Newsletter-Factory: A Thematic Newsletter Generation Tool for Business Insights

- Newsletter:
  - Periodic publication of news
  - Based on a specific theme
  - Personalized stories

- Benefits for enterprises:
  - Track evolving business & tech.
  - Contribute to success & growth
  - Reduce information overload

- Newsletter-Factory Tool:
  - Automation tool
  - Near-duplicate detection
  - Metadata classification
  - Fine-grained classification using sub-categories
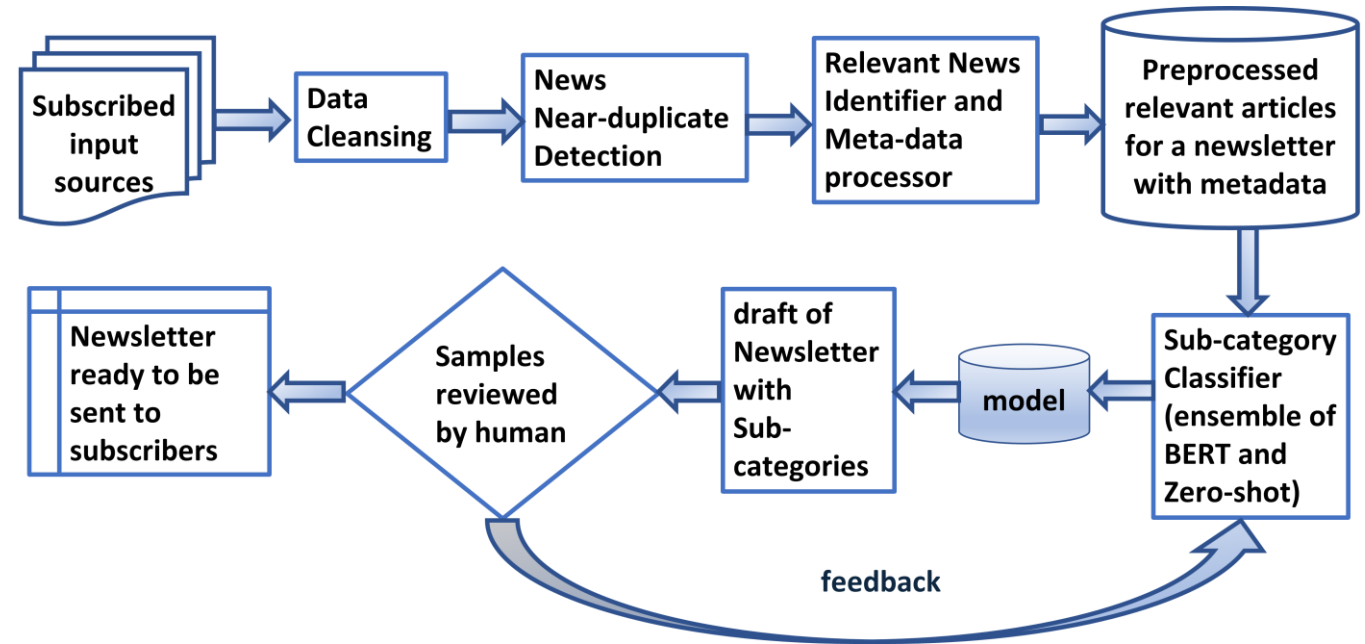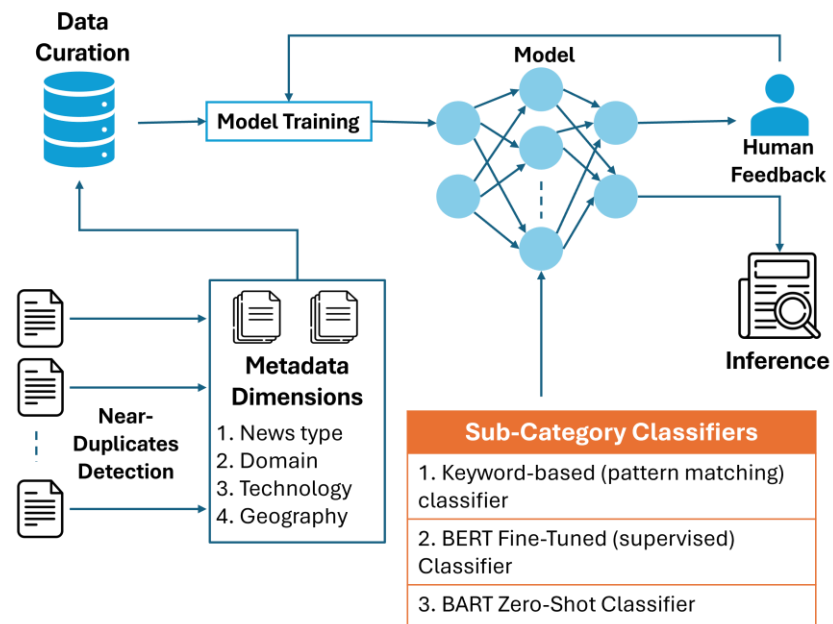 - Human-in-the-loop for quality check and model improvement

Figure: High-level block diagram of key components in Newsletter-Factory

tcs Research

Key Components of Newsletter-Factory Tool

- Near-duplicate detection and handling
  - Locality Sensitive Hashing
  - MinHash and Shingling

- Metadata classification
  - Pattern-based categorizers
  - ML classifiers
  - Identify newsletter theme

- Sub-category classification
  - Keyword-based classifier
  - Adapter fine-tuned BERT classifier
  - BART zero-shot classifier

- Human-in-the-loop
  - Inspection of draft
  - Ability to provide feedback
  - Model improvement using feedback

Figure: Illustrative example of sub-category classification step for enrichment of raw news data

tcs Research

tcs Research

Newsletter-Factory: A Thematic Newsletter Generation Tool for Curating Business Insights

Authors: Siddarth Tumre, Alok Kumar, Ajay Phade, Nihar Riswadkar, Sangameshwar Patil

Presenter: Ajay Phade

April 10th, 2025

Building on belief