Temporal Opinion Mining

Burnett and Bjørkelund
Norwegian University of Science and Technology

Jason Adams, Ian Clark, Will Middleton, Antoine Rigaut

Opinion-Mining in perspective

Determine the attitude of an author based on subjective content

Temporal Opinion Mining - How and why opinions change over time

Motivation:

- Large amount of unstructured text
- Business Applications
- Reviewing Products
- Reacting to changes in sentiment

Challenges

Mixed Reviews

Sentiment towards features vs. overall sentiment

Word Sense Ambiguity - costs rise vs. revenue rises

Domain Dependent Words

Objective vs. Subjective Comparisons

Sarcasm and Irony

Ground Truth hard to define

Research Problem

Customer opinion mining on hotel reviews

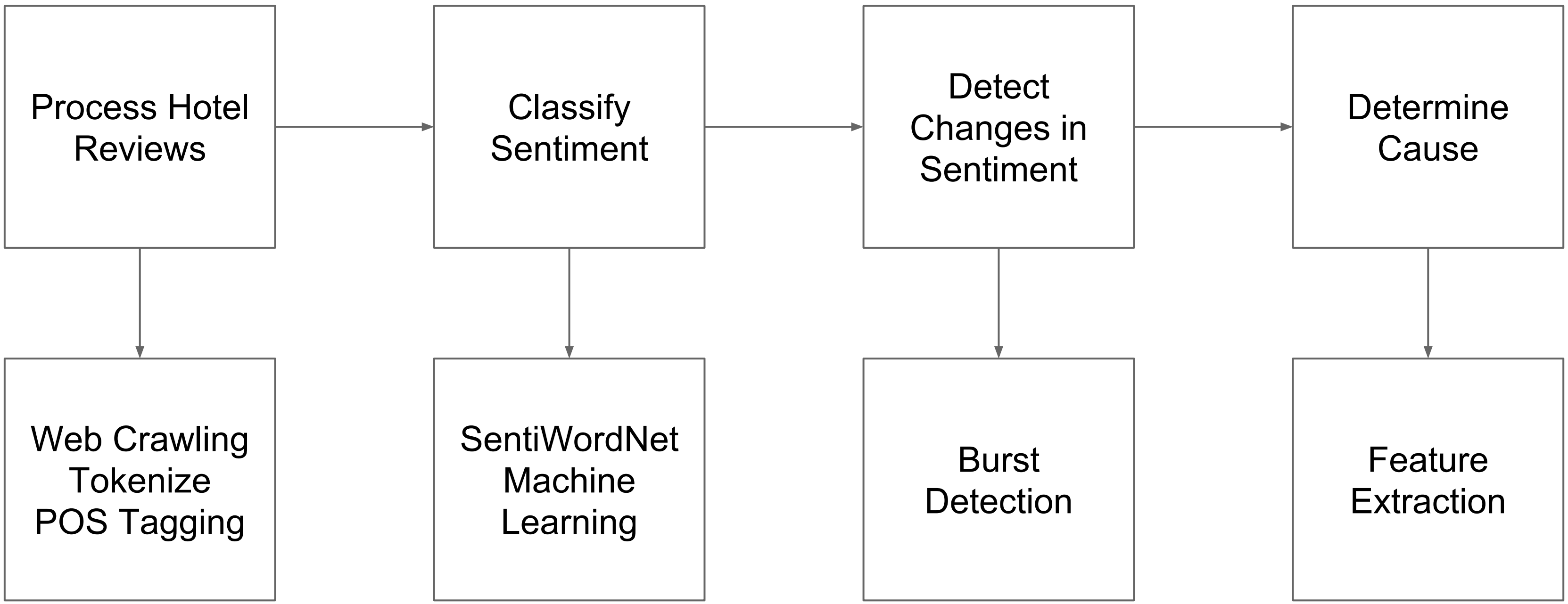
Source: Tripadvisor, Booking.com

Data: hotel location, user review text, review date, user hotel score

What do users think of hotels?

How/why have these views changed over time?

Approach


Process Hotel Reviews
Classify Sentiment
Detect Changes in Sentiment
Determine Cause
Web Crawling Tokenize POS Tagging
SentiWordNet Machine Learning
Burst Detection
Feature Extraction

Classification

Ground Truth: User rating

Classification:

- Positive/ Negative/ Neutral
- 5 Buckets (Strong/Weak Positive/Negative)

Knowledge-based approaches

SentiWord Net - classify words as positive, negative, or objective

ObjScore + PosScore + NegScore = 1

WordNet relationships used to grow known seed words

Glosses used to train a machine learner

Knowledge-based approaches

SentiWordNet as a Classifier:

1. Look up each adjective, adverb, and noun in a review
2. Sum the positive and negative scores
3. Calculate Sentiment positive/(positive + negative)
4. Classify a document as positive, negative, or neutral

Uses just the most common word sense

Also attempted to use simplified Lesk

Can determine a “degree” of sentiment

Machine Learning Approaches

Naive Bayes:

- Bayes theorem with strong independence assumptions

Language Models:

- Trained a model for each class - positive, negative, neutral
- Classified into the most probable class

