

1. Text summarization

- a. Problem: Whether for work or pleasure, we are often tasked with reading long, time consuming articles, messages or reports. The purpose of a text summarization model is to provide a brief, abstract summary instead of a whole article. This could free up significant time for greater productivity or leisure.
- b. Data: The data would come from the [CNN / DailyMail dataset](#) found on huggingface.

2. Financial News Sentiment Analysis

- a. Problem: Investors need to have a pulse on market sentiment to make informed investing decisions. However, it can be very time consuming to read through many articles to get a feel for market sentiment. A sentiment analysis could provide a concise assessment of market sentiment without having to read through many articles and reports.
- b. Data: The data would come from the [Sentiment Analysis for Financial News](#) dataset found on Kaggle.

3. Credit Risk Modeling

- a. Problem: Banks and other lenders need to assess credit risk of potential borrowers to minimize loss from defaults. An effective credit risk model can help identify borrowers that have a high risk of default.
- b. Data: The data would come from the [Loan Prediction Dataset 2025](#) found on Kaggle.