Sentiment Analysis

by

Asif Ekbal

AI-NLP-ML Group

Department of Computer Science and Engineering

IIT Patna

Patna, India-800013

Email: asif@iitp.ac.in

asif.ekbal@gmail.com

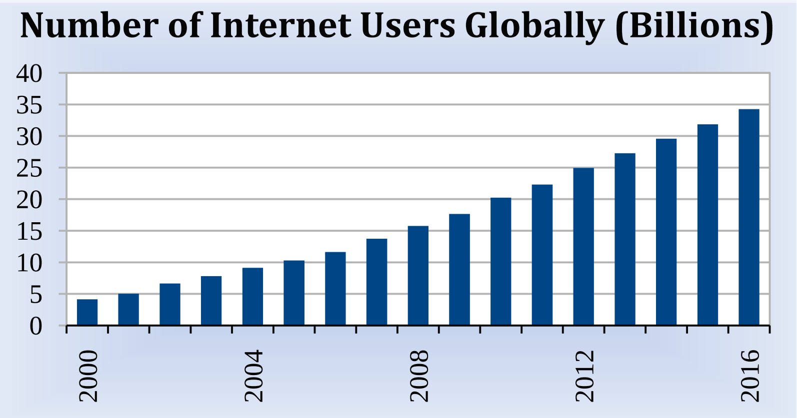
Outline

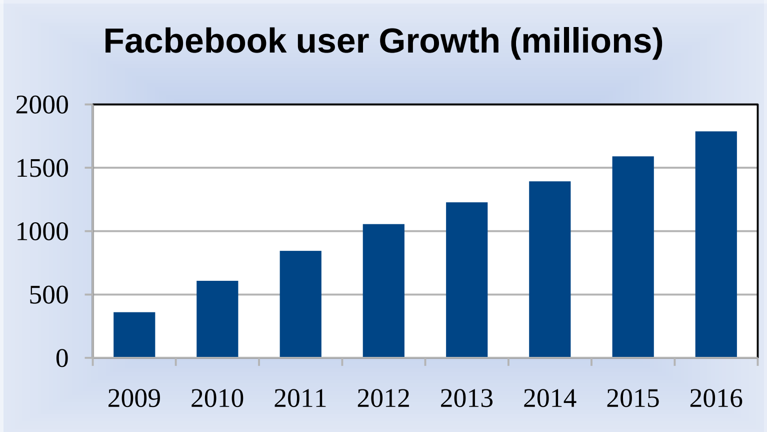
- **Background & Definition**
- Different Levels
- Feature Selection and Ensemble Construction: A Two-step Method for ABSA
  - Features
    - Aspect Term Extraction
    - Polarity Identification
  - Datasets and Experiments
- Sentiment Analysis in Indian Languages

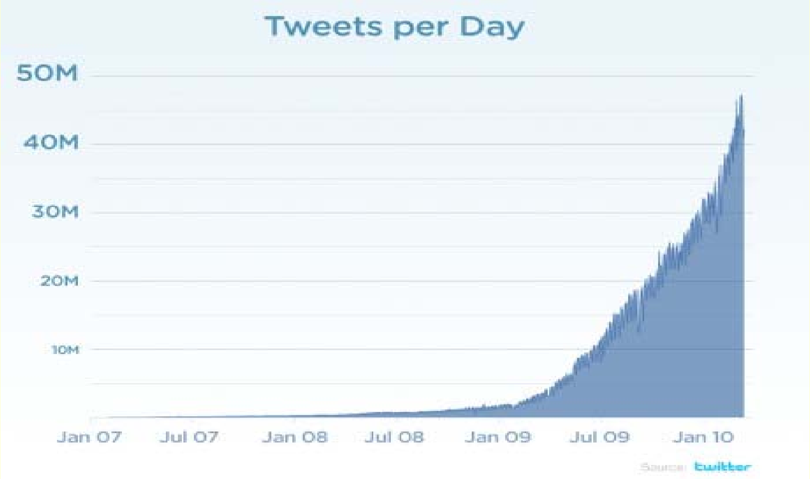
Social Media: Phenomenal Growth in Today's Web

Social Media Today

SOCIAL MEDIA: GROWTH AND IMPACT







- Advanced to 2 billion social media users in 2016, led by India
- Unprecedented volumes of user-generated contents have created new opportunities to understand social behavior and build socially intelligent systems

How Pervasive is Social Media?

People used BBM to organize marches that led to the **London Riots** (2011)

Twitter was used extensively by our politicians during the **2014 General Elections**

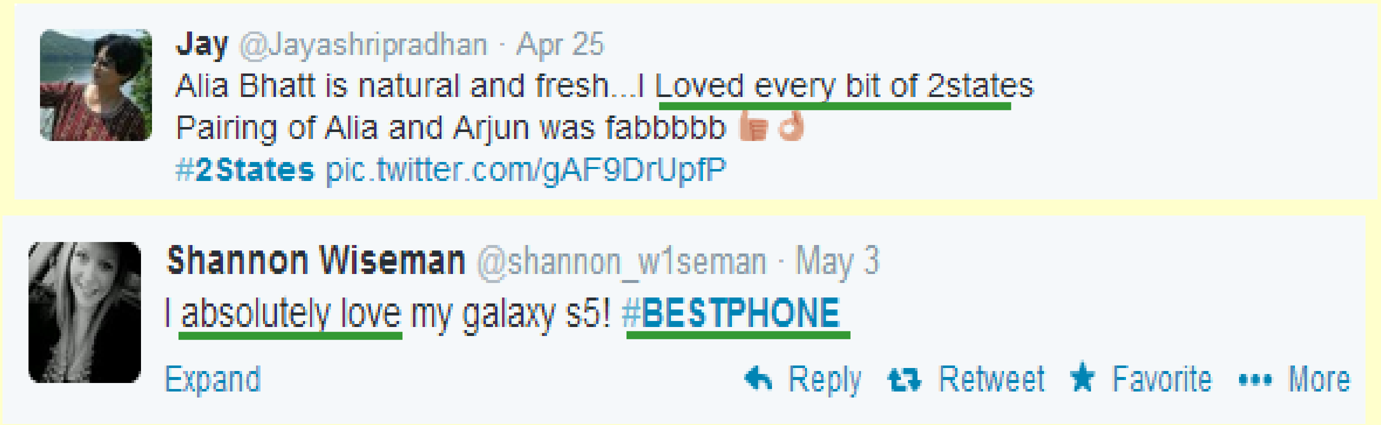
Facebook played a major role in mobilising support for the **Arab Spring**

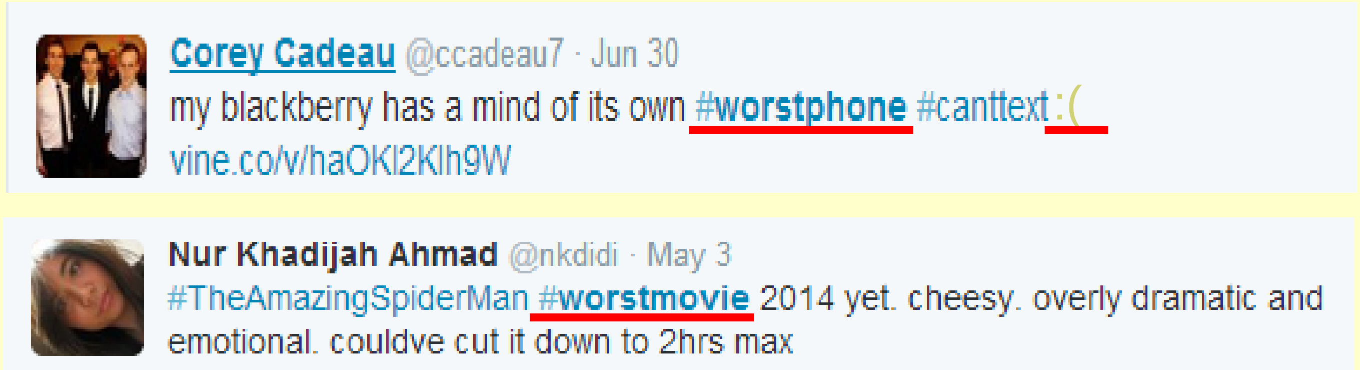
There is a dire need among governments and private organizations for accurate and <u>timely analysis</u> of data obtained from social media networks!