Burst detection

Determine cause of sentiment change

Rule based on rate of change for user review scores

Extract features from previous period

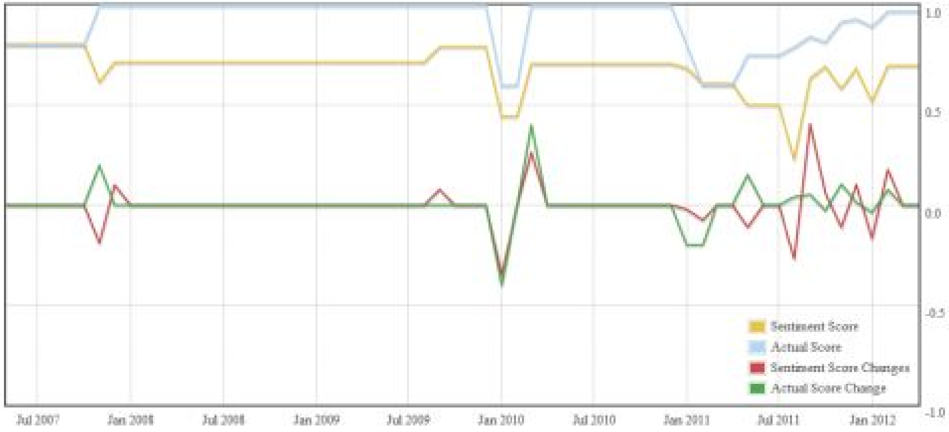


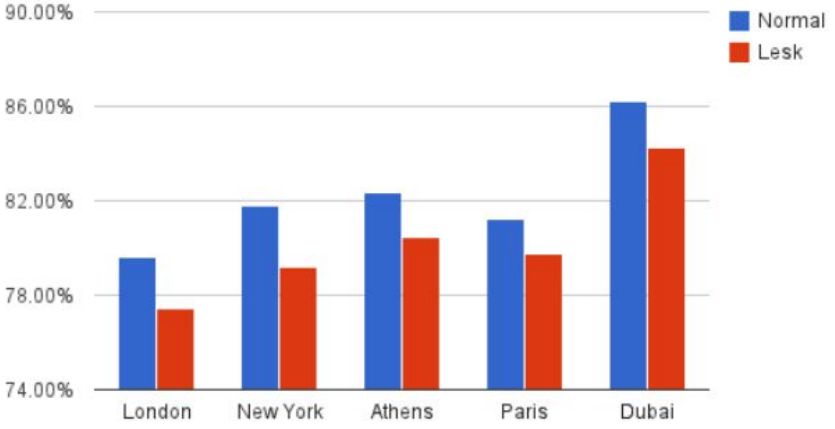
Figure 7.5: Graph from 40 reviews about San Domenico House in London, UK

Feature	# mentioned
dangerous toilet	2
poor insect	1
mouldy bed	2
infested hotel	2
unprofessional staff	1

Results -- SentiWordNet

Table 7.1: SentiWordNet accuracy results

Source	Location	# correct	# reviews	% correct
TripAdvisor	London	133,492	167,655	79.62
	New York	130,049	158,950	81.82
	Athens	14,479	17,579	82.37
	Paris	99,776	122,883	81.20
	Dubai	29,327	34,016	86.22
Booking.com	London	100,082	147,076	68.05
	New York	27,059	39,485	68.53
	Athens	8,476	11,327	74.83
	Paris	40,812	58,363	69.93
	Dubai	26,683	37,628	68.26



Results -- Machine learners

Table 7.4: Result table for machine learning algorithms

Source	Algorithm	Correct	Total	Percentage
TripAdvisor	Naive Bayes	8,924	10,000	89.24%
	Dynamic LM Classifier	9,003	10,000	90.03%
Booking.com	Naive Bayes	6,411	10,000	64.11%
	Dynamic LM CLassifier	6,592	10,000	65.92%

Table 7.5: Result table for machine learning algorithms with five categories

Source	Algorithm	Correct	Total	Percentage
TripAdvisor	Naive Bayes	5,742	10,000	57,42%
	Dynamic LM Classifier	5,712	10,000	57,12%
Booking.com	Naive Bayes	4,408	10,000	44.08%
	Dynamic LM CLassifier	4,618	10,000	46.18%

Results--Burst Detection

Table 7.6: Burst detection results with different thresholds

Threshold	Location	# correct	# bursts	% correct
1.0	London	380	712	53.37
1.0	New York	239	389	61.44
1.0	Athens	39	90	43.33
1.0	Paris	232	478	49.79
1.0	Dubai	125	246	50.81
1.0	Average			51.75
1.5	London	74	137	54.01
1.5	New York	44	66	66.67
1.5	Athens	12	23	52.17
1.5	Paris	68	124	54.84
1.5	Dubai	34	61	55.74
1.5	Average			56.69
2.0	London	18	28	64.29
2.0	New York	12	19	63.16
2.0	Athens	4	7	57.14
2.0	Paris	18	32	56.25
2.0	Dubai	8	10	80.00
2.0	Average			64.17

Conclusion

Limitations:

- Limited data set
- Lack of detail in evaluation

Applications:

- Anticipate changes in customer opinions through machine learners and burst detection.