

# Labeling Emotion in Bengali Blog Corpus – A Fine Grained Tagging at Sentence Level

**Dipankar Das**

Department of Computer Science  
& Engineering,  
Jadavpur University

dipankar.dipnil2005@gmail.com

**Sivaji Bandyopadhyay**

Department of Computer Science  
& Engineering,  
Jadavpur University

sivaji\_cse\_ju@yahoo.com

## Abstract

Emotion, the private state of a human entity, is becoming an important topic in Natural Language Processing (NLP) with increasing use of search engines. The present task aims to manually annotate the sentences in a web based Bengali blog corpus with the emotional components such as emotional expression (word/phrase), intensity, associated holder and topic(s). Ekman's six emotion classes (*anger, disgust, fear, happy, sad* and *surprise*) along with three types of intensities (*high, general* and *low*) are considered for the sentence level annotation. Presence of discourse markers, punctuation marks, negations, conjuncts, reduplication, rhetoric knowledge and especially emoticons play the contributory roles in the annotation process. Different types of fixed and relaxed strategies have been employed to measure the agreement of the sentential emotions, intensities, emotional holders and topics respectively. Experimental results for each emotion class at word level on a small set of the whole corpus have been found satisfactory.

## 1 Introduction

Human emotion described in texts is an important cue for our daily communication but the identification of emotional state from texts is not an easy task as emotion is not open to any objective observation or verification (Quirk *et al.*, 1985). Emails, weblogs, chat rooms, online forums and even twitter are considered as the

affective communication substrates to analyze the reaction of emotional catalysts. Among these media, blog is one of the communicative and informative repository of text based emotional contents in the Web 2.0 (Lin *et al.*, 2007).

Rapidly growing web users from multilingual communities focus the attention to improve the multilingual search engines on the basis of sentiment or emotion. Major studies on Opinion Mining and Sentiment Analyses have been attempted with more focused perspectives rather than fine-grained emotions. The analyses of emotion or sentiment require some basic resource. An emotion-annotated corpus is one of the primary ones to start with.

The proposed annotation task has been carried out at sentence level. Three annotators have manually annotated the Bengali blog sentences retrieved from a web blog archive<sup>1</sup> with Ekman's six basic emotion tags (*anger* (A), *disgust* (D), *fear* (F), *happy* (H), *sad* (Sa) and *surprise* (Su)). The emotional sentences are tagged with three types of intensities such as *high, general* and *low*. The sentences of non-emotional (*neutral*) and multiple (*mixed*) categories are also identified. The identification of emotional words or phrases and fixing the scope of emotional expressions in the sentences are carried out in the present task. Each of the emoticons is also considered as individual emotional expressions. The emotion holder and relevant topics associated with the emotional expressions are annotated considering the punctuation marks, conjuncts, rhetorical structures and other discourse information. The knowledge of rhetorical structure helps in removing the subjective discrepancies from the

---

<sup>1</sup> [www.amarblog.com](http://www.amarblog.com)

writer's point of view. The annotation scheme is used to annotate 123 blog posts containing 4,740 emotional sentences having single emotion tag and 322 emotional sentences for mixed emotion tags along with 7087 *neutral* sentences in Bengali. Three types of standard agreement measures such as Cohen's *kappa* ( $\kappa$ ) (Cohen, 1960; Carletta, 1996), Measure of Agreement on Set-valued Items (*MASI*) (Passonneau, 2004) and *agr* (Wiebe *et al.*, 2005) metrics are employed for annotating the emotion related components. The relaxed agreement schemes like *MASI* and *agr* are specially considered for fixing the boundaries of emotional expressions and topic spans in the emotional sentences. The inter annotator agreement of some emotional components such as sentential emotions, holders, topics show satisfactory performance but the sentences of mixed emotion and intensities of *general* and *low* show the disagreement. A preliminary experiment for word level emotion classification on a small set of the whole corpus yielded satisfactory results.

The rest of the paper is organized as follows. Section 2 describes the related work. The annotation of emotional expressions, sentential emotion and intensities are described in Section 3. In Section 4, the annotation scheme for emotion holder is described. The issues of emotional topic annotation are discussed in Section 5. Section 6 describes the preliminary experiments carried out on the annotated corpus. Finally, Section 7 concludes the paper.