Sentiment analysis

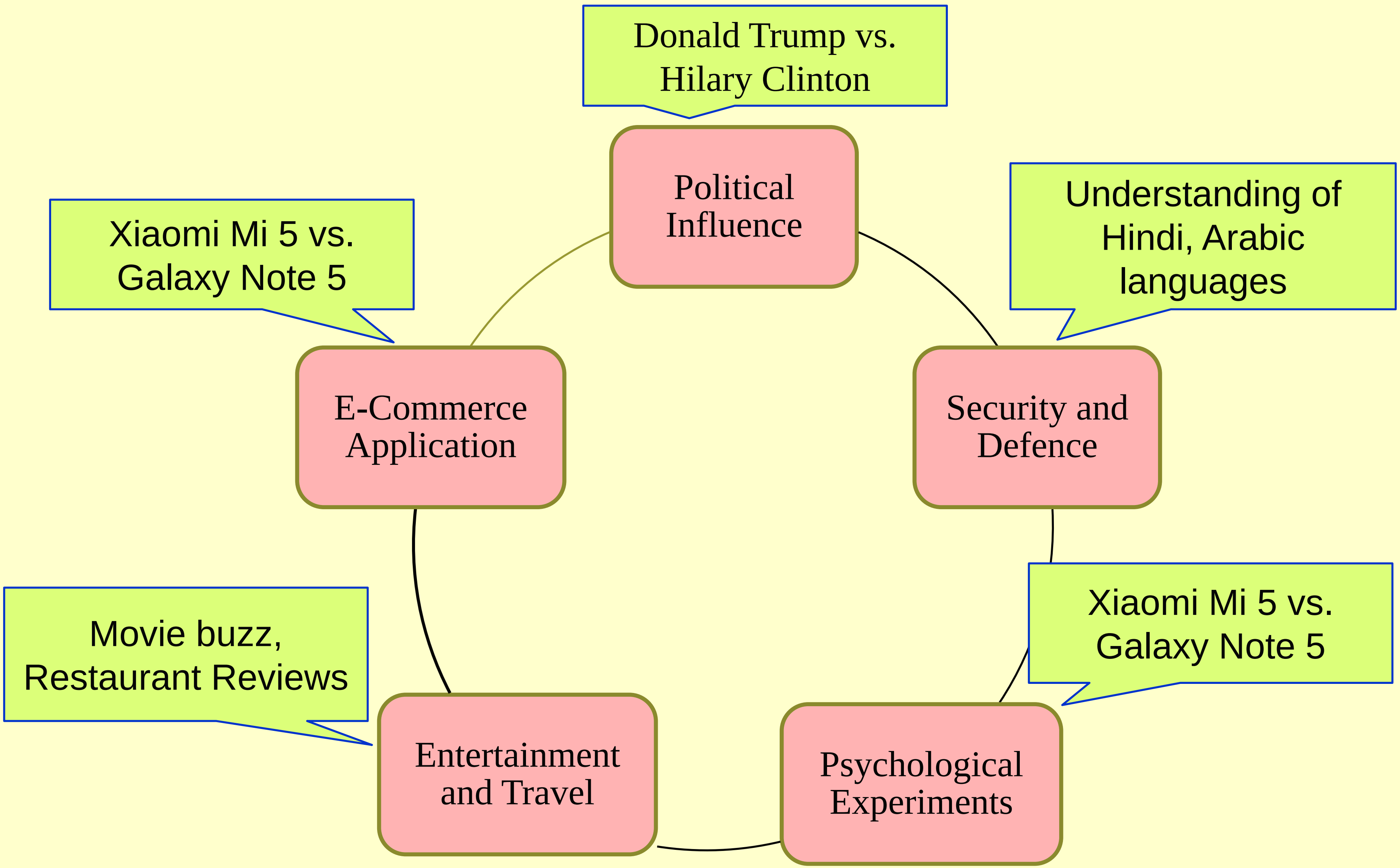
WHAT IS SENTIMENT ANALYSIS?

Sentiment analysis aims to identify the orientation of opinion in a piece of text





Why Sentiment Analysis?



“What people think?”

What others think **has always been an important piece of information**

“Which car should I buy?”

“Which schools should I apply to?”

“Which Professor to work for?”

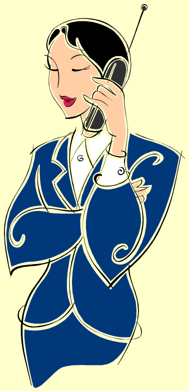
“Whom should I vote for?”

Kavita and Hyun, Tutorial on Opinion Mining

"So whom shall I ask?"

Pre Web

- Friends and relatives
- Acquaintances
- Consumer Reports



Post Web

"…I don't know who..but apparently it's a good phone. It has good battery life and…"

- Blogs (google blogs, livejournal)
- E-commerce sites (amazon, ebay)
- Review sites (CNET, PC Magazine)
- Discussion forums (*forums.craigslist.org, forums.macrumors.com*)
- Friends and Relatives (occasionally)

Kavita and Hyun, Tutorial on Opinion Mining

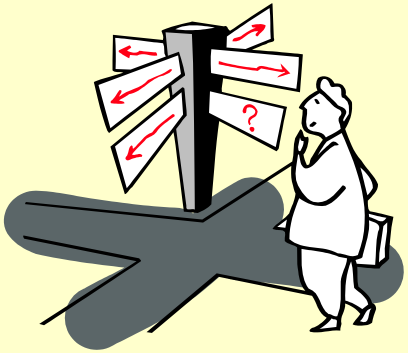
"Whoala! I have the reviews I need"

Now that I have "too much" information on one topic…I could easily form my opinion and make decisions…

Is this true?

…Not Quite

- **Searching for reviews may be difficult**
  - Can you search for opinions as conveniently as general Web search?
    eg: is it easy to search for *"iPhone vs Google Phone"*?



- **Overwhelming amount of information on one topic**
  - Difficult to analyze each and every review
  - Reviews are expressed in different ways

"the google phone is a disappointment…."

"don't waste your money on the g-phone…."

"google phone is great but I expected more in terms of…"

"…bought google phone thinking that it would be useful but…"

Kavita and Hyun, Tutorial on Opinion Mining

"Let me look at reviews on one site only…"

Problems?

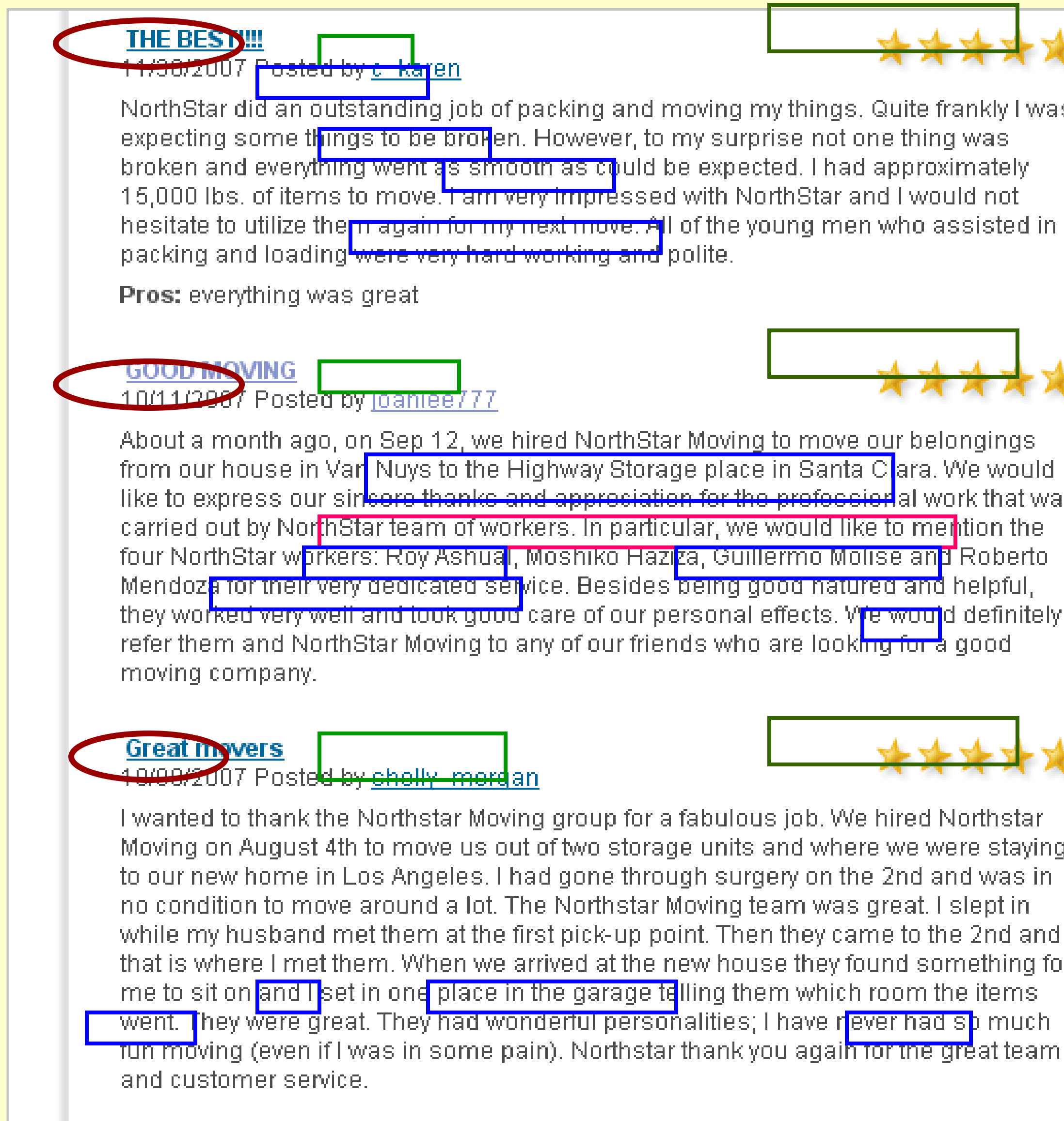
- Biased views
  - all reviewers on one site may have the same opinion
- Fake reviews/Spam (sites like YellowPages, CitySearch are prone to this)
  - people post good reviews about their own products OR services
  - some posts are plain spams

Kavita and Hyun, Tutorial on Opinion Mining

Coincidence or Fake?

Reviews for a Packers and Movers company from YellowPages

- # of merchants reviewed by each of these reviewers ➔ 1
- Review **dates close** to one another
- All rated **5 star**
- Reviewers seem to know exact names of people working in the company and **TOO many positive mentions**



THE BEST!!!!
11/30/2007 Posted by c_karen

NorthStar did an outstanding job of packing and moving my things. Quite frankly I wa
expecting some things to be broken. However, to my surprise not one thing was
broken and everything went as smooth as could be expected. I had approximately
15,000 lbs. of items to move. I am very impressed with NorthStar and I would not
hesitate to utilize them again for my next move. All of the young men who assisted in
packing and loading were very hard working and polite.

Pros: everything was great

GOOD MOVING
10/11/2007 Posted by joanlee777

About a month ago, on Sep 12, we hired NorthStar Moving to move our belongings
from our house in Van Nuys to the Highway Storage place in Santa Clara. We would
like to express our sincere thanks and appreciation for the professional work that wa
carried out by NorthStar team of workers. In particular, we would like to mention the
four NorthStar workers: Roy Ashual, Moshiko Haziza, Guillermo Molise and Roberto
Mendoza for their very dedicated service. Besides being good natured and helpful,
they worked very well and took good care of our personal effects. We would definitely
refer them and NorthStar Moving to any of our friends who are looking for a good
moving company.

