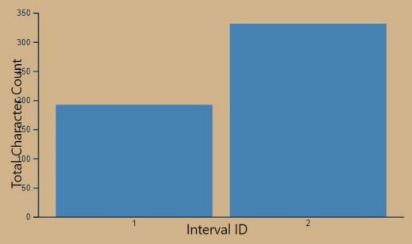

NotePad
Home
About
Insight
Summary
News
Notes
Enable Dark Mode
Welcome, Vansh Gurnani
Logout
About Us
Vansh Gurnani
Yuvraj Singh Lamba
Jay Mehta
Keshav Jha


Search Results
1. hello good
....Read More
Category: uncategorized
Timestamp: 30/09/2023, 12:25:12
2. hello stud
....Read More
Category: Study
Timestamp: 30/09/2023, 23:32:47
3. Monolithic
....Read More
Category: Development
Timestamp: 04/10/2023, 14:16:26
4. hello 121
Category: uncategorized
Timestamp: 26/09/2023, 18:27:14
5. hello book
....Read More
Category: uncategorized
Timestamp: 29/09/2023, 11:25:08
6. hello pers
....Read More
Category: Personal
Timestamp: 30/09/2023, 23:35:35
Previous 1 2 3 4 Next

NotePad
Home
About
Insight
Summary
News
Notes
Enable Dark Mode
Welcome, Vansh Gurnani
Logout
Line Chart
Length & Time Graph
Length of Content
TimeStamp
Wed 27
Thu 28
Fri 29
Sat 30
October
Mon 02
Tue 03
Wed 04

