Data Science Capstone | Project 3

Mark Hassan

December 20, 2025

Overview

The customers of Acme Communications (who own a popular messaging platform), are consistently sharing feedback that the platform is overwhelming and hard to keep up with for long threads.

Acme's product team has identified automated conversation summarization as a solution to this user experience challenge.

Stakeholders

- Acme product team: overall responsibility for the performance of the messaging product and accountable for any new features. They will effectively be the test group for this new summarization product.
- Acme legal/compliance team: provide important input on the use of new technologies, especially those that create content which could be seen as "company owned".
- Acme product users: they will determine the ultimate success of any new feature. The product team represents them during the product development process, but their feedback and usage post-launch guide sustainment decisions.

Primary Goals

My directive is to build a prototype conversation summarization tool for the Acme messaging platform. This prototype should:

1. Provide accurate and succinct summaries of long conversations/threads
2. Allow a user to stay current with more conversations
3. Demonstrate the value that AI powered features can bring to the product.

Problem Statement

The free version of Acme's messaging platform allows users to participate in 20 group chats with up to 50 participants in each. The average user of the free version has 5 active chat groups with an average of 10 people in each.

During active times (30% or more participants online), conversations often generate 100+ individual messages each hour. The average user of Acme's platform has 1 or more conversation thread with 1000+ unread messages since their last session every 5th time they log in.

The mean time to review 100 unread messages is 5 minutes. 95% of users will spend up to 10 minutes reviewing unread messages with no issues. After 10 minutes, they begin to feel overwhelmed; 90% of users will stop reviewing messages at this time.

Users that report feeling overwhelmed by the volume of unread messages on 5 or more separate instances:

- Log into the Acme platform 60% less frequently in the next 2 weeks
- Are 30% less likely to upgrade to paid memberships
- Are 45% more likely to close their account.

This results in missed opportunities to convert free users to paid users and increases pressure on customer acquisition teams to replace users that leave.

Acme Communications needs to create a feature that allows users to see what they've missed without reading every unread message. A successful implementation of this feature will improve customer retention and increase conversions from free to paid users, directly increasing the company's user base and revenues.

Proposed Solution

I will use the Bidirectional Encoder Representations from Transformers (BERT) model to build a prototype conversation summarization tool. The BERT architecture is ideally suited to this kind of task for the following reasons:

1. BERT built on transformer architecture, which is a type of neural network that can ingest entire sequences and weigh the relevance of each component each sequence.
2. BERT was developed to specialize in language understanding tasks. Some of its unique capabilities include:
    a. The ability to process text left to right and right to left, giving it robust context for any given component of a sequence
    b. Built-in training on new datasets (called masked language modeling and next sentence prediction) that enable it to quickly learn the nuances of a given text corpus.
3. There are a wide range of pre-trained BERT models that make the selection and implementation of BERT for this project relatively fast and simple.

Using BERT for this prototype project enables me to select a pre-trained model, customize it for Acme's requirements, then directly proceed to fine-tuning the model on the Acme dataset.

Success Criteria

Success for this prototype project will be measured in a few different ways.

The technical performance of this model will be measured using:

- ROUGE-N: measures the word/phrase overlap between the reference text and the summarization
- ROUGE-L: identifies the longest overlapping phrases between the reference text and the summarization

The target for both ROUGE-N and ROUGE-L is 90%. In simplified terms, this means that the model should produce summarizations that overlap by at least 90% with the original/reference text.

For the user experience, the goal is to allow users to review unread messages more quickly. Our target is a tool that outputs summaries of up to 1000 messages which users can review in 10 minutes or less. This would enable users to review threads 5x faster than the current state, reducing the risk of user attrition.

Finally, the model should be able to complete a summarization of up to 1000 messages in 10 seconds or less (a secondary measure of user experience)

Timeline

	Start	End	Hours	Notes
Planning & organization	12/22	12/27	6	Includes project timeline and milestones and defining requirements
Select and prepare the BERT model	12/28	12/28	4	
Data preparation and preprocessing	12/29	12/29	10	
Model architecture design	12/30	12/30	10	Identify and implement task-specific layers needed on top of BERT
Fine-tuning the model	12/31	1/1	10	
Evaluation and refinement	1/1	1/2	6	
Finalize and deploy	1/2	1/3	6	Includes documentation and presentation of results