

Shahjalal University of Science and Technology
Department of Computer Science and Engineering

CSE 400



**A Crowdsource Based Approach for Bangla Parts Of
Speech Tagging**

Students

Shamim Ehsan

Reg. No.: 2012331035

4th year, 1st Semester

Dept. of Computer Science & Engineering

Sadia Tasnim Swarna

Reg. No.: 2012331072

4th year, 1st Semester

Dept. of Computer Science & Engineering

Supervisor

Sabir Ismail

Assistant Professor

Dept. of Computer Science & Engineering

April 12, 2017

A Crowdsource Based Approach for Bangla Parts Of Speech Tagging



A Thesis submitted to the Department of Computer Science and Engineering,
Shahjalal University of Science and Technology, in partial fulfillment of the requirements
for the degree of B.Sc.(Engg.) in Computer Science and Engineering.

Students

Shamim Ehsan

Reg. No.: 2012331035

4th year, 1st Semester

Dept. of Computer Science & Engineering

Sadia Tasnim Swarna

Reg. No.: 2012331072

4th year, 1st Semester

Dept. of Computer Science & Engineering

Supervisor

Sabir Ismail

Assistant Professor

Dept. of Computer Science & Engineering

April 12, 2017

Recommendation of the Thesis Supervisor

The thesis entitled CrowdSourcing Data Collection for Bengali POS Tagging undertaken by the students Shamim Ehsan , Sadia Tasnim Swarna is under my supervision. I, hereby, agree that the thesis can be submitted for examination.

Signature of the Supervisor:

Sabir Ismail

Assistant Professor

Dept. of Computer Science & Engineering

Date: April 12, 2017

Students Name: Shamim Ehsan, Sadia Tasnim Swarna

Thesis Title: CrowdSourcing Data Collection for Bengali POS Tagging

This is to certify that the above mentioned thesis, submitted by the student named *above* in March, 2017 as part of the requirements of the course CSE 400, is being approved by the Department of Computer Science and Engineering as a partial fulfillment of the B.Sc.(Engg.) degrees of the above students.

Head of the Dept.

Dr Mohammad Reza Selim
Professor & Head

Chairman, Exam. Committee

M. Jahirul Islam, PhD, PEng
Professor

Supervisor

Sabir Ismail
Assistant Professor

Abstract

We focused on crowd-sourcing and use the crowd-sourced data for Bengali natural language processing purpose. In our thesis work, we build an interface for crowd-sourcing and take data from different users. We selected 337 documents from recent newspapers. We take data from PIPILIKA web crawler and select the documents according to our need. Then we collect data from the current student of CSE, SUST, and create profile matrix, consensus sequence and score from the individual sentences and tag each of the words according to them.

Keywords: Crowd Sourcing, Natural Language Processing, POS Tagging, Profile Matrix, Consensus Score

Acknowledgements

We would like to thank almighty Allah(SWT) for all His blessings and mercy. We would also like to thank our supervisor Mr. Sabir Ismail, Assistant Professor, Department of Computer Science and Engineering, Shahjalal University of Science and Technology for his continuous guiding and advices. We are also grateful to Pipilika team, for sharing their corpus with us.

Contents

Abstract	I
Acknowledgement	II
Table of Contents	III
List of Tables	V
List of Figures	VI
1 Introduction	1
1.1 Crowd-Sourcing	1
1.2 Parts of Speech	1
1.3 Parts of Speech Tagging	1
1.4 Crowdsourcing and Parts of Speech Tagging	2
1.5 Motivation	3
2 Previous Works	4
2.1 Application	4
2.1.1 Voting System	5
2.1.2 Information Sharing System	6
2.1.3 Creative System	7
2.2 Performance	7
2.2.1 User participation	7
2.2.2 Quality Management	8
2.2.3 Cheating Detection	9

3	Methodology	10
3.1	Crowd Sourcing System	10
3.2	Consensus String and Score	10
3.3	Profile Matrix	12
4	Implementation	14
4.1	Choose Dataset to Tag	15
4.2	Building the Interface	15
4.2.1	Framework Choosing	15
4.3	Functionality	17
4.4	Adding Leaderboard	22
4.5	Admin Panel	25
5	Results and Analysis	29
6	Future Work	35

