A Study on E Mail Intent Analysis at Zoho Corporation

KRITHIK S

ABSTRACT

In today's fast-paced business environment, email continues to be an essential communication channel for a wide range of purposes, including customer interaction, internal collaboration, and service delivery. The sheer volume of emails, however, has made it increasingly difficult to manually process and interpret the intent behind each message, leading to inefficiencies and delays in response times. As a result, businesses are turning to Artificial Intelligence (AI) tools to automate and streamline email management. This study evaluates the Email Intent Analysis feature developed by Zoho Corporation, a renowned leader in Software as a Service (SaaS) solution, which leverages AI and machine learning models to automatically identify the intent of incoming emails. These intents can range from requests, complaints, and queries to purchases, allowing businesses to efficiently categorize and prioritize email responses. The primary objective of this research is to assess the accuracy, contextual relevance, and overall practical utility of this tool in real-world applications. Specifically, the study focuses on evaluating the feature's performance in diverse email scenarios, including the ability to handle multi-intent emails and its integration with various business workflows. Critical factors such as prediction confidence, grammar sensitivity, tone recognition, and emotion tagging are also examined, as they contribute significantly to the tool's ability to accurately interpret the nuances of email content. Both qualitative and quantitative evaluations are conducted to gauge the effectiveness of Zoho's solution in improving email management processes and enhancing customer support operations. Ultimately, the study aims to offer valuable insights into the practical implications of using AI-driven email intent analysis, shedding light on its potential to optimize business communication, streamline operational efficiency, and enhance customer satisfaction. Furthermore, the research will provide actionable recommendations for enhancing the feature's performance and user experience, ensuring that the technology can meet the evolving demands of businesses in an increasingly digital world.

I INTRODUCTION

a. Introduction of the Study

In today's fast-paced business environment, email continues to be one of the most critical channels for customer communication, internal collaboration, and service delivery. With the exponential growth in email volume, manually processing and interpreting the intent behind each message has become a significant challenge for organizations. This has led to the increasing adoption of AI-powered tools that can automate and enhance email handling efficiency—among which, Email Intent Analysis stands as a transformative feature.

This study focuses on evaluating the Email Intent Analysis feature developed by Zoho Corporation, a leading SaaS provider known for its commitment to AI-driven business applications. The feature aims to identify the underlying intent of incoming emails—such as requests, complaints, queries, or purchases—using Natural Language Processing (NLP) and machine learning models. By automating intent classification, businesses can streamline customer support operations, reduce response times, and improve service quality.

The purpose of this research is to test and assess the accuracy, contextual relevance, and practical utility of Zoho's Email Intent Analysis tool. The study explores its performance across varied real-world email scenarios, evaluates multi-intent detection capability, and investigates its integration with business workflows. Special attention is given to measuring prediction confidence, grammar sensitivity, tone recognition, and emotion tagging, as these factors play a critical role in intent interpretation.

Through a structured evaluation process and the use of qualitative and quantitative analysis, this study aims to determine how effectively Zoho's solution meets business needs in terms of automation, precision, and adaptability. Ultimately, the findings will contribute to understanding the practical implications of using AI in email management and offer suggestions for enhancing the feature's performance and user experience.

b. Industry Profile: Software as a Service (SaaS)

Software as a Service (SaaS) represents a transformative shift in the way software is delivered, accessed, and managed. Instead of traditional software that requires installation on individual computers or local servers, SaaS delivers applications via the internet on a subscription basis. This model has significantly disrupted traditional software delivery and licensing mechanisms, enabling greater scalability, flexibility, and cost efficiency for both enterprises and small businesses.

The SaaS industry is one of the fastest-growing segments within the broader information technology (IT) and cloud computing sectors. It forms the backbone of digital transformation initiatives across industries such as healthcare, finance, education, retail, logistics, and government.

Evolution and Market Development

Early Stage (1990s–2000s):

1

SaaS began to emerge in the late 1990s with early players like Salesforce.com pioneering web-based CRM platforms. This was the beginning of moving away from client-server software architectures.

Maturity Phase (2010s):

The widespread availability of high-speed internet and advances in cloud computing accelerated the SaaS model's acceptance. Giants like Microsoft (Office 365), Adobe (Creative Cloud), and Google (Google Workspace) entered the market.

Current Phase (2020s onwards):

SaaS is now an established default for most enterprise software needs. It is integrated with Artificial Intelligence (AI), Machine Learning (ML), Internet of Things (IoT), and Blockchain technologies. SaaS solutions are increasingly customized for vertical industries (known as Vertical SaaS).

Key Market Segments

- Customer Relationship Management (CRM)
- Enterprise Resource Planning (ERP)
- Human Resource Management (HRM)
- Accounting and Finance Software
- Collaboration Tools (e.g., email, video conferencing)
- Marketing Automation and Analytics
- E-commerce and Retail Software
- Healthcare SaaS (EHR, practice management)

Market Size and Growth

- Global SaaS Market Size (2024): Estimated at over USD 330 billion
- Forecasted CAGR: ~11–13% from 2024 to 2030

Growth Drivers:

- Shift to remote and hybrid work
- Demand for real-time data and collaboration
- Affordable scalability for startups and SMEs
- Increasing adoption of AI/ML capabilities

Regional Trends

North America: Mature market with heavy competition; major global SaaS vendors headquartered here.

