Chapter No. 5
"Social Media Mining – Fundamentals"
In this package, you will find:
A Biography of the authors of the book
A preview chapter from the book, Chapter NO.5 "Social Media Mining – Fundamentals"
A synopsis of the book’s content
Information on where to buy this book

About the Authors

**Nathan Danneman** holds a PhD degree from Emory University, where he studied International Conflict. Recently, his technical areas of research have included the analysis of textual and geospatial data and the study of multivariate outlier detection.

Nathan is currently a data scientist at Data Tactics, and supports programs at DARPA and the Department of Homeland Security.

I would like to thank my father, for pushing me to think analytically, and my mother, who taught me that the most interesting thing to think about is people.

**Richard Heimann** leads the Data Science Team at Data Tactics Corporation and is an EMC Certified Data Scientist specializing in spatial statistics, data mining, Big Data, and pattern discovery and recognition. Since 2005, Data Tactics has been a premier Big Data and analytics service provider based in Washington D.C., serving customers globally.

For More Information:
Richard is an adjunct faculty member at the University of Maryland, Baltimore County, where he teaches spatial analysis and statistical reasoning. Additionally, he is an instructor at George Mason University, teaching human terrain analysis, and is also a selection committee member for the 2014-2015 AAAS Big Data and Analytics Fellowship Program.

In addition to co-authoring *Social Media Mining in R*, Richard has also recently reviewed *Making Big Data Work for Your Business* for Packt Publishing, and also writes frequently on related topics for the Big Data Republic (http://www.bigdatarepublic.com/bloggers.asp#Rich_Heimann). He has recently assisted DARPA, DHS, the US Army, and the Pentagon with analytical support.

I'd like to thank my mother who has been supportive and still makes every effort to understand and contribute to my thinking.

For More Information:
Social Media Mining with R

If you have ever been interested in social media, machine learning, data science, statistical programming, or particularly Big Data—as it relates to extracting value from the data on the Web—then this book is for you. We are excited to provide an introduction to these topics based on our applied research experience. Social Media Mining with R exposes readers to both introductory and advanced sentiment analysis techniques through detailed examples and with a large dose of rigorous social science background. Additionally, this book introduces a novel, unsupervised sentiment analysis model. These techniques can be complex, often counterintuitive, and are nearly always laden with assumptions. This book provides readers with a how-to guide for implementing these models and, most importantly, explains the techniques in depth so users can deploy them appropriately and interpret their results correctly. It explains the theoretical grounds for the techniques described and serves to bridge the potential of social media, the theoretical issues surrounding its use, and the practical necessities of its implementation. Social Media Mining with R lays out valid arguments for the value of big social media data. The book provides step-by-step instructions on how to obtain, process, and analyze a variety of socially generated data as well as a theoretical background for helping researchers interpret and articulate their findings. The book includes R code and example data that can be used as a springboard as readers undertake their own analyses of business, social, or political data. Readers are not assumed to know R or statistical analysis but are pragmatically provided with the tools required to execute sophisticated data mining techniques on data from the Web.

Overall, Social Media Mining with R provides a theoretical background, comprehensive instructions, and state-of-the-art techniques such that readers will be well equipped to embark on their own analyses of social media data.

Thank you for reading!

What This Book Covers

Chapter 1, Going Viral, introduces the readers to the concept of social media mining, sentiment analysis, the nature of contemporary online communication, and the facets of Big Data that allow social media mining to be such a powerful tool. Additionally, we provide some evidence of the potential and pitfalls of socially generated data and argue for the use of quantitative approaches to social media mining.

Chapter 2, Getting Started with R, highlights the benefits of using R for social media mining. Readers are then walked through the processes of installing, getting help for, and using R. By the end of this chapter, readers would become familiar with data import/export, arithmetic, vectors, basic statistical modeling, and basic graphing using R.

For More Information:  
Chapter 3, Mining Twitter with R, explains that an obvious prerequisite to gleaning insight from social media data is obtaining the data itself. Rather than presuming that readers have social media data at their disposal, this chapter demonstrates how to obtain and process such data. It specifically lays out a technical foundation for collecting Twitter data in order to perform social data mining and provides some foundational knowledge and intuition about visualization.

Chapter 4, Potentials and Pitfalls of Social Media Data, highlights that measurement and inference can be challenging when dealing with socially generated data, including social media data. This chapter makes readers aware of common measurement and inference mistakes and demonstrates how these failures can be avoided in applied research settings.