List of Tables

3.1	Consensus String and Score	11
3.2	Profile Matrix	12
4.1	Comparison between different PHP frameworks	16
5.1	Summery of the dataset	29
5.2	Most frequent words	31
5.3	Effects of not stemming	31
5.4	Profile Matrix of অস্ট্রেলিয়ার উদ্বোধনী জুটি ভাঙেন অশ্বিন	33
5.5	Profile Matrix of গত অর্থবছর থেকে আমদানিতে স্থবিরতা দেখা দেয়	33
5.6	Profile Matrix of বৈদ্যুতিক শর্টসার্কিট থেকে আগুনের সূত্রপাত হয়েছে বলে ফায়ার সার্ভিসের সদস্যরা জানিয়েছে	34

List of Figures

4.1	Work flow diagram of Crowd source based POS Tagging	14
4.2	Home Page of the interface	18
4.3	Average time to tag a document in HTML select box and radio button	19
4.4	Accuracy of tagging a document in HTML select box and radio button	20
4.5	Confirmation window	21
4.6	Interface of the leaderboard	22
4.7	search box in the leaderboard	23
4.8	Average time to tag a document before adding leader-board.	24
4.9	Average time to tag a document before and after adding leader-board.	25
4.10	Admin Panel views	26
4.11	Admin panel 2.	27
4.12	Reference document	28
5.1	Total tag count of all the documents	30
5.2	Correctness of the parts of speeches	32
5.3	Profile matrix and consensus score of dataset	34

Chapter 1

Introduction

1.1 Crowd-Sourcing

Crowd-Sourcing is the act of taking a job traditionally performed by a designated employee and outsourcing it to an undefined, generally large group of people in the form of an open call. So its like “I have a task to be done, but I don’t want to hire a specific person to do it, rather I put my task on the world wide web with a reward, let random people solve it and finally getting my task done.”

1.2 Parts of Speech

All words in sentences are classified into some groups which have some common grammatical properties, they are called parts of speech. Generally the words having same parts of speech that are assigned to the same word part of speech play similar roles inside the linguistic structure of sentences.

1.3 Parts of Speech Tagging

Parts of Speech tagging means a system tags the parts of speech of a sentence automatically. One may assume that it can be easy for a computer by just looking into the dictionary for the appropriate part of speech. But a word can fall into different part of speech in the context of different sentence. Consider the sentences:

- She feels *well*.
- She went to a *well* woman doctor.
- *Well!* What on earth is going on?
- She drew some water from the *well*.

it is clearly seen that the word "well" has four different meaning as well as different parts of speeches in the four sentences (verb, adjective, interjection and noun). Unfortunately, there is no fixed rules to say "well" belongs to which parts of speech in the sentences. What we can tell with just our common sense can be really really tough for a machine. Machine can learn from experience, but data is another factor here. Parts of Tagging system needs a dataset with all the sentences tagged with their corresponding parts of speech. But these kind of dataset if a bit hard to find.

1.4 Crowdsourcing and Parts of Speech Tagging

As it is mentioned earlier, tag a word in a sentence is pretty simple for a human, but very tough for a machine. Machine can solve this problem provided with enough data set. Here comes crowdsourcing. Everyone with a minimum grammatical knowledge can tag sentence, and it dont take much skills. So tagging sentences using the general crowd can be a pretty good option to create a huge dataset of tagged sentences. And it is pretty effective also in the fields of natural language processing like speech recognition, traslation etc. Translation feature of tech giant Google Inc is heavily depended on crowd sourcing from users all around the world, and they have better accuracy than typical statistical machine translation. So we can say that crowdsource can solve the problem of automated parts of speech tagging

1.5 Motivation

Automated Parts of Speech Tagging plays a vital role in the natural language processing research. Most of the NLP research needs a good POS tagger. For example, if anyone wants to build an automated question answering system, he needs a POS tagger to understand the question. And for summarizing text, document classification all of them need a good POS tagger. Although there have been some good researches in Bangla POS tagging, but they are all discrete and no online POS tagger or application is yet to be found for Bangla language. So, we think this field has a very good scope to contribute for Bangla Natural language processing research.

Bangladesh is the most populous country in the world, so, finding the manpower will be the least concern when the job takes absolutely no additional skill rather than having the simplest knowledge of Bangla grammar. As we have fought for the right of our mother language, we always have a soft corner for Bangla language, and when we are asked to help enrich our language, we'll definitely try if it's feasible. For example, a Facebook event was arranged to remove offensive suggestions from Google search in Bangla fonts. Nobody got any benefit from it, but that was an issue of prestige and people participated spontaneously. Google realized our feeling for Bangla language, and launched a campaign to collect data for their Bangla-English translation. And we gave them thousands and thousands of data only because the campaign motto was like that "ভাষার মাসে বাংলাভাষা কে সমৃদ্ধ করুন." So, they got a huge dataset using our emotion. POS tagging is like translation, difficult for a machine, but very easy for a human. There are some papers related to Bangla POS tagging, but a system is yet to be built for Bangla language. So, if we can motivate people about the importance of automated Bangla POS tagger, we hope to get enough data for crowd-sourced data for tagging Bangla parts of speech.