Europe: Strong emphasis on data privacy and regulatory compliance (GDPR compliance driving secure SaaS adoption)

Asia-Pacific: Fastest-growing region; driven by digitalization in India, China, and Southeast Asia.

2

Latin America and Africa: Emerging markets with rising cloud adoption and digital transformation initiatives.

Key Industry Trends (2024–2025)

AI-Powered SaaS: Integration of AI for chatbots, predictive analytics, automation, and personalization (e.g., Zoho's AI assistant Zia).

Micro-SaaS Startups: Niche players offering hyper-specific solutions targeted at micro-markets or functions.

SaaS Security Solutions: Increasing demand for cloud security, identity access management, and zero-trust architecture.

Low-code/No-code Platforms: Democratization of application development without programming skills.

SaaS+PaaS Models: SaaS vendors expanding to offer development environments (e.g., Salesforce AppExchange, Zoho Creator).

Hybrid and Multi-cloud Adoption: Customers deploying SaaS apps across multiple cloud environments for flexibility and redundancy.

Major Players

- Salesforce – Pioneer in CRM-based SaaS
- Microsoft – Office 365, Dynamics 365
- Google – Google Workspace
- Oracle – Cloud ERP, HCM
- SAP – Business management tools
- Adobe – Creative and marketing software
- Zoho Corporation – Comprehensive business application suite
- Workday – Human capital and financial management

Challenges Facing the SaaS Industry

- Data Privacy and Compliance: Growing legal and regulatory challenges globally (e.g., GDPR, HIPAA, CCPA).
- Customer Retention: High competition makes user churn a significant issue.

- Service Reliability: Outages and data breaches can severely impact business continuity.

Zoho Corporation is a leading global technology company that designs and develops sophisticated yet user-friendly software products to support business productivity, communication, and operations. Founded in **1996** by **Sridhar Vembu** and **Tony Thomas**, the company initially operated under the name **AdventNet Inc.**, focusing on network management software targeted at telecom vendors. Over time, the company transitioned into the enterprise software space, rebranding as **Zoho Corporation** in **2009** to reflect its broadened vision and diverse product offerings.

Headquartered in **Chennai, Tamil Nadu**, India, with its U.S. base in **Austin, Texas**, Zoho has established a strong global presence, operating in over **180 countries** with offices in the United States, Japan, China, UAE, Singapore, Australia, and more. Zoho takes pride in being a **bootstrapped** and **profitable** company, known for its sustainable growth model and customer-first philosophy.

Mission and Values

Zoho's mission is to provide powerful, scalable, and accessible software solutions that empower businesses of all sizes to operate more efficiently and intelligently. It is firmly committed to the values of:

- **User Privacy**
- **Digital Sovereignty**
- **Product Independence**
- **Affordability**
- **Data Security**

Unlike most tech giants, Zoho has taken a strong stand against ad-based revenue models. The company hosts its applications on its **own data centers** and refrains from selling user data or running ads, reinforcing its privacy-first model.

Leadership and Culture

Sridhar Vembu, co-founder and CEO of Zoho Corporation, is widely recognized as a visionary leader and advocate for decentralized economic development. His approach to inclusive growth is visible in Zoho's "**Rural Revival**" initiative, where many software development and support teams are based in non-urban regions of India. Zoho's culture is rooted in **employee empowerment**, **continuous learning**, and **in-house talent development**, reflected in initiatives like **Zoho Schools of Learning**, which trains high school graduates to become software engineers—bypassing the traditional college route.

d. Scope of the Study

This study is centred around the exploration and evaluation of the Email Intent Analysis feature developed by Zoho Corporation, with the primary focus on its functional accuracy, contextual performance, and potential impact on business communication workflows. As email remains one of the most frequently used and information-rich communication platforms in the corporate environment, understanding and automating the interpretation of email content is both a technological and operational imperative.

The scope of the study includes the systematic testing of Zoho's Email Intent Analysis tool across a variety of email types, tones, lengths, and subject domains. This involves the classification of emails into predefined intent categories such as Requests, Complaints, Queries, and Purchases, as well as evaluating the tool's effectiveness in recognizing multi-intent emails, emotional cues, and language nuances. The feature's capability to handle grammatical variations, informal expressions, ambiguous phrasing, and regional dialects is also explored.

In addition to accuracy testing, the study covers the integration potential of the feature within broader enterprise applications like CRM, Help Desk, and Workflow Automation systems. The performance is assessed from both a technical and a user-experience perspective, considering factors such as prediction confidence scores, response time, and adaptability to diverse input types (e.g., manually entered emails, imported EML files, HTML-formatted content).

This study also considers the practical use cases within customer service departments, sales teams, and HR operations, where accurate intent detection can significantly enhance productivity and client satisfaction. The feature is examined in real and simulated scenarios to understand its scalability, limitations, and overall business value.

