TRIANZ℠

Intelligent Email Automation:

Leveraging Gen AI & Machine Learning for Smarter Categorization

Ravishankar Savita
Vice President - D&AI Practice Head

Kiran Maripelli
Director - D&AI Practice Leader

Table of Contents

1	**EXECUTIVE SUMMARY**	4
2	**INTRODUCTION: TAMING THE EMAIL BEAST**	5
	• The Relentless Growth of Enterprise Email: A Crisis in Control	5
	• Intelligent Email Triage: The Key to Unlocking Efficiency and Insight	5
	• Harnessing AI for Smarter Email Management: An Overview	5
3	**UNVEILING THE MARKET NEED AND STRATEGIC USE CASES**	6
	• From Chaos to Clarity: Real-World Applications of AI-Powered Email Triage	6
	• Revolutionizing Email Classification with GenAI	6
	• Optimizing Email Triage with Generative AI	6
	Transforming Pain Points into Opportunities: Reimagining Email Workflows	6
4	**CUSTOM ML MODEL FOR EMAIL CLASSIFICATION**	7
	• Deep Dive: The Power of Machine Learning for Accurate Classification	7
	• Key Features: Engineering a Solution for Scalable Email Understanding	7
	• Technical Architecture: A High-Level View of Model Training and Deployment on AWS Cloud	7
	• Performance Insights: Achieving Scalability and Accuracy	9
5	**CONTEXT-AWARE INTELLIGENCE WITH GENERATIVE AI**	10
	• Unlocking Context: How Gen AI Revolutionize Email Understanding	10
	• Key Features: Embracing Natural Language for Smarter Triage	10
	• Technical Architecture: Leveraging LLMs and Knowledge Bases	10
	• Performance Insights: Elevating Accuracy and Nuance	11

Table of Contents

6	**COMPARATIVE ANALYSIS: CHOOSING THE RIGHT ENGINE**	13
	• Traditional Approaches vs. AI-Powered Solutions	13
	• Custom ML Model vs Gen AI-based solution – Technical Showdown	14
	• Architecture: Design and Infrastructure	14
	• Training: Data Requirements and Model Development	14
	• Adaptability: Handling Evolving Needs	14
	• Language: Understanding Complexity	14
	• Resources: Cost and Efficiency	15
	• Performance Benchmarking: Accuracy, Speed, and Handling Ambiguity	15
	• Making the Right Choice: Key Considerations for Decision-Making	15
7	**QUANTIFYING SUCCESS: UNLOCKING TANGIBLE BENEFITS AND ROI**	16
	• Boosting Efficiency: Streamlining Operations and Reducing Costs	16
	• Elevating Customer Experiences: Delivering Faster, Smarter Responses	16
8	**CONCLUSION**	16
9	**CONTRIBUTORS**	17

1. Executive Summary

In today's digital landscape, enterprises face an unprecedented email management crisis, with teams spending up to **28%** of their work week managing inboxes. This overwhelming influx of emails hampers productivity leads to critical communication bottlenecks, and results in missed opportunities.

Email overload not only consumes valuable time but also introduces significant challenges in maintaining consistent categorization, ensuring timely responses, and adhering to regulatory compliance standards. The sheer volume of daily emails—often exceeding thousands a day for large organizations—makes it increasingly difficult to scale operations efficiently. This leads to scalability issues, inconsistent categorization, and delayed responses, further exacerbating the problem.

This white paper is articulated to cover the following topics:

1. **The Email Dilemma**: Organizations face numerous challenges in processing over **10,000** daily emails.
2. **AL/ML Driven Solutions**: This whitepaper introduces two solutions designed to transform email management from a daunting task into a strategic advantage.
    - First, leveraging a machine learning model on SageMaker, it is meticulously trained and deployed to handle structured data with precision.
    - Second, employing an advanced language model on Bedrock, it processes emails to discern intricate patterns and contexts.
3. **Comparative Analysis:** The ML model excels in efficiency and accuracy for structured data, while the LLM approach offers flexibility and nuanced understanding for complex queries.
4. **Quantifiable Benefits:** Implementing AI-powered email triage can achieve:
    - **95%** accuracy in email categorization
    - **60%** reduction in response times through automated routing
    - Enhanced compliance with regulations like GDPR, CCPA, and HIPAA
5. **Seamless Integration and Scalability**: Both solutions integrate smoothly with existing systems and are built on robust AWS services, ensuring scalability to meet growing demands.

As of 2025, AI email management has become increasingly sophisticated, with features like smart compose, intelligent reply suggestions, and advanced search capabilities. These solutions align with this trend, offering auto-classification and email summarization.

This whitepaper equips decision-makers with insights to select and implement the most suitable email triaging solution based on specific organizational needs, enabling them to effectively tackle email overload while enhancing efficiency and customer satisfaction in an ever-evolving digital landscape.

2. Introduction: Taming the Email Beast

The Relentless Growth of Enterprise Email: A Crisis in Control

The digital age has ushered in an era of unprecedented email communication. By 2025, the number of worldwide email users is projected to reach **4.6 billion**, with over **376 billion** emails sent daily. This staggering volume presents a critical challenge for enterprises, as employees spend an average of **28 minutes** on email every day. The sheer scale of email traffic has created a crisis of control, with businesses struggling to manage, prioritize, and extract value from their inboxes.

Intelligent Email Triage: The Key to Unlocking Efficiency and Insight

To combat email overload, intelligent email triage has emerged as a crucial solution. AI-powered triage systems can automatically classify, prioritize, and route emails, significantly reducing the time spent on manual sorting. This approach not only enhances productivity but also improves efficiency, allowing teams to focus on high-value tasks rather than drowning in their inboxes.

