

Techniques and Applications for Sentiment Analysis

Communications of the ACM | April 2013 | Vol. 56 | nº 4

DOI:10.1145/2436256.2436274

Ronen Feldman

Hebrew University

ronen.feldman@huji.ac.il

Airton Bordin Junior

airtonbjunior@gmail.com

Prof. Dr. Nádia Félix Felipe da Silva

nadia@inf.ufg.br



Roteiro

- Introduction
 - Context
 - Problem
 - Objective
- Sentiment Analysis
- Sentiment Lexicon
- Applications
- Research Issues
- Conclusion



Context

**JAN
2017**

GLOBAL DIGITAL SNAPSHOT

KEY STATISTICAL INDICATORS FOR THE WORLD'S INTERNET, MOBILE, AND SOCIAL MEDIA USERS

TOTAL
POPULATION



7.476
BILLION

URBANISATION:
54%

INTERNET
USERS



3.773
BILLION

PENETRATION:
50%

ACTIVE SOCIAL
MEDIA USERS



2.789
BILLION

PENETRATION:
37%

UNIQUE
MOBILE USERS



4.917
BILLION

PENETRATION:
66%

ACTIVE MOBILE
SOCIAL USERS



2.549
BILLION

PENETRATION:
34%

we
are
social

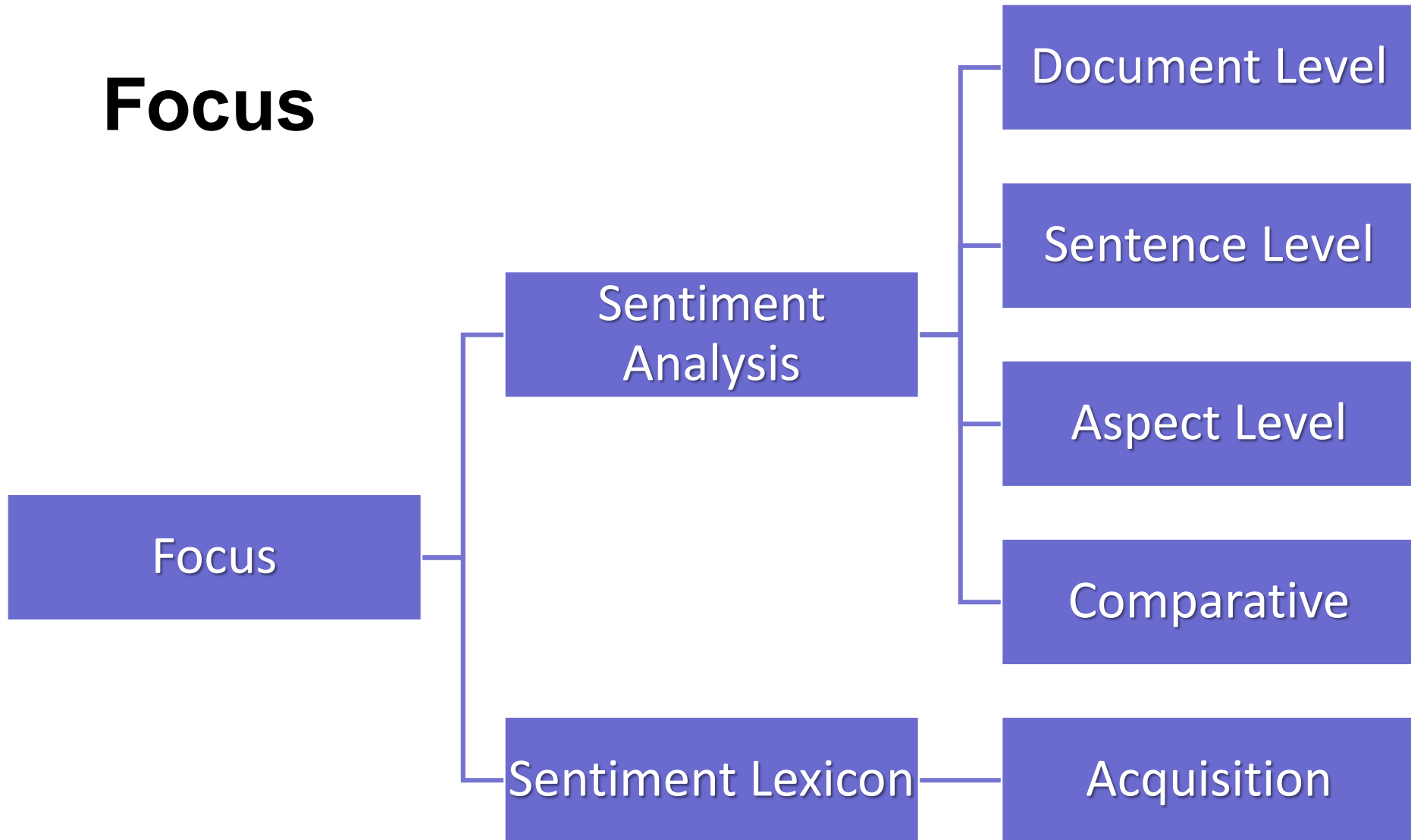
we
are
social



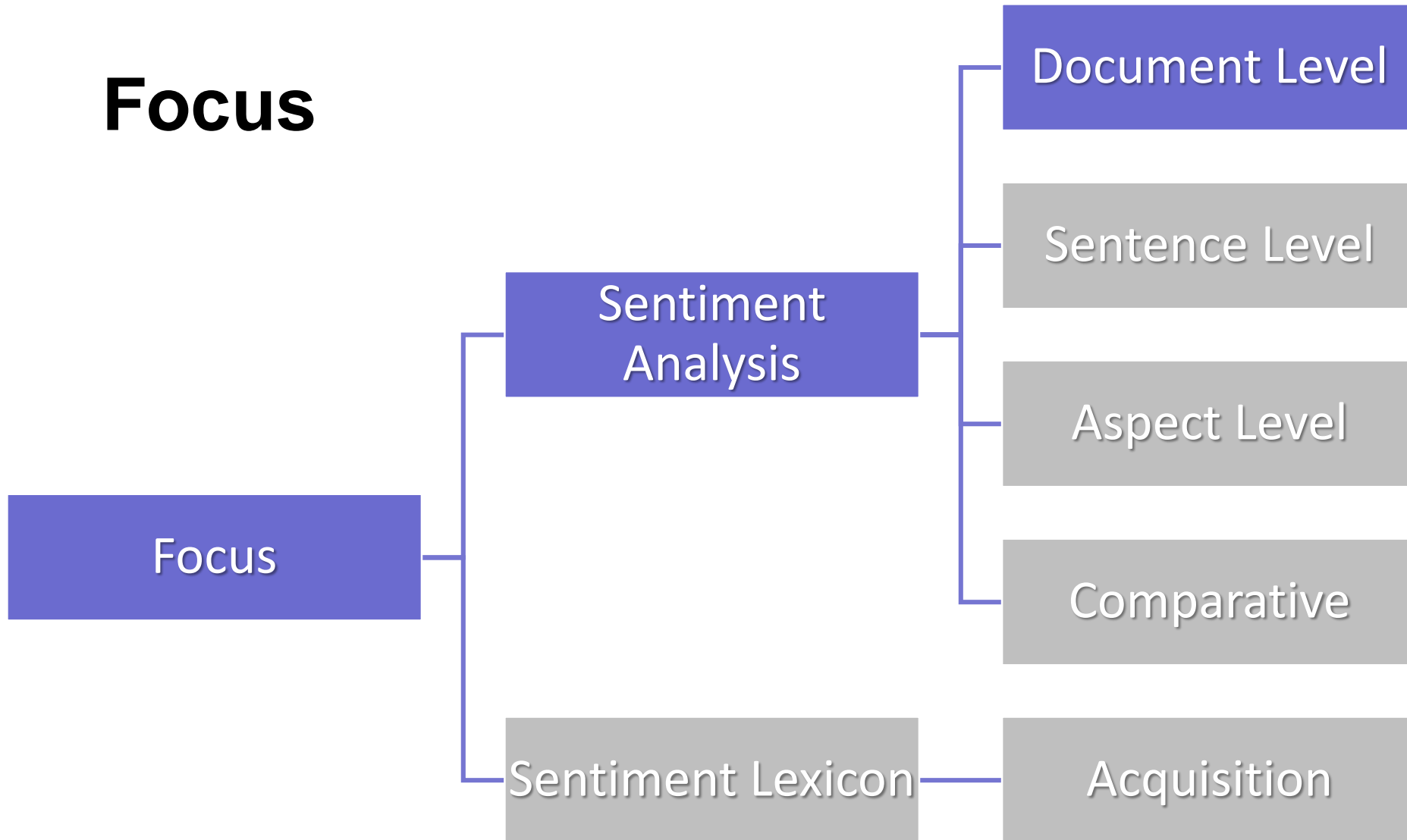
Objective

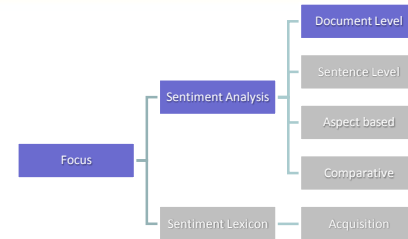
- Present the main research problems related to Sentiment Analysis (SA) and some of the techniques used to solve them
- Review some of the major application areas where sentiment analysis is being used today

Focus



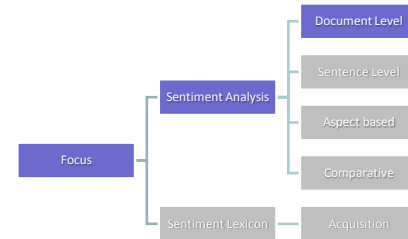
Focus





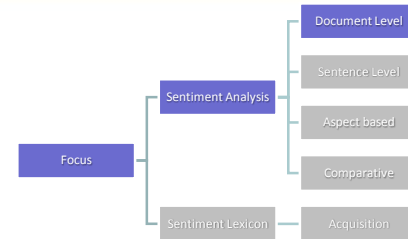
SA – Document Level

- Assumed that the document contains an opinion on one main object expressed by the author of the document
- Two main approaches
 - Supervised learning
 - Unsupervised learning