Chapter 2

Previous Works

Crowdsourcing systems enlist a large number of humans to solve a wide variety of problems. crowdsourcing was proposed to make efficient use of manpower and resources,[1]. Crowdsourcing is a distributed model which is used for problem-solving and business production. The Internet is growing very fast and the global accessibility of the internet has forwarded the research activity in crowdsourcing. In recent decade, numerous work has concentrated on various aspects of crowdsourcing, say for different performance analysis and computational techniques and. Fig. 1 shows a taxonomy of crowdsourcing.

2.1 Application

As web 2.0 technology is getting popular, websites that are using crowdsourcing are gaining much attentions right now [2][3].Users of a crowdsourcing system is of two groups:workers and requesters.A list of currently available tasks are published in the crowdsourcing site by the requesters. They are associated with a time period and a reward. During the time slot, workers give their best for the reward and compete to the site. But it is not the case every time; in those cases, workers are not trying to get money from crowd sourcing, rather they work only for fun or do it as a duty towards something[4] . Mainly Crowdsourcing applications can be classsified into three categories. first is voting system. Information sharing system is another subsection of crowdsourcing application and creative system.

2.1.1 Voting System

Among many crowdsourcing systems, Amazon Mechanical Turk (or MTurk) is one of them[5]. It has ability to support voting tasks in a large scale. In these voting system, a crowdsourcing worker has some options, and he has to select from the options. Most voted answer is said to be correct. Authorities can use voting system as a gadget to discover the justification from the general crowd.

Natural language processing

Natural language processing is a work that may be difficult for automated process but comparatively easy for humans. In recent times, as a quick alternative of expert annotations, researchers found Mechanical Turk of Amazon [4][6] [7][8][9][10]. Akkaya et al. [11] demonstrated that crowdsourcing would be good for subjectivity word sense comment. Callison-Burch and Dredze [12] demonstrated that they make information for discourse and dialect applications with a minimal effort. Gao and Vogel [7] showed that crowdsourcing laborers perform well on word arrangement assignments as far as arrangement blunder rate. Jha et al. [8] demonstrated that an exact prepositional expression connection corpus can be developed by crowdsourcing specialists. Parent and Eskenazi [9] posposed a system to decompose a task for the meaning of dictionary words in MTurk. Skory and Eskenazi [10] discussed ways to calculate the quality of the results of MTurk worker's tasks.

Named entity recognition

NER is utilized to recognize and arrange printed references to objects on the planet, for instance, individuals and associations. MTurk is an amazingly encouraging pack for commenting on massive corpuses, for example, Twitter information, nicknames, email datasets and named substances.[13][14][15][16].

Opinions

Sentiments of the general individuals can be effectively assembled utilizing crowdsourcing framework. Mellebeek et al. [17] utilized the crowdsourcing idea to order Spanish shopper remarks.

Commonsense

Many research exercises are done concentrating on gathering rational information in MTurk [6][18].

Relevance evaluation

To decide the purpose of an arrangement of test questions, people need to examine each report in a corpus. Alonso et al. [19] suggested that each crowdsourcing work carry out a little assessment undertaking.

Spam recognition

Spam email detection is impossible without the undertaking of understanding substance By people. Some hostile to spam frameworks, for example, Vipul'sRazor1 utilize human votes to characterize spam messages.

2.1.2 Information Sharing System

Some crowd sourcing systems mainly focuses on just sharing different types of informations throuout the internet. Choffnes et al.[20] proved that crowdsourced contributions can be used to check network events in the service-level. Beyond those systems, different information sharing systems which are popular among the crowd were launched on the Internet as shown in the following:

- Yahoo! Answers
- Wikipedia
- Yahoo! Suggestion Board
- The website43Things
- Yahoo's flickr
- del.icio.us

2.1.3 Creative System

No other advanced technologies can replace the role of human in creativity. To reduce the cost of production, researchers proposed to crowdsource from workers for doing some creative tasks in order to reduce the total cost of production. A web-based artworks example is the Sheep Market. It uses hundreds and thousands of online workers to create huge database of hand drawn images or drawings. Another example is Threadless9. Threadless9 is a platform where community contribute to the collection of graphic t-shirt designs. Different people may make distinctive thoughts, for example, designing a T-shirt [21]. Leimeister et al. [22] discussed a theory that to motivate more users to support and participate in crowdsource software development tasks as ideas competitions. Some theory and works were proposed to coordinate among many people for more complex human computation tasks. [23] [24].

