WHAT IS TEXT CLASSIFICATION?

Text classification plays a pivotal role in digitizing a wide variety of modern industries. Also sometimes referred to as text tagging or text categorization, text classification describes the process of arranging text into specific, organized groups by assigning text a label or class.

Using text classifications helps automate many business processes, such as customer support, survey analysis, sentiment analysis, and document summarization, and more.

Text classification has drastically evolved over time, shifting away from traditional machine learning (ML) models that need large amounts of data to Large Language Models (LLM) that require only a handful of examples for model training.

Text classification is a machine learning technique that assigns a set of predefined categories to Open – ended text. Text classifiers can be used to organize, structure, and categorize pretty much any kind of text – from documents, medical studies and files, and all over the web.

Text classification is one of the fundamental tasks in natural language processing with broad applications such as sentiment analysis, topic labelling, spam detection, and intent detection.

WHY IS TEXT CLASSIFICATION IMPORTANT ?

It's estimated that around 80% of Information is unstructured, with text being one of the most common types of unstructured data. Because of the messy nature of text, analysing, understanding, organizing, and sorting through text data is hard and time-consuming, so most companies fail to use it to its full potential.

This is where text classification with machine learning comes in. Using text classifiers, companies can automatically structure all manner of relevant text, from emails, legal documents, social media, chatbots, surveys, and more in a fast and cost-effective way. This allows companies to save time analyzing text data, automate business processes, and make data-driven business decisions

REASONS FOR USING MACHINE LEARNING TEXT CLASSIFICATION :

- Scalability
- Real time analysis
- Consistent Criteria

WAYS TO APPLY TEXT CLASSIFICATION:

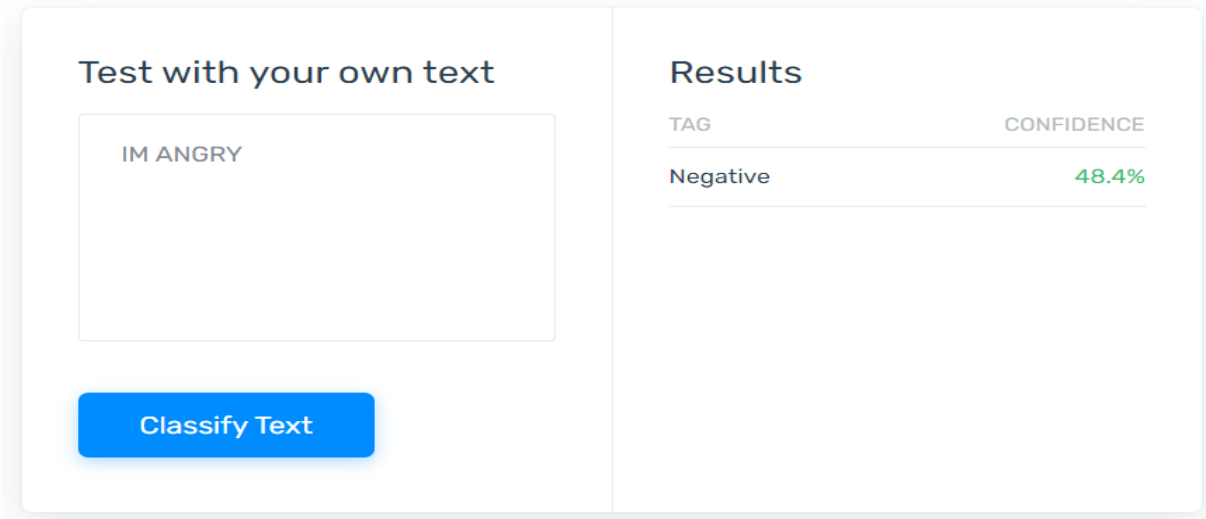
THERE ARE 2 WAYS:-

1. MANUAL - Manual text classification involves a human annotator, who interprets the content of text and categorizes it accordingly. This method can deliver good results but it's time-consuming and expensive.
2. AUTOMATIC - Automatic text classification applies machine learning, natural language processing and other AI guided techniques to automatically classify text in a faster, more cost-effective, and more accurate manner.

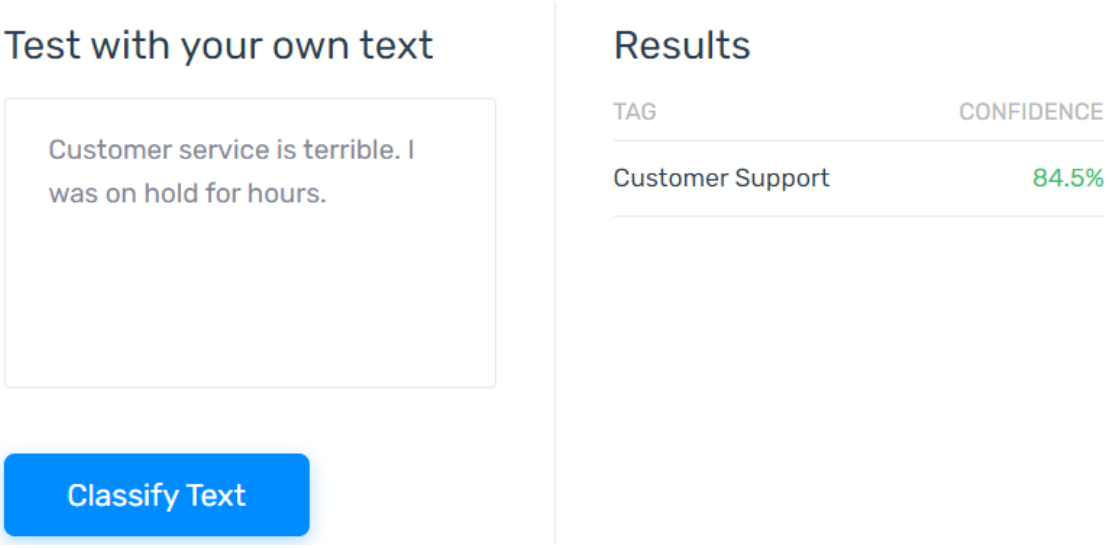
SOME EXAMPLES OF TEXT CLASSIFICATIONS ARE:-

Some of the most well-known examples of text classification include sentiment analysis, topic labelling, language detection, and intent detection.

