IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Sentimental Analysis of WhatsApp Chat

[1]Shivani Bhatt, [2] Muskan Gupta, [3] Dipali Birajdar, [4]Sayali Kshirsagar, [5] Prof. Dr. Ramchandra Pujeri
[1]UG Student, [2]UG Student, [3]UG Student, [4]UG Student, [5]Pro Vice Chancellor
Department of Computer Science Engineering,
MIT School of Computing, Pune, India

Abstract: In the digital age, the distinction between personal and professional communication is increasingly blurred, necessitating effective methods for categorizing messages across various platforms. This paper presents a comprehensive study on the application of sentiment analysis for categorizing messages into personal and professional domains. Leveraging natural language processing techniques, our research explores the development and implementation of a sentiment analysis framework capable of accurately discerning the emotional tone of text data. Through a meticulous research methodology encompassing data collection, preprocessing, feature extraction, model development, and evaluation, we demonstrate the efficacy of our approach in categorizing messages into business chats and professional chats. Our findings highlight the potential of sentiment analysis to enhance communication management, productivity, and user experience in digital environments. Additionally, ethical considerations such as privacy protection and bias mitigation are addressed to ensure the responsible deployment of sentiment analysis technologies. This research contributes to the advancement of automated message categorization systems, offering practical insights and methodologies for leveraging sentiment analysis in diverse communication contexts.

INTRODUCTION

In today's digital age, the exponential growth of communication channels has led to an overwhelming influx of messages across various platforms. As individuals and businesses interact through emails, social media, instant messaging, and other mediums, the need to efficiently categorize these messages into personal and professional domains becomes increasingly essential. Sentiment analysis, a subfield of natural language processing (NLP), offers a promising solution to this challenge by analyzing the emotional tone of text data to distinguish between personal and professional communication.

1.1 Background and Context

The distinction between personal and professional communication holds significant importance in both individual and organizational contexts. Personal messages typically involve informal language, expressions of emotion, and intimate discussions, while professional messages adhere to formal etiquette, focus on business-related matters, and maintain a professional tone. However, with the advent of digital communication platforms, the lines between personal and professional communication have blurred, necessitating automated methods for effective message categorization.

1.2 Research Problem

The research problem addressed in this paper revolves around the development and implementation of sentiment analysis techniques to categorize messages as either personal or professional. The primary challenge lies in accurately discerning the underlying emotional nuances of text data and leveraging this information to classify messages into appropriate categories.

1.3 Objectives

The primary objective of this research is to investigate the efficacy of sentiment analysis in categorizing personal and professional messages. Specifically, the study aims to:

- Develop a comprehensive understanding of sentiment analysis techniques and their applicability to message categorization.
- Explore the underlying emotional characteristics of personal and professional communication.
- Design and implement a sentiment analysis model capable of accurately distinguishing between personal and professional messages.
- Evaluate the performance of the developed model through rigorous experimentation and comparative analysis.

1.4 **Significance and Benefits**

The adoption of sentiment analysis for categorizing personal and professional messages offers several compelling benefits:

1.4.1 **Enhanced Communication Management**: By automatically categorizing incoming messages, individuals and organizations can streamline communication management processes, ensuring that personal messages receive appropriate attention without encroaching on professional commitments.

1.4.2 **Improved Productivity**: Efficient message categorization facilitates better time management by enabling users to prioritize their responses based on the nature of incoming messages. This leads to enhanced productivity and reduced cognitive load associated with message triaging.

1.4.3 **Customized User Experience:** Digital platforms can leverage sentiment analysis insights to personalize user experiences, delivering tailored content and services based on individual preferences and communication patterns.

1.4.4 **Scalability and Automation**: Sentiment analysis algorithms can be deployed at scale across diverse communication channels, offering automated message categorization capabilities that adapt to evolving communication patterns and user preferences.

In summary, this research endeavors to explore the potential of sentiment analysis in addressing the challenges associated with categorizing personal and professional messages. By harnessing the power of NLP techniques, we aim to develop robust models that not only accurately classify messages but also contribute to a more efficient, personalized, and insightful communication ecosystem.

PROBLEM STATEMENT

The objective of this project is to design and develop an innovative messaging application experience that encourages users to adopt it as their primary messaging platform. This initiative aims to address critical challenges hindering user engagement and adoption, with a specific focus on implementing robust message segregation capabilities for personal and professional communication.