Harnessing AI for Smarter Email Management: An Overview

Artificial Intelligence is revolutionizing email management through advanced techniques. AI-driven email classification and summarization leverages natural language processing to understand email content, intent, and context. These systems can categorize emails based on urgency, department relevance, or custom criteria, ensuring that important messages receive prompt attention. By 2025, AI based email management is expected to significantly reduce email processing time, with some solutions achieving up to **95%** accuracy in categorization.

As enterprises grapple with the ever-increasing volume of digital communication, harnessing AI for email management is no longer a luxury—it is a necessity for maintaining operational excellence and competitive edge in the modern business landscape.

3. Unveiling the Market Need and Strategic Use Cases

From Chaos to Clarity: Real-World Applications of AI-Powered Email Triage

Leveraging AI for email triaging is already implemented in various industries, showcasing its potential. Some use cases where it has delivered value are provided below.

Revolutionizing Email Classification with GenAI

A leading insurance provider partnered with AWS's GenAI Innovation Center to automate email classification using foundation models, replacing manual processes with a high-accuracy, scalable solution.

Solution:

- Amazon Textract to extract key data from email text and PDFs.
- Anthropic's Claude models to classify emails with precision via Amazon Bedrock.
- A serverless architecture ensuring cost efficiency and seamless scaling.
- Integrated retrieval-augmented generation (RAG) for knowledge base searches and automated ticketing.
- Enabled explainability and feedback loops to refine AI-driven responses continuously.

Value Ad:

With **91%** accuracy achieved through advanced prompt engineering, the solution significantly reduced manual effort, enhancing operational efficiency, and adapting seamlessly to future needs—setting a new benchmark for AI-driven automation in document processing.

Optimizing Email Triage with Generative AI

A global enterprise leveraged AI-powered email classification to streamline customer support, enhance response efficiency, and improve customer satisfaction.

Solution:

- Implemented a GenAI-driven email triage system using Amazon Bedrock, AWS Lambda, LangChain, and ChromaDB on Amazon EC2.
- Automated email context extraction, classification, and sentiment analysis to prioritize customer inquiries.

Value Ad:

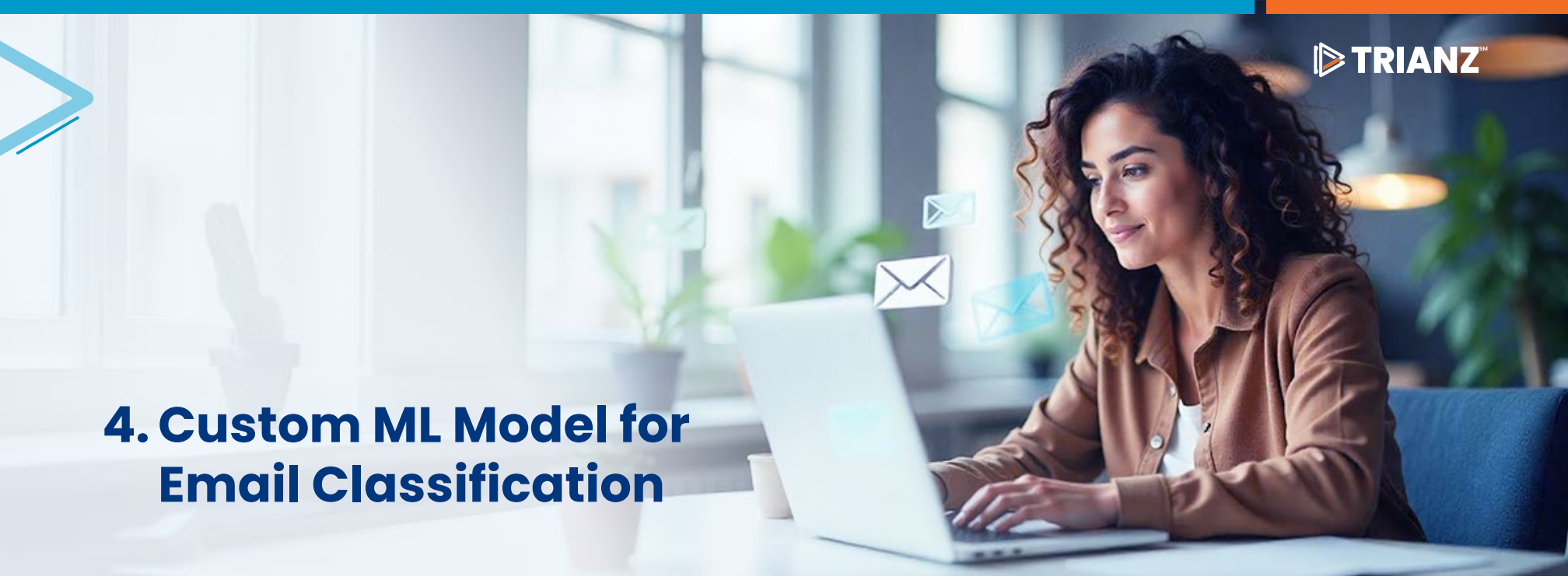
By reducing response times by **30%** and improving customer satisfaction (NPS +20 points), the AI solution enhanced operational efficiency, improved scalability, and freed up agents for complex customer interactions.

Transforming Pain Points into Opportunities: Reimagining Email Workflows

AI based email triage will revolutionize email management by:

- **Automated Categorization:** Using NLP and machine learning to assess email content, urgency, and importance, with some solutions achieving over **90%** accuracy in categorization.
- **Contextual Summarization:** Condensing lengthy emails to highlight key points, allowing for quicker comprehension and response.
- **Smart Prioritization:** Identifying and flagging critical communications while filtering out spam and low-priority messages.
- **Time Savings:** Reducing time spent on email management by up to **60%** through smart templates and automated responses for routine inquiries.
- **Enhanced Compliance:** Helping businesses navigate stricter privacy regulations by implementing robust data protection measures and transparent data usage policies.

By reimagining email workflows through AI, organizations can transform overwhelming email volumes into optimized, efficient communication channels, boosting productivity and improving customer experiences across industries.