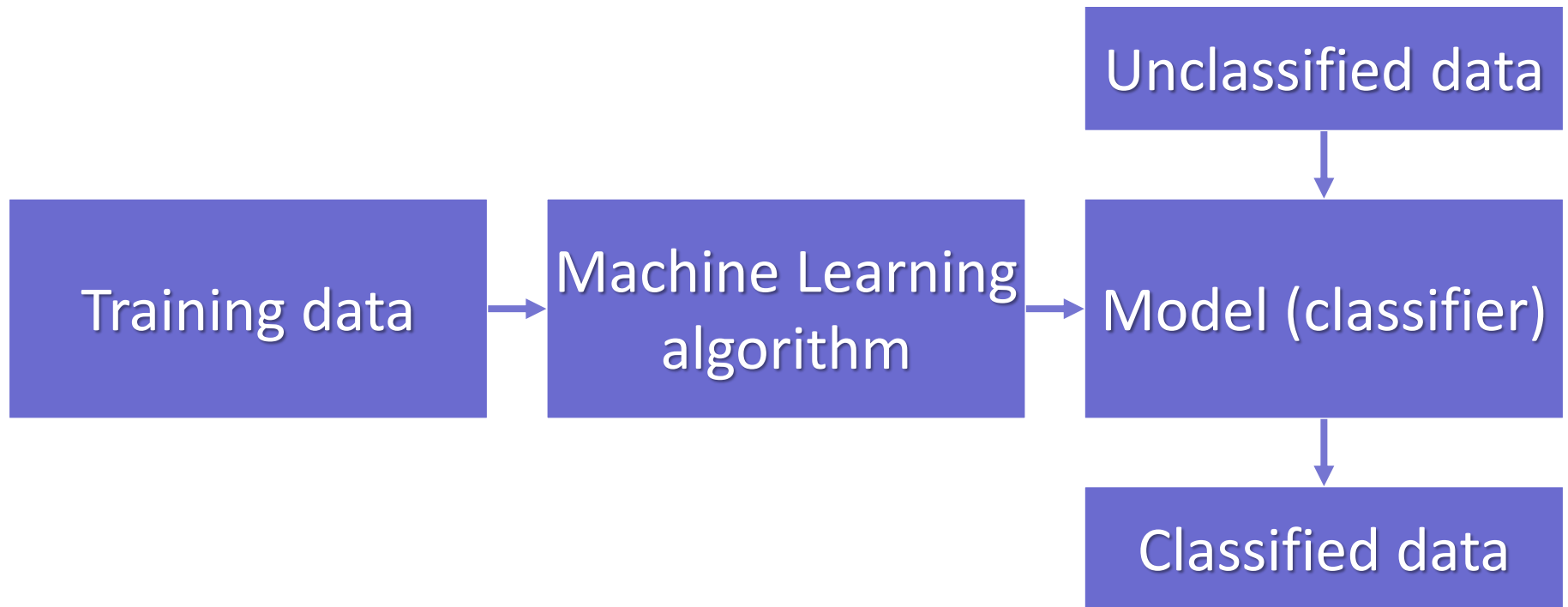


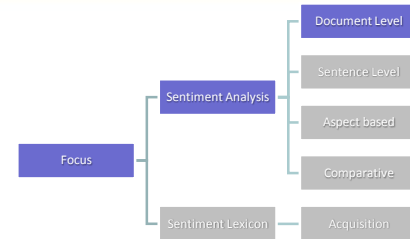
SA – Document Level

- Supervised learning
 - Finite set of classes
 - Training data is available
 - Classes
 - Positive/negative/neutral
 - Numeric scale (stars)
 - SVM, KNN, Naïve Bayes, Logistic Regression, Genetic programming



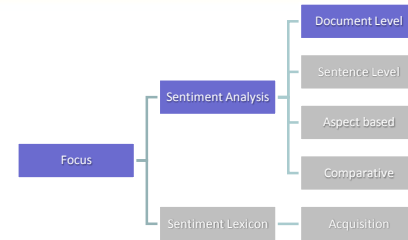
SA – Document Level





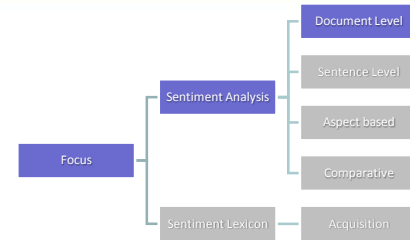
SA – Document Level

- Important representations
 - Bag of words
 - TFIDF
 - Part of Speech (PoS)
 - Sentiment Lexicons