Chapter 5, Social Media Mining – Fundamentals, aims to develop theory and intuition over the models presented in the final chapter. These theoretical insights are provided prior to the step-by-step model building instructions so that researchers can be aware of the assumptions that underpin each model, and thus apply them appropriately.

Chapter 6, Social Media Mining – Case Studies, helps to bring everything together in an accessible and tangible concluding chapter. This chapter demonstrates canonical lexicon-based, and supervised sentiment analysis techniques as well as laying out and executing a novel unsupervised sentiment analysis model. Each class of model is worked through in detail, including code, instructions, and best practices. This chapter rests heavily on the theoretical and social science information provided earlier in the book, but can be accessed right away by readers who already have the requisite understanding.

Appendix, Conclusions and Next Steps, wraps everything up with our final thoughts, the scope of the data mining field, and recommendations for further reading.

For More Information:  
Social Media Mining – Fundamentals

Techniques used to extract sentiment from social media data are complex, at times counterintuitive, and often laden with assumptions. Before providing readers with a how-to guide to implement these models, we think it is critical to explain the techniques in depth so users can deploy them appropriately. This chapter explains the theoretical grounds for the techniques developed in the next chapter and serves as a bridge between the discussion of the pitfalls of social media mining and the execution of that mining.

Key concepts of social media mining

We find it useful to situate social media mining within the context of traditional social science research. While defining social science is difficult, Jean Anyon’s perspective is a nice starting point. She suggests that socially explicit theory, and thus social science, should be empirically constructed, theoretically defensible, and socially critical. More generally, social science’s main aims are to generate theories that explain individual-and group-level behaviors and then to examine the veracity of those theories with evidence. Generally, these theories are more valuable insofar as they allow a deeper understanding of human behavior, and especially so if they provide an understanding sufficient to allow for intervention. Our approach to social media mining strives to take this challenge to heart; thus, throughout this book, we use social media data to ask and answer questions of pressing social relevance.

For More Information:
Traditional social science not only focuses on important questions, but also seeks to uncover relationships that are interesting and unexpected. The world around us is full of complex social behavior; though identifying mundane facts is sometimes helpful in the name of basic research, it does little to help us understand social behavior. We take to heart the mandate to find interesting relationships as we mine social media data—a particularly complex and rich source.

At heart, however, social science is not a focus on the important or the interesting. It is science, which means that it is a set of methods and practices designed to generate and verify facts. The logic of science, regardless of whether it proceeds quantitatively or in a qualitative fashion, is fundamentally about knowledge discovery and accumulation. This logic helps mitigate several shortcomings in reasoning that frequently hinder our ability to make correct inferences. Some examples include illusory correlations (perceiving correlations that do not exist), selective observation (inadvertently cherry-picking data), illogical reasoning, and over or under generalizing (assuming that facts discovered in one domain apply to others as well). Generally, the scientific process helps avoid the discovery of false truths often arrived at through deduction, speculation, justification, and groupthink.

**Good data versus bad data**

Traditional social science data differs markedly from social media data in several respects. First and foremost, traditional social science data is most often collected in targeted and rigorous ways. For instance, the US census targets nearly the entire US population and has a strong methodology for attaining this target. Researchers interested in the sentiments of particular demographics can target them specifically through surveys or polls, and can additionally tune survey instruments to carefully elicit the information they desire.

The steep downside to these classes of data sources is that they are often extremely limited in their geographic or temporal coverage. As such, they do not allow for broad generalizations or comparisons across place and time. Broader surveys, such as the US census, capture information about a large number of people, but usually only capture cursory descriptive information. Furthermore, this information is captured infrequently and in ways that are incomparable across borders. Narrower surveys, such as those fielded by researchers and firms, obviously are limited in their ability to support inferences about broad populations.

For More Information:

These data sources, despite their shortcomings in terms of coverage, are held in high regard due to their focused and authoritative nature. These sources are used due to the fear that bad input data will yield low-quality inferences, or "garbage in, garbage out," as the saying goes. But, what is garbage data? Should we consider social media data garbage because of its unfocused nature? Also, under what circumstances might we be willing to use social media data?

