Ideation Phase
Literature Survey

Date	07 October 2022
Team ID	PNT2022TMID44561
Project Name	Project – News Tracker Application
Maximum Marks	4 Marks

ABSTRACT : A news application is a major intelligent data set that recounts a report. Think about it like you would some other piece of news coverage. It simply utilizes programming rather than words and pictures. As our lives are exceptionally bustling nowadays, we frequently feel we want in excess of 24 hrs. a day to adapt up to all that we have in our timetable. Indeed, that is unrealistic yet diminishing the time by changing the traditional strategy for perusing news can help. Simply let us know what market news you're keen on and get a fast look for the afternoon. Just read what you feel is applicable and save yourtime. This application assists you with questioning for all data about Indices, Commodities, Currencies, Future Rates, Bonds, and so on… as on true sites.

INTRODUCTION : The client wanted to create a solution to facilitate work for media professionals through a location-based assignment managementtool. The main challenge was to create a system on which journalists and other staff remained connected, and could locate or contact each other swiftly to relay stories faster**.**We built NewsTracker to collect a variety of sources of misinformation on Facebook and aggregate the stories published so that we couldhave a clearer sense of the different types of misinformation out there. We wantedto know how frequently it was published and what strategies and narratives were employed to engage audiences on Facebook.

Literature Survey:

(1)News Keyword Extraction for Topic Tracking

This paper presents a keyword extraction technique that can be used for tracking topicsover time. In our work, keywords are a set of significant words in an article that gives high-level description of its contents to readers. Identifying keywords from a large amount of on-line news data is very useful in that it can produce a short summary of news articles. As on-line text documents rapidly increase in size with the growth of WWW, keyword extraction has become a basis of several text mining applications such as search engine, text

categorization, summarization, and topic detection. Manual keyword extraction is an extremely difficult and time consuming task; in fact, it is almostimpossible to extract keywords manually in case of news articles published in a single day due to their volume. For a rapid use of keywords, we need to establish an automated process that extracts keywords from news articles. We propose an unsupervised keyword extraction technique that includes several variants of the conventional TF-IDF model with reasonable heuristics.

(2) Breaking News Detection and Tracking in Twitter

Twitter has been used as one of the communication channels for spreading breaking news. We propose a method to collect, group, rank and track breaking news in Twitter. Since short length messages make similarity comparison difficult, we boost scores on proper nouns to improve the grouping results. Each group is ranked based on popularity and reliability factors. Current detection method is limited to facts part of messages. We developed an application called “Hotstream” based on the proposed method. Users can discover breaking news from the Twitter timeline. Each story is provided with the information of message originator, story development and activity chart. This provides a convenient way for people to follow breaking news and stay informed with real-time updates.

(3) Learning approaches for detecting and tracking news events

The authors extend existing supervised-learning and unsupervised-clustering algorithms to allow document classification based on the information content and temporal aspectsof news events. They've adapted several IR and machine learning techniques for effective event detection and tracking. The article discusses our research using manually segmented documents.

(4) Using Cloud Computing Capabilities On The Example Of Implementing A NewsApplication-Function

The possibilities of cloud computing technologies are considered on the example ofthe application implementation, which is a function that receives a news feed through the NewsApi service. The cloud computing model FaaS (Function as a Service), the Microsoft Azure cloud platform and the Azure Functions solution are used for implementation.

(5) Explaining the News Feed Algorithm: An Analysis of the "News Feed FYI" Blog

Facebook uses algorithmic curation---automated selection and ranking of content--- to present a personalized News Feed to each user for consumption. However, the News Feed user interface provides little information to help users understand how the ranking algorithm works. We analyzed the company's "News Feed FYI' blog seriesto better understand the degree to which Facebook employs "how'and "why' explanations of its News Feed algorithm. These types of explanations have been used in other recommendation and intelligent systems as a means of promoting userunderstanding and acceptance. Our findings show that the "News Feed" FYI' blog posts focus more on explanations that justify why the algorithm works the way it does, and less on explanations that describe how the system works. These findings suggest that the "News Feed" FYI' series would be more helpful for increasing users' confidence in the system, but not improving their trust in the system.

(6) Android News App

As world's technology is rapidly growing we has fast connection and network to instantly connect to other person. Day to day use in mobile, tablets and laptop is increasing, most of the people already have this facilities. In this fast and information oriented world we need tostay updated with every incidents and news too. This News app is android mobile application where user have access to latest news from 120+ newspapers from 50+ countries. The main focus of this application is to connect news articles from all around the world and deliver it to user as fast as possible in best visualize way.