SA – Document Level

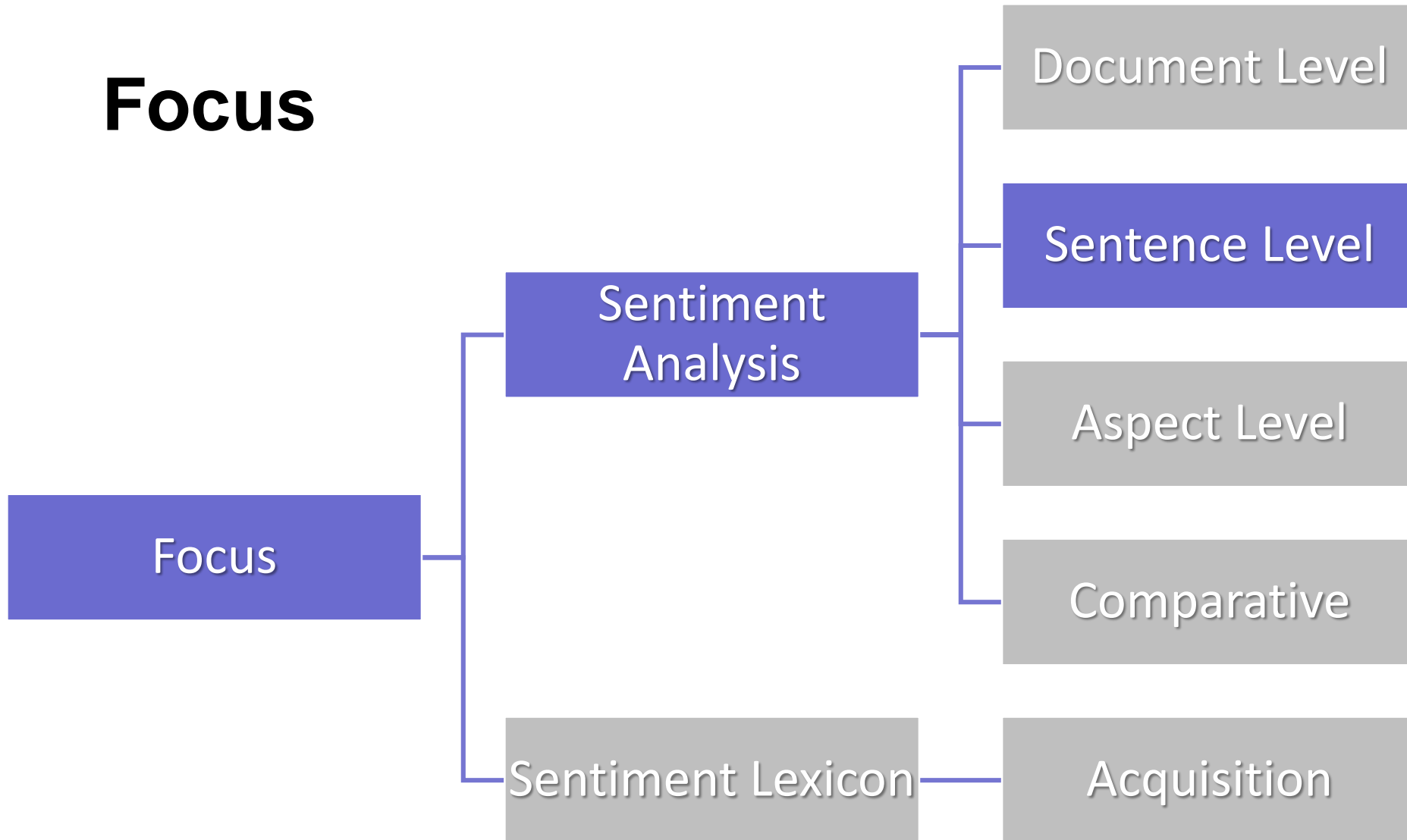
- Unsupervised learning
 - Semantic orientation (SO) specific phrases
 - PMI (Pointwise Mutual Information) of the phrase with two sentiment words
 - $PMI(P, W)$
 - Statistical dependence between phrase P and word W based on their co-occurrence in a corpus

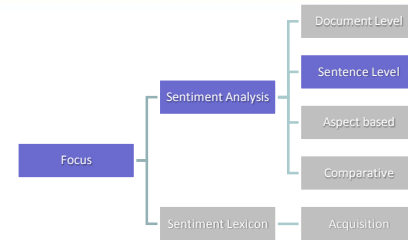


SA – Document Level

- Unsupervised learning
 - The two words used in Turney (2002) are ‘excellent’ and ‘poor’
 - The SO measures whether P is closer in meaning to the positive word (‘excellent’) or the negative word (‘poor’)

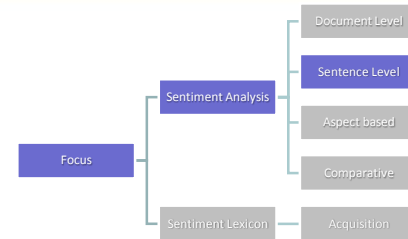
Focus





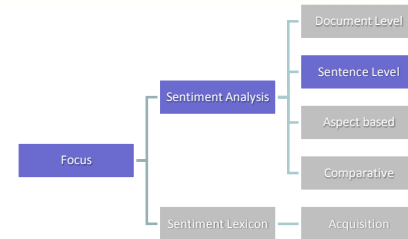
SA – Sentence Level

- Assume
 - We know the identity of the entity discussed in the sentence
 - There is a single opinion in each sentence