Our view is that focused, purpose-collected data is the best option when it is available. This statement may come as a surprise in a book on social media mining, but the linchpin of the statement is the phrase \textit{when it is available}. For the vast majority of emerging questions related to business, politics, and social life, purpose-collected social science data-sets simply do not exist. As such, we take the pragmatic position that social data, due to its broad coverage and large volume, makes a nice fallback to targeted data. Social data is bad in the sense that much of it will be inapplicable to any particular question; however, limited applicability is certainly better than utter absence of data. The reality is that we live in an imperfect world, which will consequently yield imperfect data. Our job, as data analyst is to work with data in responsible ways.

This book does not cover how to handle poor, dirty, missing, or incorrect data in a comprehensive manner. However, we do wish to promote the use of social media data and its utility in cases where traditional social science data-sets do and do not exist and where there are low and high barriers to targeted collection.

Traditional social science modeling techniques tend to require data-sets in which observations are independent of one another. However, data gleaned from social media outlets, such as Twitter, is almost certainly not independent. That is, data is not randomly sampled from a larger population and thus each observation is likely to be related to observations that are nearby in some sense. For example, tweets about a large public event arise around the same time and from the same area. Also, many may express similar views. This nonindependence has implications for how you handle tweets given their degree of centrality, shared geography, and repetition through retweets. Although we do not often study sentiment polarity explicitly in terms of networks, doing so may prove useful for future researchers. We anticipate research in that direction will produce better measures and predictions, localized lexicons, and other advantages.
Understanding sentiments
Social media mining can and should have broad interpretation. It is not the intent of
the authors to confine social media mining to sentiments or opinions, but rather we
suggest that a sentiment or opinion is a useful tool for many research pursuits.

Until recently, sentiment was understood as a ubiquitous and constant part of the
human experience, with variations in sentiments changing only slightly up or down.
Klaus Scherer (2000) developed a working definition as follows:

"Emotions (sentiments) are episodes of coordinated changes in several components
in response to external and internal events of major significance to the organism."

It is our intent to understand, measure, and interrelate these changes in a sentiment.
Scherer's typology of emotions is a useful grounding point for the understanding of
sentiments, and as a jumping-off point for a discussion of the difficulty in measuring
sentiment-laden text.

Scherer's typology of emotions
Scherer's typology of emotions is briefly explained as follows:

- **Emotion**: This is a brief, organically synchronized evaluation of a major
event, for example, being angry, sad, joyful, ashamed, proud, or elated
- **Mood**: This is a diffused, non-caused, low-intensity, long-duration change
  in subjective feeling, for example, being cheerful, gloomy, irritable, listless,
depressed, or buoyant
- **Interpersonal stance**: This is an affective stance towards another person in
  a specific interaction, for example, being friendly, flirtatious, distant, cold,
  warm, supportive, or contemptuous
- **Attitude**: This is enduring, affectively colored beliefs or dispositions towards
  objects or persons, for example, being liking, loving, hating,
  valuing, or desiring
- **Personality traits**: These are stable personality dispositions and typical
  behavior tendencies, for example, being nervous, anxious, reckless, morose,
  hostile, or jealous

Generally, when we try to measure a sentiment, we talk about Scherer's emotions;
though, in some situations, we might try to capture longer-term phenomena such
as moods.

For More Information:
Anchoring the neo-social science approach using Twitter data versus other types of social media data is important as well because not all data is equal. Twitter data differs from data derived from sites such as Yelp and Google Reviews due to the simple fact that Twitter does not have ratings or explicit targets. If we want to know the sentiment of a given source or topic, whether it is the iPhone 5S or something more sensitive such as social policy, we have to discover that signal in a corpus of other signals. However, Yelp and Google Reviews (just two examples of many) have explicitly accounted for the source or topic by design and have ratings designed to measure sentiments.

A tweet is what Twitter users send to each other and to the Twittersphere. A tweet is sometimes a sentence and other times not, but it is restricted to 140 characters or approximately 11 words. Twitter therefore provides sentence-level sentiment analysis as opposed to reviews on Yelp or Google Review, which usually constitute the entire documents.

**Sentiment polarity – data and classification**

Social media mining primarily involves the following two steps:

1. Identifying and retrieving content related to the topic of interest.
2. Measuring the polarity of each datum.

The first step, message retrieval, requires some a priori insight into the topic of interest. The goal of message retrieval is to seek out only the messages or pieces of text that contain sentiment-laden content related to a particular topic. This topic could be almost anything of interest, subject to the constraint that information exists about it on public social media. For instance, in Chapter 3, *Mining Twitter with R*, we examined the topic Big Data, and in Chapter 6, *Social Media Mining – Case Studies*, we delve into social issues such as abortion and the economy. Atmospherics, that is, data gathered in an effort to track local sentiments with regards to economic, cultural, or political topics, can also be analyzed, as we do in *Chapter 6, Social Media Mining – Case Studies*. Lest readers think that social media is too diffuse to be useful, as of the writing of this book, at least one hedge fund uses atmospherics gleaned from Twitter to gauge stock prices.