## 2 Related Work

One of the most well known tasks of annotating the private states in texts is carried out by (Wiebe *et al.*, 2005). They manually annotated the private states including emotions, opinions, and sentiment in a 10,000-sentence corpus (the MPQA corpus) of news articles. The opinion holder information is also annotated in the MPQA corpus but the topic annotation task has been initiated later by (Stoyanov and Cardie, 2008a). In contrast, the present annotation strategy includes the fine-grained emotion classes and specially handles the emoticons present in the blog posts.

(Alm *et al.*, 2005) have considered eight emotion categories (angry, disgusted, fearful,

happy, sad, positively surprised, negatively surprised) to accomplish the emotion annotation task at sentence level. They have manually annotated 1580 sentences extracted from 22 Grimms' tales. The present approach discusses the issues of annotating unstructured blog text considering rhetoric knowledge along with the attributes, e.g. negation, conjunct, reduplication etc.

Mishne (2005) experimented with mood classification in a blog corpus of 815,494 posts from

Livejournal (<http://www.livejournal.com>), a free weblog service with a large community. (Mihalcea and Liu, 2006) have used the same data source for classifying the blog posts into two particular emotions – *happiness* and *sadness*. The blog posts are self-annotated by the blog writers with *happy* and *sad* mood labels. In contrast, the present approach includes Ekman's six emotions, emotion holders and topics to accomplish the whole annotation task.

(Neviarouskaya *et al.*, 2007) collected 160 sentences labeled with one of the nine emotions categories (anger, disgust, fear, guilt, interest, joy, sadness, shame, and surprise) and a corresponding intensity value from a corpus of online diary-like blog posts. On the other hand, (Aman and Szpakowicz, 2007) prepare an emotion-annotated corpus with a rich set of emotion information such as category, intensity and word or phrase based expressions. The present task considers all the above emotion information during annotation. But, the present annotation task additionally includes the components like emotion holder, single or multiple topic spans.

The emotion corpora for Japanese were built for recognizing emotions (Tokuhisa *et al.*, 2008). An available emotion corpus in Chinese is Yahoo!'s Chinese news (<http://tw.news.yahoo.com>), which is used for Chinese emotion classification of news readers (Lin, *et al.*, 2007). The manual annotation of eight emotional categories (expect, joy, love, surprise, anxiety, sorrow, angry and hate) along with intensity, holder, word/phrase, degree word, negative word, conjunction, rhetoric, punctuation and other linguistic expressions are carried out at sentence, paragraph as well as document level on 1,487 Chinese blog documents (Quan and Ren, 2009). In addition

to the above emotion entities, the present approach also includes the annotation of single or multiple emotion topics in a target span.

Recent study shows that non-native English speakers support the growing use of the Internet<sup>2</sup>. This raises the demand of linguistic resources for languages other than English. Bengali is the fifth popular language in the World, second in India and the national language in Bangladesh but it is less computerized compared to English. To the best of our knowledge, at present, there is no such available corpus that is annotated with detailed linguistic expressions for emotion in Bengali or even for other Indian languages. Thus we believe that this corpus would help the development and evaluation of emotion analysis systems in Bengali.

### 3 Emotion Annotation

Random collection of 123 blog posts containing a total of 12,149 sentences are retrieved from Bengali web blog archive<sup>3</sup> (especially from comics, politics, sports and short stories) to prepare the corpus. No prior training was provided to the annotators but they were instructed to annotate each sentence of the blog corpus based on some illustrated samples of the annotated sentences. Specially for annotating the emotional expressions and topic(s) in emotional sentences, the annotators are free in selecting the texts spans. This annotation scheme is termed as relaxed scheme. For other emotional components, the annotators are given items with fixed text spans and instructed to annotation the items with definite tags.

#### 3.1 Identifying Emotional Expressions for Sentential Emotion and Intensity

The identification of emotion or affect affixed in the text segments is a puzzle. But, the puzzle can be solved partially using some lexical clues (e.g. discourse markers, punctuation marks (*sym*), negations (*NEG*), conjuncts (*CONJ*), reduplication (*Redup*)), structural clues (e.g. rhetoric and syntactic knowledge) and especially some direct affective clues (e.g.

emoticons (*emo\_ icon*)). The identification of structural clues indeed requires the identification of lexical clues.