4. Custom ML Model for Email Classification

Deep Dive: The Power of Machine Learning for Accurate Classification

Machine learning techniques have revolutionized email classification, offering powerful solutions for accurate and efficient email management. Recent studies have shown that custom ML models can achieve remarkable accuracy in spam detection and email categorization.

Key findings include:

- Logistic Regression, Convolutional Neural Network (CNN), and XGBoost classifiers achieving up to **98%** accuracy in spam detection.
- Random Forest technique reaching **99%** accuracy in email classification.

These results demonstrate the potential of ML models to significantly improve email triage and management processes.

Key Features: Engineering a Solution for Scalable Email Understanding

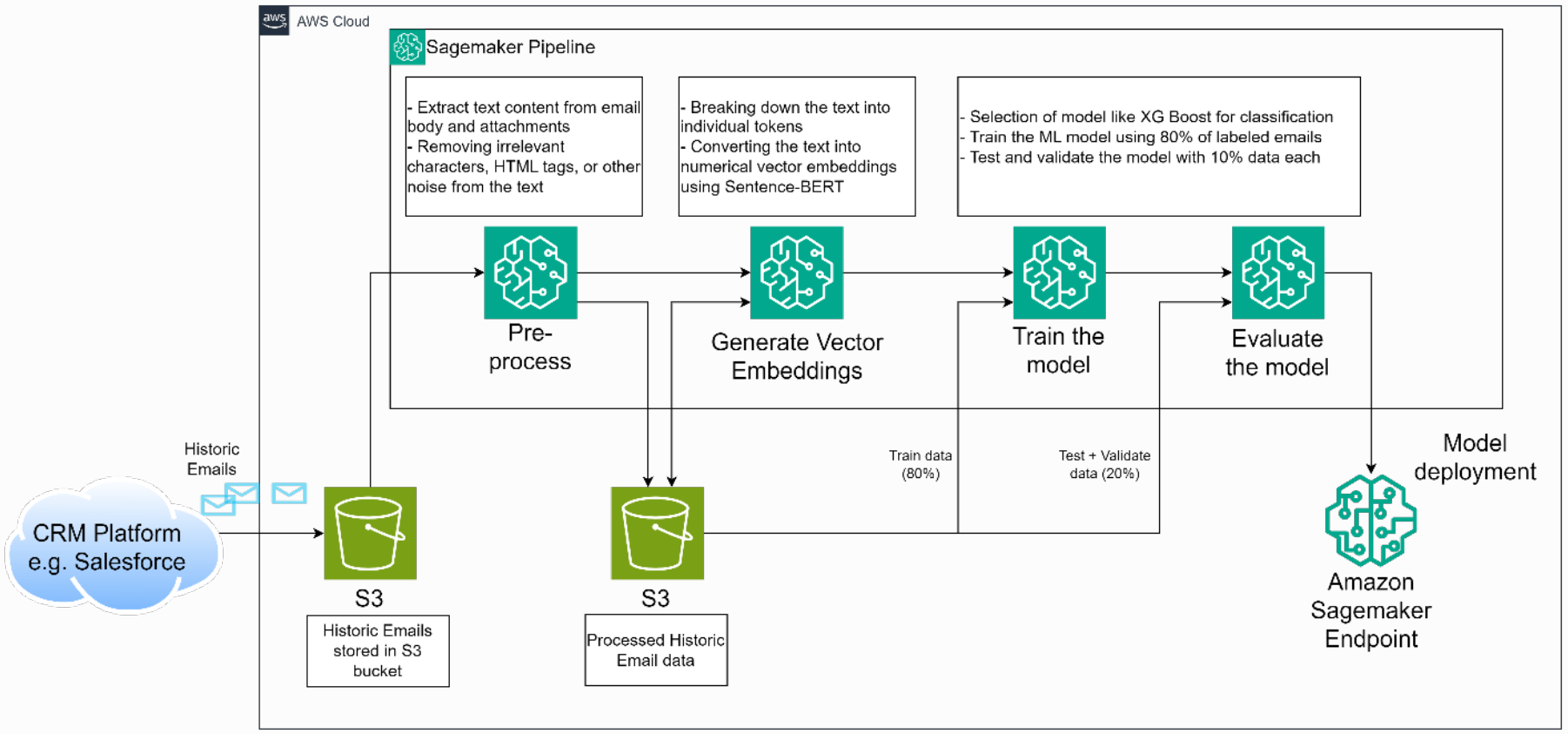
ML models for email classification incorporate several key features to ensure scalable and accurate email classification:

1. **SBERT-Based Classification:** Utilizing Sentence-BERT (SBERT) to capture contextual meaning, enabling more accurate email categorization by understanding semantic nuances.
2. **Multi-Model Optimization:** Employing an ensemble approach that evaluates multiple models, with Random Forest emerging as the top performer, achieving **98.81%** accuracy.
3. **Class Imbalance Handling:** Implementing up sampling techniques to ensure fair representation of all email categories, addressing the challenge of imbalanced datasets.

