Data Mining for Tweet Sentiment Classification

By Roy de Groot

LAP Lambert Academic Publishing Nov 2012, 2012. Taschenbuch. Book Condition: Neu. 221x152x12 mm. Neuware - The goal of this work is to classify short Twitter messages with respect to their sentiment using data mining techniques. Twitter messages, or tweets, are limited to 140 characters. This limitation makes it more difficult for people to express their sentiment and as a consequence, the classification of the sentiment will be more difficult as well. The sentiment can refer to two different types: emotions and opinions. This research is solely focused on the sentiment of opinions. These opinions can be divided into three classes: positive, neutral and negative. The tweets are then classified with an algorithm to one of those three classes. Known supervised learning algorithms as support vector machines and naive Bayes are used to create a prediction model. Before the prediction model can be created, the data has to be pre-processed from text to a fixed-length feature vector. The features consist of sentiment-words and frequently occurring words that are predictive for the sentiment. The learned model is then applied to a test set to validate the model. 108 pp. Englisch.

READ ONLINE
[1.1 MB]

Reviews

It becomes an incredible book that we actually have possibly study. It really is rally exciting throgh studying period of time. I am very easily could get a satisfaction of reading through a written book.
-- Gianni Hoppe

A really awesome pdf with perfect and lucid reasons. It is actually rally fascinating throgh reading period of time. Your lifestyle period will probably be transform as soon as you total looking over this ebook.
-- Alford Kihn