Rhetorical Structure Theory (RST) describes the various parts of a text, how they can be arranged and connected to form a whole text (Azar, 1999). The theory maintains that consecutive discourse elements, termed *text spans*, which can be in the form of clauses, sentences, or units larger than sentences, are related by a relatively small set (20–25) of *rhetorical relations* (Mann and Thompson, 1988). RST distinguishes between the part of a text that realizes the primary goal of the writer, termed as *nucleus*, and the part that provides supplementary material, termed *satellite*. The separation of *nucleus* from *satellite* is done based on punctuation marks (, ! @?), emoticons, discourse markers (যেহেতু *jehetu* [as], যেমন *jemon* [e.g.], কারণ *karon* [because], মানে *mane* [means]), conjuncts (এবং *ebong* [and], কিন্তু *kintu* [but], অথবা *athoba* [or]), causal verbs (ঘটায় *ghotay* [caused]) if they are explicitly specified in the sentences.

Use of emotion-related words is not the sole means of expressing emotion. Often a sentence, which otherwise may not have an emotional word, may become emotion bearing depending on the context or underlying semantic meaning (Aman and Szpakowicz, 2007). An empirical analysis of the blog texts shows two types of emotional expressions. The first category contains explicitly stated emotion word (*EW*) or phrases (*EP*) mentioned in the *nucleus* or in the *satellite*. Another category contains the implicit emotional clues that are identified based on the context or from the metaphoric knowledge of the expressions. Sometimes, the emotional expressions contain direct emotion words (*EW*) (কৌতুক *koutuk* [joke], আনন্দ *ananda* [happy], আশ্চর্য *ashcharjyo* [surprise]), reduplication (*Redup*) (সন্দ সন্দ *sanda sanda* [doubt with fear], question words (*EW\_Q*) (কি *ki* [what], কবে *kobe* [when]), colloquial words (ক্ষমা *kshyama* [pardon]) and foreign words (থ্যাংকু *thanku* [thanks], গোস্যা *gossya* [anger]). On the other hand, the emotional expressions contain indirect emotion words e.g. proverbs, idioms (ভাসের ঘর *taser ghar* [weakly built], গৃহদাহ *grihadaho* [family disturbance]) and emoticons (☺, ☹).

<sup>2</sup> <http://www.internetworldstats.com/stats.htm>

<sup>3</sup> [www.amarblog.com](http://www.amarblog.com)

A large number of emoticons (*emo\_icon*) present in the Bengali blog texts vary according to their emotional categories and slant. Each of the emoticons is treated as individual emotional expression and its corresponding intensity is set based on the image denoted by the emoticon. The labeling of the emoticons with Ekman's six emotion classes is verified through the inter-annotator agreement that is considered for emotion expressions.

The intensifiers (খুব *khub* [too much/very], অনেক *anek* [huge/large], ভীষণ *bhishon* [heavy/too much]) associated with the emotional phrases are also acknowledged in annotating sentential intensities. As the intensifiers depend solely on the context, their identification along with the effects of negation and conjuncts play a role in annotating the intensity. Negations (না *na* [no], নয় *noy* [not]) and conjuncts freely occur in the sentence and change the emotion of the sentence. For that very reason, a crucial analysis of negation and conjuncts is carried out both at intra and inter phrase level to obtain the sentential emotions and intensities. An example set of the annotated blog corpus is shown in Figure 1.

<ES_Sa><holder>গেদুচাঁচা:</holder>	ভাইও
<sym>!</sym>	<EW_D>বেজার</EW_D>
হইছেন?</ES_Sa>	
<ES_A><ES_Su>তিনি	<EW_Su><EW_Q>
কি</EW_Q></EW_Su>	বুঝলেন
<EW_Su><EW_Q>	কে</EW_Q></EW_Su>
জানেন<EW_Su>!</EW_Su>	<Redup><EW_A>রেগেমেগে</EW_A></Redup>
ভিতরে	চলে গেলেন। আমাকে
<EW_F>বোধহয় </EW_F>	চিনতে
<NEG>নি</NEG>!</ES_Su></ES_A>	পারেন
<ES_H>আপনার	<topic>কবিতাটা</topic>
পড়তএ	গিয়া এই
<EW_H>কউতুকটা</EW_H>	মনে
পড়ছিলো </ES_H>	

“Figure 1. Annotated sample of the corpus”

### 3.2 Agreement of Sentential Emotion and Intensity

Three annotators identified as A1, A2 and A3 have used an open source graphical tool to carry out the annotation<sup>4</sup>. As the Ekman's emotion classes and intensity types belong to some definite categories, the annotation

agreement for emotion and intensities are measured using standard Cohen's *kappa* coefficient ( $\kappa$ ) (Cohen, 1960). The annotation agreement for emoticons is also measured using the *kappa* metric. It is a statistical measure of inter-rater agreement for qualitative (categorical) items. It measures the agreement between two raters who separately classify items into some mutually exclusive categories.