1. **Inadequate Financial Management:**
   Existing financial management solutions often rely on manual record-keeping methods, leading to inefficiencies and errors. Businesses struggle to gain a comprehensive understanding of their financial patterns and lack efficient budgeting tools to optimize their spending.

2. **Data Security Concerns:**
   Handling financial data poses inherent security risks. Without proper safeguards, user data is susceptible to unauthorized access and breaches, potentially compromising sensitive financial information.
3. **Lack of User-Friendly Solutions:**
   Many available financial tracking applications lack intuitive interfaces, making them less appealing and challenging to use for a wider audience. Businesses require a solution that prioritizes user experience to ensure widespread adoption and usability.

4. **Limited Data Analysis Capabilities:**
   Traditional financial management tools offer limited data analysis and insight capabilities, hindering users' ability to identify trends and make informed financial decisions. Businesses need advanced analytics features to extract actionable insights from their financial data.

5. **Accessibility and Scalability Requirements:**
   To accommodate a growing user base and ensure accessibility across multiple devices, a scalable web-based solution is essential. The application must be easily accessible and adaptable to different screen sizes and devices, providing users with seamless access to financial data anytime, anywhere.

In light of these challenges, this research project aims to develop a comprehensive Finance Dashboard Application using modern technologies and user-centric design principles. By addressing these key issues, the application intends to empower businesses with the tools they need to optimize financial management, enhance decision-making processes, and drive strategic growth.

LITERATURE REVIEW

Paper1: Influence of WhatsApp on communication in internal organizational

Author: S M Nazmuz Sakib

School of Business and Trade; Dhaka International University (DIU); Sonargaon University; Faculty of Medicine, University of Dhaka

Abstract: The survey found that WhatsApp doesn't meet users' needs for internal communication within the organization. It fails to satisfy emotional, cognitive, and social integration needs, and users don't find it essential for staying connected with colleagues. Demographics don't impact satisfaction. Researchers concluded that WhatsApp lacks effectiveness in facilitating organizational communication and doesn't motivate employees for work-related communication.

Paper2: Survey Analysis on the usage and Impact of WhatsApp Messenger

Author: Naveen Kumar and Sudhansh Sharma

SOCIS, Indira Gandhi National Open University, New Delhi-110030, Delhi, India; naveenkumar@ignou.ac.in, sudhansh@ignou.ac.in

Abstract: The study examined the usage and impact of WhatsApp in Northern India, collecting 460 responses with 136 deemed suitable for analysis. It noted a swift transition of users from social networking sites to WhatsApp. Results indicate a significant positive impact on user relationships, with 66% reporting improvement, and over 63% not considering it harmful. Various analyses based on age and gender were conducted. The findings suggest implications for academia and research, highlighting WhatsApp's potential in education, social services, and governance.

Paper3: Classification and Separation of WhatsApp Images Using Machine Learning

Author: Publisher: IEEE Thasneem Rafath Sk; V. Komalatha; G. Sandeep; K. Bhanu Rekha; S. Ravi Kishan

Abstract: The research employs machine learning, specifically Convolutional Neural Networks (CNNs) via Python's Keras library, to automate the categorization and management of images received on WhatsApp. It distinguishes study notes, notices, screenshots, and photos, organizing them into separate folders for easier reference and deletion. This approach streamlines the process of handling WhatsApp media content, enhancing efficiency and user experience.

Paper4: Using WhatsApp messenger for health systems research:

Author: Karima Manji, Johanna Hanefeld, Jo Vearey, Helen Walls, Thea de Gruchy

Abstract:

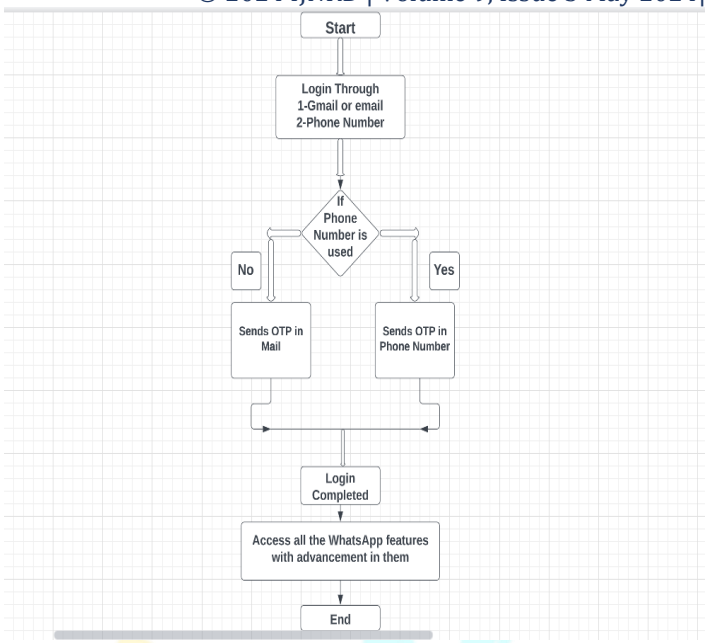
The study investigates WhatsApp's potential as a tool for health research, especially in LMICs with growing mobile phone usage. It found that WhatsApp is primarily used for data collection, but ethical considerations, particularly regarding data protection, are often overlooked. The study provides recommendations for researchers to address these issues when using WhatsApp for health research.

PROPOSED SYSTEM

The proposed system is a Message segregation Application for the users which helps users to find difference between personal and professional messages. It encourages users to adopt our messaging app as their primary platform by offering robust message segregation between personal and professional conversations. It is developed using the K-means Clustering Algorithm with the help of programming language Python.

Key features of the system include:

1. Automated Message Segregation: Implement a machine learning model using K-means clustering to categorize messages as personal or professional based on content and context.
2. Enhanced User Engagement: Provide a seamless and intuitive user experience for managing segregated messages
3. Privacy and Security: Ensure robust data security and privacy measures for user messages.
4. User can send and receive variety of media: text, photos, videos, documents, And location, as well as voice calls and video calls with multiple filters.



Flowchart of the proposed system

RESEARCH METHODOLOGY

1. Data Collection:
- Source Selection: Gather messages from diverse sources including emails, social media platforms, and messaging apps.
- Dataset Diversity: Ensure the dataset encompasses a wide range of communication styles, cultural backgrounds, and topics.

2. Data Preprocessing:
- Cleaning: Remove noise such as special characters, emojis, and HTML tags from the messages.
- Tokenization: Break down messages into individual words or phrases.
- Normalization: Standardize text by converting uppercase letters to lowercase and removing punctuation.

3. Feature Extraction:
- Word Embeddings: Represent words as dense vectors capturing semantic relationships.
- TF-IDF Vectorization: Assign importance scores to words based on their frequency in messages.
- N-grams: Extract sequences of adjacent words to capture contextual information.

4. Model Development:
- Algorithm Selection: Explore various sentiment analysis algorithms including KMeans, Naive Bayes, and neural networks.
- Model Training: Train the selected model on the pre-processed dataset.
- Hyper parameter Tuning: Optimize model parameters to improve performance.
- Ensemble Methods: Combine multiple models to enhance classification accuracy.

5. Evaluation Metrics:
- Accuracy: Measure the proportion of correctly classified messages.
- Precision and Recall: Assess the model's ability to minimize false positives and false negatives.
- F1-score: Harmonic mean of precision and recall, providing a balanced measure of model performance.

6. Experimental Setup:
- Cross-Validation: Divide the dataset into training and testing subsets to assess generalization performance.
- Parameter Optimization: Use techniques like grid search or random search to find optimal hyperparameters.
- Baseline Comparison: Compare the proposed model against simple heuristics or rule-based approaches.

7. Ethical Considerations:
- Privacy Protection: Safeguard personal and professional data in compliance with privacy regulations.
- Bias Awareness: Address potential biases in the dataset and model predictions to ensure fairness.
- Transparency: Provide clear explanations of model decisions to promote trust and accountability.

8. Deployment and Validation:
- Real-world Application: Implement the trained model in communication platforms for practical use.
- User Feedback: Gather feedback from users to validate model effectiveness and usability.
- Iterative Improvement: Continuously refine the model based on user input and evolving communication patterns.