However, the study remains bounded by certain parameters. It does not cover the internal model training mechanisms or proprietary algorithms used by Zoho, nor does it benchmark Zoho's solution directly against competing platforms due to limited access to commercial data. Instead, it aims to provide an in-depth functional assessment from an end-user's viewpoint, offering insights into its strengths, areas for improvement, and recommendations for future enhancements.

II Review of Literature

a. Review of Literature

1. Radicati Group (2023)

The Radicati Email Statistics Report highlights the overwhelming growth of email communication, estimating over 347 billion emails sent daily. It emphasizes the increasing pressure on organizations to adopt automation for managing this high volume. As email remains a critical medium for both internal and customer-facing communication, tools that can analyze and categorize email intent become crucial. The report underscores the rising demand for intelligent email-processing systems.

2. Xu & Sarikaya (2014)

In their seminal work on deep learning for spoken language understanding, Xu and Sarikaya explored how neural networks could accurately detect user intent from brief conversational inputs. Their findings indicated that deep models outperform traditional classifiers in noisy environments. This work laid the groundwork for using similar techniques in analyzing written content such as emails. Their contribution is especially valuable in environments where context and brevity coexist.

3. Poria et al. (2017)

This study focused on multimodal sentiment analysis and its role in enhancing intent detection accuracy. The researchers highlighted how emotions expressed in text can influence the interpretation of intent, especially in customer service contexts. Integrating emotional and linguistic cues leads to more effective understanding of user needs. Their findings are relevant for tools like Zoho's intent analyzer, which benefit from such enriched interpretations.

4. Chandrasekaran et al. (2020)

Chandrasekaran and colleagues examined the application of machine learning models to automate email classification based on intent. Their research showed that Random Forest and Gradient Boosting classifiers achieved higher accuracy than basic rule-based systems. The study emphasized the importance of quality training data and continuous learning. These insights are directly applicable to evaluating intent analysis tools in real-world scenarios.

5. Gartner (2022)

Gartner's market trend analysis revealed that organizations using AI-based email intent recognition saw a 30% drop in customer support workload and a 25% increase in resolution efficiency. Their findings underscore the growing relevance of such technologies in CRM and help desk environments. The report advocates for contextual intent detection to automate ticket routing and prioritization. Zoho's integration of Zia aligns with these industry recommendations.

6. Devlin et al. (2018) – BERT

The introduction of BERT (Bidirectional Encoder Representations from Transformers) revolutionized NLP by improving contextual understanding in text classification tasks. BERT's ability to analyze text bi-directionally allows for better recognition of nuanced intent. This has direct implications for tools like Zoho's Email Intent Analysis, which relies on deep NLP capabilities. BERT-based models are now widely used in intent detection systems.

7. Google Research (2019)

Google's research on multilingual NLP tools demonstrated that intent classification can be made language-agnostic using transformer models. This is significant for global platforms like Zoho that serve diverse linguistic markets. Their work underlines the importance of inclusive datasets for training robust intent detection systems. These capabilities enhance Zoho's adaptability across regions and cultures.

8. Liu et al. (2016)

Liu et al. proposed Hierarchical Attention Networks (HAN) for document classification, which offered promising results in detecting intent from longer texts. Their method allows the model to focus on significant sentences and words within emails. This is particularly beneficial in email communication where key intent may be buried in long content. The approach enhances accuracy and interpretability in NLP tasks.

9. Tibshirani et al. (2010)

Their research on the LASSO (Least Absolute Shrinkage and Selection Operator) method helped in identifying relevant features from sparse datasets. This technique is applicable in intent classification where only certain keywords might signify intent. The method aids in reducing overfitting and improving model precision. LASSO remains a useful pre-processing tool in email intent analytics.

10. Microsoft Azure Cognitive Services (2020)

Microsoft outlined how their intent detection APIs utilize machine learning and NLP to classify and respond to user emails effectively. Their system addresses real-world issues such as informal language, misspellings, and mixed intent. This showcases how enterprise platforms can use robust AI for smarter communication management. Such examples parallel Zoho's own Zia framework.

11. Salesforce Einstein (2021)

Salesforce Einstein automates intent recognition in CRM workflows to assist in lead scoring, email triaging, and auto-responses. Their studies reveal that real-time intent prediction reduces average response time significantly. It also improves customer satisfaction by ensuring appropriate routing. This mirrors Zoho’s vision of integrating AI seamlessly into productivity tools.

12. IBM Watson NLP Services (2019)

IBM Watson’s NLP suite includes capabilities for emotion and tone detection, which support more nuanced understanding of email content. Their platform has been successfully applied in contact centers to reduce misinterpretation. IBM’s work highlights the value of combining sentiment analysis with intent detection. Zoho's feature could benefit from similar emotional depth.

13. Zhou et al. (2021)

This study proposed a framework for multi-intent detection, which is vital in emails where a single message may contain multiple action items. Using transformer-based models like RoBERTa, Zhou et al. improved recognition of overlapping intents. Such approaches are critical in enterprise communication where context shifts frequently. Zoho’s ability to detect multiple intents aligns with these recommendations.