2.2 Performance

Recently some studies have investigated the performance of crowdsourcing system. They can be grouped into different sections.

2.2.1 User participation

Every task is distributed among random population of Internet for completion in all crowd sourcing system. Ross et al. [25] discussed that the worker population is always changing after some time. he show that it is shifting from United states centered workforce and moves to an increasingly international group of well educated people, among them most of the worker are Indian. Downs et al. [26] screened all the workers of MTurk by using two screening questions which have been tested previously. The results of the experiment showed that financial workers, hourly workers are less likely to take the task seriously than those that are students currently, professionals, and non-workers. Moreover, women of all age and adult men with age over thirty were much more likely to qualify.

Economic motivation

The impact of the economic motivation were studied in [27][28][29][30]. If any task is planned properly, economic motivation can enhance the standard of the tasks [27]. When workers get low payment, they started unusable labeling [28].

Intrinsic motivation

Although financial crowdsourcing motivation is dominant among other motivations, some of the systems that are depended on crowd sourcing do not offer handsome money or rewards to the labours, for example, YouTube or quora. Huberman et al. [31] demonstrated that giving handsome rewards is not the main motivation to contribute in youtube, rather gaining more views or upvotes is enough to motivate the users to contribute. Wu et al. [32] discussed that YouTube's user who contributes in video uploading, stop their contribution when they've stopped receiving attentions.

Worker Behavior

Crowdsourcing workers can be influenced by user interfaces. It verily changes their behaviour. Ipeirotis [33] told that workers can be limited for not user friendly interfaces. To design effective tasks for document evaluation on MTurk, human factor is one of the main factors [34]. Yang et al. [35] analyzes the crowdsourcing workers activity in china and found that most people become inactive after giving few submissions.

2.2.2 Quality Management

When it comes to partition the task into several tasks, a requester have to decide how many division has to be made in a crowd sourcing system. Main challenge of a requester in a crowd-sourcing systems is how to design the task to get good feedback from the users. Watts and Mason [36] stated that increased reward value will increase the quantity, but it does not guarantee the improvement of quality. It is important to find some principles related to design for the tasks on to ensure the output quality of workers.

2.2.3 Cheating Detection

Crowdsourcing workers are anonymous and random, so mischievous labors sometimes try to maximize their profit by producing random, solution rather than trying to solve the task actually. Currently, techniques of cheating detection are sometimes based on some control questions which are being evaluated automatically and sometimes, the requester have to manually check each documents to check random solutions. Eickhoff and de Vries [37] investigated the common behavior of malicious crowdsourcing labors such as task-dependent evaluation, interface-dependent evaluation

Chapter 3

Methodology

3.1 Crowd Sourcing System

A system is called a crowd sourcing system if it depends on the opinions and answers of the general crowd. People who are participating in the crowd- sourcing are called crowd users. Some crowd sourcing system sets some questions and takes opinion from the users, and most voted opinion is considered to be correct answer. These systems are called Voting System. In some cases , there is no fixed answer to the question, users have to use their creativity to crowd source. FOr example, draw.withgoogle.com ask users to draw an object within 20 seconds. Suppose a user is asked to draw a car, he has to draw it with his own creativity, these systems are called creative systems. Our proposed crowd sourcing system is basically a voting system. There will be sentences without any proper tags, users tagged the sentences, and with the response of the users, we will tag the words.

3.2 Consensus String and Score

In the field of bioinformatics, calculated order of the most frequent nucleotides is called the consensus sequence. Consensus String is that string in which each position is occupied by the most frequent neuclutides in that position from all the motifs. So basically each position of consensus string is the most popular symbol in each columns, if we arrange each gene sequences in a row. Table 3.1 shows the consensus sequence and score of different motifs.