By following this methodology, we aim to develop an accurate and robust sentiment analysis framework for categorizing personal and professional messages, contributing to more efficient communication management and user experience.

IMPLEMENTATION APPROACH

The development of the Finance Dashboard Application using the MERN stack follows a structured implementation scheme involving several key components and steps.

This section provides an overview of the processes and technologies utilized in the development process:

Scope Definition

Objectives:

Develop a messaging application that automatically categorizes incoming messages into personal and professional categories using unsupervised learning techniques. Provide a user-friendly interface to interact with the categorized messages.

Features:

- User authentication and registration.
- Message ingestion and processing.
- Automatic message categorization using the K-Means clustering algorithm.
- Display of categorized messages in a dashboard-like interface.
- User settings for customization and preferences.
  Technology Stack Selection

Backend:

- scikit-learn library for implementing the K-Means clustering algorithm.
- MongoDB or SQLite for storing categorized messages.
- Frontend: HTML, CSS, JavaScript for building the frontend interface.

Development Phases

- **User Authentication and Registration:**

  Implement the K-Means clustering algorithm using scikit-learn on the pre-processed message data. Determine the optimal number of clusters based on message data characteristics. Categorize incoming messages into personal and professional categories based on cluster labels. Store categorized messages in a database.

- **Iterative Development and Testing:**

  Adopt an iterative development approach with continuous integration and deployment (CI/CD) practices. Conduct regular testing, including unit testing, integration testing, and user acceptance testing, to ensure functionality, performance, and reliability. Gather user feedback throughout the development process to identify and address usability issues and feature enhancements.

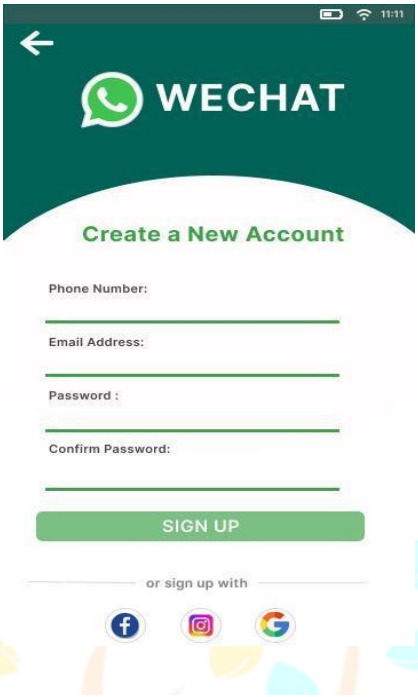
- **Security and Compliance:**

  Implement robust security measures, including end-to-end encryption for message transmission and storage, to ensure data privacy and confidentiality. Adhere to relevant security standards and regulations for handling user data and privacy.

- **Documentation and Knowledge Transfer:**

  Maintain comprehensive documentation covering design decisions, implementation details, and usage instructions. Conduct knowledge transfer sessions to share insights and best practices with stakeholders and future maintainers of the messaging app.

4.1 SignUp Page



Screenshot of the SignUP Page

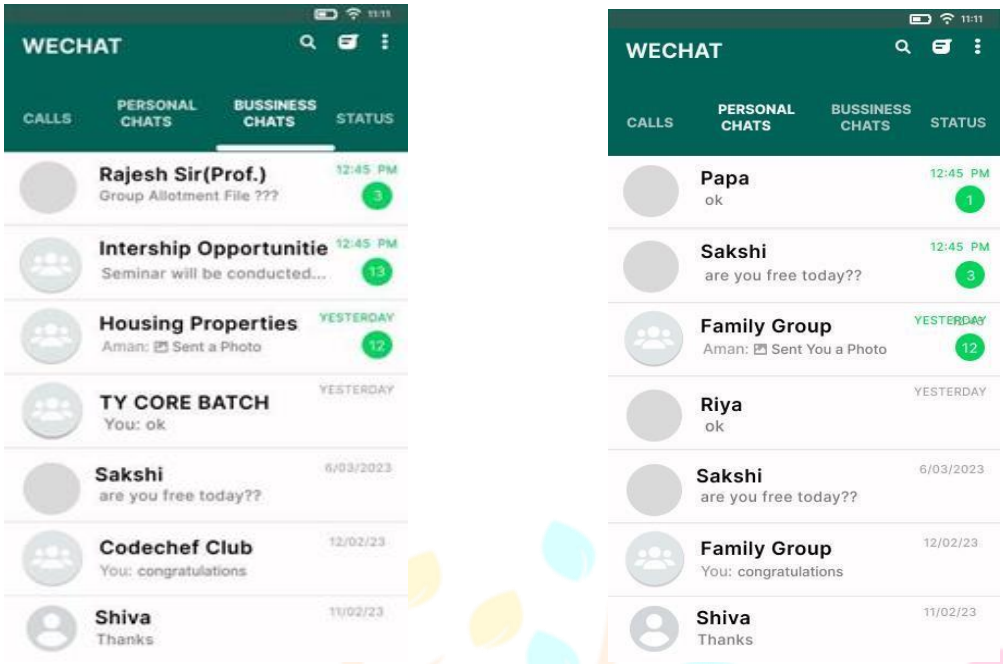
This page is designed to capture important information such as users' full names, email addresses, and strong passwords. By assuring data accuracy through real-time validation checks and delivering a straightforward, userfriendly account setup experience, the registration procedure establishes the framework for successful and secure exchange of messages.