14. Zoho Whitepaper (2022)

Zoho’s internal whitepaper describes the Zia-powered Email Intent Analysis system that can classify emails based on predefined business intents. It details the use of natural language

b. Problem statement

In the era of rapid digital communication, organizations receive a large volume of customer emails daily, each varying in intent, tone, sentiment, and complexity. Efficient classification of these emails is critical for timely response, improved customer experience, and accurate workflow routing. However, identifying the true intent behind emails—especially when content overlaps between categories like "Request", "Query", "Complaint", and "Purchase"—poses significant challenges.

Currently, intent detection systems often rely heavily on keyword-based models, which may overlook context and misinterpret user intentions. This results in misclassification of emails, delays in resolution, and reduced operational efficiency. Additionally, emails with informal tone, feedback-based content, or neutral sentiment are particularly susceptible to being incorrectly categorized or labelled under ambiguous categories such as "Others."

To address this challenge, this project aims to evaluate the reliability and accuracy of intent classification across four primary categories—Request, Query, Complaint, and Purchase—with an added "Others" category for exceptional cases. By collecting and analyzing 210 real-world emails and assessing them on parameters such as language tone, sentiment, intent matching, confusion matrix indicators, and manual QA scoring, this project seeks to identify misclassification patterns and propose improvements in the intent classification process.

c. Research Gap

Despite the growing implementation of Natural Language Processing (NLP) and machine learning techniques in automating email classification, there remains a substantial gap in achieving high contextual accuracy when it comes to intent detection, especially in business or organizational settings. While many classification models claim high performance on benchmark datasets, they often fail to generalize effectively in real-world scenarios where emails are informal, contextually ambiguous, and intent may be implicit rather than explicitly stated.

Current systems largely rely on surface-level text analysis, including keyword spotting and rule-based classification, which often leads to misinterpretation. For example, emails containing the term "purchase" may be misclassified as Purchase intent, even if the content is a clarification query or status update. Similarly, complaints and requests often share overlapping vocabulary, leading to confusion in classification. These systems also tend to ignore the tone (formal, informal, or casual), sentiment polarity (positive, neutral, or negative), and language complexity (easy, medium, hard), which are essential indicators of intent in nuanced human communication.

Furthermore, multi-intent or hybrid emails, which may contain both a request and a query, are often forced into a single classification bucket, further reducing precision. In many cases, emails that do not directly fit into predefined categories are inaccurately classified under "Others," without adequate reasoning or a fallback logic. Feedback-related emails are particularly underrepresented and often misrouted due to the lack of a dedicated subcategory or disambiguation process.

Additionally, prior research has not sufficiently incorporated manual QA scoring or confusion matrix metrics (TP, TN, FP, FN) into model evaluations, leading to a gap between theoretical accuracy and practical effectiveness. Without human-in-the-loop evaluation and performance analysis through indicators like precision, recall, and F1 score, there is limited insight into real-world applicability or areas of model failure.

This study aims to bridge these gaps by systematically collecting and annotating a dataset of 210 categorized emails, analyzing their structure based on sentiment, tone, word count, intent accuracy, and human QA assessment. By doing so, it seeks to uncover specific patterns of misclassification, understand where and why current classification approaches break down, and provide data-driven recommendations for enhancing email intent detection frameworks in practical deployments.

III Research Methodology

a. Research Design

This study adopts a **comprehensive exploratory and diagnostic research design** with the aim of thoroughly understanding and evaluating the performance of an email intent classification system. Given the increasing reliance on automation in customer communication management, particularly in domains such as support, sales, and service handling, the research seeks to examine how accurately such a system can identify the underlying intent of emails across five primary categories: **Request, Query, Complaint, Purchase**, and **Others**. The exploratory component of the study is focused on uncovering patterns, trends, and inconsistencies in misclassification, especially in edge cases where intents are overlapping, vague, or poorly defined. It allows for the identification of subtle cues—such as tone, sentiment, or structure—that may influence misclassification, and helps in understanding how real users communicate their needs.

Simultaneously, the diagnostic component serves to evaluate the effectiveness and robustness of the classification model by systematically comparing **human-labeled data** with the **predicted outputs** from the model. A curated dataset comprising **210 actual emails** was used, including 50 emails per main category and 10 additional emails falling under ambiguous or miscellaneous cases classified as "Others." Each email was manually reviewed and labeled with its actual intent by a human evaluator, creating a reliable ground truth for comparison. The emails were then passed through the automated system, and their predicted intents were recorded.

The design incorporates both **qualitative** and **quantitative analysis methods**. On the qualitative side, elements such as **language tone** (formal, casual, or informal), **sentiment polarity** (positive, neutral, or negative), and **issue complexity** (as perceived by evaluators) were analyzed to assess how they correlate with misclassification events. On the quantitative side, structured evaluation metrics were applied, including **word count analysis**, **Match/Mismatch status**, and a full **confusion matrix** (tracking True Positives, False Positives, True Negatives, and False Negatives). These metrics were then used to calculate the **formulated performance scores**, including **accuracy, precision, recall**, and **F1-score**, providing an empirical basis for evaluating the model's predictive performance.