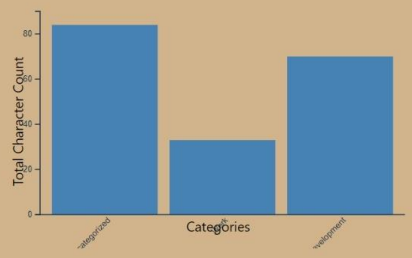
Bar Graph

Total Character Count for Each 10-Note Interval



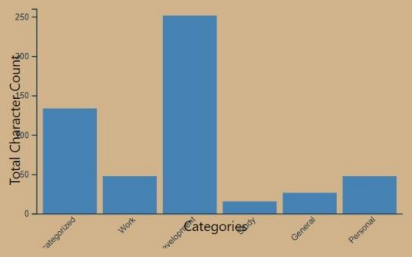
Bookmark Graph

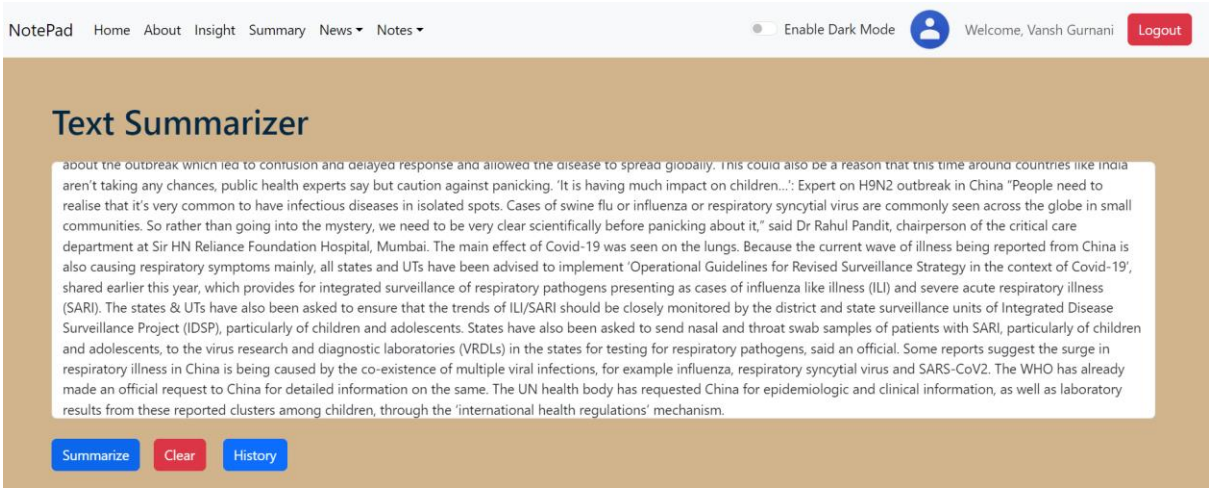
Category Graph for Bookmarked Notes

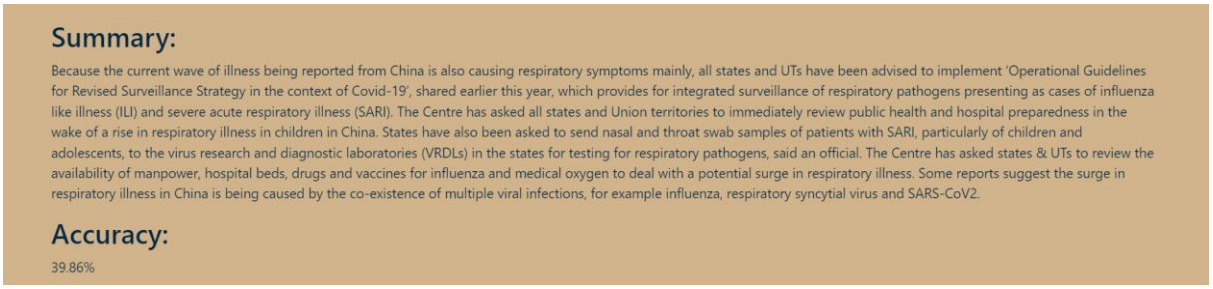


Bar Chart

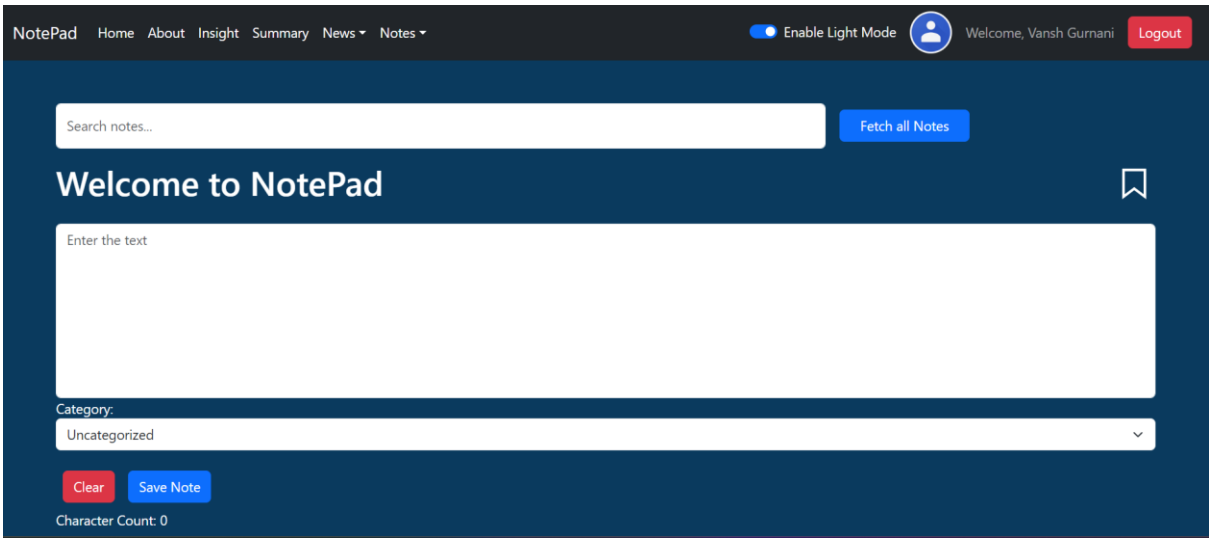
Category Graph

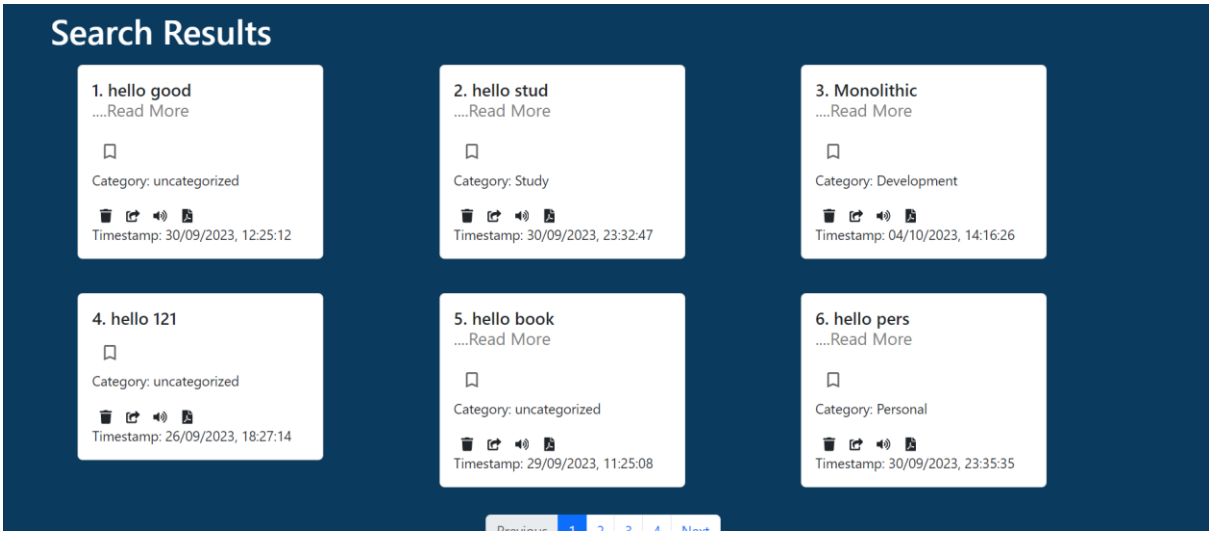


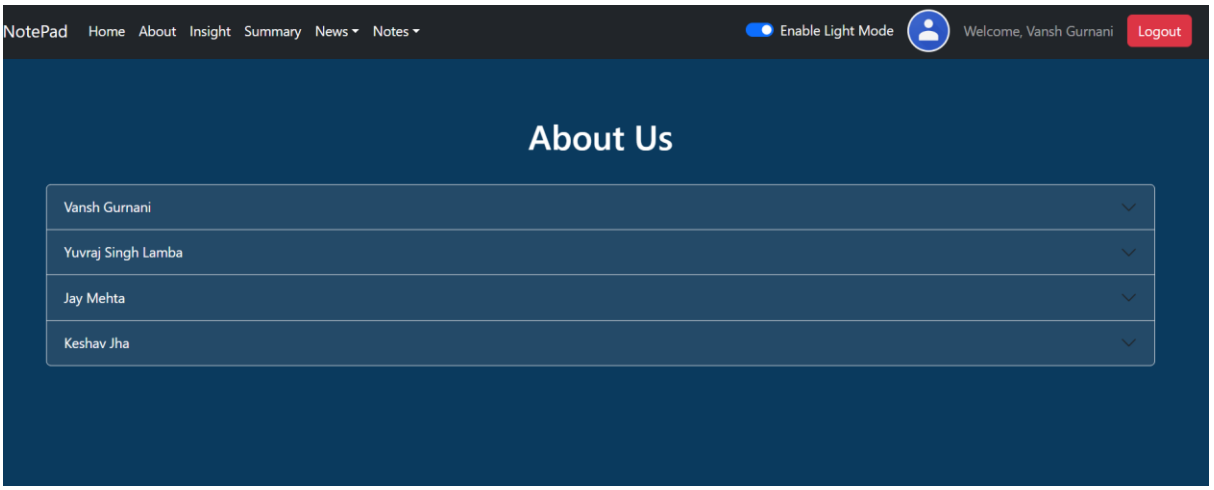

NotePad
Home
About
Insight
Summary
News
Notes
Enable Dark Mode
Welcome, Vansh Gurnani
Logout
Text Summarizer
about the outbreak which led to confusion and delayed response and allowed the disease to spread globally. This could also be a reason that this time around countries like India aren't taking any chances, public health experts say but caution against panicking. 'It is having much impact on children...': Expert on H9N2 outbreak in China "People need to realise that it's very common to have infectious diseases in isolated spots. Cases of swine flu or influenza or respiratory syncytial virus are commonly seen across the globe in small communities. So rather than going into the mystery, we need to be very clear scientifically before panicking about it," said Dr Rahul Pandit, chairperson of the critical care department at Sir HN Reliance Foundation Hospital, Mumbai. The main effect of Covid-19 was seen on the lungs. Because the current wave