

CHAPTER 3

Document Sentiment Classification

Starting from this chapter, we discuss the current major research directions or topics and their core techniques. Sentiment classification is perhaps the most extensively studied topic (also see the [Pang and Lee, 2008](#)). It aims to classify an opinion document as expressing a positive or negative opinion or sentiment. The task is also commonly known as the *document-level sentiment classification* because it considers the whole document as a basic information unit. A large majority of research papers on this topic classifies online reviews. We thus also define the problem in the review context, but the definition is also applicable to other similar contexts.

Problem definition: Given an opinion document d evaluating an entity, determine the overall sentiment s of the opinion holder about the entity, i.e., determine s expressed on aspect GENERAL in the quintuple

$$(_, GENERAL, s, _, _),$$

where the entity e , opinion holder h , and time of opinion t are assumed known or irrelevant (do not care).

There are two formulations based on the type of value that s takes. If s takes categorical values, e.g., positive and negative, then it is a classification problem. If s takes numeric values or ordinal scores within a given range, e.g., 1–5, the problem becomes regression.

To ensure that the task is meaningful in practice, existing research makes the following implicit assumption ([Liu, 2010](#)).

Assumption: Sentiment classification or regression assumes that the opinion document d (e.g., a product review) expresses opinions on a single entity e and contains opinions from a single opinion holder h .

In practice, if an opinion document evaluates more than one entity, then the sentiments on the entities can be different. For example, the opinion holder may be positive about some entities

and negative about others. Thus, it does not make practical sense to assign one sentiment orientation to the entire document in this case. It also does not make much sense if multiple opinion holders express opinions in a single document because their opinions can be different too.

This assumption holds for reviews of products and services because each review usually focuses on evaluating a single product or service and is written by a single reviewer. However, the assumption may not hold for a forum and blog post because in such a post the author may express opinions on multiple entities and compare them using comparative sentences.

Below, we first discuss the classification problem to predict categorical class labels and then the regression problem to predict rating scores. Most existing techniques for document-level classification use supervised learning, although there are also unsupervised methods. Sentiment regression has been done mainly using supervised learning. Recently, several extensions to this research have also appeared, most notably, *cross-domain sentiment classification* (or *domain adaptation*) and *cross-language sentiment classification*, which will also be discussed at length.

3.1 SENTIMENT CLASSIFICATION USING SUPERVISED LEARNING

Sentiment classification is usually formulated as a two-class classification problem, *positive* and *negative*. Training and testing data used are normally product reviews. Since online reviews have rating scores assigned by their reviewers, e.g., 1–5 stars, the positive and negative classes are determined using the ratings. For example, a review with 4 or 5 stars is considered a positive review, and a review with 1–2 stars is considered a negative review. Most research papers do not use the neutral class, which makes the classification problem considerably easier, but it is possible to use the neutral class, e.g., assigning all 3-star reviews the neutral class.

Sentiment classification is essentially a text classification problem. Traditional text classification mainly classifies documents of different topics, e.g., politics, sciences, and sports. In such classifications, topic-related words are the key features. However, in sentiment classification, sentiment or opinion words that indicate positive or negative opinions are more important, e.g., *great*, *excellent*, *amazing*, *horrible*, *bad*, *worst*, etc.

Since it is a text classification problem, any existing supervised learning method can be applied, e.g., naïve Bayes classification, and support vector machines (SVM) (Joachims, 1999; Shawe-Taylor and Cristianini, 2000). Pang et al. (2002) was the first paper to take this approach to classify movie reviews into two classes, positive and negative. It was shown that using unigrams (a bag of words) as features in classification performed quite well with either naïve Bayes or SVM, although the authors also tried a number of other feature options.

In subsequent research, many more features and learning algorithms were tried by a large number of researchers. Like other supervised machine learning applications, the key for sentiment