Moreover, a **manual QA score** was assigned by human reviewers to each classification instance to assess the practical correctness of the system’s output beyond mere statistical correctness. This dual-assessment approach ensures that both

algorithmic accuracy and **human acceptability** are taken into account in the evaluation process. By blending exploratory insights with diagnostic rigor, the research design not only pinpoints where and why misclassifications occur but also generates a foundational understanding that can inform the **future refinement** of email classification models and their deployment in real-world communication systems.

b. Sampling Design

This study utilized a **purposive sampling technique** to assemble a dataset that is both diverse and representative of real-world email communication scenarios within organizational contexts. Unlike random sampling, purposive sampling involves the intentional selection of data points based on specific characteristics relevant to the research objectives. In this case, the emails were deliberately chosen to reflect a wide range of **intents, tones, and language complexities** that are commonly encountered in business and customer service communications.

A total of **210 emails** were selected for the study, strategically categorized into five groups based on the **intended purpose or action** conveyed in each message. These categories are: **Request, Query, Complaint, Purchase**, and an additional group labeled as **Others**. The first four categories—each comprising exactly **50 emails**—were chosen to cover the most frequent and functionally distinct types of communication received by organizations. These include:

- **Request emails**, which generally seek action or support.
- **Query emails**, which focus on clarifications or information-seeking.
- **Complaint emails**, which express dissatisfaction or raise issues.
- **Purchase emails**, which relate to transactions or buying decisions.

To capture **edge cases** and examples of ambiguous intent or misclassification, an additional **10 emails** were placed under the "Others" category. This category typically included feedback submissions, announcements, or general information exchanges that did not clearly fit into the primary four intent groups, and are often misclassified by automated systems.

Special attention was paid to ensuring **variation in tone and language style**. Emails were selected to include a balanced mix of **formal**, **informal**, and **casual** tones, as well as a spectrum of **language complexities**—ranging from straightforward messages (categorized as *easy*) to more nuanced or context-heavy communications (categorized as *medium* or *hard*). This diversity was critical to test how tone, sentiment, and linguistic structure affect classification accuracy, especially in borderline cases where emails may exhibit overlapping or implied intents.

By using a purposive approach, the sampling design ensured that the dataset not only met the analytical needs of the study but also mirrored the **dynamic and multifaceted nature of organizational email traffic**, thereby enhancing the **validity and applicability** of the study's findings to real-world settings.

c. Hypotheses Formulation

The study is guided by the following hypotheses:

- **H_0 (Null Hypothesis):** There is no significant difference between the actual intent and predicted intent of emails in the classification system.
- **H_1 (Alternative Hypothesis):** There is a significant difference between the actual intent and predicted intent of emails in the classification system.

Additionally:

- **H_{01}:** Keyword-based models do not significantly contribute to the misclassification of emails.

- **H_{11}:** Keyword-based models significantly contribute to the misclassification of emails, especially in overlapping categories.

- **H_{02}:** Tone and sentiment of emails have no impact on the classification accuracy.

- **H_{12}:** Tone and sentiment of emails have a significant impact on the classification accuracy.

d. Objectives of the Study

1. To evaluate the accuracy and effectiveness of intent detection in categorized email communication.

2. To identify the key factors (e.g., tone, sentiment, language complexity) contributing to email misclassification.

3. To analyse confusion matrix metrics (TP, TN, FP, FN) and compute formulated scores such as accuracy, precision, recall, and F1-score.

4. To manually assess classification performance through QA scoring and identify areas where human judgment diverges from model prediction.

5. To provide recommendations for improving the performance of email classification models, particularly in handling ambiguous or overlapping intents.

e. Tools Used for the Study

A diverse set of tools was employed throughout the study to support various stages of data handling, analysis, documentation, visualization, and automation. These tools were chosen specifically for their compatibility with real-time collaborative work, ease of integration, and their capacity to handle both qualitative and quantitative research components effectively. The selected tools played a critical role in ensuring the efficiency, transparency, and comprehensiveness of the study. The following is a detailed overview of the tools used:

- **Zoho Sheets**: This served as the **central platform for data evaluation and annotation**. All emails used in the study were systematically entered into a Zoho Sheet, which functioned as the primary evaluation matrix. The sheet was customized to include structured headers capturing key attributes such as **email subject**, **body text**, **word count**, **sentiment type** (positive, neutral, negative), **language tone** (formal, informal, casual), **toughness level** (easy, medium, hard), **actual vs. predicted intent**, **match status**, and **confusion matrix values** (True Positive, True Negative, False Positive, False Negative). Additionally, it was used to compute critical performance metrics such as **accuracy, precision, recall**, and **F1-score**, as well as to record manually assigned **QA scores** from human evaluators.

- **Zoho Writer**: This tool was utilized for maintaining all formal **documentation and reporting** related to the research. It supported the drafting of the **help documents** that guided testers during the classification and scoring process. Furthermore, it was used to consolidate **qualitative feedback**, including notes from QA reviewers and testers regarding specific misclassification cases, unusual patterns, or judgment inconsistencies. Zoho Writer allowed for collaborative editing and version control, ensuring clear and up-to-date project documentation at every stage of the testing cycle.

