Natural Language Processing Applications

Week 4: Opinion mining

❏ Introduction
❏ Approaches
❏ Basic methods
❏ Evaluation

Opinion mining

Introduction

❖ Which Laptop should I buy?

❖ Which Restaurant should I go to?

❖ Which Food do I need to order?

❖ Which Service do I need to use?

Opinion mining

❖ Everyday there is a huge number of opinions regarding documents on the Internet.
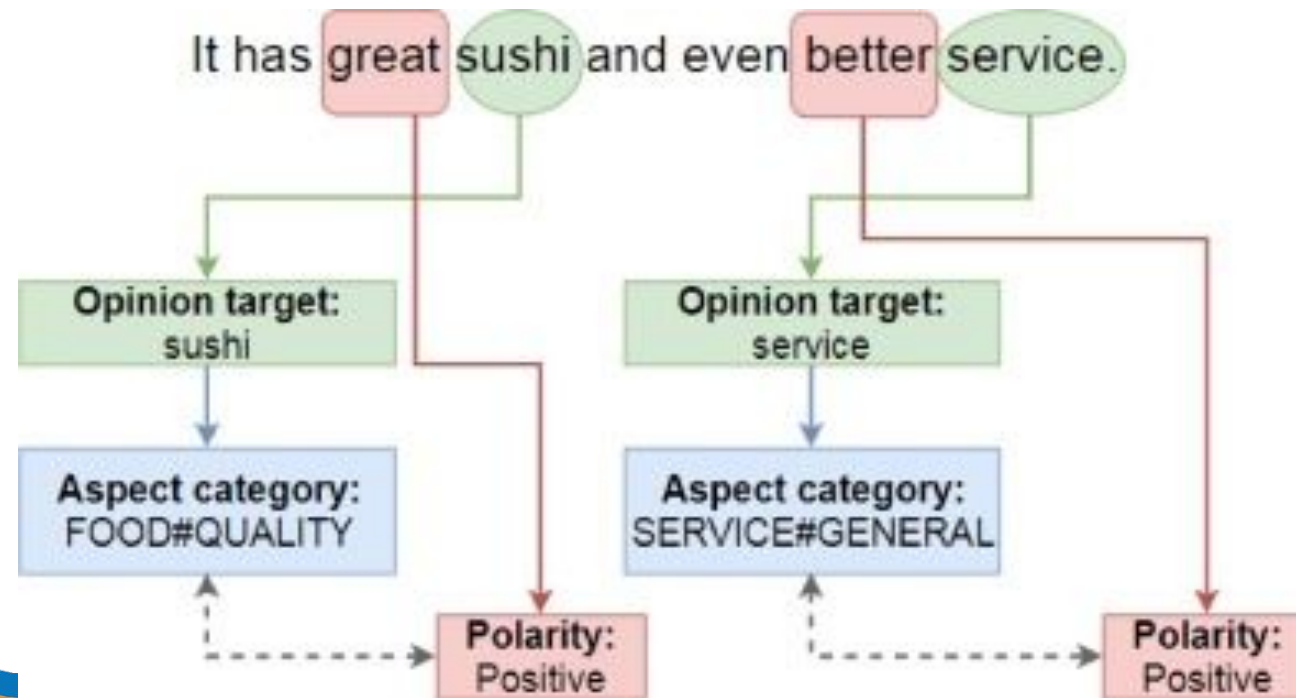
➢ Product reviews

➢ Comments

➢ ...

Opinion mining

❖ Opinion mining, or Sentiment analysis, intends to determine the **attitude** of **speaker** about a topic.

Aspect-based sentiment analysis

❖ **Aspect based** Sentiment Analysis: determine aspects of an entity and an opinion about that aspect.

Aspect-based sentiment analysis

❖ Aspect extraction:

➢ The Shrimp was awesome, but over-priced
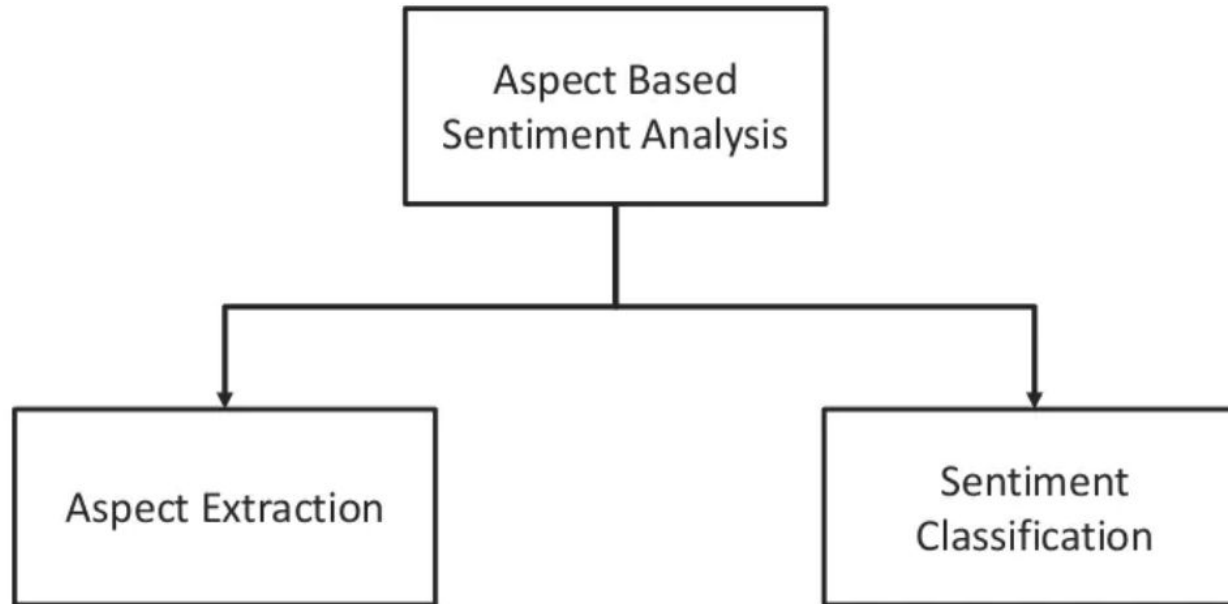
➢ {Entity#Attribute} -> {Food#Quality, Food#Prices}

❖ Opinion:

➢ The Shrimp was awesome, but over-priced

➢ {Entity#Attribute, Polarity} -> {Food#Quality, Positive}

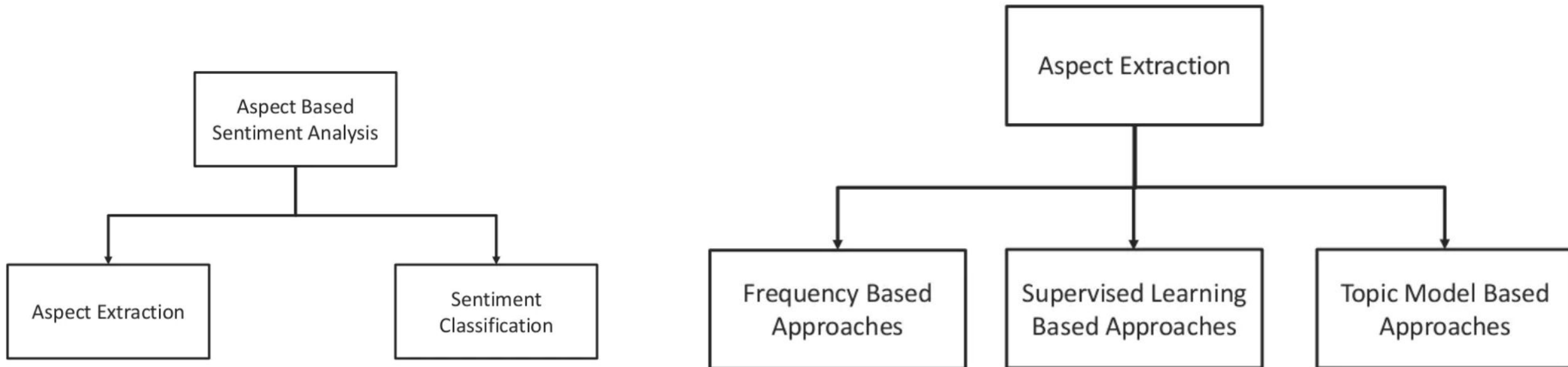➢   {Food#Prices, Negative}
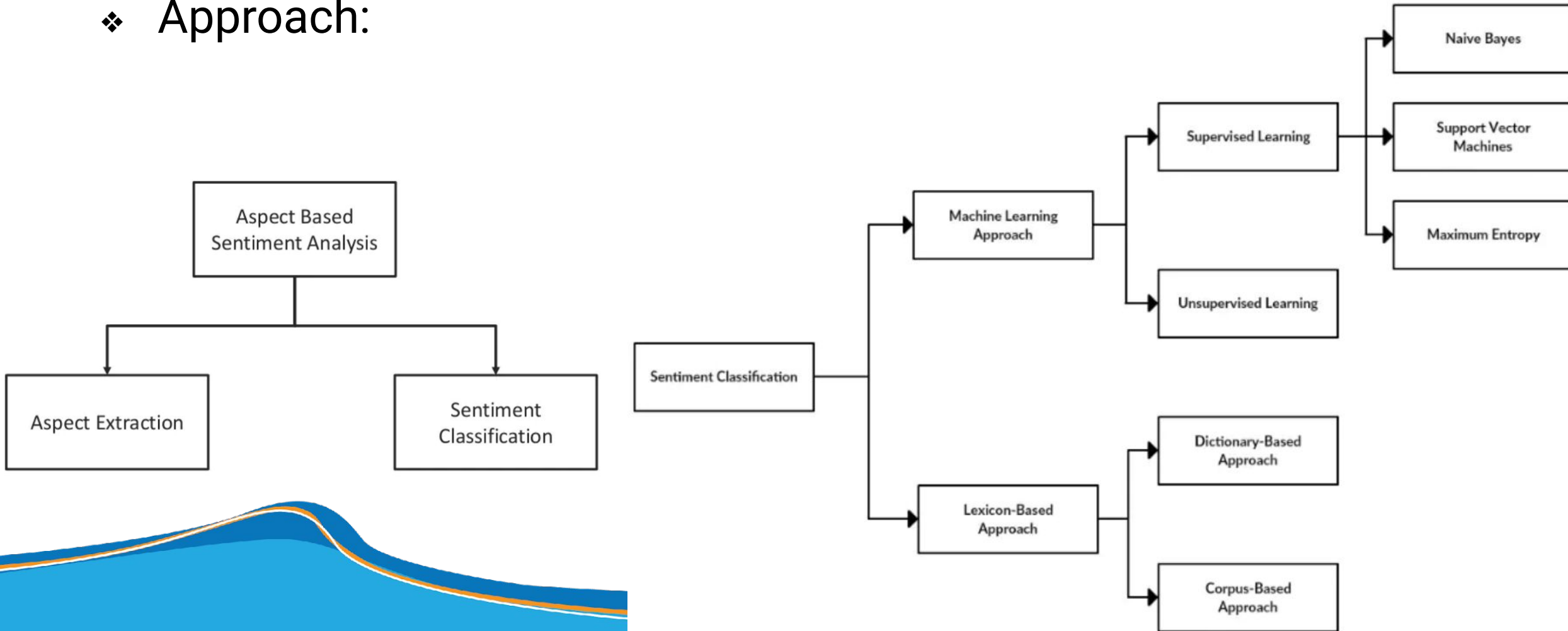
Aspect-based sentiment analysis

❖ Approach:

Aspect-based sentiment analysis

❖ Approach:

Aspect-based sentiment analysis

❖ Approach:

Aspect extraction

System	Technique	Model	Features
NLANGP	Supervised	SVM	Word Clusters, Pos tags, Head words
Sentinue	Supervised	MaxEnt	Text words and lemmas
Hu and Liu	Unsupervised	-	Noun Frequency Association Rule Mining

Sentiment classification

System	Technique	Model	Features
Wagner J. et al.	Supervised	SVM	• SentiWordNet, General Inquirer, Bing Liu (2004). • Normalized the lexicon scores
Sentinue	Supervised	MaxEnt	• Lexical features • Lexicon features • Domain specific featues
B. Pang Study	Supervised	SVM, Naïve Bayes, MaxEnt	• Unigrams, Bigrams, Adjectives, Poistion of words
Harb et al. Stuy	Unsupervised	Association Rule	• Adjectives and Adverbs