of illness being reported from China is also causing respiratory symptoms mainly, all states and UTs have been advised to implement 'Operational Guidelines for Revised Surveillance Strategy in the context of Covid-19', shared earlier this year, which provides for integrated surveillance of respiratory pathogens presenting as cases of influenza like illness (ILI) and severe acute respiratory illness (SARI). The states & UTs have also been asked to ensure that the trends of ILI/SARI should be closely monitored by the district and state surveillance units of Integrated Disease Surveillance Project (IDSP), particularly of children and adolescents. States have also been asked to send nasal and throat swab samples of patients with SARI, particularly of children and adolescents, to the virus research and diagnostic laboratories (VRDLs) in the states for testing for respiratory pathogens, said an official. Some reports suggest the surge in respiratory illness in China is being caused by the co-existence of multiple viral infections, for example influenza, respiratory syncytial virus and SARS-CoV2. The WHO has already made an official request to China for detailed information on the same. The UN health body has requested China for epidemiologic and clinical information, as well as laboratory results from these reported clusters among children, through the 'international health regulations' mechanism.
Summarize
Clear
History

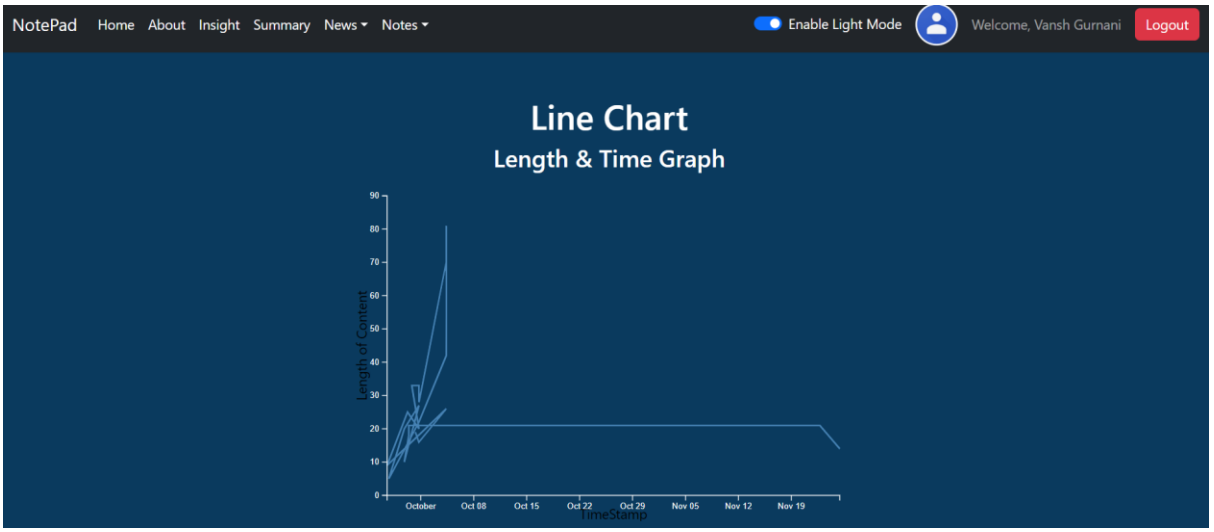

Summary:
Because the current wave of illness being reported from China is also causing respiratory symptoms mainly, all states and UTs have been advised to implement 'Operational Guidelines for Revised Surveillance Strategy in the context of Covid-19', shared earlier this year, which provides for integrated surveillance of respiratory pathogens presenting as cases of influenza like illness (ILI) and severe acute respiratory illness (SARI). The Centre has asked all states and Union territories to immediately review public health and hospital preparedness in the wake of a rise in respiratory illness in children in China. States have also been asked to send nasal and throat swab samples of patients with SARI, particularly of children and adolescents, to the virus research and diagnostic laboratories (VRDLs) in the states for testing for respiratory pathogens, said an official. The Centre has asked states & UTs to review the availability of manpower, hospital beds, drugs and vaccines for influenza and medical oxygen to deal with a potential surge in respiratory illness. Some reports suggest the surge in respiratory illness in China is being caused by the co-existence of multiple viral infections, for example influenza, respiratory syncytial virus and SARS-CoV2.
Accuracy:
39.86%

