Text Summarization Task

Annotation Guidelines (Version 1.0)

---

1. Introduction

This document provides detailed instructions for annotating **summaries of Hindi financial text**.

The objective of this task is to **verify the quality of an existing summary** with respect to the provided source content and, when necessary, **edit the summary to correct errors**.

You will evaluate whether a summary:

- accurately reflects the key information in the source text, and

- does not introduce incorrect, misleading, or unsupported information.

This task focuses on **verification and targeted editing**, not writing a new summary from scratch.

You should read this document carefully before beginning the annotation.

---

2. Onboarding

You will be given access to the annotation workspace during the annotation period. Click on the Summarization page. Here you will find the link to the platform and guidelines.

There will also be a table where you will find your name. When you hover over your name, you will find the 'open' option. That will lead you to your page (image below)
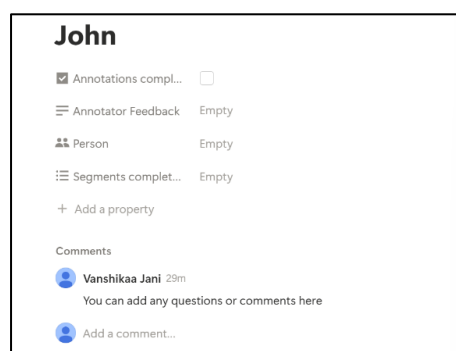
From the workspace, you will find a link to the **Text Summarization annotation platform** (with password), this guideline document, and a checklist for tracking your assigned segments.

For communication, you may email questions or completed annotations to Vanshikaa.Jani@mbzuai.ac.ae or leave comments in the shared workspace (Notion)

3. Annotation Workflow

You will complete the task using a dedicated online interface.
The workflow is as follows:

1. Login

- Enter your name and password on the login screen.

2. Review Guidelines

- Read this document within the interface.

- Confirm that you have read and understood the guidelines before proceeding.

3. Segment Selection

- Data are divided into predefined segments.

- Begin with the pilot segments.

- Complete all pilot segments before starting remaining segments.

4. Summary Verification

For each annotation item:

- Read the **Hindi source content** carefully.

- Read the **provided summary**.

- Complete the verification steps described in Section 4.

5. Navigation

- **Previous** allows revisiting earlier items.

- **Save & Next** records the annotation and advances.

- A **Jump to Item** option in the sidebar enables direct navigation.

- A **progress indicator** shows how many items in the segment are completed.

6. Translation Option (Optional)

- Translation buttons are available for:
    - the source content
    - the summary

- Use translation only when necessary for understanding.

- All translation usage is logged automatically.

7. Download

- After completing a segment, download the CSV file.

- Send the file to the designated email or upload it to the workspace.

---

4. Summary Verification Task

Each annotation item consists of:

- **Source Content** (Hindi text)

- **Provided Summary** (Hindi)

Your task has two steps.

4.1 Step 1: Summary Quality

You must decide whether the provided summary is **faithful and adequate** with respect to the source content.

Select one of the following:

- **Yes**
  The summary accurately represents the key information in the source content. It does not contain factual errors or misleading statements.

- **No**
  The summary contains errors, omissions of critical information, or misleading content.

- **Uncertain**
  It is unclear whether the summary is fully correct due to:

    o insufficient clarity in the source content, or

    o ambiguity in how the information is summarized.

This decision should be made **before editing the summary**.

4.2 Step 2: Summary Editing (If Applicable)

This step applies **only if you selected "No" in Step 1**.

- You will be shown the existing summary in an editable text box.

- Edit **only what is necessary** to correct errors or omissions.

- Do **not** rewrite the summary entirely unless absolutely required.

- Preserve the original style and structure as much as possible.

If you selected **Yes** or **Uncertain**, do **not** edit the summary.

---

5. What Makes a Good Summary (Evaluation Criteria)

A summary marked **Yes** should satisfy all of the following:

1. **Faithfulness**

   - All statements are supported by the source content.

   - No new facts are introduced.

2. **Coverage**

   - Major points in the source content are represented.

   - Minor details may be omitted.

3. **Clarity**

   - The summary is understandable and coherent.

   - Information is not distorted by wording.

---

6. Use of "Uncertain"

The **Uncertain** option should be used when you are not sure if this is an adequate summary for the given passage.

---

7. Principles for Annotation

1. Base all decisions strictly on the source content provided.
   Do not use external knowledge.

2. Do not infer missing information.
   If information is not clearly present, treat it as unavailable.

3. Judge faithfulness before judging completeness.

4. Use translation only as a comprehension aid.
   All decisions must be grounded in the Hindi text.

5. Apply criteria consistently across all segments.

8. Interface Reference Guide

8.1 Source Content Panel

- Displays the full Hindi source text

- Source translation button

8.2 Summary Panel

- Displays the provided summary

- Summary translation button

8.3 Verification Panel

- Step 1: Summary quality (Yes / No / Uncertain)

- Step 2: Summary editing (if applicable)

8.4 Sidebar Tools

- Jump to Item

- Segment progress indicator

8.5 Segment Completion

- Download CSV file after completion

- Partial progress can be downloaded if needed

- Multiple CSVs may be submitted if annotation is split across sessions

9. Quality Expectations

You are expected to:

- Apply summary evaluation criteria consistently

- Avoid unnecessary edits

- Use Uncertain appropriately

- Complete all assigned segments

10. Contact and Support

For questions or clarification, contact:
Vanshikaa.Jani@mbzuai.ac.ae

Or use the comment section in the shared workspace on Notion.

---

11. Completion

Click on Download CSV to download after each segment **(your work will not get saved automatically).**

All changes are saved locally in the session until explicitly downloaded.

A manageable size of 10 per segment is chosen but if you have to stop early then you may download the current progress. When you return, use the jump option to skip to the last annotated fact in the segment. In such cases you can upload multiple CSVs.

Upload the datasets here

Example:

pilot_segment_1_Vana.csv 6.8 KiB

Or email it.

To upload it on Notion:

1. Click on + option

   Write, press 'space' for AI, '/' for commands...

2. Then scroll down till you find "file"

   Image
   Video
   Audio
   Code
   File
   Web bookmark
   Database

   Type '' on the page   esc

   Type to filter...