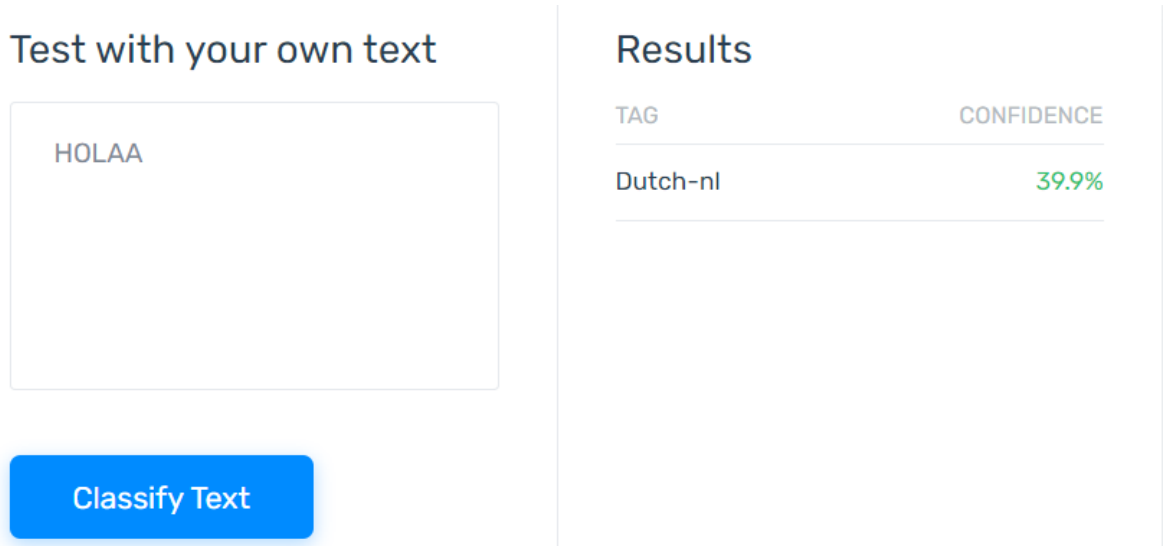
1. SENTIMANT ANALYSIS:



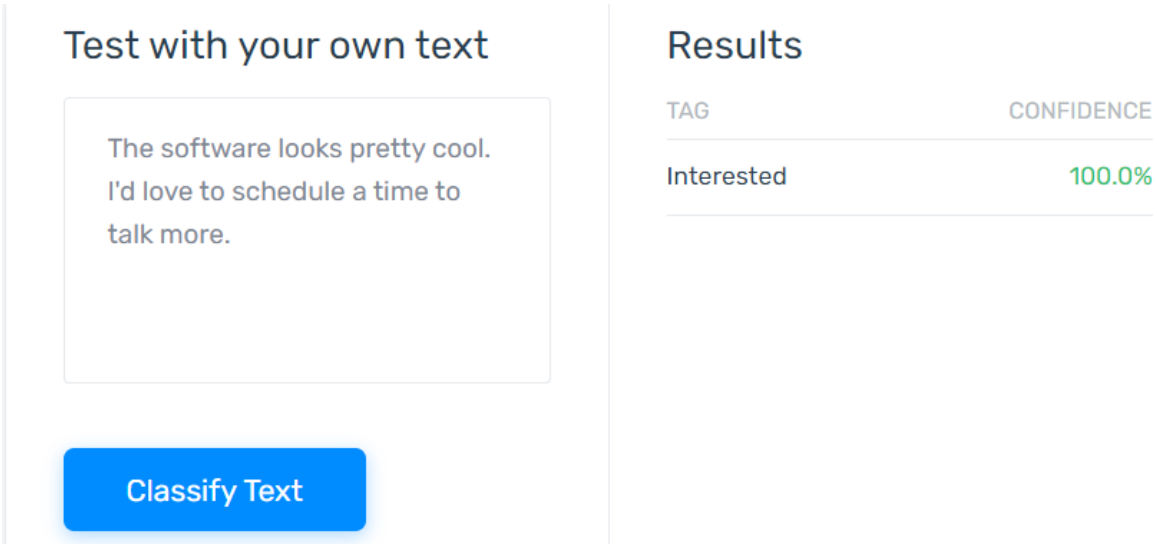
2. TOPIC LABELLING:



3. LANGUAGE DETECTION:



4. INTENT DETECTION:



RESOURCES USED IN TEXT CLASSIFICATION: -

1. DATASETS –
   - → REVIEWS
   - → SPAMS
   - → NEWS DATASET
2. TEXT CLASSIFICATION TOOLS –
   - → OPEN-SOURCE LIBRARIES
   - → SAAS APIs
     - PYTHON
     - JAVA
     - R
     - MONKEYLEARN
     - IBM WATSON
     - GOOGLE CLOUD NLP