Technical Architecture: A High-Level View of Model Training and Deployment on AWS Cloud

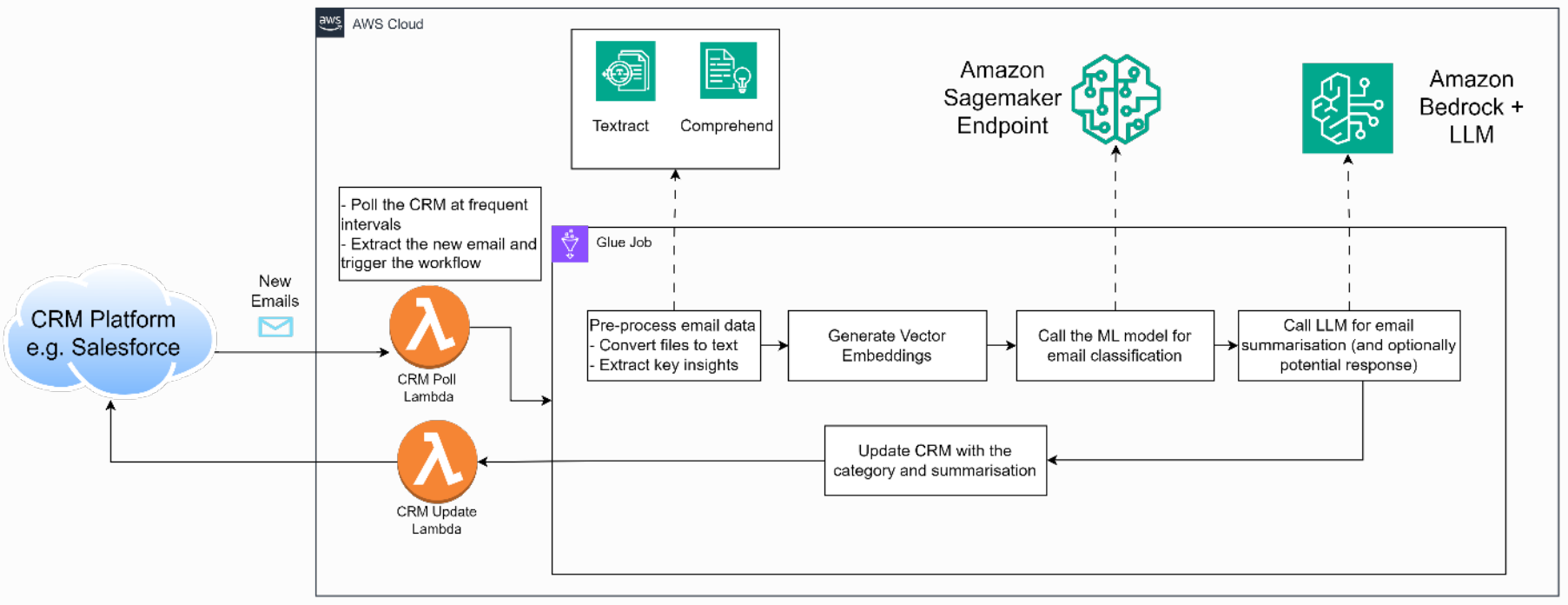
The architecture of a custom ML model for email classification and LLM for summarization typically follows these key steps:

- **Data Gathering:** Downloading the historic and new emails from any CRM system like salesforce.
- **Data Preprocessing:** Cleaning and normalizing email data using Amazon SageMaker Data Wrangler
- **Feature Selection:** Applying correlation techniques to select optimal features, such as using Amazon SageMaker Data Wrangler's built-in analysis for feature correlation.
- **Up-Sampling:** Performing up-sampling on underrepresented categories since the dataset contained class imbalance.
- **Vectorization:** Using SBERT (Sentence-BERT) Embeddings to convert textual data into numerical form and capture semantic meaning beyond simple keyword matching.
- **Model Selection:** Choosing the best-performing classifier based on accuracy metrics, which can be facilitated by Amazon SageMaker's model evaluation capabilities.
- **Model Training:** Using a large dataset (often 35000+ emails) to train multiple classifiers with Amazon SageMaker.
- **Model Deployment:** Integrating the model into the email processing pipeline on AWS.



- **New Email Processing:** Polling the CRM using Lambda at regular intervals. For any new emails, triggering the Glue job to retrieve and process the new email with following steps
    - Extract email content and attachments.
    - Use Amazon Textract to convert attachments to text.
    - Call Amazon Comprehend to generate key insights, including entity recognition and sentiment analysis.
    - Generate vector embeddings for the email content.
    - Call trained ML model on Sagemaker Endpoint for classification.
- **Summarization:** Utilizing email content, extract of attachments, key insights from comprehend, and calling LLM on Bedrock to generate summary of email, enhancing the overall efficiency of email management.
- **CRM Update:** Updating the source system with ML-generated categories and LLM generated summaries using Lambda function.

This architecture leverages various AWS services to create a robust and scalable email triaging solution, enabling businesses to efficiently manage and automate their email workflows.