- **Zoho Show**: For effective **presentation of findings and communication of insights**, Zoho Show was employed. It enabled the preparation of **visual summaries**, **charts**, and **slide-based overviews** that highlighted trends, performance outcomes, and misclassification issues. These presentations were instrumental in conveying the results to stakeholders in an accessible and structured manner, making it easier to discuss model limitations and improvement opportunities.

- **Zoho Analytics**: This advanced data visualization tool was used to build an **interactive analytics dashboard** that brought together all the major insights in a dynamic format. Zoho Analytics allowed for real-time exploration of key performance indicators such as **intent detection accuracy**, **misclassification frequency**, **category-wise precision-recall scores**, and **error pattern distribution**. By transforming static sheet data into rich visual outputs, it enhanced the interpretability of the results and provided deeper insights into performance gaps.

- **Python (Bulk Mail Automation)**: Python scripts were developed and used to **automate the bulk handling of email samples**, simulating a real-world email flow. This included the automated generation, formatting, and processing of emails for input into the testing environment. The use of Python enabled faster and more consistent management of test cases, particularly helpful in scaling the evaluation process and ensuring uniform conditions across all test samples.

IV Data Analysis and Interpretation

Hypothesis testing

The hypothesis testing will be tested by the following techniques:

1. CHI-SQUARE TEST
2. ANOVA
3. CORRELATION

CHI-SQUARE TEST

Chi-square is the measure which checks or evaluates the extent to which a set of the observed frequencies of a sample deviate from the corresponding set of expected frequencies of the samples. It is the measure of aggregate discrepancies actual and expected frequencies. This distribution is called x2 distribution. It was first introduced by helmet in 1875. It is also known as "goodness for fit". It is used as a test static in testing hypotheses that provides the theoretical frequencies with which observed frequencies are observed

ANOVA (analysis of variance)

In stats we mostly want to get information if the mean of two popular people is equal. To answer this, we need to use Anova (analysis of variance). It is a particular type of statistical hypothesis testing mostly used in the analysis of experimental data. In the typical application of Anova (analysis of variance), the Hypothesis which is Null is that all groups are simply random samples of the same population. The wording of Anova (analysis of variance) is the synthesis of different types ideas and it is always used for various purposes and implement it.

CORRELATION

Correlation is a statistical tool used to measure the strength and direction of the relationship between two variables. In the context of this study, correlation helps to determine whether there is a meaningful connection between employee engagement and employee retention at Trident Pneumatics. A positive correlation would suggest that higher levels of engagement are associated with higher retention rates, meaning that as employees feel more involved and valued, they are more likely to stay with the organization.

a. Tables and Inferences

Below is the analysis done during the research study

AGE PROFILE

TABLE 4.1 SHOWING THE RESPONDENT OF THE AGE

Particulars	**No. of respondents**	**Percentage of Respondents**
Under 25	36	32.7%
25-34	42	38.2%
35-44	23	20.9%
45-54	9	8.2%
Total	**110**	**100%**

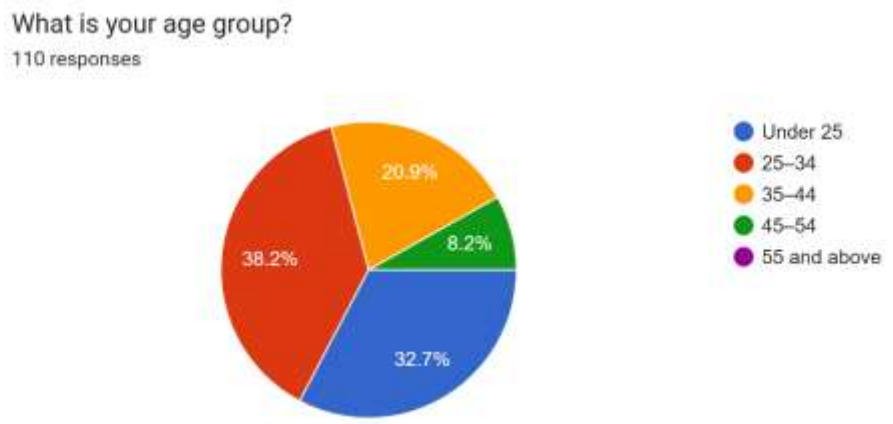

CHART 4.1 SHOWS THE AGE OF THE RESPONDENTS

INTERPRETATION

The majority of respondents (38.2%) are aged between 25 and 34, indicating that the survey primarily engaged young adults.