The agreement of classifying sentential intensities into three classes (*high*, *general* and *low*) is also measured using *kappa* ( $\kappa$ ). The intensities of mixed emotional sentences are also considered. Agreement results of emotional, non-emotional and mixed sentences, emoticons, along with results for each emotion class, intensity types are shown in Table 1. Sentential emotions with *happy*, *sad* or *surprise* classes produce comparatively higher *kappa* coefficient than the other emotion classes as the emotional expressions of these types were explicitly specified in the blog texts. It has been observed that the emotion pairs such as “sad-anger” and “anger-disgust” often cause the trouble in distinguishing the emotion at sentence level. Mixed emotion category, *general* and *low* intensity types give poor agreement results as expected. Instead of specifying agreement results of emoticons for each emotion class, the average results for the three annotation sets are shown in Table 1.

### 3.3 Agreement of Emotional Expressions

Emotional expressions are words or strings of words that are selected by the annotators. The agreement is carried out between the sets of text spans selected by the two annotators for each of the emotional expressions. As there is no fixed category in this case, we have employed two different strategies instead of *kappa* ( $\kappa$ ) to calculate the agreement between annotators. Firstly, we chose the measure of agreement on set-valued items (MASI) (Passonneau, 2006) that was used for measuring agreement on co reference annotation (Passonneau, 2004) and in the semantic and pragmatic annotation (Passonneau, 2006). MASI is a distance between sets whose value is 1 for identical sets, and 0 for disjoint sets. For sets A and B it is defined as:  $MASI = J * M$ , where the Jaccard metric is:

<sup>4</sup> <http://gate.ac.uk/gate/doc/releases.html>

$$J = \frac{|A \cap B|}{|A \cup B|}$$

Monotonicity ( $M$ ) is defined as,

$$1, \text{ if } A = B$$

$$2/3, \text{ if } A \subset B \text{ or } B \subset A$$

$$1/3, \text{ if } A \cap B \neq \phi, A - B \neq \phi, \text{ and } B - A \neq \phi$$

$$0, \text{ if } A \cap B = \phi$$

Secondly, the annotators will annotate different emotional expressions by identifying the responsible text anchors and the agreement is measured using  $agr$  metric (Wiebe *et al.*, 2005). If  $A$  and  $B$  are the sets of anchors annotated by annotators  $a$  and  $b$ , respectively,  $agr$  is a directional measure of agreement that measures what proportion of  $a$  was also marked by  $b$ . Specifically, we compute the agreement of  $b$  to  $a$  as:

$$agr(a || b) = \frac{|A_{\text{matching}B}|}{|A|}$$

The  $agr(a || b)$  metric corresponds to the recall if  $a$  is the gold standard and  $b$  the system, and to precision, if  $b$  is the gold standard and  $a$  is the system. The results of two agreement strategies for each emotion class are shown in Table 1. The annotation agreement of emotional expressions produces slightly less values for both  $kappa$  and  $agr$ . It leads to the fact that the relaxed annotation scheme that is provided for fixing the boundaries of the expressions causes the disagreements.

#### 4 Identifying Emotion Holder

The source or holder of an emotional expression is the speaker or writer or experiencer. The main criteria considered for annotating emotion holders are based on the nested source hypothesis as described in (Wiebe *et al.*, 2005). The structure of Bengali blog corpus (as shown in Figure 2) helps in the holder annotation process. Sometimes, the comments of one blogger are annotated by other bloggers in the blog posts. Thus the holder annotation task in user comments sections was less cumbersome than annotating the holders inscribed in the topic section.