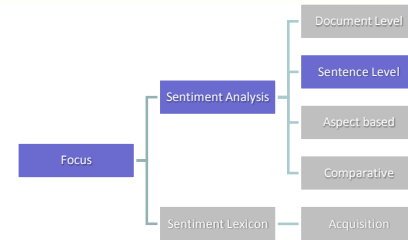
SA – Sentence Level

- Before analyzing the polarity of the sentences we must determine if the sentences are subjective or objective
 - Only subjective sentences will be analyzed



SA – Sentence Level

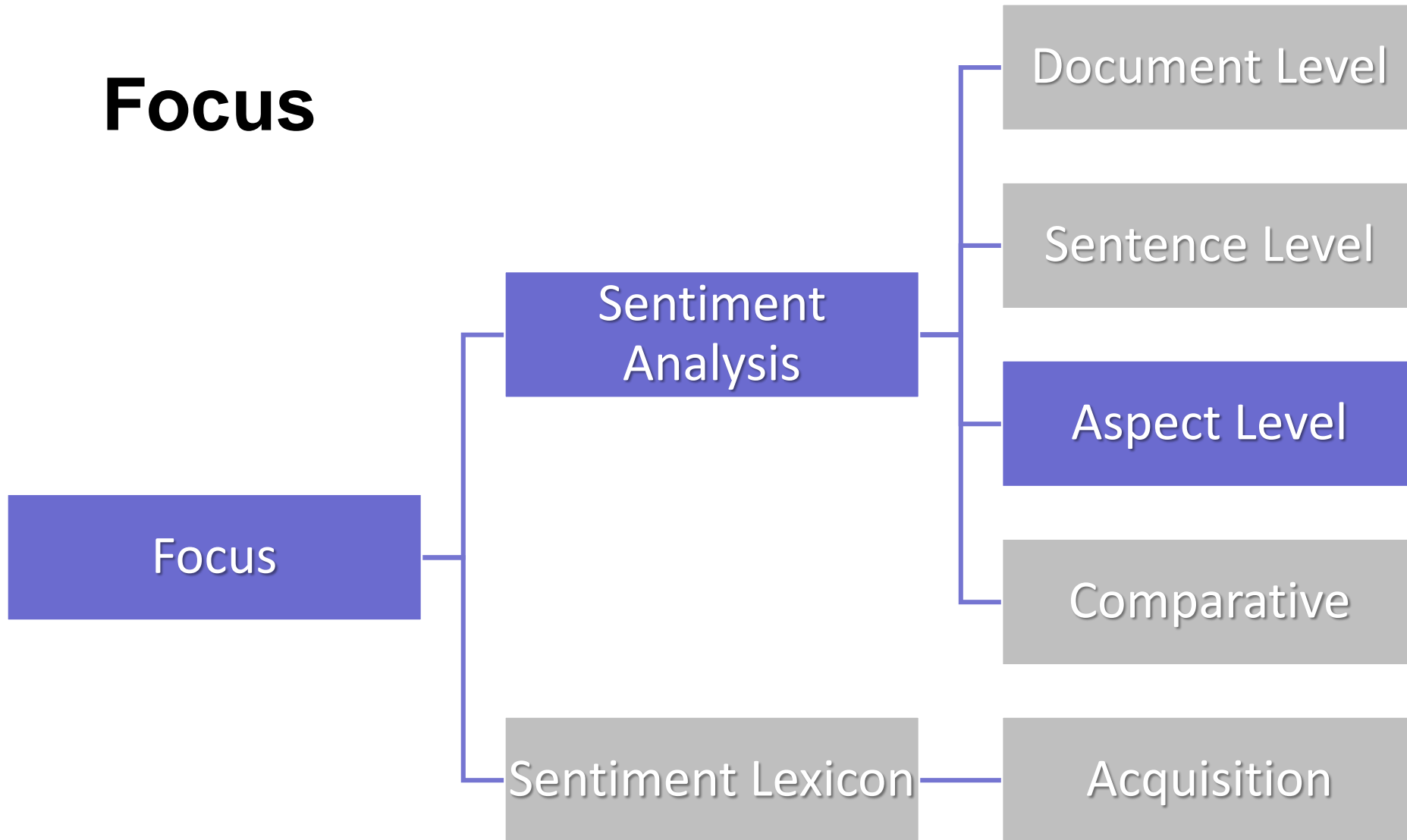
- Most approaches are based on supervised learning
- Unsupervised approach is similar of Turney (2002)

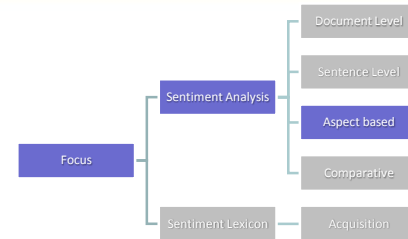


SA – Sentence Level

- Handle different types of sentences by different strategies
- Sentences that need unique strategies
 - Conditional
 - Question
 - Sarcastic