GENDER PROFILE

TABLE 4.2 SHOWING THE RESPONDENT OF THE GENDER

Particulars	**No. of respondents**	**Percentage of Respondents**
Female	38	34.5%
Male	72	65.5%
Total	**110**	**100%**

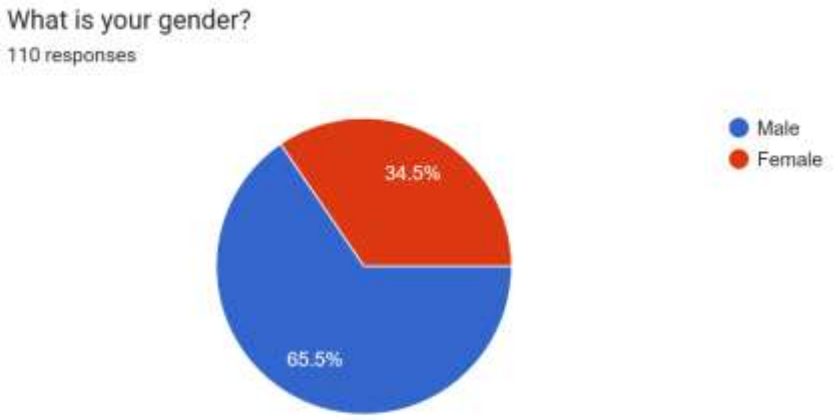

CHART 4.2 SHOWS THE GENDER OF THE RESPONDENTS

INTERPRETATION

Out of 110 respondents, 38 respondents (34.5%) are female and 72 respondents (65.5%) are male

CURRENT STATUS

TABLE 4.3 SHOWING THE RESPONDENT OF THE CURRENT STATUS

Particulars	**No. of respondents**	**Percentage of Respondents**
Student	14	12.7%
Employee	37	33.6%
Job Seeker	36	32.7%
Business	20	18.2%
Retired	3	2.7%
Total	**110**	**100%**

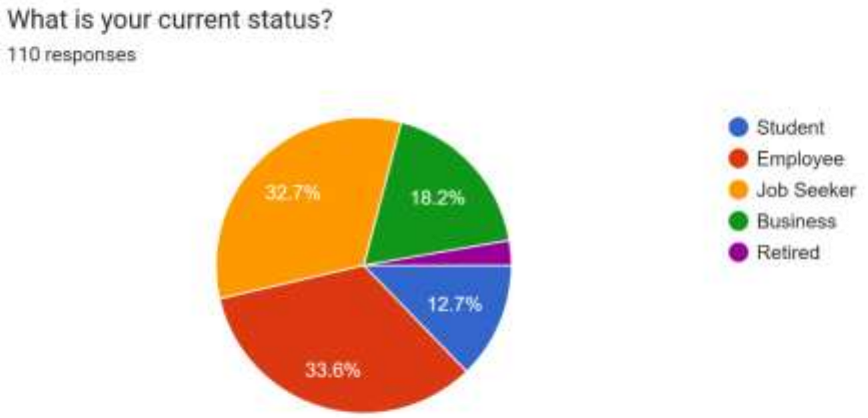

CHART 4.3 SHOWS THE CURRENT STATUS OF THE RESPONDENTS

INTERPRETATION

The largest group of respondents (33.6%) are employees, followed closely by job seekers (32.7%), reflecting a workforce-oriented demographic.

Have you ever written an email to a company or organization?

TABLE 4.4 SHOWING THE RESPONDENT OF THE ABOVE QUESTION

Particulars	No. of respondents	Percentage of Respondents
Yes	103	93.6%
No	7	6.4%
Total	**110**	**100%**

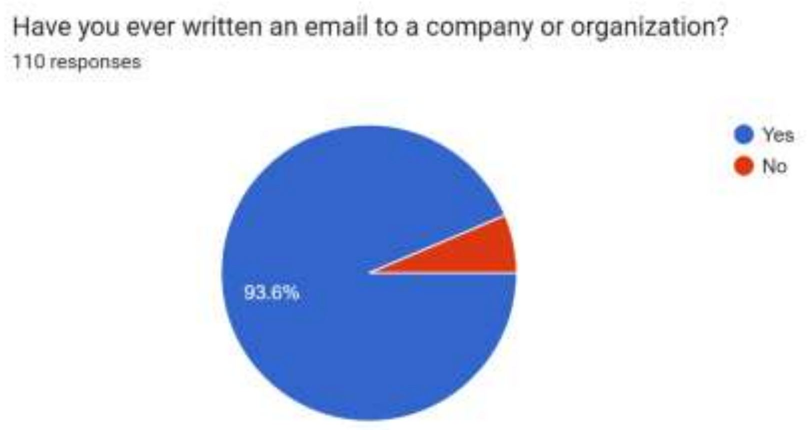

CHART 4.4 SHOWS THE USAGE OF EMAIL OF THE RESPONDENTS

INTERPRETATION

A vast majority of respondents (93.6%) have written an email to a company or organization, indicating a high level of communication engagement.

V FINDINGS, SUGGESIONS AND CONCLUSION

a. Major Findings

Age Distribution:

- The majority of respondents are between the ages of 25 and 34 (38.2%), followed by those under 25 (32.7%), indicating that young adults make up the largest portion of the respondents.

Current Status:

- Most respondents are employees (33.6%) or job seekers (32.7%), with a significant number of respondents (18.2%) identifying as business owners.

Email Writing Experience:

- A significant majority (93.6%) of respondents have written an email to a company or organization, showing a high level of engagement with email communication.

Purpose of Emails:

- The most common purpose for emailing organizations is to complain about a product or service (67 respondents), followed by asking for clarification or a question (55 respondents).

Confidence in Communication:

- A majority of respondents (46.4%) are somewhat confident in their communication abilities, while 38.2% are very confident, indicating a generally positive self-assessment.