Great movers
10/09/2007 Posted by shelly_morgan

I wanted to thank the Northstar Moving group for a fabulous job. We hired Northstar
Moving on August 4th to move us out of two storage units and where we were staying
to our new home in Los Angeles. I had gone through surgery on the 2nd and was in
no condition to move around a lot. The Northstar Moving team was great. I slept in
while my husband met them at the first pick-up point. Then they came to the 2nd and
that is where I met them. When we arrived at the new house they found something fo
me to sit on and I set in one place in the garage telling them which room the items
went. They were great. They had wonderful personalities; I have never had so much
fun moving (even if I was in some pain). Northstar thank you again for the great team
and customer service.

Outline

- Background & Definition
- **Different Levels**
- Feature Selection and Ensemble Construction: A Two-step Method for ABSA
    - Features
        - Aspect Term Extraction
        - Polarity Identification
    - Datasets and Experiments
- Sentiment Analysis in Indian Languages
- Conclusions

Different Levels: Sentiment Analysis

Levels of Sentiment Analysis

- Document level
  - Sentiment of the whole document
- Sentence level
  - Sentiment of each sentence in a document
- Phrase level
  - Sentiment *w.r.t.* given phrase in a sentence
- Aspect level
  - Sentiment *w.r.t.* attributes of a product discussed in a sentence

Increasing level of information

Levels of Sentiment Analysis

- Document level
  - Sentiment of the whole document

Document 1	Sentence 1. Sentence 2. … … Sentence n	**Positive**
Document 2	Sentence 1. Sentence 2. … … Sentence n	**Negative**
	…	
Document n	Sentence 1. Sentence 2. … … Sentence n	**Positive**

Levels of Sentiment Analysis

- Sentence level
  - Sentiment of each sentence in a document

Sentence 1	**Positive**
Sentence 2	**Negative**
Sentence 3	**Negative**
…	**…**
Sentence n	**Positive**

- Phrase level
  - Sentiment *w.r.t.* given phrase in a sentence

Sentence 1	w1 w2 <u>w3 w4 w5</u> w6 …	**Positive**
Sentence 2	<u>w1 w2 w3 w4</u> w5 w6 …	**Negative**
Sentence 3	w1 <u>w2 w3</u> w4 w5 w6 …	**Negative**
…	…	**…**
Sentence n	w1 w2 <u>w3 w4 w5 w6</u> …	**Positive**

Levels of Sentiment Analysis

- Aspect level
  - Sentiment *w.r.t.* attributes of a product or service discussed in a sentence

Sentence 1	w1 w2 <u>w3</u> w4 w5 w6 …	**Positive**
Sentence 2	<u>w1</u> w2 w3 w4 w5 w6 …	**Negative**
Sentence 3	w1 <u>w2 w3</u> w4 w5 w6 …	**Negative**
…	…	**…**
Sentence n	w1 w2 w3 w4 w5 <u>w6</u> …	**Positive**

Sentiment Analysis: Recent Trends

- Aspect based Sentiment Analysis (product/service reviews level)
- Sentiment Analysis in Twitter
- Sentiment Analysis of Figurative Languages in Twitter
- Sarcasm Detection

- International Workshop on Semantic Evaluation – SemEval
  - SemEval 2014 (http://alt.qcri.org/semeval2014/)
  - SemEval 2015 (http://alt.qcri.org/semeval2015/)
  - SemEval 2016 (http://alt.qcri.org/semeval2016/)

Outline

- Background & Definition
- Different Levels
- **Resources**
- Feature Selection and Ensemble Construction: A Two-step Method for ABSA
  - Features
    - Aspect Term Extraction
    - Polarity Identification
  - Datasets and Experiments
- Sentiment Analysis in Indian Languages
- Conclusions

Main resources

- **Lexicons**
  - General Inquirer (Stone et al., 1966)
  - OpinionFinder lexicon (Wiebe & Riloff, 2005)
  - SentiWordNet (Esuli & Sebastiani, 2006)
  - NRC Lexicon (http://saifmohammad.com/WebPages/lexicons.html)

- **Annotated corpora**
  - Used in statistical approaches (Hu & Liu 2004, Pang & Lee 2004)
  - MPQA corpus (Wiebe et. al, 2005)
  - SemEval datasets

- **Tools**
  - Algorithm based on minimum cuts (Pang & Lee, 2004)
  - OpinionFinder (Wiebe et. al, 2005)

Sentiment Lexicons

- Sentiment (Opinion) words are the most important indicators of sentiments
  - Positive words
    - Good, wonderful, amazing etc.
    - For e.g.: "*This camera is wonderful*"
  - Negative words
    - Bad, poor, terrible etc.
    - For e.g.: "*Battery life is too poor.*"
- A list of such words is called sentiment lexicon
- Sentiment lexicons are necessary but not sufficient for sentiment analysis
  - "*The food is very **cheap** here.*" vs "*The service is very **cheap** here.*"

Sentiment Lexicons

- Bing-Liu lexicons
  - Positive words: 2006
  - Negative words: 4783
  - Useful properties: includes mis-spellings, morphological variants, slang, and social-media mark-up
- MPQA (Multi-Perspective Question Answering) subjectivity lexicons
  - ~8000 words
  - Created from news articles from a wide variety of news sources manually annotated for opinions
- Twitter specific lexicons
  - NRC-Sentiment140
  - NRC-Hashtags-Emoticons

Sentiment Lexicons

- **SentiWordNet**
  - Positive and negative real-valued sentiment scores to WordNet synsets
  - Very widely used!

Why Aspect Level Sentiment Analysis?

Why Aspect based Sentiment Analysis?

- Most of the works focus on sentiment analysis of a text or a span of text
- **Document & Sentence level analysis** do not discover what exactly people liked and did not like!
- **Aspect Level Sentiment Analysis**
  - Aspect refers to the attribute or feature of a product or service
  - Opinion consists of a sentiment (positive, negative or neutral) and target of opinion
- *Opinion targets* helps us to understand the sentiment analysis problem better

  An example: ***Although the service is <u>not that great</u>, I still <u>love</u> this restaurant***

  - *Positive* about the *restaurant* but *negative* about the *service*

Document Level vs. Aspect Level Sentiment Analysis

Camcorder X

- The zoom is excellent, but the LCD is blurry.
- Great value for the price.
- Although the display is poor the picture quality is amazing.
- Batteries drain pretty quickly.
- I love this camera but for short battery life is definitely a pain.
- It is good camera for the price.
- ..

Product	Rating
Camcorder X	3.1

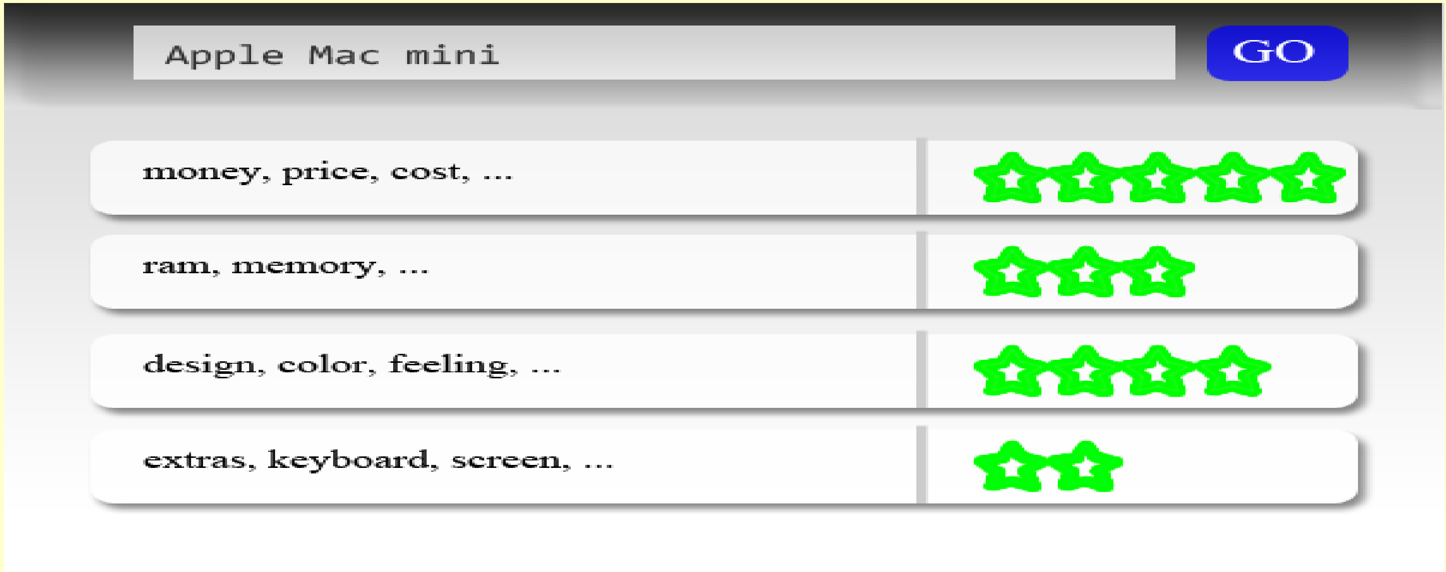
Documents level sentiment analysis

Aspect Term	Rating
Zoom	5
Price	4
Picture quality	4
Battery life	2
Screen	1
…	…

Aspect based sentiment analysis

Aspect Based Sentiment Analysis (ABSA)

- Aspect: an ***attribute or component of the product that has been commented*** on in a review
- ABSA: primarily focuses on mining relevant information from the thousands of online reviews available for a popular product or service



Aspect Term Extraction

Given a set of sentences with pre-identified entities (e.g., restaurants), identify the aspect terms present in the sentence and return a list containing all the distinct aspect terms

“ I liked the *service* and the *staff* , but not the *food* ” →

{ *service, staff , food* }

“*Ambiance* and *music* funky, which I enjoy” →

{ *Ambiance, music* }

“Awesome *form factor* and great *battery life*” →

{ *form factor, battery life* }

Polarity Identification

- For a given set of aspect terms within a sentence, determine whether the polarity of each aspect term is *positive, negative, neutral* or *conflict* (i.e., both positive and negative)

" I liked the *service* and the *staff* , but not the *food* "

$\longrightarrow$ { **service: *Positive*, staff: *Positive*, food: *Negative*** }

" I did add a *SSD drive* and *memory* "

$\longrightarrow$ { **SSD drive: *Neutral*, memory: *Neutral*** }

" The *RAM memory* is good but should have splurged for 8Mb instead of 4Mb "

$\longrightarrow$ { **RAM memory: *Conflict*** }

Outline

- Motivation, Background & Definition
- **Feature Selection and Ensemble Construction: A Two-Step Method for ABSA**
- Features
  - Aspect Term Extraction
  - Polarity Identification
- Datasets and Experiments
- Conclusions
- Future Works

***S**. Akhtar, D. Gupta, A. Ekbal and P. Bhattacharyya (2017). Feature Selection and Ensemble Construction: A Two-step Method for Aspect based Sentiment Analysis. Knowledge based Systems, Elsevier*

What we present?