Classes (# Sentences or Instances)	Agreement (pair of annotators)			
	A1-A2	A2-A3	A1-A3	Avg.
Emotion / Non-Emotion (5,234/7,087)	0.88	0.83	0.86	0.85
Happy (804)	0.79	0.72	0.83	0.78
Sad (826)	0.82	0.75	0.72	0.76
Anger (765)	0.75	0.71	0.69	0.71
Disgust (766)	0.76	0.69	0.77	0.74
Fear (757)	0.65	0.61	0.65	0.63
Surprise (822)	0.84	0.82	0.85	0.83
Mixed (322)	0.42	0.21	0.53	0.38
High (2,330)	0.66	0.72	0.68	0.68
General (1,765)	0.42	0.46	0.48	0.45
Low (1345)	0.21	0.34	0.26	0.27
Emoticons w.r.t six Emotion Classes (678)	0.85	0.73	0.84	0.80
Emoticons w.r.t three In- tensities	0.72	0.66	0.63	0.67
Emotional Ex- pressions (7,588) [MASI]	0.64	0.61	0.66	0.63
Emotional Ex- pressions (7,588) [agr]	0.67	0.63	0.68	0.66

Table 1: Inter-Annotator Agreements for sentence level Emotions, Intensities, Emoticons and Emotional Expressions

<pre> -&lt;DOC docid = xyz&gt;   -&lt;Topic&gt;.... &lt;/Topic&gt;   -&lt;User Comments&gt;     -&lt;U uid=1&gt;... &lt;/U&gt;     -&lt;U uid=2&gt;... &lt;/U&gt;     -&lt;U uid=3&gt;....     -&lt;U uid=1&gt;... &lt;/U&gt; ...&lt;/U&gt;...   &lt;/User Comments&gt; &lt;/DOC&gt; </pre>
---

“Figure. 2. General structure of a blog document”

Prior work in identification of opinion holders has sometimes identified only a single opinion per sentence (Bethard *et al.*, 2004),



and sometimes several (Choi *et al.*, 2005). As the blog corpus has sentence level emotion annotations, the former category is adopted. But, it is observed that the long sentences contain more than one emotional expression and hence associated with multiple emotion holders (*EH*). All probable emotion holders of a sentence are stored in an anchoring vector successively according to their order of occurrence.

The annotation of emotion holder at sentence level requires the knowledge of two basic constraints (*implicit* and *explicit*) separately. The *explicit* constraints qualify single prominent emotion holder that is directly involved with the emotional expression whereas the *implicit* constraints qualify all direct and indirect nested sources as emotion holders. For example, in the following Bengali sentences, the pattern shown in **bold** face denotes the emotion holder. In the second example, the appositive case (e.g. **রামের** সুখ (*Ram's pleasure*)) is also identified and placed in the vector by removing the inflectional suffix (-এর in this case). Example 2 and Example 3 contain the emotion holders **রাম** (*Ram*) and **নাসরিন সুলতানা** (*Nasrin Sultana*) based on *implicit* constraints.

Example 1. *EH\_Vector*: < সায়ণ >

সায়ণ ভীষণ আনন্দ অনুভব  
(**Sayan**) (bhishon) (anondo) (anubhob)  
করেছিল  
(korechilo)  
*Sayan felt very happy.*

Example 2. *EH\_Vector*: < রাশেদ, রাম >

রাশেদ অনুভব করেছিল যে রামের  
(*Rashed*) (anubhob) (korechilo) (je) (**Ramer**)  
সুখ অন্তহীন  
(sukh) (antohin)  
*Rashed felt that Ram's pleasure is endless.*

Example 3. *EH\_Vector*: < গেদু চাচা, নাসরিন সুলতানা >

গেদু চাচা বলে : আমি নাসরিন সুলতানার  
(GeduChaCha) (bole) (ami) (**Nasrin Sultanar**)  
দুঃখের কথাতে কেন্দে ফেলি।  
(dookher) (kathate) (kende) (feli)  
*Gedu Chacha says: No my sister, I fall into cry on the sad speech of Nasrin Sultana*

#### 4.1 Agreement of Emotion Holder Annotation

The emotion holders containing multi word Named Entities (*NEs*) are assumed as single emotion holders. As there is no agreement discrepancy in selecting the boundary of the single or multiple emotion holders, we have used the standard metric, Cohen's *kappa* ( $\kappa$ ) for measuring the inter-annotator agreement. Each of the elementary emotion holders in an anchoring vector is treated as a separate emotion holder and the agreement between two annotators is carried out on each separate entity. It is to be mentioned that the anchoring vectors provided by the two annotators may be disjoint.

