**ABSTRACT**

**Text Classification using NLP**

Text classification is the process of assigning tags or categories to text according to its content. It’s one of the fundamental tasks in [Natural Language Processing](https://monkeylearn.com/blog/definitive-guide-natural-language-processing/) (NLP) with broad applications such as sentiment analysis, topic labeling, spam detection, and intent detection.

What is Text Classification?

Text classification is the task of assigning a set of predefined categories to free-text. Text classifiers can be used to organize, structure, and categorize pretty much anything. We can classify Emails into spam or non-spam, news articles into different categories like Politics, Stock Market, Sports, etc.

Text classification can be done in two different ways:

1. Manual classification.
2. Automatic classification.

Text Classification Algorithms

Some of the most popular machine learning algorithms for creating text classification models include the naive Bayes family of algorithms, support vector machines, and deep learning.

How Text Classification Can Be Used:-

It can be explained by the project in which we will be predicting the movie genres (Action, Adventures, Horror, Comedy, Romantic, etc.) on the base of the overview of the movies.

In this project we will be doing following steps:

1. At first we take a dataset containing movie name, movie genres, and movie overviews.
2. Then take genre column and check no of factors in that column we must be making that factor as a column attribute.
3. Now here comes the text classification
4. First take overview column.
5. Now separate the sentences by considering every single words.
6. Remove stop words using Stop Words Removal.
7. Left words are key words.
8. Make a dictionary of genres n related words to it.
9. Compare the keywords and dictionary words.
10. According to the similarities between them we can predict the generes

And for this whole step 3 we must be using different text classifying algorithms.

So in this way text classification is done.