Feature selection and ensemble construction for ABSA
(*Aspect Term Extraction and Sentiment Classification*)

Based on

Particle Swarm Optimization(PSO)

- Evolves the search space based on some criterion

Automatically determines

- Most relevant set of features *w.r.t.* some objective function(s)
- Most relevant set of classifiers for ensemble *w.r.t.* some objective function(s)

Why the techniques?

Feature selection and ensemble construction for ABSA

(*Aspect Term Extraction and Sentiment Classification*)

Based on

Particle Swarm Optimization(PSO)

Evolves the search space based on some criterion

Automatically determines

- The most relevant set of features *w.r.t.* some objective function(s)
- The most relevant set of classifiers for ensemble *w.r.t.* some objective function(s)

Why??

Why the techniques?

Feature selection and ensemble construction for ABSA

(*Aspect Term Extraction and Sentiment Classification*)

Based on

Why??

Particle Swarm Optimization(PSO)

Evolves the search space based on some criterion

Automatically determines

- The most relevant set of features *w.r.t.* some objective function(s)
- The most relevant set of classifiers for ensemble *w.r.t.* some objective function(s)

➢ Why Feature Selection?

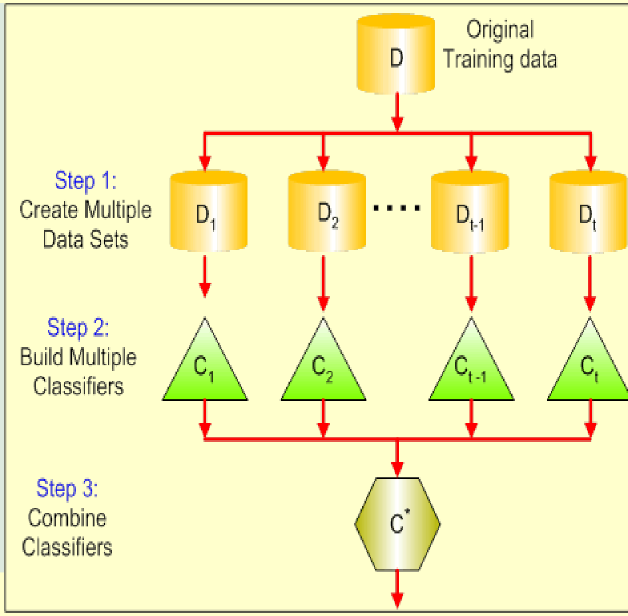
- •Improves predictive accuracy
- •Reduces complexity of the training model (Space and time both)
- **•No existing works for automated feature selection in ABSA**

➢Why Ensemble?

❑The "**best**" classifier not necessarily the ideal choice

❑Ensemble provides a way to combine predictions of multiple models

- ❑Advantage: Improves predictive accuracy
- ❑Disadvantage: it is difficult to understand an ensemble of classifiers!



➢ Why PSO?

❑ PSO **converges much faster** compared to the evolutionary optimization techniques such as Genetic Algorithm

❑ **Less computational overhead as few parameters to tune!**

Classifier Ensemble

Drawbacks of Single Classifier

- **The "best" classifier not necessarily the ideal choice**
- **For solving a classification problem, many individual classifiers with different parameters are trained**
  - The "best" classifier is selected according to some criteria e.g., *training accuracy* or *complexity of the classifiers*
- **Problems: Which one is the best?**
  - Maybe more than one classifiers meet the criteria (e.g. same training accuracy), especially in the following situations:

    -Without sufficient training data

    – Learning algorithm leads to different local optima easily

Drawbacks of Single Classifier

- Potentially valuable information may be lost by discarding the results of less-successful classifiers

  E.g., the discarded classifiers may correctly classify some samples

- **Other drawbacks**
  - Final decision must be wrong if the output of selected classifier is wrong
  - Trained classifier may not be complex enough to handle the problem

Ensemble Learning

- **Employ multiple learners and combine their predictions**
- **Methods of combination**
  - Bagging, boosting, voting
  - Error-correcting output codes
  - Mixtures of experts
  - Stacked generalization
  - Cascading
  - …
- **Advantage: improvement in predictive accuracy**
- **Disadvantage: it is difficult to understand an ensemble of classifiers**

Why Do Ensembles Work?

Dietterich(2002) showed that ensembles overcome three problems:

- ***Statistical Problem*- arises when the hypothesis space is too large for the amount of available data. Hence, there are many hypotheses with the same accuracy on the data and the learning algorithm chooses only one of them! There is a risk that the accuracy of the chosen hypothesis is low on unseen data!**
- ***Computational Problem*- arises when the learning algorithm cannot guarantee finding the best hypothesis.**
- ***Representational Problem*- arises when the hypothesis space does not contain any good approximation of the target class(es).**

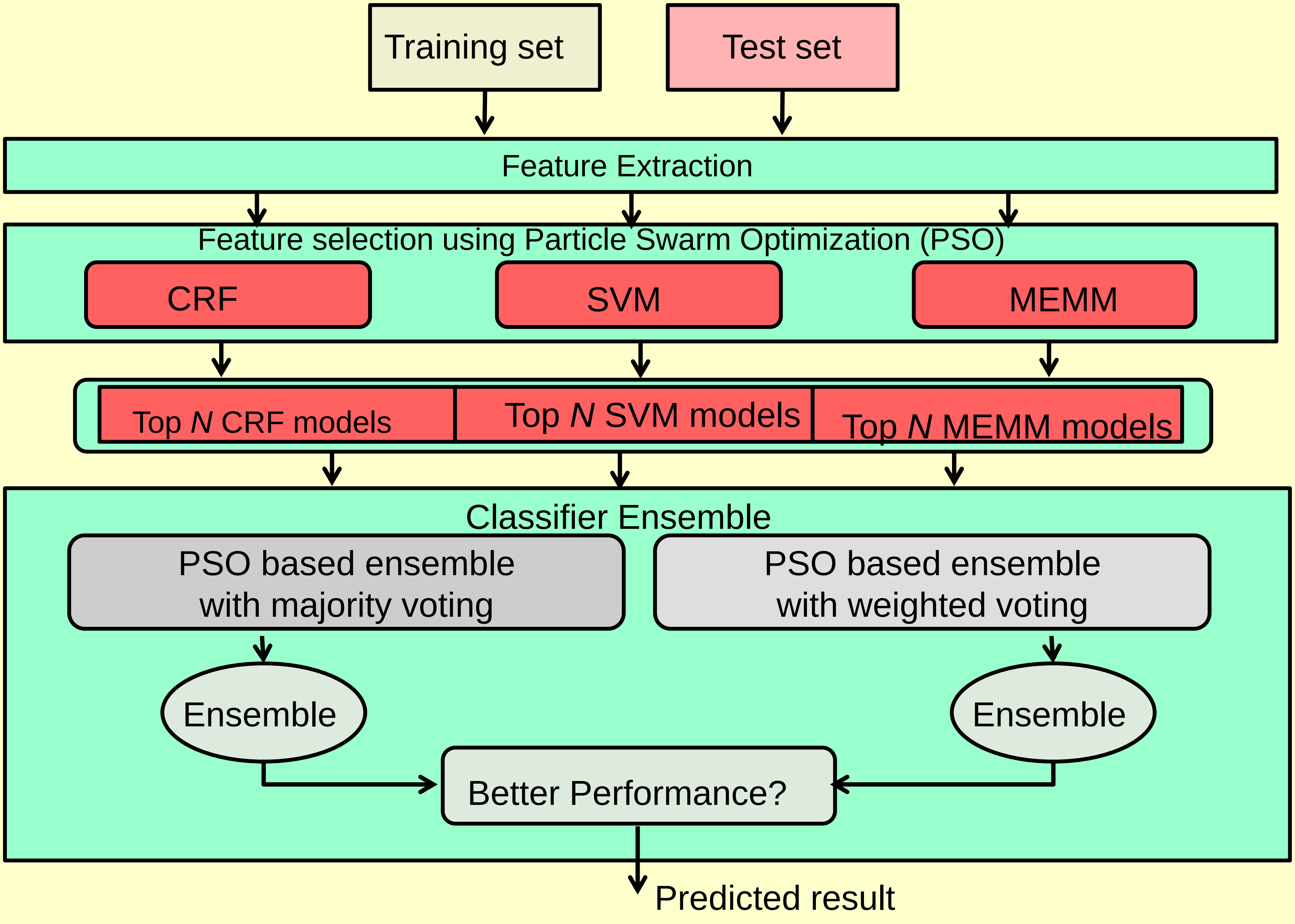
T.G. Dietterich, Ensemble Learning, 2002