Emotional Responses:

- Most respondents feel disappointed (50%) or frustrated (27.3%) when emailing companies, highlighting negative emotional responses to email interactions.

Email Writing Challenges:

- Emotional expression is the most commonly noted challenge (66 respondents), followed by the tone of writing and complex language (44 respondents each), with clear keyword usage being a notable issue for 38 respondents.

c. Conclusion

This study has explored the performance of an email intent classification system designed to categorize emails into specific intents, such as Request, Query, Complaint, and Purchase, as well as to identify edge cases under the "Others" category. Through a structured methodology that involved purposive sampling of 210 real-world emails, the study aimed to evaluate the accuracy and effectiveness of the system in correctly predicting email intent. A variety of tools, including Zoho Sheets, Zoho Writer, Zoho Show, Zoho Analytics, and Python automation scripts, were employed to process, analyze, and present the data.

The findings of the study revealed that the system performed with an overall accuracy of 66.67%, which, while functional, highlighted several areas for improvement. Misclassifications were commonly observed, particularly with ambiguous emails and those containing overlapping intents. Specific issues identified included the misclassification of feedback requests as "Others," confusion between Request and Query intents, and bulk purchase inquiries being categorized as Requests. Additionally, the model struggled with handling emails that exhibited mixed or nuanced tones, further exacerbating the misclassification rates.

The study's analysis of confusion matrix metrics and formulated scores (including accuracy, precision, recall, and F1-score) provided valuable insights into the strengths and weaknesses of the classification system. Manual QA scoring revealed discrepancies between human judgment and model predictions, emphasizing the need for continuous improvement and validation.

Based on these findings, several recommendations were made to enhance the performance of email intent analysis models. These included addressing ambiguous intents by incorporating multi-label classification, improving the handling of mixed-tone emails, expanding the training dataset, and leveraging more advanced NLP techniques such as **BERT** for contextual understanding. Additionally, incorporating a feedback loop for continual retraining and model updates was advised to ensure the system adapts to changing language and communication patterns over time.

In conclusion, while the email intent classification system provides a solid foundation for automating email categorization, further improvements are necessary to achieve higher accuracy and better handle complex, ambiguous, and overlapping intents. The study provides actionable insights that can guide future advancements in email classification systems, with a focus on refining algorithmic performance and enhancing user trust through better transparency and feedback mechanisms.

d. Limitations

- The sample consisted of only 210 emails, which may not capture the full variability of real-world email communication.
- Actual intents were manually labeled, introducing subjectivity and potential inconsistencies in classification standards.
- The classification was confined to only five categories, limiting the system's ability to recognize nuanced or multi-intent emails.
- The model's reliance on specific keywords often led to misclassifications, especially in cases where tone or context altered the intended meaning.
- The system used a single-label classification approach, making it ineffective in handling emails that expressed more than one intent.

e. Scope for future study

The current study provides valuable insights into the performance and limitations of intent detection in categorized email communication. However, several avenues remain open for further exploration and improvement. Future research can expand the dataset to include a wider variety of industries, organizational communication styles, and languages, thereby improving the generalizability of the findings. Introducing **multi-intent classification models** would allow systems to better handle emails that contain more than one underlying purpose, such as a combination of a complaint and a request.

Additionally, the use of **advanced NLP techniques**—such as transformer-based models like BERT or GPT—could improve context recognition, tone understanding, and semantic analysis. These models may help overcome the current system's overreliance on surface-level keywords. Future studies could also incorporate **real-time user feedback loops** to continuously refine the classification logic based on actual usage patterns and corrections.

There is also scope for integrating **emotional and sentiment analysis more deeply** into the classification process, enabling the model to better distinguish between complaints and emotionally neutral queries. Lastly, future studies may evaluate the practical impact of classification performance on organizational outcomes, such as customer satisfaction, response time, and operational efficiency.

REFERENCES

1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019).
BERT: Pre-training of deep bidirectional transformers for language understanding.
arXiv preprint arXiv:1810.04805.

2. Zhang, Y., & Wallace, B. C. (2017).
A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification.
Proceedings of the 8th International Conference on Learning Representations (ICLR).

3. McCallum, A., & Nigam, K. (1998).
A comparison of event models for Naive Bayes text classification.
AAAI-98 Workshop on Learning for Text Categorization.

4. Bird, S., Klein, E., & Loper, E. (2009).
Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit.
O'Reilly Media, Inc.

5. Mihalcea, R., & Strapparava, C. (2005).
Making computers laugh: Investigations in automatic humor recognition.
Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing.

6. Jurafsky, D., & Martin, J. H. (2021).
Speech and Language Processing (3rd ed. draft).
Retrieved from https://web.stanford.edu/~jurafsky/slp3/

7. Zoho Corporation. (n.d.).
Zoho Sheet, Zoho Writer, Zoho Show, and Zoho Analytics Documentation.
Retrieved from https://www.zoho.com/

8. Pang, B., & Lee, L. (2008).
Opinion mining and sentiment analysis.
Foundations and Trends® in Information Retrieval, 2(1–2), 1–135.