Performance Insights: Achieving Scalability and Accuracy

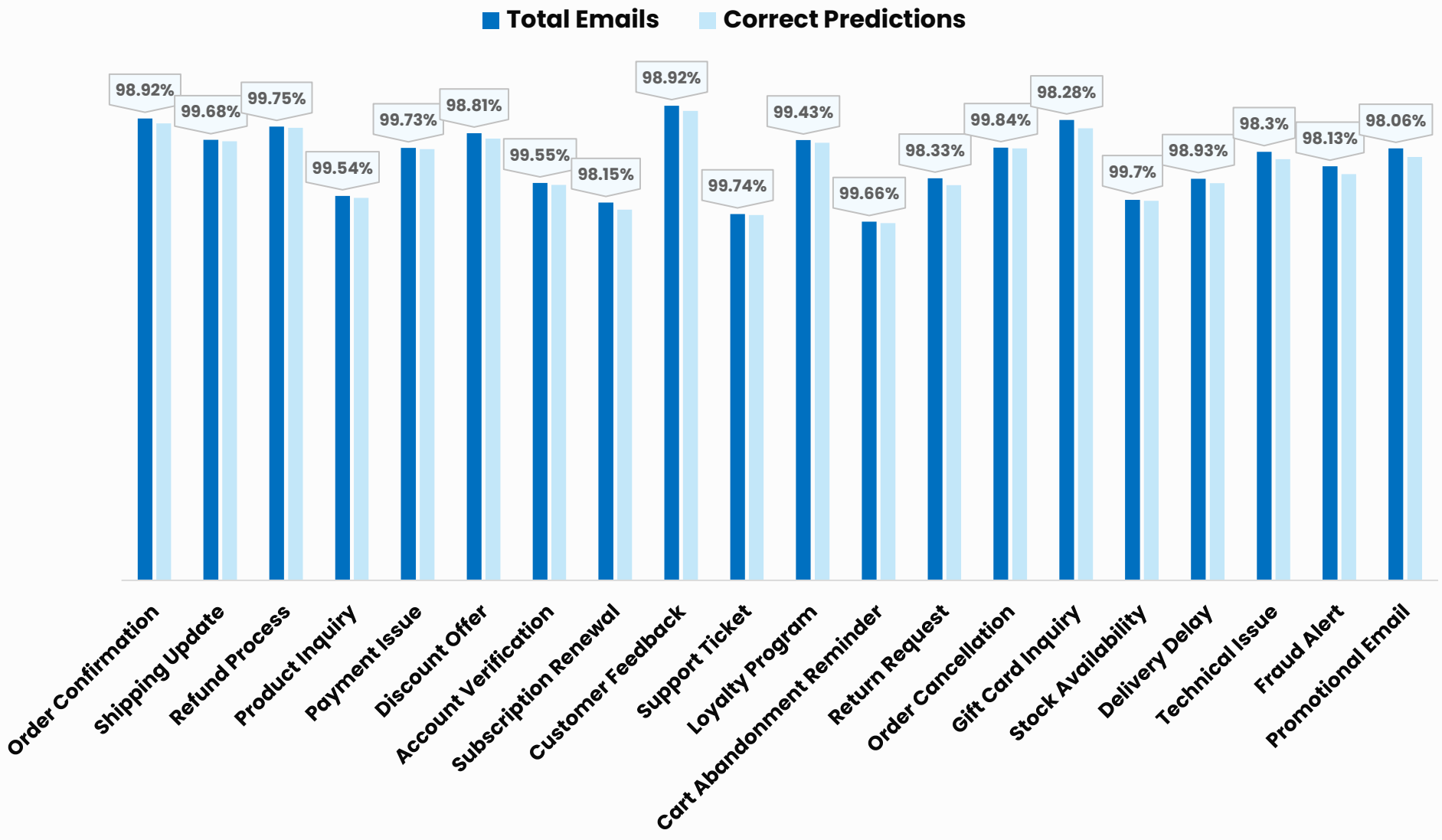
In the benchmarking exercise, the different models were benchmarked for the email classification task and model performance was evaluated using Accuracy, Precision, Recall, and F1-score. The results highlight the effectiveness of different models in achieving high performance in email classification tasks.

This shows that trained ML models can deliver impressive performance and provide a powerful, data-driven approach to email classification, providing organizations with tools to efficiently manage their email communications.

The following diagram shows performance ML Model developed with Random Forest algorithm (pretrained on large corpus of labelled emails) in classifying emails by department and product in an ecommerce scenario.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	98.81%	98.85%	98.81%	98.88%
XGBoost	98.55%	98.60%	98.55%	98.54%
Logistic Regression	95.44%	95.45%	95.44%	95.34%
Deep Learning	98.23%	98.29%	98.23%	98.22%

Correct Prediction percentage for the eCommerce Email Categories



5. Context-Aware Intelligence with Generative AI

Unlocking Context: How Gen AI Revolutionize Email Understanding

Large Language Models (LLMs) have transformed email understanding by offering superior context awareness and natural language processing capabilities. Unlike traditional rule-based systems, LLMs can parse and understand nuanced language without dedicated prior training, saving time and resources. This flexibility allows LLMs to overcome limitations of rule-based AI detection systems, enabling them to identify subtle cues in emails that may evade traditional detection methods.

Key Features: Embracing Natural Language for Smarter Triage

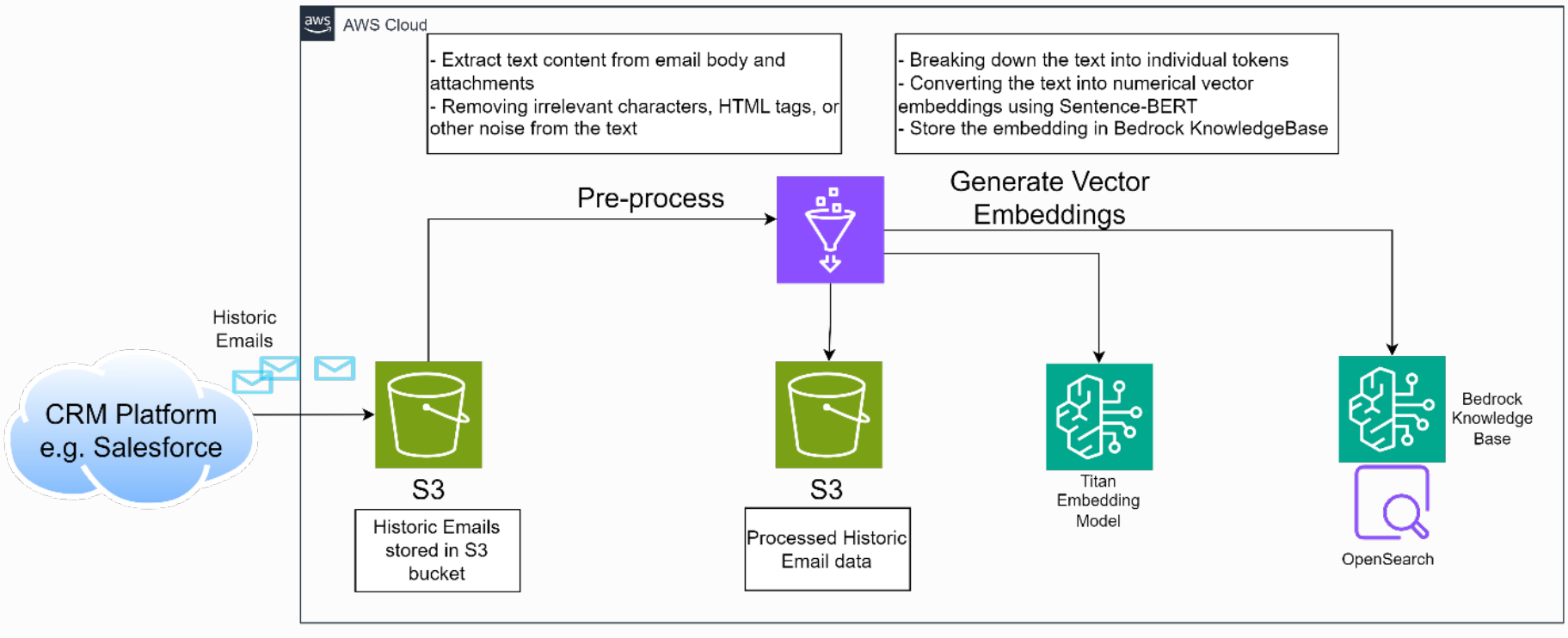
LLM-powered email triage systems offer several key features:

- **Natural Language Processing (NLP):** Understanding the context and intent behind emails, allowing for more accurate categorization, summarization and routing.
- Sentiment Analysis: Evaluating the tone and urgency of messages to prioritize accordingly.
- Content Relevance: Grouping emails by content similarity and importance.

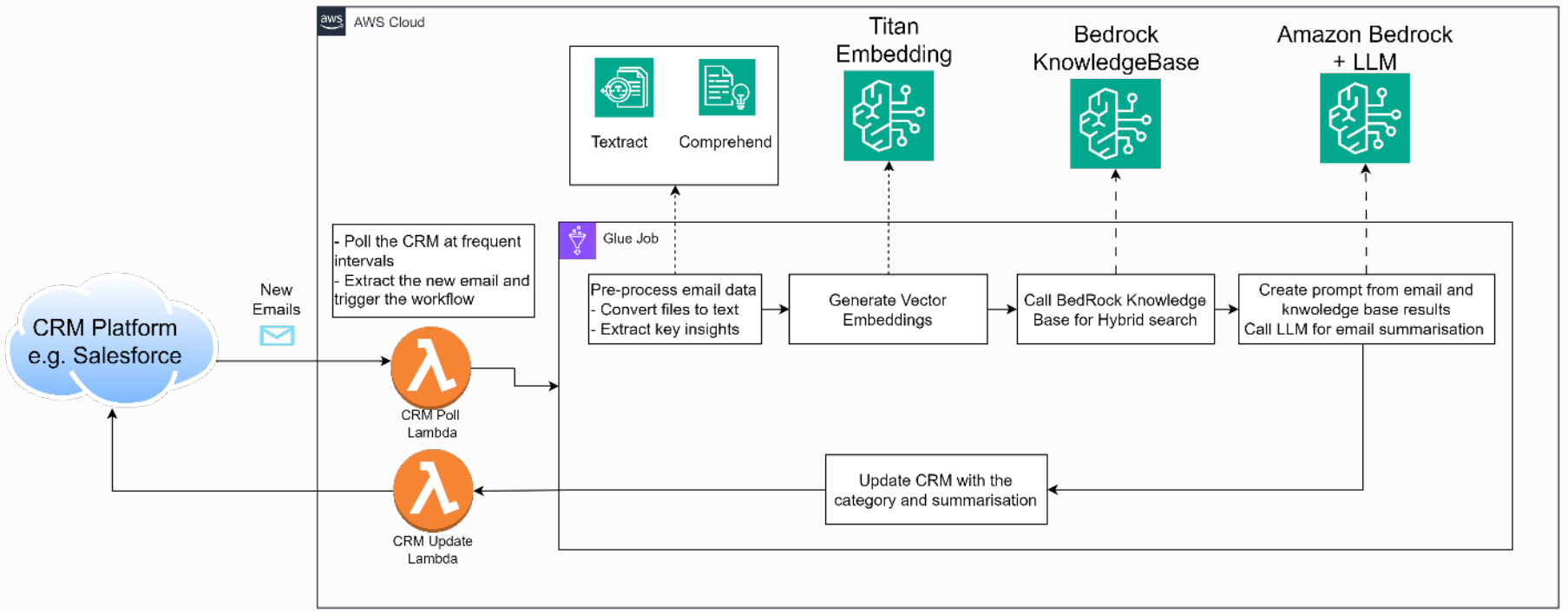
Technical Architecture: Leveraging LLMs and Knowledge Bases

The technical architecture for an LLM-based email triage system typically includes:

- **Data Ingestion:** Retrieving historical from CRM systems like Salesforce.
- **Vector Embedding:** Converting email content to vector embeddings using Amazon Titan Text Embeddings V2.
- **Knowledge Base Storage:** Storing vector embeddings in Amazon Bedrock Knowledge Base, leveraging its high-performance Retrieval-Augmented Generation (RAG) capabilities.



- **Content Processing:** Implementing an orchestrator using Glue job that processes every new email using AWS Textract to convert PDFs and images to text and using AWS Comprehend to generate key insights.
- **Hybrid Search:** Generating vector embeddings for the email content using Titan embedding model on Bedrock, calling Bedrock Knowledgebase for Hybrid search for any related emails and classification. This provides additional content to LLM during classification tasks.
- **LLM Analysis:** Utilizing Amazon Bedrock for email content analysis, categorization, and summary generation, achieving over **90%** accuracy in email categorization.
- **CRM Update:** Updating the source system with LLM-generated categories and summaries based on the analysis results using Lambda function.