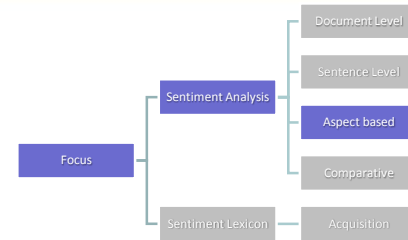
Focus





SA – Aspect Level

- In many cases, people talk about entities that have many aspects (attributes) and they have a **different opinion about each of the aspects**
- Often happens in reviews about products or in discussion forums



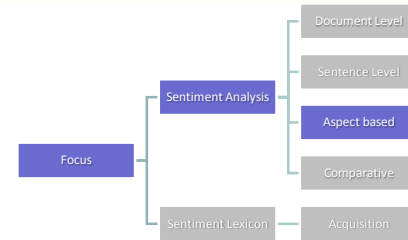
SA – Aspect Level

★★★★☆ I want to love it, I really do. But I can't.

By [waetherman](#) on November 15, 2011

Verified Purchase

“As a long-time Kindle fan I was eager to get my hands on a Fire. There are some great aspects; the device is quick and for the most part dead-simple to use. The screen is fantastic with good brightness and excellent color, and a very wide viewing angle. But there are some downsides too; the small bezel size makes holding it without inadvertent page-turns difficult, the lack of buttons makes controls harder, the accessible storage memory is limited to just 5GB.”



SA – Aspect Level

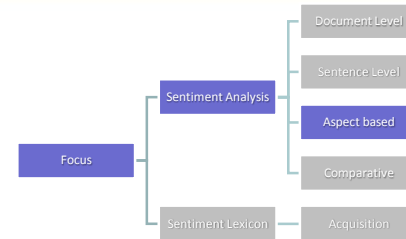
★★★★☆ I want to love it, I really do. But I can't.

By [waetherman](#) on November 15, 2011

Verified Purchase

*“As a long-time Kindle fan I was eager to get my hands on a Fire. There are some great aspects; the **device** is **quick** and for the most part **dead-simple to use**. The **screen** is **fantastic** with **good brightness** and **excellent color**, and a very **wide viewing angle**. But there are some downsides too; the **small bezel** size makes **holding** it without inadvertent page-turns **difficult**, the lack of **buttons** makes **controls harder**, the accessible **storage memory** is **limited** to just 5GB.”*