For More Information:
To gather data, we generally collect content that contains a manually specified (set of) keyword(s). This is called the target. For example, the target for presidential approval would use the topic keyword \texttt{obama}. We may wish to add context to analyses done on particular keywords by adding additional opposing or specifying keywords. For example, in addition to \texttt{obama}, we could add \texttt{romney} to provide a counterpoint if we were studying the 2012 presidential election campaigns. Depending on the purpose of our analysis, we could jointly search for, say, \texttt{obama} and \texttt{economy} to target more specific subjects.

Topic models represent a second, more sophisticated, and potentially more thorough way of capturing bits of text that are relevant to a particular analysis. These models take very large sets of documents as their inputs and group them probabilistically into estimated topics. That is, each document is proclaimed to be a mixture of one or more topics that are themselves estimated from the data. This allows users to find texts that are related to a topic, though they may not explicitly use a particular keyword. The details of this class of statistical models are outside the scope of this text; however, in \textit{Appendix, Conclusions and Next Steps}, we point readers to references on the theory and estimation of this exciting new class of tools.

Social data mining is the detection of attitudes, and the easiest way to understand it is through the following structure:

\[
\text{sentiment} = \{\text{data source, source, target, sentiment, polarity}\}
\]

The parameters are explained in detail as follows:

- **Data source**: This relates to understanding the source of the data; that is, is the source a sentence or an entire document? Twitter or a blog?
- **Source or holder**: This is the one that expresses a sentiment or an opinion,
- **Target or aspect**: The target or aspect is what or to whom the sentiment is directed toward.
- **Type of sentiment**: This is the type(s) of emotion(s) expressed, that is, like, love, hate, value, desire, and so on.
- **Polarity**: These are juxtapositional sentiments on a dimension, that is, positive or negative.
Chapter 5

The following examples highlight these components and also some of the challenges involved in sentiment analysis. We have parts of two reviews: one about Steven Spielberg and another about John Carpenter. In both examples, the data source is the Internet Movie Database (IMDB) that considers itself the world's most popular and authoritative source for movie, TV, and celebrity content. The holder is the one who wrote the review, and the targets are Steven Spielberg and John Carpenter respectively. However, the target is complicated by mentions of various movies over time. Also, complicating matters is the variety of sentiment types and polarities.

Steven Spielberg's second epic film on World War II is an unquestioned masterpiece. Spielberg, ever the student on film, has managed to resurrect the war genre by producing one of its grittiest and most powerful entries. He also managed to cast this era's greatest answer to Jimmy Stewart, Tom Hanks, who delivers a performance that is nothing short of an astonishing miracle for about 160 out of its 170 minutes; Saving Private Ryan is flawless, literally!

There was a time when John Carpenter was a great horror director. Of course, his best film was 1978's masterpiece Halloween; however, he also made The Fog in the 1980s and 1987's underrated Prince of Darkness. Even, Heck made a good film, In the mouth of madness, in 1995. However, something terribly wrong happened to him in 1992 with the terrible comedy Memoirs of an Invisible Man.

Supervised social media mining – lexicon-based sentiment

Lexicon-based sentiment classification is perhaps the most basic technique for measuring the polarity of the sentiment of a group of documents (that is, a corpus). Lexicon-based sentiment measurement requires a dictionary of words (a lexicon) and each word's associated polarity score. For example, a lexicon may contain the word excellent, which might have a score of positive two. Similarly, the word crummy may score negative one and a half. In the simplest implementation of lexicon-based sentiment analysis, all of the words in a document are compared to the words in the lexicon. Every time a word is used that is in the lexicon, the associated score is added to that text's overall sentiment score. For example, the sentence "I found the customer assistance to be excellent," would score a positive two.

For More Information:  
A lexicon-based sentiment often entails merely counting the opinion words from a subset of data from a particular source. This approach certainly has errors, as does perhaps all natural language processing; however, in aggregate, the lexicon-based approach has proven to be fairly robust, even when only used on subsets. Additionally, there are many possible ways to aggregate the sentiment scores of each word, but most commonly, they are simply summed up to form an overall score for a document.