Motifs	G	C	T	G	G	G	g	T	T	T	T	t
	c	C	T	G	t	G	C	c	T	T	a	C
	G	t	T	G	G	G	C	c	T	T	C	C
	G	C	T	G	G	G	C	T	T	c	a	t
	G	C	T	G	G	G	C	T	T	c	C	t
	G	a	T	G	G	G	C	a	c	T	a	C
	G	C	T	G	G	t	C	T	a	a	C	C
	G	C	T	G	G	G	C	T	T	T	C	C
Consensus String	G	C	T	G	G	G	C	T	T	T	C	C
Consensus Score	1 +	2 +	0 +	0 +	1 +	1 +	1 +	3 +	2 +	3 +	3 +	3 = 20

Table 3.1: Consensus String and Score

Profile(Motifs)	A	0	1	0	0	0	0	7	1	1	1	3	0
	C	1	6	0	0	0	0	0	2	1	2	4	5
	G	7	0	0	8	7	7	1	0	0	0	0	0
	T	0	1	8	0	1	1	0	5	6	5	1	3

Table 3.2: Profile Matrix

3.3 Profile Matrix

A profile matrix, also known as position weight matrix (PWM) is another frequently used representation of sequence patterns in biological sequences. From the example motifs shown in table 3.1, we can calculate the profile matrix of it which is shown in table 3.2. There are four rows in the profile matrix, each cell contains frequency of nucleotide i in column j . profile matrix is shown here.

We can use the concept of consensus score and profile matrix to tag the parts of speech. instead of using A T G and C, we use twelve parts of speech tags in each position, and our profile matrix shall have 12 rows. After crowdsourcing from the users, we'll have a sentence tagged differently by the users, like different motifs.

Chapter 4

Implementation

Objective of our thesis is to build a crowd-sourcing system, collect data from users and finally build a system which tags Bengali parts of speeches automatically. Parts of speech tagging can be easy for a general crowd than to teach a machine, so parts of speech tagging can be a good example to test a crowd sourcing system. Flow chart of our work is shown in figure 4.1

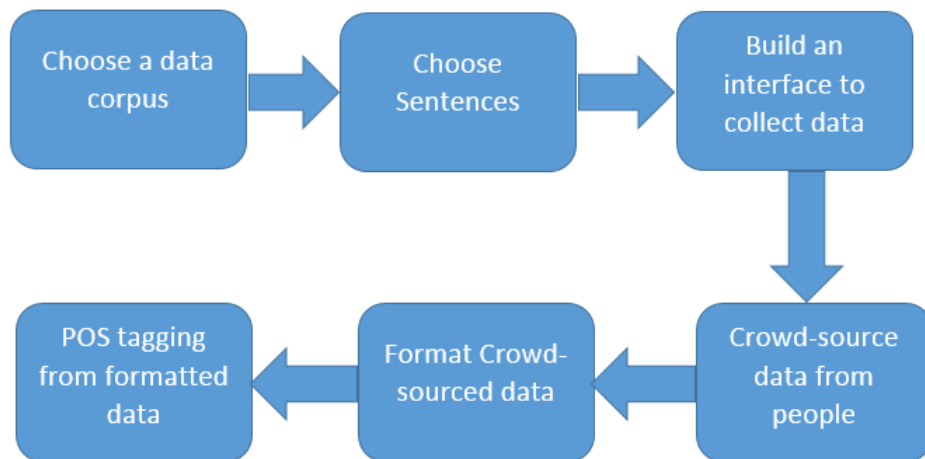


Figure 4.1: Work flow diagram of Crowd source based POS Tagging

4.1 Choose Dataset to Tag

First step is to select a group of sentence as our primary dataset. We cannot just choose some random sentence and ask users to tag them, and obviously we dont need to write all the sentences all by ourselves, rather we choose a Bengali data corpus and select sentences from it by hand or by another code. Team Pipilikas research team has shared their corpus with us, Pipilika corpus has 10000 documents in their corpus. Then we select the documents from the corpus according to our need. Pipilika crawler does not guarantee 100 percent error free document, so We have to manually check every sentence because there are some irrelevant or some merged sentences in a single document like this

আর দুদক পদ্মা সেতুর দুর্নীতির যে মামলা করেছে, তাতে আবুল হোসেনকে বাদ দেয়ার কারণে আরেফিন রুমি বলেছেন দ্বিতীয় ভলোবাসা

There is also some unfinished sentence due to crawling error, we've also exclude them from our dataset. And finally we run a code to remove very large sentences from our initial dataset. User satisfaction should be the main priority in a crowd-sourcing system, so if the user gets bored with very large sentences, that will prevent them from tagging more documents. The sentences with word length more than 30 are considered long sentences in our case. After formatting, we have 337 documents with 3 sentences in each document, with average length of 22 words per sentence in our dataset.