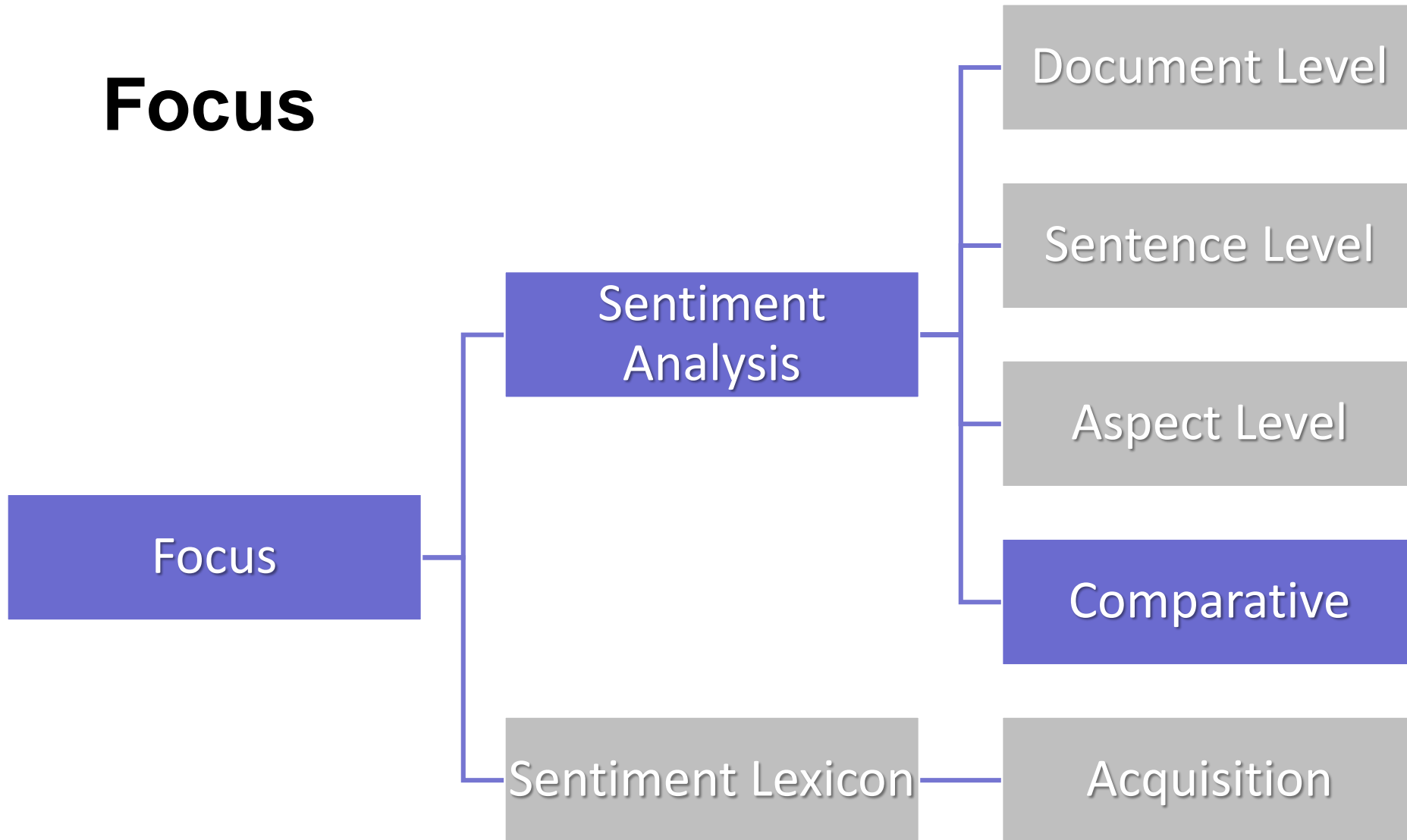


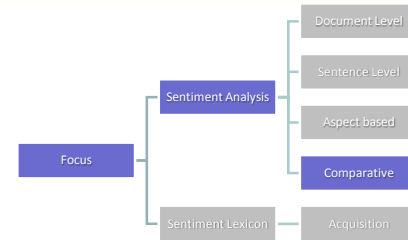


SA – Aspect Level

- Can view the problem of aspect identification as an information extraction problem
 - Use a tagged corpus to train a sequence classifier such as a Conditional Random Field (CRF)
- Implicit aspects

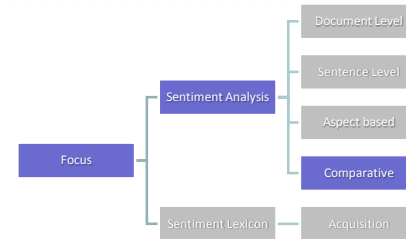
Focus





SA – Comparative

- In many cases users do not provide a direct opinion about one product but instead provide comparable opinions
- The goal of SA in this case is to identify the sentences that contain comparative opinions, and to extract the preferred entity(-ies) in each opinion.



SA – Comparative

- Comparative adjectives adverbs
 - ‘more,’ ‘less,’ and words ending with –er (for example, ‘lighter’)
- Superlative adjectives and adverbs
 - ‘most,’ ‘least,’ and words ending with –est (for example, ‘finest’)
- Additional phrases
 - ‘favor,’ ‘exceed,’ ‘outperform,’ ‘prefer,’ ‘than,’ ‘superior,’ ‘inferior,’ ‘number one’

Focus

Focus

Sentiment
Analysis

Document Level

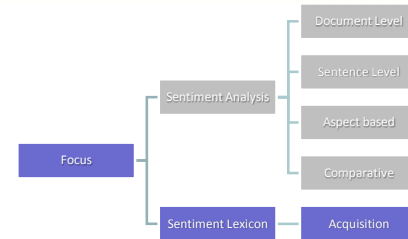
Sentence Level

Aspect Level

Comparative

Sentiment Lexicon

Acquisition



Sentiment Lexicon

- Most crucial resource for most sentiment analysis algorithms
- Acquisition
 - Manual
 - Dictionary
 - Sentiwordnet
 - Corpus
 - Sentiment consistency

Applications

Chesapeake Energy Corporation (NYSE:CHK)

0 אהבתי [Add to Portfolio](#)