fit@hcmus
Method

Input

Preprocessing → Aspect Category Extraction → Sentiment Analyzer

Aspect Category
{Entity#Attrubute}

Polarity
Positive Negative
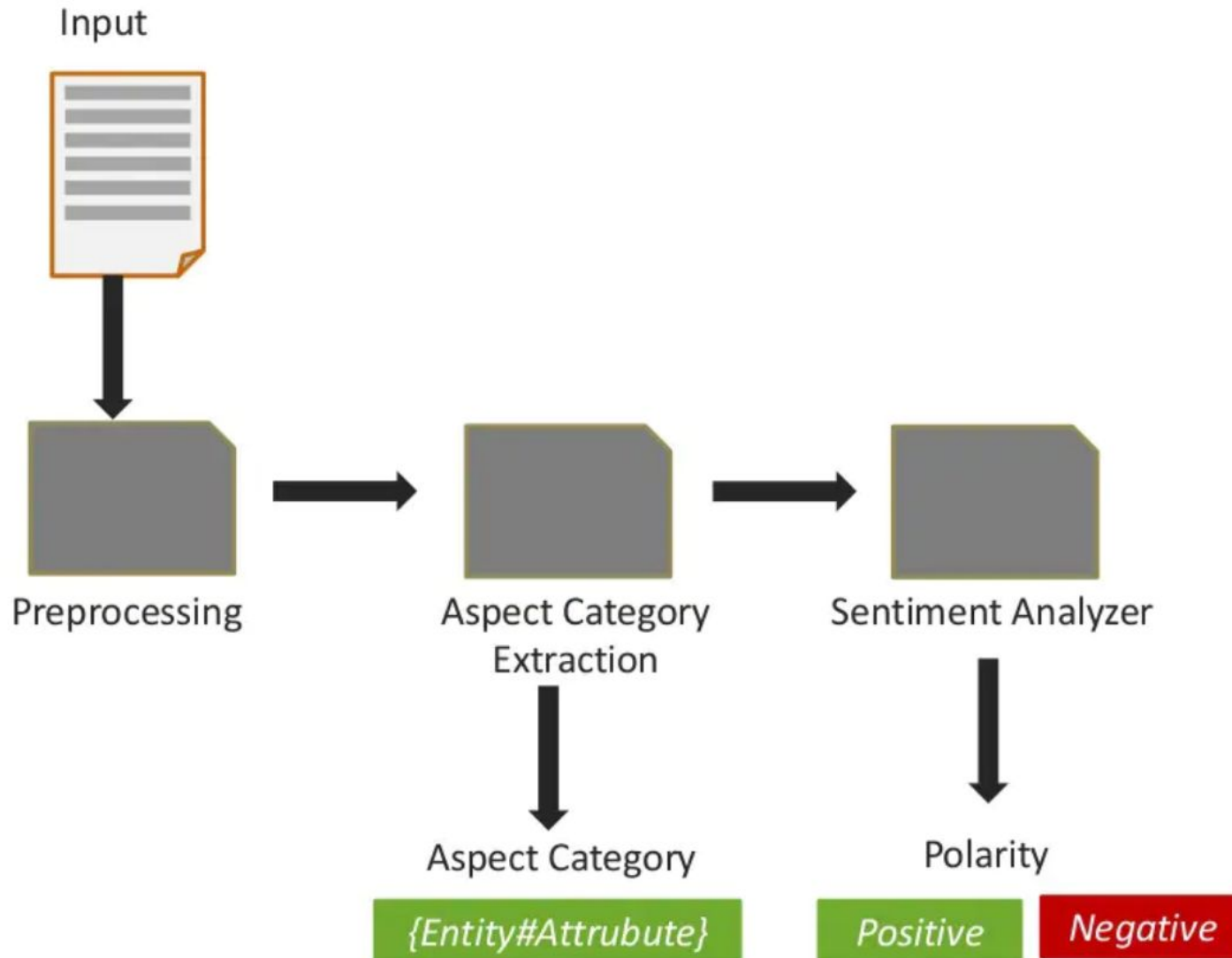
Method (cont.)

❖ Pre-processing:

➢ Whitespace, punctuation…

➢ Remove stop words

➢ Stemming

➢ Dealing with negations

Method (cont.)

❖ Aspect extraction:

  ➢ A limit number of classifiers known in advance: classification

  ➢ A set of classifiers for each aspect.

  ➢ Common models:

    ■ SVM, Maximum Entropy

Method (cont.)

❖ Sentiment analysis: Binary classification

  ➢ Models: SVM, MaxEnt, Naive Bayes

  ➢ Features classification:

   ■ Fields

   ■ Parts of speech

   ■ …

Text corpus

❖ https://github.com/yangheng95/ABSADatasets
  ➢ MAMS https://github.com/siat-nlp/MAMS-for-ABSA
  ➢ SemEval 2014:
    https://alt.qcri.org/semeval2014/task4/index.php?id=data-and-tools
  ➢ SemEval 2015:
    https://alt.qcri.org/semeval2015/task12/index.php?id=data-and-tools
  ➢ SemEval 2016:
    https://alt.qcri.org/semeval2016/task5/index.php?id=data-and-tools
  ➢ Shampoo: brightgems@github
  ➢ MOOC: jmc-123@github
  ➢ Twitter: https://dl.acm.org/doi/10.5555/2832415.2832437
  ➢ Television & TShirt: https://github.com/rajdeep345/ABSA-Reproducibility

Demo

❖ https://github.com/yangheng95/PyABSA