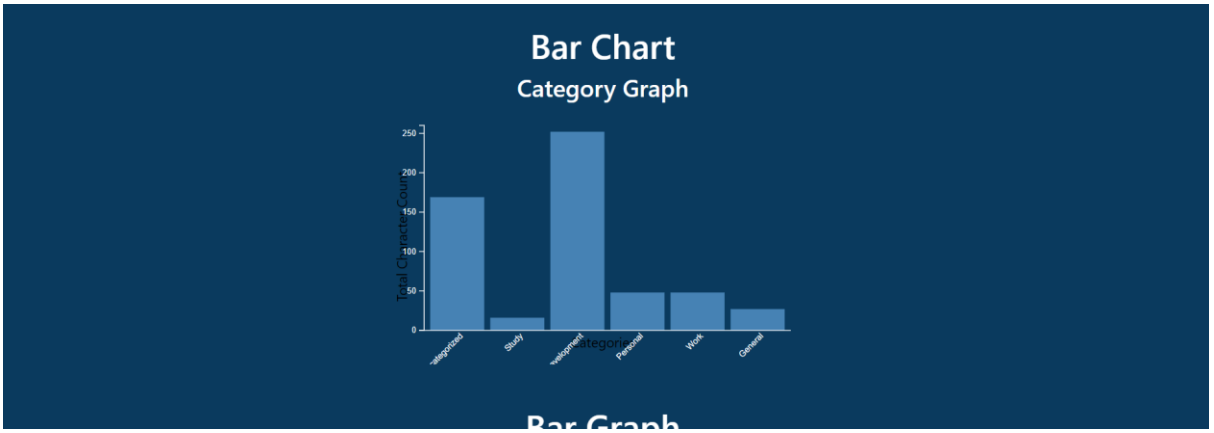
Dark Mode :-

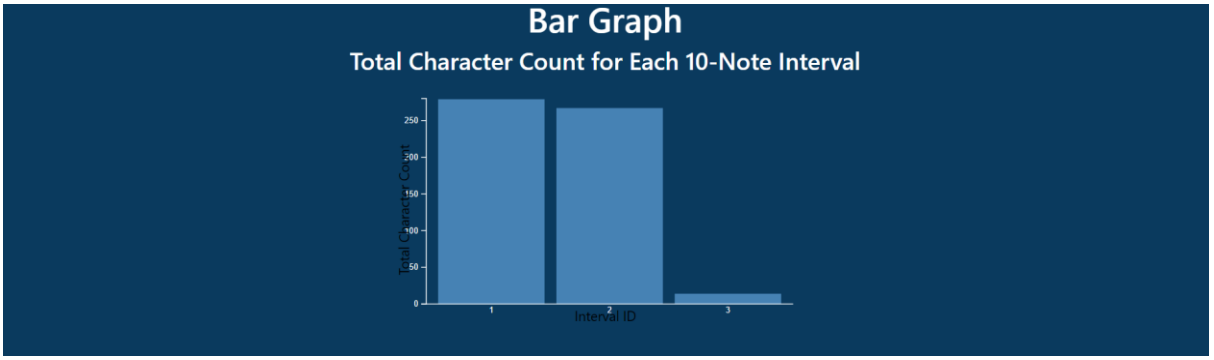

NotePad
Home
About
Insight
Summary
News
Notes
Enable Light Mode
Welcome, Vansh Gurnani
Logout
Search notes...
Fetch all Notes
Welcome to NotePad
Enter the text
Category:
Uncategorized
Clear
Save Note
Character Count: 0