Despite lexicon-based sentiment classification being considered here as basic, it is still difficult. This is primarily due to the fact that when counting words with positive or negative valences, one must decide which words to count as each. Different dictionaries of positive and negative words can generate different sentiment scores for the same sentences. Some words with perceived sentiment are more neutral, while others have perceived neutrality, but are in fact more extreme. This challenge arises, in part, due to varied usages of words within and across contexts.

Preassembled lexicons are incredible resources and are applicable for a wide variety of problems—we use several in Chapter 6, Social Media Mining – Case Studies. Despite subtle differences, they are all good starting points, but they are just that, starting points and not end points. Rather than utilizing a preassembled lexicon indiscriminately, researchers should often develop lexicons that are sensitive to the domain they are analyzing. For instance, a lexicon that is useful for economic atmospherics (where moderate and stable are positive) may prove useless for examining political leanings. Preassembled domain-specific lexicons exist as well and two popular economic lexicons will be used later. There are many approaches to extending both generic preassembled lexicons as well as domain-specific preassembled lexicons, and we will describe two rather intuitive ones, dictionary-based lexicons, and corpus-based lexicons in addition to preassembled lexicons.

Both dictionary-based and corpus-based approaches augment preassembled lexicons in one of the following two ways:

- Using a dictionary (that is, synonyms and antonyms) to add keywords external to our corpus to enhance our preassembled lexicon(s)
- Using the corpus directly to add words already internal to our corpus that are keywords but are not accounted for by preassembled lexicons

Merging preassembled lexicons, dictionary-based lexicons, and corpus-based lexicons offers the best chance to successfully estimate sentiment.
The two approaches (dictionary and corpus) produce empirically constructed lexicons that seek to calibrate the underlying sentiment by adding to the preassembled lexicons. The intuition is words appearing in the complete collection of lexicons (preassembled, dictionary, and corpus) within our set of documents returned from a narrow topic (the search set) are more likely to objectively describe sentiment information. In the lexicon approach, it is sufficient to simply count the frequency of words from our lexicons with the set of documents returned from our topic of interest and sum the results over time, by space or by product. The next chapter outlines this process in detail and sums the results' set over time where the target is the US economy.

Supervised social media mining – Naive Bayes classifiers

Methods to extract sentiments from documents can be broadly classified into supervised and unsupervised approaches (semisupervised approaches are also available but are outside the scope of this text. Interested readers can consult Abney (2007)). Supervised methods are those that utilize data that has been tagged or labeled. In the parlance of statistics, these approaches utilize observations with both independent and dependent variables. For instance, the following Naive Bayes classifier approach involves a training dataset of documents that have already been scored as having positive or negative sentiment; a statistical model based on these forms the basis of scoring further documents. In contrast, unsupervised learning algorithms do not require a dependent variable to be provided. For instance, the IRT-based method described later in this chapter scales documents along a continuum of sentiments with no need to provide a labeled training set. Additionally, lexicon-based approaches mentioned earlier can also function without prelabeled observations.

The Naive Bayes classifier, in spite of its unfortunate name, turns out to be a highly useful tool for sentiment analysis. At the most general level, the Naive Bayes classifier is exactly that: a classifier. Classifiers are statistical tools that are used for, among other things, predicting which of two or more classes a new observation belongs to. In our case, we want to train our classifier to be able to distinguish documents featuring positive sentiment from those featuring negative sentiment (the two types or classes of interest). To do so, we feed the algorithm a large set of documents that are already coded as containing positive or negative sentiments about a particular topic. Then, if all goes as planned, we can pass new documents to the model and have it predict the direction of their sentiment, or valence, for us. The downside to this and other supervised techniques is having to handcode a sufficient set of initial training data.

For More Information:
So, where did the naive part of the name come from, and why is this method useful in spite of its self-assumed simplicity? The goal of any classifier is to determine which class or type a new observation belongs to based on its characteristics and previous examples from both types that we have seen before (that is, from existing data). Some types of classifiers can account for the fact that the characteristics we use for this prediction may be correlated. That is, if we are trying to predict whether an e-mail is spam or not by looking at what words and phrases the e-mail contains, the words easy and money are likely to co-occur (and are likely to be indicative of spam messages). The Naive Bayes classifier does not try to account for correlations between characteristics. It just uses each characteristic separately to try to determine each new observation's class membership.