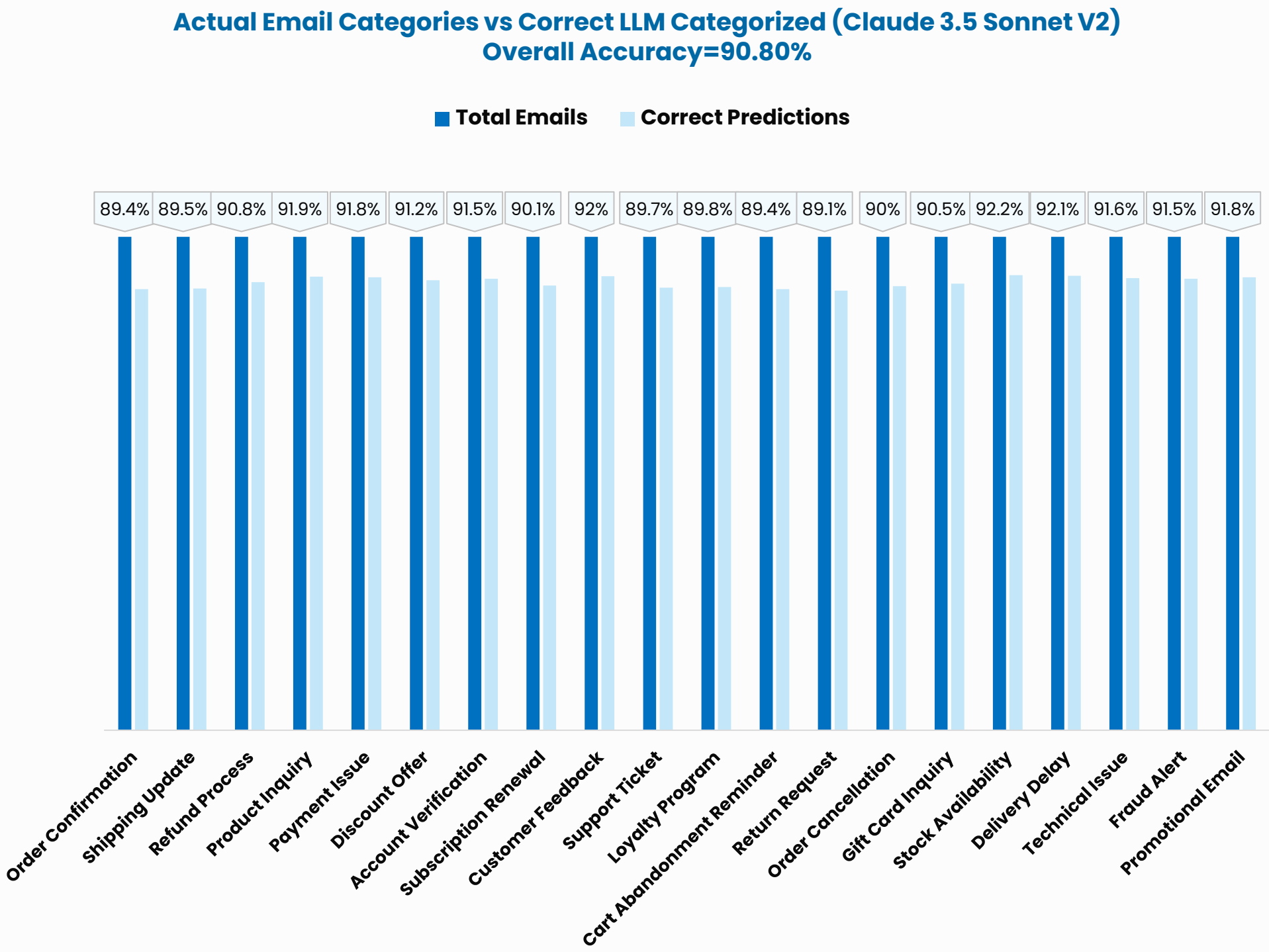
Performance Insights: Elevating Accuracy and Nuance

In the benchmarking exercise, Gen AI-based email triage systems have demonstrated impressive performance in managing and categorizing emails, offering significant proficiency in accuracy, contextual understanding, and adaptability. These systems, particularly those powered by advanced language models like Claude 3.5 Sonnet V2, can set new standards in email management efficiency.

The following table presents the performance metrics of various language models for email categorization:

Parameter	Claude 3.5 Sonnet V2	Llama 3.3 70B Instruct	Mistral Large (24.02)
Accuracy (%)	**90.80%**	**85.72%**	**85.45%**
Zero-shot Performance	High	Medium	Medium
Handling Similar Categories	High	Medium	Medium

The following diagram shows Claude 3.5's performance in classifying emails by department and product in an ecommerce scenario.



6. Comparative Analysis: Choosing the Right Engine

Traditional Approaches vs. AI-Powered Solutions

Starting with Traditional solutions, organizations face a critical choice between traditional methods and advanced AI/ML-powered solutions. This section explores the key differences between these two approaches, highlighting the advantages that ML and GenAI technologies bring to email processing.

Traditional Approach	AI-Powered Solution
Traditional systems rely on static, rule-based filtering, which can be rigid and prone to errors when encountering complex or unexpected scenarios.	AI models utilize data-driven classification techniques, leveraging large datasets to improve categorization accuracy. These models can recognize complex patterns in emails that traditional rules might overlook, offering a more robust and adaptable solution.
Traditional systems typically require manual intervention to prioritize emails, which can be time-consuming and prone to human error.	AI-driven predictive triaging eliminates this need by automatically assigning priority levels based on learned patterns and historical data.
Traditional approaches often involve fixed workflows that lack flexibility and cannot adapt to changing circumstances or evolving communication patterns.	AI-enabled systems continuously learn from feedback loops, allowing them to adjust workflows in real-time.
As new categories emerge, traditional systems may need drastic changes	As the Gen AI solution work based on the NLP engine and contextual understanding, they can adapt to the changes in classification requirements

Custom ML Model vs Gen AI-based solution – Technical Showdown

This comparison is specifically aimed at classification tasks, while summarization in both instances utilize LLMs.
A Side-by-Side Comparison of Key Capabilities of the two options are given in the table below.