Categories of Ensemble Learning

- **Methods for Independently Constructing Ensembles**
  - Bagging
  - Randomness Injection
  - Feature-Selection Ensembles
  - Error-Correcting Output Coding
- **Methods for Coordinated Construction of Ensembles**
  - Boosting
  - Stacking
  - Co-training

Overall Architecture of the Proposed Approach

Overall Architecture of the Proposed System



Feature Selection: Problem Formulation

- Given a set of features $F = \{ f_1, \ldots, f_i, \ldots, f_n \}$

  Feature selection can be formulated as follows:

  Find a subset $F' \subseteq F$ that "maximizes the learners ability to classify patterns"

  Formally *F'* should maximize some scoring function

  $\Theta : \Gamma \rightarrow \mathbb{R}$ (where Γ is the space of all possible feature subsets of *F*), i.e.

$$\boxed{F' = \arg\max_{G \in \Gamma} \{ \Theta(G) \}}$$

Aspect term extraction: maximize F-measure

Polarity identification: maximize Accuracy

Ensemble Construction: Problem Formulation

- Given a set of classifiers $C = \{c_1, \ldots, c_i, \ldots, c_m\}$

  Ensemble construction can be formulated as follows:

  Find a subset $C' \subseteq C$ that **when combined together** "maximizes the learners ability to classify patterns"

  Formally *C'* should maximize some scoring function

  $\Theta : \Gamma \to \square$ (where Γ is the space of all possible classifier subsets of *C*), i.e.

$$\boxed{C' = \arg\max_{G \in \Gamma} \{\theta(G)\}}$$

Aspect term extraction: maximize F-measure

Polarity identification: maximize Accuracy

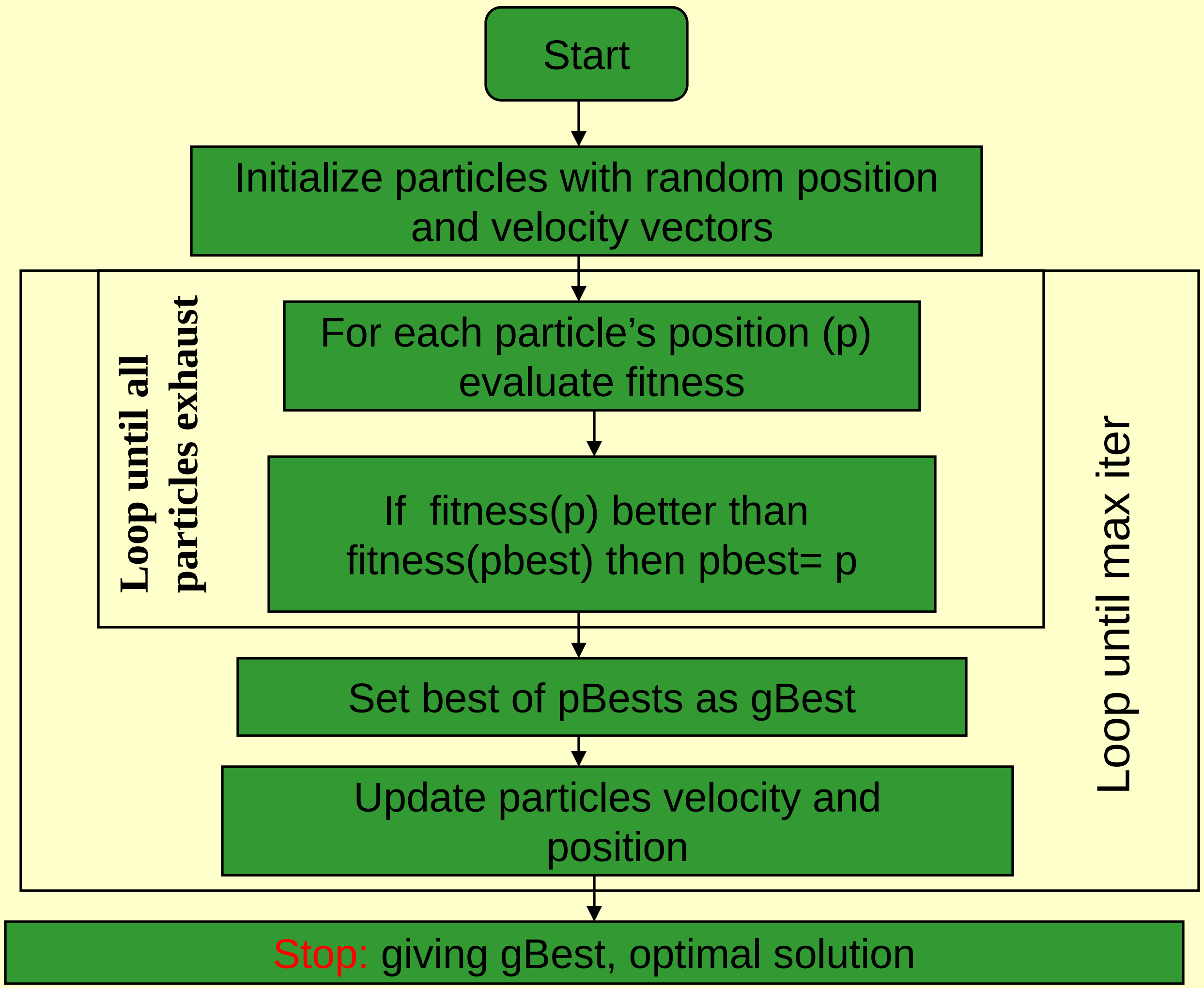
Particle Swarm Optimization (PSO)

- Robust stochastic optimization technique based on the movement and intelligence of swarms
- Applies the concept of social interaction to problem solving
- Uses a number of agents (particles) that constitute a swarm
  - Moves around in the search space looking for the best solution
- Particle
  - a point in a N-dimensional space
  - Point adjusts "flying" according to its own flying experience & the flying experience of other particles

Particle Swarm Optimization (PSO)

- Particle keeps track of the best solution in the search space
- Best solution determined by the *fitness* function
  - Personal best, ***pbest***
  - Global best solution, ***gbest***
    - Best value obtained so far by any particle in its neighborhood
- Accelerate each particle toward its *pbest* and the *gbest* locations with random weights at each time step

Particle Swarm Optimization: Work Flow



Proposed Approach: Particle Encoding

010110111110011111

- Total number of available features/classifiers: M
- **0:** at position $\boldsymbol{i}$ – i^{th} feature/classifier does not participate in classifier's training/ensemble construction
- **1** at position $\boldsymbol{i}$ – i^{th} feature/classifier participates in classifier's training/ensemble construction

Bits are randomly initialized to 0 and 1

Proposed Approach: Fitness calculation

- Feature Selection
  - Train a classifier with the features present in the particle (positions having values 1)
  - Compute the average F-measure or Accuracy based on 3-fold cross-validation
- Ensemble Construction
  - Combine predictions of selected classifiers
    - Majority Voting
    - Weighted Voting (*F-measure*)
  - Compute the average F-measure or Accuracy based on 3-fold cross-validation

Proposed Approach: Updating position

- **Velocity Update:** $v_{(i,d)} = w * v_{(i,d)} + \mu_1(b_{(i,d)} - x_{(i,d)}) + \mu_2(g_{(d)} - x_{(i,d)})$

w: weight (lies between 0 and 1; controls *global* and *local* exploration)

$x_{(i,d)}$: current position; $b_{(i,d)}$: previous best position; $g_{(d)}$: global best position; μ_1 and μ_2: cognitive and social scale parameters

- **Sampling**

$$x_{(i,d)} = \begin{cases} 1 & \text{if} \quad r < S(v_{(i,d)}) \\ 0 & \text{otherwise} \end{cases}$$

r- an uniform random number

$$S(v_{(i,d)}) = \frac{1}{1 + \exp(-v_{(i,d)})}$$

Outline

- Motivation, Background & Definition
- Feature Selection and Ensemble Construction: A Two-Step Method for ABSA
- **Features**
  - Aspect Term Extraction
  - Polarity Identification
- Datasets and Experiments
- Conclusions
- Future Works

Features: Aspect Term Extraction

- Word and Local Context
- PoS information
- Head word of the noun phrase (*word that denotes the syntactic type of the phrase- "movie"* in *"good movie"*)
- PoS of the head word
- Chunk
- Lemma
- Stop word or not
- Length
- Prefixes and suffixes of fixed length character sequences
- Word appears in the frequent aspect term list or not
- Dependency feature
    - relation when the current token is the governor ("*amod*", *"nsubj" and "dep"*)
    - relation when the current token is the dependent ("*nsubj*", *"dobj" and "dep*)

Features: Aspect Term Extraction

- Noun Synsets (e.g. noun : food) of the current token
- Named entity information: NE and aspect term often introduce ambiguities

  Ex-1: Certainly not the best *sushi* in New York

  Ex-2: I trust the people at *Go Sushi*, it never disappoints

  Aspect term in Ex-1, but *NE* in Ex-2
- Character n-grams: Character n-grams of sizes 1, 2, 3, 4 and 5
- Aspect term list (Toh and Wang, 2014):
  - Prepare two aspect term lists from each domain (high precision lists)
- Word clusters induced from the training data
- Orthographic features: Checks whether the token starts with a capital letter or starts with a digit

Features: Sentiment classification

- Words, PoS, Chunk, Prefix and Suffix features
  - Defined similar to the aspect term extraction
- Aspect term
  - Surface as well as normalized forms
- Local context: Preceding five and next five tokens surrounding the aspect term
- Lexicon based features
  - MPQA lexicon (*one feature*)
  - Bing Liu lexicon (*two features*)
  - SentiWordNet lexicon (*one feature*)
- Domain-specific words that sentiment lexicons do not cover
  - For e.g., *mouth watering, yummy* and *over cooked*
- PMI scores

Outline

- Motivation, Background & Definition
- Feature Selection and Ensemble Construction: A Two-Step Method for ABSA
- Features
  - Aspect Term Extraction
  - Polarity Identification
- **Datasets and Experiments**
- Conclusions
- Future Works

Datasets: Restaurant and Laptop Domains

Restaurants Data Sets	Training Set	Test Set
No. of Reviews	3,044	800
No. of Aspect Term	3,699	1,134

Laptop Data Sets	Training Set	Test Set
No. of Reviews	3,045	800
No. of Aspect Term	2,358	654

Results of Ensemble: Aspect Term Extraction

Selection criterion	Voting	Restaurant			Laptop		
		P	R	F	P	R	F
F-measure	Majority	86.58	81.39	83.90	83.98	65.75	73.75
	Weighted	86.27	82.01	84.09	84.70	66.05	74.22
Precision & Recall	Majority	87.07	82.01	84.46	84.99	66.67	74.72
	Weighted	87.09	82.10	84.52	85.49	66.70	74.93
Proposed Method		**87.09**	**82.10**	**84.52**	**85.49**	**66.70**	**74.93**

- ~ 2% improvements over the PSO based feature selection (4-7% overall)
- Effect of feature selection: Reduction of features significantly
- Heuristics vs. PSO: MEMM (83 vs. 41), CRF (83 vs. 44) and SVM (67: 27)
- Improvement achieved through feature selection (2-5% approximately)

Results of Ensemble: Sentiment Classification

.