From: 04/21/2012

To: 05/22/2012



Show

W

M

3M

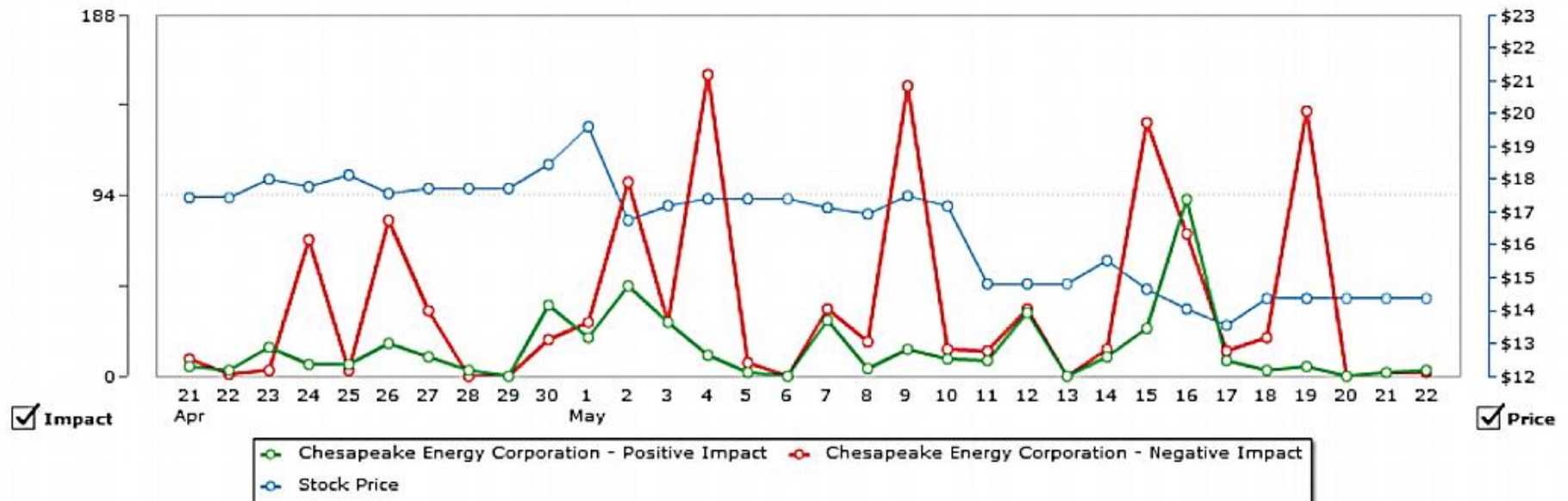
6M

Y

Upside: 31.42% Latest Target Price: 22

Impact

Price



Applications

Articles

Events

Filter

By Event Type:

- ☒ All
- ☒ Acquisition
- ☒ Analyst Recomm...
- ☒ Deals
- ☒ Employment
- ☒ Financial
- ☒ Legal
- ☒ Partnerships
- ☒ Product
- ☒ Stock Price Change

Set Filter

Sort by: Most Recent

⊖	Analyst Rating Moody's downgrades Chesapeake outlook to negative 5/9/2012 bizjournals.com
⊖	Analyst Rating Moody's Lowers Chesapeake Outlook To Negative On Spending Gap 5/9/2012 online.wsj.com
⊖	Stock Price Change Chesapeake Energy Corp (NYSE:CHK) shares were trading at \$16.93 while reporting down 1.17% from yesterdays close. 5/9/2012 themumbaitimes.com
⊖	Lawsuit Chesapeake Sued Over Claims on Executives?? Jet Use 5/9/2012 bloomberg.com
⊖	Investigation The Securities and Exchange Commission has opened an informal inquiry into Chesapeake's well program and the transactions involving McClendon. 5/8/2012 finance.yahoo.com
⊖	Investigation The Securities and Exchange Commission has opened an informal inquiry into Chesapeake's well program and the transactions involving McClendon. 5/8/2012 finance.yahoo.com

Article Preview

Send this story to a friend Email address of friend (insert comma between multiple addresses): Your email address: Copy Me Add a brief note: Send Email Processing... **Moody's downgrades Chesapeake outlook to negative** Houston Business Journal by Olivia Pulsinelli, Web producer Date: Wednesday, May 9, 2012, 1:35pm CDT Moody's Investors Service has downgraded Chesapeake Energy Corp.'s rating outlook from stable to negative, citing the company's spending funding gap and CEO Aubrey McClendon's potential effect on Chesapeake's business. Moody's Investors Service has downgraded Chesapeake Energy Corp.'s (NYSE: CHK) rating outlook from stable to negative, citing the company's spending funding gap and CEO Aubrey McClendon's potential effect on the oil and gas driller's business prospects. Moody's said Wednesday it also downgraded the ratings outlooks for Chesapeake Midstream Partners LP (NYSE: CHKM) and Chesapeake Oilfield Operating LLC from stable to negative. The downgrades are due to Chesapeake's first quarter results, which indicate a larger spending funding gap for 2012 because of lower natural gas prices and increased

Research Issues

- Automatic entity resolution
- Sarcasm
- Noisy texts
- Sentiment to objective statements



Conclusion

- Reviewed some of the main research problems within the field of SA
- Discussed ways to solve each of these problems
- Described some of the major applications
- Provided a few major open challenges

Techniques and Applications for Sentiment Analysis

Communications of the ACM | April 2013 | Vol. 56 | nº 4

DOI:10.1145/2436256.2436274

Ronen Feldman

Hebrew University

ronen.feldman@huji.ac.il

Airton Bordin Junior

airtonbjunior@gmail.com

Prof. Dr. Nádia Félix Felipe da Silva

nadia@inf.ufg.br



Introduction



“The king suite was spacious, clean, and well appointed. The reception staff, bellmen, and housekeeping were very helpful. Requests for extras from the maid were always provided. The heating and air conditioning functioned well; this was good as the weather was variable. The sofa bed was the best I’ve ever experienced. The king size bed was very comfortable. The building and rooms are very well soundproofed. The neighborhood is the best for shopping, restaurants, and access to subway. Only “complaint” has to do with high-speed Internet access. It’s only available on floors 8–12.”



Introduction



★★★★★

Odiei mt ruim trava tudo 5 estrela



★★★★★

Mto engraçado mas n consegui dar nenhuma risada ate agora.



★★★★★

Ruin



★★★★★

Rhian Se for bom eu aviso mais se for ruim falo