

# Text Classification Using R

## What makes a good post on StackOverflow?

Julian Aviles

### 1. Introduction

Text classification, also known as text tagging, is the process of categorizing text into predefined groups. This area of text mining is significant in many real-world applications. Some tasks, such as manually labeling thousands of documents, are labor intensive and require specialized training. Using text classification techniques, we can eliminate some of these costs, while at the same time often reducing the error rate. Prior research has been done in this area, for example . Another interesting application is .

### 2. Method

This project will focus on using machine learning techniques to classify user-submitted post on the website StackOverflow.com. The posts are