ML-based solutions excel in efficiency and accuracy for well-defined categories, achieving **95-98%** accuracy in specific classification tasks. They are particularly effective for structured data and repetitive tasks, making them ideal for standardized email processing.

Capability	ML-based Solution	LLM-based Solution
Data Handling	Excels with structured data and specific tasks	Proficient with unstructured data and diverse language inputs
Language Understanding	Limited to predefined patterns	Advanced comprehension of context, nuance, and complex language
Adaptability	Requires retraining for new patterns	Easily handles new scenarios without extensive retraining
Task Specificity	Highly effective for well-defined, specific tasks	Versatile, capable of performing a wide range of language-related tasks
Setup and Training	Extensive setup and training phase needed	Quicker deployment with minimal setup for diverse tasks
Resource Requirements	Generally, more cost-effective for specific tasks	Requires more computational resources but offers greater flexibility
Error Handling	May struggle with variations or errors in data input	Can handle variations and even correct minor errors due to contextual understanding

LLM-based solutions offer superior flexibility and natural language understanding. They can comprehend complex queries and provide nuanced responses, making them adept at handling ambiguous cases. LLMs can extract relevant information from unstructured data sources, such as lengthy emails or complex PDFs, with impressive accuracy.

Architecture: Design and Infrastructure

ML-based solutions typically utilize specific models like Support Vector Machines (SVM) or Random Forest, deployed on cloud platforms such as AWS SageMaker. LLM-based solutions, on the other hand, leverage more complex neural network architectures, often using services like AWS Bedrock for deployment.

Training: Data Requirements and Model Development

ML models require large datasets of labelled emails (often **5000+**) for training. LLMs are pre-trained on vast amounts of text data and can be fine-tuned for specific tasks with smaller datasets.

Adaptability: Handling Evolving Needs

ML models need retraining to adapt to new patterns, while LLMs can more easily handle new scenarios without extensive retraining due to their broader knowledge base.

Language: Understanding Complexity

LLMs excel in understanding context, nuance, and complex language structures, including metaphors and cultural references. ML models are more limited in their language understanding capabilities

Resources: Cost and Efficiency

ML models are generally more cost-effective for specific, well-defined tasks. LLMs require more computational resources but offer greater flexibility.

Performance Benchmarking: Accuracy, Speed, and Handling Ambiguity

ML models can achieve **95-98%** accuracy in specific classification tasks. LLMs offer superior performance in handling ambiguous cases and understanding context, potentially reducing email processing time by up to **28%** of the workday.

Scalability: Meeting the Demands of Enterprise Growth

Both solutions leverage cloud platforms like AWS, ensuring scalability. ML models can efficiently handle large volumes of structured data, while LLMs are better suited for processing diverse, unstructured content.

Making the Right Choice: Key Considerations for Decision-Making

When choosing the most suitable solution for your business, it is essential to take into account several critical factors:

- **Task Specificity:** ML models for well-defined, specific tasks; LLMs for complex, varied language processing needs.
- **Data Availability:** ML models require large, labelled datasets, however, in return they offer better accuracy in classification of high domain specific content. In contrast LLMs can work with smaller, domain-specific data for fine-tuning providing reasonable accuracy.
- **Resource Allocation:** Consider volume of content for categorization, computational requirements (which tends to be higher for ML model hosting) and long-term costs.
- **Adaptability Needs:** LLMs offer greater flexibility for evolving language patterns.
- **Compliance and Security:** Evaluate data handling practices, especially for sensitive information.

By carefully considering these factors, organizations can select the AI engine that best aligns with their email management needs and overall business objectives.

7. Quantifying Success: Unlocking Tangible Benefits and ROI

Boosting Efficiency: Streamlining Operations and Reducing Costs

Implementing AI-powered email management solutions can lead to significant operational efficiencies and cost reductions:

- **Time Savings:** AI-driven email triage can reduce email processing time by up to **28%** of the workday, freeing up valuable employee time for more strategic tasks.
- **Increased Productivity:** Automated categorization and routing of emails can improve overall team productivity by **20-30%.**
- **Cost Reduction:** By automating routine email tasks, organizations can potentially save **30-40%** on email management costs.

Elevating Customer Experiences: Delivering Faster, Smarter Responses

AI-powered email solutions significantly enhance customer experience through:

- **Faster Response Times:** AI can reduce average response times by up to **60%**, improving customer satisfaction.
- **Personalization:** AI can analyze customer history and preferences to provide more tailored responses, increasing engagement by up to **25%**.
- **24/7 Availability:** Automated responses ensure round-the-clock customer service, improving customer satisfaction scores by up to **15%**.

8. Conclusion

The future of email management is poised for significant transformation, driven by advancements in AI and intelligent automation. By 2025, we can expect to see:

- AI-powered systems that can understand complex email content, context, and sentiment, enabling more accurate and nuanced responses.
- Automated email triage systems that can categorize, prioritize, and route emails with high accuracy, significantly reducing manual workload.
- Advanced language models capable of generating personalized, context-aware responses to routine inquiries, freeing up human resources for more complex tasks.

These innovations promise to revolutionize email management, making it more efficient, accurate, and responsive to evolving business needs.

Contributors

Apurva Anand - Solutions Lead

Tarun Sheoraj –Architect

Vijayprakash Mundunuru – Architect

Naresh Jampala - Solutions Lead

Yaseen Mohammed - Principal Architect

Ramandeep Singh - Presales Lead

Arvind Chopra – Client Partner

Ravishankar Savita
Vice President –
D&AI Practice Head

Kiran Maripelli
Director –
D&AI Practice Leader

Contact Information

Trianz D&AI Practice

D-AI-Innovation-Team@Trianz.com