Criterion	Voting	Restaurant	Laptop
		Accuracy	Accuracy
Accuracy	Majority	79.98	74.00
	Weighted	80.07	75.22
Proposed Method		**80.07**	**75.22**

- Overall performance improvement: 5-6%
- Significant feature pruning
- Heuristics vs. PSO (Restaurant): MEMM (38 vs. 20), CRF (38 vs. 16) and SVM (25: 11)
- Heuristics vs. PSO (Laptop): MEMM (22 vs. 13), CRF (22 vs. 11) and SVM (20: 11)
- Improvement achieved through feature selection (2-3% approximately)

Most Prominent Features

Most Prominent Features: Aspect Term Extraction

- **Word and context** [-2+2]
  - Sequence labelling task
- **PoS tags**
  - Most of the aspect terms are noun (Restaurant: 84% & Laptop: 80%)
- **Brown cluster** – Groups semantic similar words
  - Most of the aspect terms belong to similar group: 21(Restaurant) & 20 (Laptop) clusters
- **WordNet synset**
  - To handle unseen aspect terms (Restaurant: 31% & Laptop: 41%)

Most Prominent Features: Sentiment Classification

- **Word and context** **[-1,+1]**
  - Sentiment bearing words normally occurs nearby aspect terms
- **Lexicons**
  - BingLiu, BingLiu Direct
    - More prominent for Laptop domain (Electronics domain corpus)
  - SentiWordNet

PSO v/s Other feature selection algorithms

Classifier	Method	Restaurant				Laptop			
		Aspect		Sentiment		Aspect		Sentiment	
		F1	f_n	Acc	f_n	F1	f_n	Acc	f_n
MEMM	**PSO**	72.86	38	74.95	20	59.39	41	66.81	13
	PCA	71.33	48	68.78	24	53.96	50	63.60	18
	InfoGain	-	-	72.66	25	-	-	62.99	17
CRF	**PSO**	83.11	35	78.65	16	72.75	44	72.17	11
	PCA	81.51	49	67.27	25	70.18	52	65.96	17
	InfoGain	-	-	75.22	25	-	-	67.58	17
SVM	**PSO**	81.76	29	77.24	11	72.78	27	66.97	11
	PCA	75.73	57	74.51	25	62.88	64	64.22	17
	InfoGain	-	-	74.86	25	-	-	62.84	17

PSO v/s Other feature selection algorithms

- **PSO – Evolutionary Algorithm**
  - Randomized – Each attributes have equal chance of being selected in the feature subset.
- **PCA**
  - It selects top *k* attribute
    - Set of top 1-*k* attributes may be less relevant than set of attributes 1 and *k*+*i*
  - It ignores attributes that have low Eigen values
    - Ignored attributes may be more informative or sensitive than the selected attributes
- **Information gain**
  - Evaluate attribute by measuring information gain w.r.t. each class

PSO v/s Other ensemble techniques

Method	Restaurant		Laptop	
	Aspect	**Sentiment**	**Aspect**	**Sentiment**
PSO	84.52	80.07	74.93	75.22
Bagging	65.69	73.28	42.62	62.84
AdaBoost	65.69	70.01	42.81	59.02
Stacking	63.45	73.98	41.05	63.60
Voting	64.15	71.64	43.67	62.07

Outline

- Background & Definition
- Different Levels
- Resources
- Feature Selection using PSO for ABSA
  - Features
    - Aspect Term Extraction
    - Polarity Identification
  - Datasets and Experiments
- Sentiment Analysis in Indian Languages
- Conclusions

Aspect Based Sentiment Analysis in Hindi: Resource Creation and Evaluation

Joint Works with

Shad Akhtar and Pushpak Bhattacharyya

Challenges: Sentiment analysis in Indian languages

Challenges in Indian Languages

- Free word order nature: Difficult to locate aspect terms based on the positions

  - *कैमरा अच्छा है इस मोबाइल का।* (***kaimaraa Achchhaa hai Is mobaaIl kaa..***)
  - *अच्छा कैमरा है इस मोबाइल का।* (***Achchhaa kaimaraa hai Is mobaaIl kaa.***)
  - *इस मोबाइल का कैमरा अच्छा है।* (***Is mobaaIl kaa kaimaraa Achchhaa hai.***)

Equivalent English: This mobile has good camera

Challenges in Indian Languages

- **Scarcity of various NLP tools and resources**
  - PoS tagger
  - Chunker
  - Dependency Parser
  - Sentiment Lexicons
- **Absence of benchmark datasets**
  - Quantity of reviews – few 100s [Balamurali et. al.,2012]
  - Quality of reviews –Translated reviews [Bakliwal et. al., 2012]
  - SAIL (http://amitavadas.com/SAIL/index.html)-Twitter

Research on sentiment analysis at the aspect level is rare

Resource Creation: Data Collection

- Crawled various news, blogs, e-commerce websites[1]
- Reviews across 12 domains
  - *Mobiles, Laptops, Tablets, Cameras, Smart watches, Home Appliances, Head Phones, Speakers, Televisions, Mobile Apps, Travels, Movies*
- Total collected reviews: 8,000

[1] List of sources at the end.

Resource Creation: Data Pre-processing (1/2)

- Removed irrelevant reviews
- Dropped off many unprintable characters
- Corrected obvious spelling mistakes
- Corrected mismatched braces and quotes
- Appended missing sentence end marker

Resource Creation: Data Pre-processing (2/2)

	Review Text
Devanagari	स्क्रीन का रेज्यूलूशन 1024 गुणा 600 है , जो काफी अच्छ है
Transliterated	*skreen kaa rejyoolooshan 1024 guNNaa 600 hai , jo kaaphee Achchh hai*
Corrected (Devanagari)	स्क्रीन का रेज्यूलूशन 1024 गुणा 600 है , जो काफी अच्छा है।
Corrected (Transliterated)	*skreen kaa rejyoolooshan 1024 guNNaa 600 hai , jo kaaphee Achchhaa hai .*

Resource Creation: Data Annotation (1/3)

- Annotation Guidelines: SemEval 2014 Annotation Principle [2]
- Data Format: XML
- Aspect term extraction
    - Mark group of token(s) as aspect term
- Sentiment classification
    - Classify identified aspect terms to one of the following four classes: Positive, Negative, Neutral and Conflict

[2] http://alt.qcri.org/semeval2014/task4/

Resource Creation: Data Annotation (2/3)

```
<sentences>
  <sentence id="lap_1">
    <text> इसकी ऑडियो कालिटी शानदार है। </text>
      <aspectTerms>
        <aspectTerm from="5" to=“18“ term=" ऑडियो कालिटी" polarity=“positive" />
      </aspectTerms>
  </sentence>
  <sentence id="lap_2”>
    <text> यह बहुत महंगा है। </text>
  </sentence>
        ..
        ..
</sentences>
```

Resource Creation: Data Annotation (3/3)

- Annotators: Three native language speakers
- Cohen's Kappa Coefficient: Statistical measure of inter-rater agreement

$$K = (Pr(a) - Pr(e)) / \ 1 - Pr(e)$$

Pr(a): Agreement by observation *Pr(e): Agreement by chance*

- Average Agreement: 95.18%

Resource Creation: Data Statistics (1/2)

- Domains:12
- Review sentences: 5,417
- Aspect terms: 4,509
  - Positive aspects: 1,986
  - Negative aspects: 569
  - Neutral aspects: 1,914
  - Conflict aspects: 40

Resource creation: Data Statistics (2/2)

Domains	# Tokens	# Sentences	Aspect Terms				
			# Pos	# Neg	# Neu	# Con	Total
Laptops	6419	348	185	33	169	1	388
Mobiles	21923	1141	600	210	578	28	1416
Tablets	25323	1244	418	157	479	2	1056
Cameras	3097	150	107	11	64	1	183
Headphones	835	43	20	8	19	0	47
Home appliances	1746	84	10	0	34	0	44
Speakers	726	47	20	3	25	0	48
Smart watches	5709	330	47	22	149	2	220
Televisions	2179	135	41	3	99	1	144
Mobile apps	4577	229	98	20	46	0	164
Travels	14157	776	273	19	98	0	390
Movies	13588	890	167	83	154	5	409
Overall	100279	5417	1986	569	1914	40	4509

Methodology

- **Aspect Term Extraction**
  - Sequence labeling task
  - CRF
  - Tokenized and marked aspect terms in BIO encoding

Review Text	इसकी	ऑडियो	कालिटी	शानदार	है	।
Transliterated	*Isakee*	*AWDiyo*	*kvaaliTee*	*shaanadaara*	*hai*	.
BIO encoding	**O**	**B**	**I**	**O**	**O**	**O**