(7) Self-Hosted Kubernetes: Deploying Docker Containers Locally With Minikube

Containerization is a cutting-edge DevOps technology which unifies the IT operations and Development domains. In recent times, virtualization using Virtual Machines has become an overkill for its large overhead on systems. As a lightweight alternative, containerization offers containers that constitute a package of an application along withall its dependencies that is required for it to execute. Containerization platforms help in building containers from images. Docker is a widely popular containerization platform. Containerization Orchestration tools manage these containers. Kubernetes is the front- runner of the emerging market of container orchestration tools. These software work together seamlessly in order to successfully implement containerization both locally andon the cloud. In this paper, we aim to deploy the container orchestration tool Kubernetes on a local system with a Docker sample container. The purpose of this is to ensure that all the configurations and management needed for a Docker container is setsuccessfully on the local system before it is deployed onto the cloud or on the premise. The on-premise deployment use case is very important in domains such as finance and healthcare where

organizations hesitate to upload confidential information on to the clzoud for security reasons but still require scaling of their applications

(8) Research on Topic Detection and Tracking for Online News Texts

With the rapid development of the Internet, the amount of data has grown exponentially. On the one hand, the accumulation of big data provides the basic support for artificial intelligence. On the other hand, in the face of such huge data information, how to extract the knowledge of interest from it has become a matter of general concern. Topic tracking can help people to explore the process of topic development from the huge and complex network texts information. By effectively organizing large- scale news documents, a method for the evolution of news topics over time is proposed in this paper to realize the tracking and evolution of topics in the news text set. First, the LDA (latent Dirichlet allocation) model is used to extract topics from news texts and the Gibbs Sampling method is used to speculate parameters. The topic mining using the K- means method is compared to highlight the advantages of using LDA for topic discovery. Second, the improved single-pass algorithm is used to track news topics. The JS (Jensen- Shannon) divergence is used to measure the topic similarity, and the time decay function is introduced to improve the similarity between topics with the similar time.
Finally, the strength of the news topic and the content change of the topic in different time windows are analyzed. The experiments show that the proposed method can effectively detect and track the topic and clearly reflect the trend of topic evolution

(9) A Cloud-based Framework for COVID-19 Media Classification, InformationExtraction, and Trends Analysis

The coronavirus COVID-19 pandemic has become the center of concern worldwide and hence the focus of media attention. Checking the coronavirus-related news and updates has become a daily routine of everyone. Hence, news processing and analytics become key solutions to harvest the real value of this massive amount of news. This conscious growth of published news about COVID-19 makes it hard for a variety of audiences to navigate through, analyze, and select the most important news (e.g., relevant information about the pandemic, its evolution, the vital precautions, and the necessary interventions). This can be realized using current and emerging technologies including Cloud computing, Artificial Intelligence (AI) and Deep Learning (DL). In this paper, we propose a framework to analyze the massive amount of public Covid-19 media reports over the Cloud. This framework encompasses four modules, including text preprocessing, deep learning, and machine learning-based news information extraction, and recommendation. We conducted experiments to evaluate three modules of our framework and the results we have obtained prove that combining derived information from the news reports provides the policymakers, health authorities, and the public, a complete picture of the way this virus is proliferating. Analyzing this data swiftly is a powerful tool to provide imperative answers to questions that are relevant to public health.

Reference:

- **News Keyword Extraction for Topic Tracking** (https://ieeexplore.ieee.org/document/4624203)
- **Breaking News Detection and Tracking in Twitter** (https://ieeexplore.ieee.org/abstract/document/5616930)
- **Learning approaches for detecting and tracking news events** (https://ieeexplore.ieee.org/abstract/document/784083)
- **USING CLOUD COMPUTING CAPABILITIES ON THE EXAMPLE OF IMPLEMENTING ANEWS APPLICATION-FUNCTION** (https://elib.psu.by/handle/123456789/31517)
- **Explaining the News Feed Algorithm: An Analysis of the "News Feed FYI" Blog** (https://dl.acm.org/doi/abs/10.1145/3027063.3053114)
- **Android News App** (https://www.ripublication.com/ijaer18/ijaerv13n11_78.pdf)
- **SELF-HOSTED KUBERNETES: DEPLOYING DOCKER CONTAINERS LOCALLY WITHMINIKUBE** (https://ieeexplore.ieee.org/abstract/document/9170208)
- **Research on Topic Detection and Tracking for Online News Texts** (https://ieeexplore.ieee.org/document/8703401)
- **A Cloud-based Framework for COVID-19 Media Classification, Information Extraction, and Trends Analysis** (https://ieeexplore.ieee.org/document/9658709)