4.2 LOGIN Page

Screenshot of the Login Page

If the individual already has an account, they can log in directly. By entering their registered email address and password in the fields prominently displayed on this page, users can safely login to there account. The design prioritizes user pleasure and adds password validation checks to improve security. The Login Page enhances the user experience by offering quick and secure access to the program's dashboard, where users can manage their expenses and income efficiently.

4.3 Application Page



Screenshot of the Application Page

The primary application page described above employs sentiment analysis techniques to categorize incoming messages into two distinct categories: business chats and professional chats. This categorization is achieved by analyzing the emotional content of the messages, enabling the system to differentiate between communication intended for business purposes and those of a more personal nature. Through the utilization of advanced natural algorithms, the application effectively discerns the underlying sentiments expressed in each message, thereby facilitating accurate classification into the appropriate category. This functionality serves to streamline communication management processes, allowing users to prioritize their responses and allocate their time more efficiently based on the nature and context of incoming messages.

CONCLUSION

In conclusion, the introduction of the ability to segregate personal and professional chats on WeChat offers significant advantages for users seeking to maintain distinct boundaries between their work and personal lives. This feature not only facilitates professionalism in workplace communications but also safeguards the privacy of personal conversations. By enabling users to compartmentalize their chats into separate sections, this enhancement promotes organizational efficiency and reduces the likelihood of sending messages to unintended recipients. Users can enjoy enhanced control and clarity over their communications, leading to a more streamlined and secure messaging experience. Ultimately, the option to separate personal and professional chats on WeChat represents a valuable solution for individuals who leverage the app for both personal and business purposes. This feature underscores WeChat's commitment to user-centric design, providing users with the flexibility and tools necessary to manage their communications effectively in diverse contexts.

REFERENCES

1. www.tyntec.com
2. http://www.accenture.com/SiteCollectionDocuments/PDF/Technology/accenture-mobileweb-watch-2013-survey-new-persuaders.
3. Ibrahim.I. Ima, et.al.(2013)"Antecedent Stirring Purchase Intention of Smartphone among Adolescents in Perlis", International Journal of Academic Research in Business and Social Sciences, Vol. 3, No. 12 (http://www.hrmars.com/journals) \
4. http://www.iabeurope.eu/files/1913/8574/2118/deloitte-nl-global-mobile-consumer-survey.
5. Ham, Y.J,et. al., 2014, "Android Mobile Application System Call Event Pattern Analysis for Determination of Malicious Attack", International Journal of Security and Its Applications Vol.8, No.1, pp.231-246.
6. (The) Times of India (2014), "Facebook to pay $19bn to buy WhatsApp" ,Times of India, 21 February, p12.
7. Barnes, N. G., & Lescault, A. M. (2011). Social media adoption soars as higher-ed experiments and reevaluates its use of new communication tools. . North Dartmouth: Center for Marketing Research, University of Massachusetts Dartmouth, . Bouhnik, D., & Deshen, M. (2014). WhatsApp Goes to school: Mobile Instant Messaging Between Teachers and students. journal of Information technology Education Research, 217-231.