To emphasize the fact, a different technique is employed to measure the annotation agreement. If **X** is a set of emotion holders selected by the first annotator and **Y** is a set of emotion holders selected by the second annotator for an emotional sentence containing multiple emotion holders, inter-annotator agreement **IAA** for that sentence is equal to quotient of number of emotion holders in **X** and **Y** intersection divided by number of emotion holders in **X** and **Y** union:

$$IAA = X \cap Y / X \cup Y$$

Two types of agreement results per emotion class for annotating emotion holders (*EH*) are shown in Table 2. Both types of agreements have been found satisfactory and the difference between the two agreement types is significantly less. The small difference indicates the minimal error involved in the annotation process. It is found that the agreement is highly moderate in case of single emotion holder, but is less in case of multiple holders. The disagreement occurs mostly in the case of satisfying the implicit constrains but some issues are resolved by mutual understanding.

#### 5 Topic Annotation and Agreement

Topic is the real world object, event, or abstract entity that is the primary subject of the opinion as intended by the opinion holder (Stoyanov and Cardie, 2008). They mention that the topic identification is difficult within the single target span of the opinion as there are multiple potential topics, each identified

with its own topic span and the topic of an opinion depends on the context in which its associated opinion expression occurs. Hence, the actual challenge lies on identification of the topics spans from the emotional sentences. The writer’s emotional intentions in a sentence are reflected in the target span by mentioning one or more topics that are related to the emotional expressions. Topics are generally distributed in different text spans of writer’s text and can be distinguished by capturing the rhetorical structure of the text.

Emotion Classes [# Sentences, # Holders]	Agreement between pair of annotators ( $\kappa$ ) [IAA]			
	A1-A2	A2-A3	A1-A3	Avg.
Happy [804, 918]	(0.87) [0.88]	(0.79) [0.81]	(0.76) [0.77]	(0.80) [0.82]
Sad [826, 872]	(0.82) [0.81]	(0.85) [0.83]	(0.78) [0.80]	(0.81) [0.81]
Anger [765,780]	(0.80) [0.79]	(0.75) [0.73]	(0.74) [0.71]	(0.76) [0.74]
Disgust [766, 770]	(0.70) [0.68]	(0.72) [0.69]	(0.83) [0.84]	(0.75) [0.73]
Fear [757, 764]	(0.85) [0.82]	(0.78) [0.77]	(0.79) [0.81]	(0.80) [0.80]
Surprise [822, 851]	(0.78) [0.80]	(0.81) [0.79]	(0.85) [0.83]	(0.81) [0.80]

Table 2: Inter-Annotator Agreement for Emotion Holder Annotation

In blog texts, it is observed that an emotion topic can occur in *nucleus* as well as in *satellite*. Thus, the whole sentence is assumed as the scope for the potential emotion topics. The text spans containing emotional expression and emotion holder can also be responsible for being the candidate seeds of target span. In Example 3 of Section 4, the target span (নাসরিন সুলতানার দুঃখের কথাতে ‘sad speech of Nasrin Sultana’) contains emotion holder (নাসরিন সুলতানার ‘Nasrin Sultana’) as well as the emotional expression (দুঃখের কথাতে ‘sad speech’) For that reason, the annotators are instructed to consider the whole sentence as their target span and to identify one or more topics related to the emotional expressions in that sentence.

As the topics are multi word components or string of words, the scope of the individual topics inside a target span is hard to identify. To accomplish the goal, we have not used the

standard metrics Cohen’s *kappa* ( $\kappa$ ). We employed MASI and *agr* metric (as mentioned in Section 3) for measuring the agreement of topic spans annotation. The emotional sentences containing single emotion topic has shown less disagreement than the sentences that contain multiple topics. It is observed that the agreement for annotating target span is ( $\approx 0.9$ ). It means that the annotation is almost satisfactory. But, the disagreement occurs in annotating the boundaries of topic spans. The inter-annotator agreement for each emotion class is shown in Table 3. The selection of emotion topic from other relevant topics causes the disagreement.