The naive assumption that all of the characteristics of an observation are unrelated is always wrong. In predicting whether or not to extend a loan to an individual, a bank may look at their credit score, whether or not they own a home, their income, and their current debt level. Obviously, all of these things are likely to be correlated. However, ignoring the correlations between predictive characteristics allows us to do two things that would otherwise be problematic. First, it allows us to include a huge number of characteristics, which becomes important. This is because in text analysis, individual words often have predictive characteristics, and documents often contain thousands of unique words. Other models have a difficult time accommodating this number of predictors. Secondly, the Naive Bayes classifier is fast, thus allowing us to use large training sets to train a model and to generate results quickly.

Unsupervised social media mining – Item Response Theory for text scaling

The techniques set out earlier for scaling or classifying sentiments in texts are fairly robust; that is, they tend to work well under a wide variety of conditions such as heterogeneous text lengths and topic breadths. However, each of these methods requires substantial analyst input, such as labeling training data or creating a lexicon. Item Response Theory (IRT) is a theory, but will be used in this text to refer to a class of statistical models that rely on that theory, providing a way to scale texts according to sentiment in the absence of labeled training data. That is, IRT models are unsupervised learning models.

For More Information:
IRT models were developed by psychologists for scoring complex tests and were then picked up by political scientists who employ them for scaling legislators. We will briefly explain the legislative context as that will help readers build intuition over how models work when applied to scaling texts. Consider a set of $V$ voters, such as US Senators, who, over the course of a year, vote on $B$ bills. For simplicity, assume each voter can only vote yes or no. We could then put all of the data into a matrix, where each row represents a voter and each column a bill. Each cell then represents a particular voter's decision on a particular bill, that is, yes (1) or no (0). Now, we need to make two related assumptions. The first is that all or most of these voters can be described as lying along a single underlying continuum. The second more trivial assumption is that this position influences their votes, at least on some bills. With these assumptions in place, we can estimate a statistical model that describes the probability of each cell in our data matrix being a one or a zero. The model is a function of each bill's difficulty of being voted for (that is, how controversial it is), each voter's position on the underlying scale, and how strongly each bill is affected by voters' locations on the scale. Technically, we estimate a logistic regression as follows:

$$\text{pr}(y_{vb}=1) = \logit(b_1 \cdot x_v - b_0)$$

Here, $x$ is the scaled position of each voter ($v$), $b0$ is the difficulty of voting yes for each bill ($b$), and $b1$ is the degree to which each voter's position affects their proclivity to vote in favor of each bill ($b$). Positive values of $b1$ mean that voters to the right are more likely to vote in favor of a bill, and negative values of $b1$ mean that senators to the left are more likely to vote for a bill.

As you will see, we apply the previous assumption to the analogous case of text scaling. To do so, we create a matrix with rows representing authors or documents (instead of voters) and columns representing words or phrases, (instead of bills). Each cell represents whether or not a particular author used a particular word or phrase. We modify the previous assumptions: authors lie along a sentiment continuum, and their placement affects their pattern of word use. The first part of this assumption is limiting. We can only apply the method to sets of documents that are sufficiently narrow to be usefully described by a single underlying continuum, and that continuum must essentially be the sentiment we are trying to measure. The results of this analysis are a continuous scaled measure of author (or document) location ($x$) as well as estimates of the weights for each word or phrase ($b1$). This scaled measure of location ($x$) represents the author's sentiment towards the topic under study if the previous assumptions are met.

For More Information:
The IRT-based method described here has mixed properties. It requires no training data, little subject matter expertise to employ, is language agnostic (that is, could function on any language), and generates a quantitative (instead of merely a binary) measure of a sentiment. However, this model can only be applied to documents that are all about the same topic, can only estimate a single underlying dimension, can be slow to estimate, and is not guaranteed to converge.

**Summary**

In this chapter, we learned key concepts related to sentiment analysis. Sentiment was defined and difficulties related to its mining were covered. We then walked through the theoretical underpinnings of three different models for sentiment analysis. We intentionally separated the details of implementation from theoretical concerns in hopes of giving readers an appreciation for the methods, including their strengths and weaknesses.

The next chapter delves into the details of implementing the classes of models described earlier.

---

For More Information:
Where to buy this book


Free shipping to the US, UK, Europe and selected Asian countries. For more information, please read our shipping policy.

Alternatively, you can buy the book from Amazon, BN.com, Computer Manuals and most internet book retailers.