- **Sentiment Classification**
  - Multi-class classification problem
  - SVM

Features: Aspect Term Extraction

- Word and its Context
  - Surface word + local context words (-3, -2, -1, 0, +1, +2, +3)
- Part-of-Speech (PoS) Tag[3]
  - PoS tag of surface and local context words
- Chunk Information[3]
  - Helpful in identifying multiword aspect terms
- Suffixes and prefixes
  - Fixed length character sequences stripped from the beginning or end position of word

[3] http://ltrc.iiit.ac.in/showfile.php?filename=downloads/shallow_parser.php

Features: Sentiment Classification

- Target aspect term
- Local context
  - Sentiment bearing words are normally closer to the aspect terms
- Word bigram
  - Captures the co-occurrence behavior of the words

- Semantic Orientation (SO)
  - Measures association of a token towards positive and negative sentiments

$$\mathrm{SO}_t = PMI(t, Reviews_{Positive}) - PMI(t, Reviews_{Negative})$$

$PMI(t, Reviews_{Positive})$: Point-wise mutual information of t towards positive reviews

Experiments: Aspect Term Extraction

- **Classifier**
  - CRF-CRF++ toolkit [4]
- **Experimental Setup**
  - **3 fold cross validation**
- **Result (Overall)**
  - Precision: 61.96
  - Recall: 30.72
  - F-measure: 41.07

[4] http://taku910.github.io/crfpp/

Results: Aspect Term Extraction

Domain	Aspect term extraction		
	Precision	Recall	F-measure
Laptops	74.59	56.87	64.53
Mobiles	67.48	44.42	53.57
Tablets	61.50	33.67	43.52
Cameras	60.0	31.76	41.53
Headphones	100.0	27.78	43.47
Home Appliances	100.0	16.67	28.57
Speaker	83.33	22.72	35.71
Smart watch	50.0	41.50	45.36
Television	75.60	42.46	54.38
Mobile Apps.	50.0	18.0	26.47
Travels	32.60	9.77	15.03
Movies	70.14	58.02	63.51
Overall	**61.96**	**30.72**	**41.07**

Experiments: Sentiment Classification

- **Classifier**
  - Support Vector Machine (SVM) : TinySVM toolkit [5]
- **Experimental Setup**
  - 3-fold cross validation
- **Result**
  - Accuracy: 54.05% (Overall)

[5] https://chasen.org/taku/software/TinySVM/

Results: Sentiment Classification

Domain	Sentiment classifictaion
	Accuracy
Laptops	50.98
Mobiles	54.07
Tablets	57.19
Cameras	59.06
Headphones	46.15
Home Appl.	79.23
Speaker	53.84
Smart watch	64.70
Television	65.47
Mobile Apps.	61.53
Travels	68.78
Movies	39.23
Overall	**54.05**

Error Analysis: Aspect Term Extraction

- **Presence of preposition and conjunction**
  - Presence of preposition & conjunction inside aspect term confuses the system to correctly identify its boundary

Review Text	…	डिस्प्ले	की	व्यइंग	एंगल	और	ब्राइटनेस	…
Transliterated	…	*Disple*	*kee*	*vyaINg*	*ENgal*	*AOra*	*braaITanes*	…
True class	**O**	**B**	**I**	**I**	**I**	**I**	**I**	**O**
Predicted class	**O**	**B**	**I**	**I**	**I**	**O**	**O**	**O**

Error Analysis: Aspect Term Extraction

- **Noun phrase in neighborhood**
  - When a noun phrase precedes or succeeds an aspect term the system marks the neighboring phrase as aspect term along with the target aspect term

Review Text	*3*	*मेगापिक्सेल*	*रियर*	*कैमरा*	*720 पी*	*वीडियो*	*रिकॉर्डिंग*	…
Transliterated	*3*	*megaapiksel*	*riyara*	*kaimaraa*	*720 pee*	*veeDiyo*	*rikawrDiNg*	…
True class	**O**	**O**	**B**	**I**	**O**	**O**	**O**	**O**
Predicted class	**B**	**I**	**I**	**I**	**I**	**I**	**I**	**O**

Error Analysis: Sentiment classification

Distant sentiment words

- If sentiment words occurs at a far distant from target aspect term, the system fails to capture the correct sentiment

Review Text	*पेरिस्कोप एप्प में रिज्यूमे नोटिफिकेशन फीचर दिया गया है जो बहुत ही खास है।*
Transliterated	*periskop Ep meN rijyoom noTiphikeshan pheechara diyaa gayaa hai jo bahut hee khaas hai.*
Target aspect term and its true class	*रिज्यूमे नोटिफिकेशन (rijyoom noTiphikeshan) : Positive*
Sentiment bearing word	*खास (khaas)*
Distance	**8 words**

Sentiment Analysis In Indian Language Using Lexical Acquisition

Joint works with
Ayush Kumar, Sarah Kohail and
Chris Biemann

Sentiment Analysis in Indian Languages

Domain: Twitter

To identify the polarity of the given tweet in positive, negative or neutral classes

Languages: Hindi and Bengali

Positive: happiness, satisfaction, surprise, trust

केन्द्र सरकार ने बिहार को दिया विशेष पैकेज, 100 करोड़ रुपये की आर्थिक सहायता।

Negative: grief, hatred, despair

इस धोनी की वजह से ही आज युवी टीम से बाहर है!

Neutral: no opinion

@rkl @viratt_23 कार्टून बनाता है क्या?

Lexical Acquisition

"You shall know a word by the company it keeps."

- Distributional Hypothesis: Words that occur in same context tend to have similar meanings
- Our method employs the same concept to expand existing lexicon using computation on external dataset
- Useful in capturing and overcoming rare and unseen words

Lexical Expansion

- Distributional Thesaurus (DT)
  - An automatically computed resource that relates words according to their similarity (morphological variations/synonyms/closely associated words)
  - For every sufficiently frequent word, the most similar words as captured over the holing operations

अतुलनीय (atulnIya)	अद्भुत (adabhuta)	महान (mahAna)	शानदार (shAnadAra)
तर्कसंगत (tarkasangata)	उचित (uchita)	सही (sahI)	गलत (galata)
धार्मिक (dhArmika)	सामाजिक (sAmAjika)	राजनीतिक (rAjanItika)	हिंदू (hindU)
ऊँची (UNchI)	ऊंची (UnchI)	लंबी (lambI)	छोटी (ChotI)

Computation of DT entries

- Holing Operation and Similarity Calculation

- **Extract Jo’s and Bim’s**
  - Jo’s: word, lemma, ngram … (call it as ***Terms***)
  - Bim’s: neighbouring words, dependency parsers, … (call is as ***Features***)

- **Similarity Computation**: Compute similarity between Jo’s (and also Bim’s)

Holing Operations

- **Example:** ***Sachin plays good Cricket***

- **Trigram @@ Operation**

JO (Term)	BIM (Feature)	POSSIBLE EXPANSIONS
• Sachin	Trigram(_, @, plays)	Dravid, Sehwag, Virat …
• Plays	Trigram(Sachin, @, good)	play, enjoys, supports …
• Good	Trigram(plays, @, cricket)	passionate, beautiful, excellent …
• Cricket	Trigram(good, @, _)	restaurant, camera, cricket …

DT Computation

Output of Holing Operation

Term	Feature
Sachin	T(of, @, is)
Jadeja	T(_, @ , bowled)
IPL	T(of, @, final)
huge	T(a, @, six)
…	…

Term Counts

Term	Count
Sachin	5782
huge	9827
Jadeja	1281
Injury	968
…	…

Feature Counts

Term	Count
T(of, @, is)	798
T(took, @, wicket)	592
T(an, @, on)	123
T(a, @, six)	704
…	

Term - Feature Counts

Term	Feature	Count
huge	T(a, @, six)	296
huge	T(a, @, ground)	123
huge	T(a, @, wicket)	189
Sachin	T(by, @, _)	
…	…	…

Compute Significance Scores		
Term	**Feature**	**Score**
huge	T(a, @, six)	547.1
big	T(a, @, six)	259.8
big	T(a, @, wicket)	458.6
long	T(too, @, _)	128.3
…	…	…

Aggregate Term Per Feature	
Feature	**Term**
T(a, @, six)	big, long, huge, flat, …
T(plays, @, cricket)	good, honest, beautiful, …
T(_, @, cricket)	big, huge, priceless, good, …
….	….

Count Similar Terms		
Term	**Feature**	**Count**
big	long	232
big	flat	18
huge	big	193
good	big	10
…	…	…

DT Expansion

DT Expansions	
Term	**Expansion**
good	great, decent, brilliant, well-balanced, …
big	long, huge, mega, gigantic, massive, …
sponsor	sponsors, promotor, organizer, …
wicket	wickets, pitch, stump, innings, …
….	….

DT Computation

1. Get Jo-Bim pairs from Holing Operations
2. Compute significance score for each Jo-Bim pair
3. Sort and prune the list
4. Aggregate Jos per Bim basket
5. Sort similar Jos based on the number of baskets they occur together

Corpus Overview

- Corpus obtained from Leipzig site: http://corpora.informatik.uni-leipzig.de
- Hindi: 2,358,708 sentences; 45,580,789 tokens (Newspaper corpus from 2011)
- Bengali: 109,855 sentences; 1,511,208 tokens (Newspaper corpus from 2011)

Lexical Expansion (contd.)

- **Co-Occurrence (COOC)**
  - Obtained a list of words that co-occur significantly with other words in a sentence.