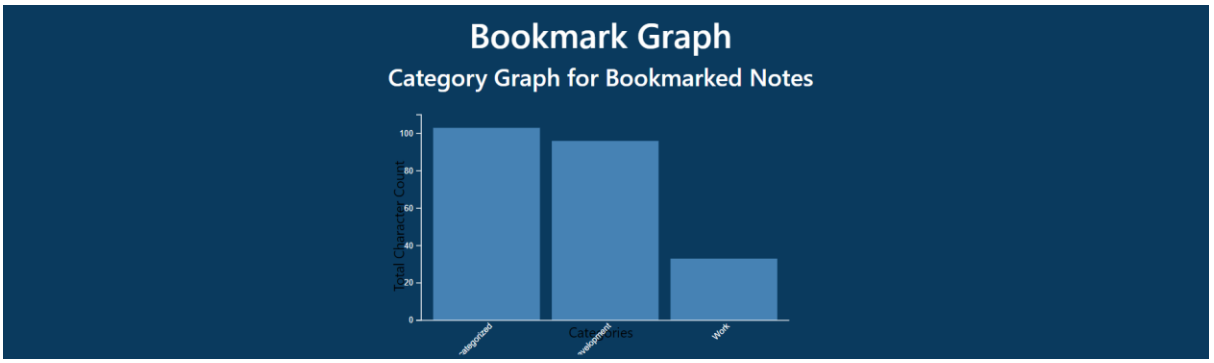

Search Results
1. hello good
....Read More
Category: uncategorized
Timestamp: 30/09/2023, 12:25:12
2. hello stud
....Read More
Category: Study
Timestamp: 30/09/2023, 23:32:47
3. Monolithic
....Read More
Category: Development
Timestamp: 04/10/2023, 14:16:26
4. hello 121
Category: uncategorized
Timestamp: 26/09/2023, 18:27:14
5. hello book
....Read More
Category: uncategorized
Timestamp: 29/09/2023, 11:25:08
6. hello pers
....Read More
Category: Personal
Timestamp: 30/09/2023, 23:35:35

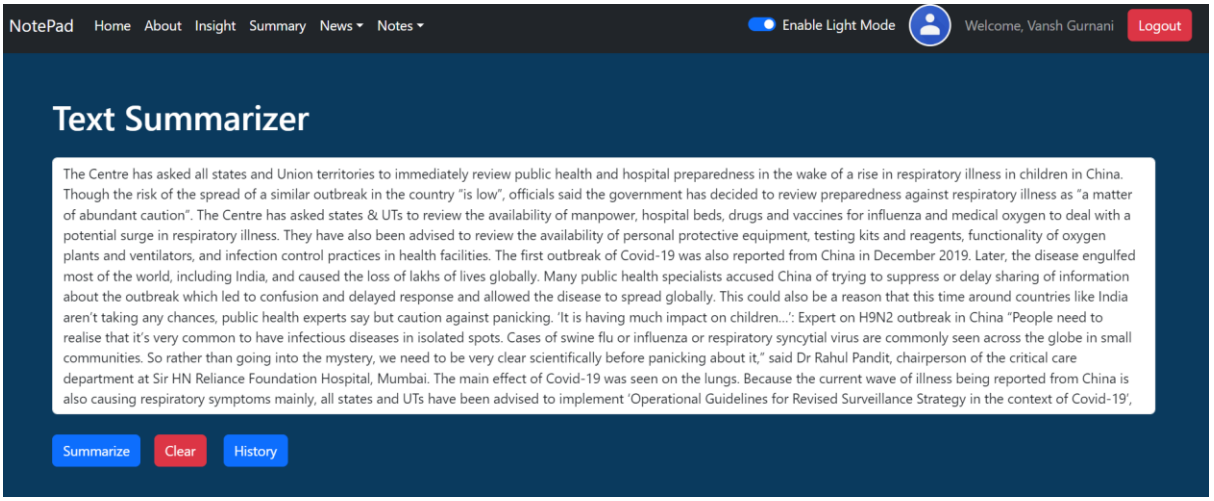

NotePad
Home
About
Insight
Summary
News
Notes
Enable Light Mode
Welcome, Vansh Gurnani
Logout
About Us
Vansh Gurnani
Yuvraj Singh Lamba
Jay Mehta
Keshav Jha