Emotion Classes [# Sentences, # topics]	Agreement between Pair of annotators (MASI) [ <i>agr</i> ]			
	A1-A2	A2-A3	A1-A3	Avg
Happy [804, 848]	(0.83) [0.85]	(0.81) [0.83]	(0.79) [0.82]	(0.81) [0.83]
Sad [826, 862]	(0.84) [0.86]	(0.77) [0.79]	(0.81) [0.83]	(0.80) [0.82]
Anger [765,723]	(0.80) [0.78]	(0.81) [0.78]	(0.86) [0.84]	(0.82) [0.80]
Disgust [766, 750]	(0.77) [0.76]	(0.78) [0.74]	(0.72) [0.70]	(0.75) [0.73]
Fear [757, 784]	(0.78) [0.79]	(0.77) [0.80]	(0.79) [0.81]	(0.78) [0.80]
Surprise [822, 810]	(0.90) [0.86]	(0.85) [0.82]	(0.82) [0.80]	(0.85) [0.82]

Table 3: Inter-Annotator Agreement for Topic Annotation

## 6 Experiments on Emotion Classification

A preliminary experiment (Das and Bandyopadhyay, 2009b) was carried out on a small set of 1200 sentences of the annotated blog corpus using Conditional Random Field (CRF) (McCallum *et al.*, 2001). We have employed the same corpus and similar features (e.g. POS, punctuation symbols, sentiment words etc.) for classifying the emotion words using Support Vector Machine (SVM) (Joachims, 1999). The results on 200 test sentences are shown in Table 4. The results of the automatic emotion classification at word level show that SVM outperforms CRF significantly. It is observed

that both classifiers fail to identify the emotion words that are enriched by morphological inflections. Although SVM outperforms CRF but both CRF and SVM suffer from sequence labeling and label bias problem with other non-emotional words of a sentence. (For error analysis and detail experiments, see Das and Bandyopadhyay, 2009b).

Emotion Classes (# Words)	Test Set	
	CRF	SVM
Happy (106)	67.67	80.55
Sad (143)	63.12	78.34
Anger (70)	51.00	66.15
Disgust (65)	49.75	53.35
Fear (37)	52.46	64.78
Surprise (204)	68.23	79.37

Table 4: Word level Emotion tagging Accuracies (in %) using CRF and SVM

Another experiment (Das and Bandyopadhyay, 2009a) was carried out on a small set of 1300 sentences of the annotated blog corpus. They assign any of the Ekman’s (1993) six basic emotion tags to the Bengali blog sentences. Conditional Random Field (CRF) based word level emotion classifier classifies the emotion words not only in emotion or non-emotion classes but also the emotion words into Ekman’s six emotion classes. Corpus based and sense based *tag weights* that are calculated for each of the six emotion tags are used to identify sentence level emotion tag. Sentence level accuracies for each emotion class were also satisfactory.

Knowledge resources can be leveraged in identifying emotion-related words in text and the lexical coverage of these resources may be limited, given the informal nature of online discourse (Aman and Szpakowicz, 2007). The identification of direct emotion words incorporates the lexicon lookup approach. A recently developed Bengali *WordNet Affect* lists (Das and Bandyopadhyay, 2010) have been used in determining the directly stated emotion words. But, the affect lists covers only 52.79% of the directly stated emotion words.

The fact leads not only to the problem of morphological enrichment but also to refer the problem of identifying emoticons, proverbs, idioms and colloquial or foreign words. But, in

our experiments, the case of typographical errors and orthographic features (for e.g. হেসসস ‘*disgusting*’, বাব্বা ‘*surprising*’) that express or emphasize emotion in text are not considered.

## 7 Conclusion

The present task addresses the issues of identifying emotional expressions in Bengali blog texts along with the annotation of sentences with emotional components such as intensity, holders and topics. Nested sources are considered for annotating the emotion holder information. The major contribution in the task is the identification and fixing the text spans denoted for emotional expressions and multiple topics in a sentence. Although the preliminary experiments carried out on the small sets of the corpus show satisfactory performance, but the future task is to adopt a corpus-driven approach for building a lexicon of emotion words and phrases and extend the emotion analysis tasks in Bengali.