अतुलनीय (atulnIya)	भारतीय (bhAratIya)	अन्य (anya)	वर्ष (warSha)
तर्कसंगत (tarkasangata)	कहना (kahanA)	ज्यादा (jyadA)	काफी (kAphI)
धार्मिक (dhArmika)	परंपराओं (paramparAon)	अपितु (apitu)	संतों (santon)
ऊँची (UNchI)	इमारत (imArata)	जाति (jAti)	जगहों (jagahon)

DT_COOC Lexicon

- Used the given SentiWordNet for both languages as the seed corpus for expansion
- Obtained top 125 DT expansion for all the words in seed corpus
- Ranked each word in the complete expanded corpus with score of

(No. of occurrences/Word Count in DT Corpus) to obtain candidate terms

- Score1 of candidate words were calculated as:

(No. of positive occurrences-No. of negative occurrences) / Total occurrences

- Another score (Score2) for each word is obtained using sentence level concurrences of words using the same formula

- To get a final expanded list, only the agreement between two polarity list is matched and only those words that match in absolute polarity is taken into the final list

Expansion Statistics

- Used the given SentiWordNet for both languages as the seed corpus for expansion.
- Initial Seed Corpus:

  Hindi: 1587 negative, 1314 positive words.

  Bengali: 2534 negative, 1390 positive words.

- 1st expanded corpus:

  Hindi: 3331 negative, 3980 positive.

  Bengali: 10005 negative, 1205 positive.

- 2nd expanded corpus:

  Hindi: 3926 negative, 5521 positive.

  Bengali: ---

Balanced Set for Bengali

- 1st Expanded Corpus: 1461 negative, 7213 positive

- Did not go for 2nd expansion as the expanded is still skewed though less than the previous corpus.

Comparison with SentiWordNet

- **Hindi:**

  Total found: 1215, Match: 1018, Percentage Match: 83.78%
  Total positive words found: 569.0, Match: 493.0, Percentage Match: 86.64%
  Total negative found: 646.0, Match: 525.0, Percentage Match: 81.27%

- **Bengali:**

  Total found: 380, Match: 294, Percentage Match: 77.37%
  Total positive words found: 205.0, Match: 187.0, Percentage Match: 91.22%
  Total negative found: 175.0, Match: 107.0, Percentage Match: 61.14%

Coverage of DT_COOC Lexicon

- Only 17.57% and 25.98% of the adjectives in the training and test set appear in the Hindi SentiWordNet.

- Coverage improves to 36.56% and 42.29% adjectives using DT_COOC lexicon.

Results: Feature Ablation Study

Features	Accuracy: Hindi	Accuracy: Bengali
All	47.96	42.00
All - Word Ngram	43.25 (-4.71)	38.40 (-2.80)
All – Character Ngram	47.75 (-0.21)	42.20 (+0.20)
All - SentiWordNet	47.32 (-0.64)	41.20 (-0.80)
All – DT_COOC Lexicon	49.03 (+1.07)	42.20 (+0.20)

Thank you for your attention!

References

- Bing Liu
  Sentiment Analysis and Opinion Mining.
  Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers 2012

- S. Brody and N. Elhadad.
  An unsupervised aspect-sentiment model for online reviews.
  In Proceedings of NAACL, pages 804–812, Los Ángeles, CA, 2010

- M. Hu and Bing Liu.
  Mining and summarizing customer reviews.
  In Proceedings of the 10th KDD, pages 168–177, Seattle, WAs, 2004

- Kennedy J and Eberhart R.C.
  Swarm Intelligence.
  Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2001)

Comparisons: Aspect term extraction

	Model	F-score	Remarks
Restaurant	(Toh and Wang, 2014)- CRF	84.01 78.34 (C)	External name list, Word cluster from large dataset
	Ours	84.52	Much less features
Laptop	(Maryna Chernyshevich, 2014)- CRF	74.55	Additional resources and rule-based SA tool
	Ours	74.93	Less features and not much domain-specific resources

Comparisons: Sentiment classification

	Model	F-score	Remarks
Restaurant	(Wagner et al., 2014)-SVM-SMO	80.95	Bag of words features, rule-based system, heuristics to combine lexicons
	Ours	80.07	Less external resources, less no. of features
Laptop	a. (Wagner et al., 2014)-SVM-SMO b. (Kiritchenko et., 2014)	70.48 70.48	a. Bag of words features, rule-based system, heuristics to combine lexicons b. Extensive feature sets
	Ours	75.22	Less features and not much domain-specific resources

Different PSO Parameter Values

- Pederson (2010) suggest several ways to select parameter settings.
- For the problem at hand, we cross-validated and selected four (near)-optimal different parameter settings.

Run	# Particle	# Iteration	Inertia weight (w)	μ_1	μ_2
PSO_{Run_1}	50	100	0.3593	-0.7238	2.0289
PSO_{Run_2}			0.7298	1.49618	1.49618
PSO_{Run_3}			-0.3699	-0.1207	3.3657
PSO_{Run_4}			-0.4349	-0.6504	2.2073

Features: Aspect Term Extraction

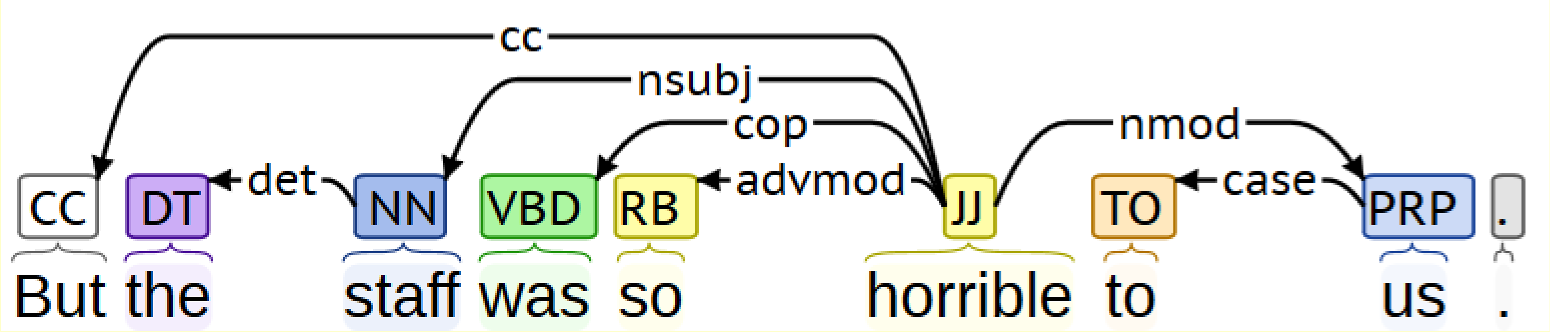
- Length
  - Longer words have a tendency of being aspect term (set to 5)
- Prefixes and suffixes of fixed length character sequences
  - Stripped up to four characters
- Word appears in the frequent aspect term list or not
  - Most frequent (at least five times) aspect terms extracted from the training data
  - Checks whether the current word appears in this list or not
- Dependency feature (Toh and Wang, 2014)
  - Stanford dependency features
  - relation when the current token is the governor ("*amod*", "*nsubj*" *and* "*dep*")
  - relation when the current token is the dependent ("*nsubj*", "*dobj*" *and* "*dep*)

Features: Aspect Term Extraction

- Dependency feature (Toh and Wang, 2014)

But the staff was so horrible to us .

CoreNLP output



Feature value for "*staff*"

1. null, null, null

2. nsub(horrible, staff), null, null

Features: Aspect Term Extraction

- Noun Synsets (e.g. noun : food) of the current token
- Aspect term list (Toh and Wang, 2014):
  - High precision aspect term list
  - Prepare two aspect term lists from each domain
  - Aspect terms appearing more than certain threshold (say, c1) kept
  - For multi-word aspect term, count the frequencies of single words
    - Keep words that appear above certain threshold value (here, c2)
    - Compute its probability of being annotated as aspect term from the training data
    - Keep those words that have probabilities above certain threshold value
  - Define two binary-valued features

Features: Aspect Term Extraction

- Word clusters induced from the training data
  - Apply Brown clustering on the training data
  - Induce 1000 clusters
  - Use a prefix (here 5) as feature for each token
  - Results can be better if clusters can be induced from the unlabelled data
- Semantic orientation scores
  - Compute PMI score of each aspect term to decide how much it is associated with the positive and how much with the negative reviews
  - Used Amazon product reviews for the same (5-star: +ve; 1-star:-ve)
- Orthographic features: Checks whether the token starts with a capital letter or starts with a digit

Features: Sentiment classification

- Words, PoS, Chunk, Prefix and Suffix features
  - Defined similar to the aspect term extraction
- Aspect term
  - Surface forms of aspect terms converted to lowercase
  - Actual surface forms + lower-cased wordform
- Local context: Preceding five and next five tokens surrounding the aspect term
- Lexicon based features
  - MPQA lexicon (*one feature*)
  - Bing Liu lexicon (*two features*)
  - SentiWordNet lexicon (*one feature*)
- Domain-specific words that sentiment lexicons do not cover
  - For e.g., *mouth watering, yummy* and *over cooked*

Features: Sentiment classification

- MPQA lexicon (*one feature*)
  - Set the following scores: 1-positive, -1-negative, 0-neutral, 2-does not appear in the list
  - Extract the words within the context of previous five and next five words
  - Sum the scores of all such words

Features: Sentiment classification

- Bing Liu lexicon (*two features*)
  - Check the word in the lexicon
  - Set 1, -1 and 2 for positive, negative and neutral words, respectively
  - First feature-sum of all the polarity scores of the words appearing in the context of previous five and next five words
  - Second feature
    - Extract the previous five and next five words
    - sum the polarity scores of only those words that have direct dependency relation with the aspect term
    - Set the following scores: 1-positive, -1-negative, 0-neutral, 2-does not appear in the list

Features: Sentiment classification

- SentiWordNet (*one feature*)
  - Extract the words within the context of previous five and next five words
  - Sum the scores of all these words as found in the lexicon
- Domain-specific words
  - Many words do not appear in the lexicon
  - For e.g.: *mouth watering, yummy, over cooked* etc.
  - List extracted from http://world-food-and-wine.com/describing-food
  - More items are added from the training data
  - Define the scores as: 1 for positive, -1 for negative and 2 for the non-listed word
  - Compute the score within the previous five and next five words