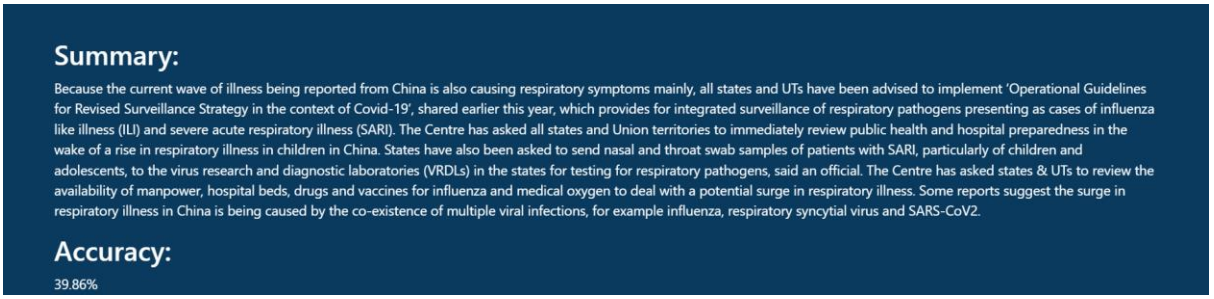

NotePad
Home
About
Insight
Summary
News
Notes
Enable Light Mode
Welcome, Vansh Gurnani
Logout
Line Chart
Length & Time Graph
Length of Content
TimeStamp
October
Oct 08
Oct 15
Oct 22
Oct 29
Nov 05
Nov 12
Nov 19


Bar Chart
Category Graph
Total Character Count
250
200
150
100
50
0
Study
Personal
Work
General
Bar Graph


Bar Graph
Total Character Count for Each 10-Note Interval
Total Character Count
250
200
150
100
50
0
1
2
3
Interval ID


Bookmark Graph
Category Graph for Bookmarked Notes
Total Character Count
100
80
60
40
20
0
Work


NotePad
Home
About
Insight
Summary
News
Notes
Enable Light Mode
Welcome, Vansh Gurnani
Logout
Text Summarizer
The Centre has asked all states and Union territories to immediately review public health and hospital preparedness in the wake of a rise in respiratory illness in children in China. Though the risk of the spread of a similar outbreak in the country "is low", officials said the government has decided to review preparedness against respiratory illness as "a matter of abundant caution". The Centre has asked states & UTs to review the availability of manpower, hospital beds, drugs and vaccines for influenza and medical oxygen to deal with a potential surge in respiratory illness. They have also been advised to review the availability of personal protective equipment, testing kits and reagents, functionality of oxygen plants and ventilators, and infection control practices in health facilities. The first outbreak of Covid-19 was also reported from China in December 2019. Later, the disease engulfed most of the world, including India, and caused the loss of lakhs of lives globally. Many public health specialists accused China of trying to suppress or delay sharing of information about the outbreak which led to confusion and delayed response and allowed the disease to spread globally. This could also be a reason that this time around countries like India aren't taking any chances, public health experts say but caution against panicking. 'It is having much impact on children...': Expert on H9N2 outbreak in China "People need to realise that it's very common to have infectious diseases in isolated spots. Cases of swine flu or influenza or respiratory syncytial virus are commonly seen across the globe in small communities. So rather than going into the mystery, we need to be very clear scientifically before panicking about it," said Dr Rahul Pandit, chairperson of the critical care department at Sir HN Reliance Foundation Hospital, Mumbai. The main effect of Covid-19 was seen on the lungs. Because the current wave of illness being reported from China is also causing respiratory symptoms mainly, all states and UTs have been advised to implement 'Operational Guidelines for Revised Surveillance Strategy in the context of Covid-19',
Summarize
Clear
History


Summary:
Because the current wave of illness being reported from China is also causing respiratory symptoms mainly, all states and UTs have been advised to implement 'Operational Guidelines for Revised Surveillance Strategy in the context of Covid-19', shared earlier this year, which provides for integrated surveillance of respiratory pathogens presenting as cases of influenza like illness (ILI) and severe acute respiratory illness (SARI). The Centre has asked all states and Union territories to immediately review public health and hospital preparedness in the wake of a rise in respiratory illness in children in China. States have also been asked to send nasal and throat swab samples of patients with SARI, particularly of children and adolescents, to the virus research and diagnostic laboratories (VRDLs) in the states for testing for respiratory pathogens, said an official. The Centre has asked states & UTs to review the availability of manpower, hospital beds, drugs and vaccines for influenza and medical oxygen to deal with a potential surge in respiratory illness. Some reports suggest the surge in respiratory illness in China is being caused by the co-existence of multiple viral infections, for example influenza, respiratory syncytial virus and SARS-CoV2.
Accuracy:
39.86%