## References

- Aman Saima and Stan Szpakowicz. 2007. Identifying Expressions of Emotion in Text. *V. Matoušek and P. Mautner (Eds.): TSD 2007, LNAI*, vol. 4629, pp.196-205.
- Azar, M. 1999. Argumentative Text as Rhetorical Structure: An Application of Rhetorical Structure Theory. *Argumentation*, vol 13, pp. 97–114.
- Bethard Steven, Yu H., Thornton A., Hatzivassiloglou V., and Jurafsky, D. 2004. Automatic Extraction of Opinion Propositions and their Holders. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Carletta Jean. 1996. Assessing Agreement on Classification Tasks: The Kappa Statistic. *Computational Linguistics*, vol. 22(2), pp.249-254.
- Choi, Y., Cardie, C., Riloff, E., and Patwardhan, S. 2005. Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns. *Human Language Technology / Empirical Method in Natural Language Processing*.
- Alm, Cecilia Ovesdotter, Dan Roth, and Richard Sproat. 2005. Emotions from text: Machine learning for text-based emotion prediction. *Human Language Technology - Empirical Method in Natural Language Processing*, pp. 579-586.



- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, vol. 20, pp. 37–46.
- Das Dipankar and Sivaji Bandyopadhyay. 2009a. Word to Sentence Level Emotion Tagging for Bengali Blogs. *Association for Computational Linguistics –International Joint Conference of Natural Language Processing-2009*, pp. 149-152. Suntec, Singapore.
- Das Dipankar and Sivaji Bandyopadhyay. 2009b. Emotion Tagging – A Comparative Study on Bengali and English Blogs. *7<sup>th</sup> International Conference On Natural Language Processing-09*, pp.177-184, India.
- Das Dipankar and Sivaji Bandyopadhyay. 2010. Developing Bengali WordNet Affect for Analyzing Emotion. *International Conference on the Computer Processing of Oriental Languages-International Conference on Software Engineering and Knowledge Engineering-2010*, USA.
- Ekman, P. 1992. An Argument for Basic Emotions. *Cognition and Emotion*. vol. 6, pp.169–200.
- Joachims, Thorsten. 1998. Text Categorization with Support Machines: Learning with Many Relevant Features. In *European Conference on Machine Learning (ECML)*,137-142
- Lin K. H.-Y., C. Yang and H.-H. Chen. 2007. What Emotions News Articles Trigger in Their Readers?. *Proceedings of SIGIR*, pp. 733-734.
- Mann, W. C. and S. A. Thompson. 1988. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization, *TEXT* 8, pp. 243–281.
- McCallum Andrew, Fernando Pereira and John Lafferty. 2001. Conditional Random Fields: Probabilistic Models for Segmenting and labeling Sequence Data. *ISBN*, 282 – 289.
- Mihalcea Rada and Hugo Liu. 2006. A corpus-based approach to finding happiness. *Association for the Advancement of Artificial Intelligence*, pp. 139-144.
- Mishne Gilad. 2005. Emotions from text: Machine learning for text-based emotion prediction. *SIGIR '05*, pp. 15-19.
- Neviarouskaya Alena, Helmut Prendinger, and Mitsuru Ishizuka. 2007. Textual Affect Sensing for Social and Expressive Online Communication. *2nd international conference on Affective Computing and Intelligent Interaction*, pp. 218-229.
- Passonneau, R. 2004. Computing reliability for coreference annotation. *Language Resources and Evaluation*, Lisbon.
- Passonneau, R.J. 2006. Measuring agreement on set-valued items (MASI) for semantic and pragmatic annotation. *Language Resources and Evaluation*.
- Quan Changqin and Fuji Ren. 2009. Construction of a Blog Emotion Corpus for Chinese Emotional Expression Analysis. *Empirical Method in Natural Language Processing- Association for Computational Linguistics*, pp. 1446-1454, Singapore
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J. 1985. A Comprehensive Grammar of the English Language. Longman, New York.
- Stoyanov, V., and C. Cardie. 2008. Annotating topics of opinions. *Language Resources and Evaluation*.
- Tokuhsa Ryoko, Kentaro Inui, and Yuji. Matsu-moto. 2008. Emotion Classification Using Massive Examples Extracted from the Web. *COLING 2008*, pp. 881-888.
- Wiebe Janyce, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, vol. 39, pp.164–210.
- Yang C., K. H.-Y. Lin, and H.-H. Chen. 2007. Building Emotion Lexicon from Weblog Corpora. *Association for Computational Linguistics*, pp. 133-136.