4.2 Building the Interface

If we want to crowd-source from the people, we have to build a platform in where people visit and give us data. It could be a web application or a smart phone app/game. We choose to build a web application.

4.2.1 Framework Choosing

As we are planning for a large scale crowd sourcing interface, we have to use a framework, there are plenty of web frameworks available, now-a-days it is not a good idea to raw code them again when a framework can save our thousand lines of codes. We choose PHP laravel framework and we have used laravel version 5.3 and for storing database we have used MySQL database.

Framework	Req/sec
CodeIgniter	187.78
Silex	120.01
Laravel	135.9
YII	123.5
Fuel PHP	116.34
Hazaar MVC	103.53
Zend 1	103.02
Cake PHP	54.97
Nette	53.48
Symphony2	39.22
Zend2	36.1

Table 4.1: Comparison between different PHP frameworks

Why laravel?

Laravel is a PHP framework designed by following the same MVC model that aids developers to develop website effortlessly. A Laravel developers get plenty of features and its latest update 5.4 is actually magnificent and flawless in concerns of advancement in web creation technologies. In short, it is a powerful PHP framework to guide developers in the rapid website development coupled with excellent applications. The popularity of Laravel can be examined based on many aspects, and among them, one is the celebrated set of Laravel community support itself. Any clarifications on documentation, coding, or even database manipulations can be smoothly managed through the extended community support. Laravel can handle concurrent requests smartly, Table 4.1 1 shows the comparison between different laravel frameworks, and we can see that laravel can handle concurrent request more than most of the PHP frameworks.

¹Table collected from www.systemsarchitect.net

we didnt choose any templates, we tried to keep the site as simple as possible. Most of the templates have to load lots of javascript and css files and loading all of the files in every page makes the site a bit slower.

4.3 Functionality

Our user interface is pretty simple, register page , login page and home page. Some people don't get motivated to enter when it requires too much information to log into a site, so we've made a guest login feature, where people gets logged in with just user names. in the home page , each user is assigned with a document and asked to tag the words of the documents. There are three sentences in a document. All the documents are coming from database with a queue fashioned manner with the least tagged document first. 12 universal tags are used. These tag sets are globally used in the parts of speech tagging research. They are:

- বিশেষ্য(NN)
- সর্বনাম(PR)
- নির্দেশক(DM)
- ক্রিয়া(V)
- বিশেষণ(JJ)
- ক্রিয়া-বিশেষণ(RB)
- অনুসর্গ(PSP)
- সংযোজক(CC)
- অব্যয়(RP)
- পরিমাণবাচক(QT)
- যতি-চিহ্ন(PNC)
- অন্যান্য(RD)

Snapshot of our homepage is given in fig 4.2:

Bengali POS Tagging [View leaderboard](#) Welcome, ehsan ▾

You have Tagged 5 documents yet

Tag the sentences:

" গত ডিসেম্বরে আন্তর্জাতিক ক্রিকেট থেকে অবসর নেন পন্টিং । সিপিএল ছাড়াও ইন্ডিয়ান প্রিমিয়ার লিগে মাতাবেন তিনি । প্রতিযোগিতার ষষ্ঠ আসরে আইপিএলের দল মুম্বাই ইন্ডিয়ান্সের অধিনায়ক হিসেবে দেখা যাবে তাকে । "

Legend:
NN বিশেষ্য PR সর্বনাম
DM নির্দেশক V ক্রিয়া
JJ বিশেষণ RB ক্রিয়া বিশেষণ
PSP অনুসর্গ CC সংযোজক
RP অব্যয় QT পরিমানবাচক
PNC যতিচিহ্ন RD অন্যান্য

গত	ডিসেম্বরে	আন্তর্জাতিক	ক্রিকেট	থেকে	অবসর	নেন	পন্টিং
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

সিপিএল	ছাড়াও	ইন্ডিয়ান	প্রিমিয়ার	লিগ	মাতাবেন	তিনি
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

প্রতিযোগিতার	ষষ্ঠ	আসরে	আইপিএলের	দল	মুম্বাই	ইন্ডিয়ান্সের	অধিনায়ক	হিসেবে
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Please select an item in the list.

Figure 4.2: Home Page of the interface

We've planned to make use three different types of input fields of html (select box, radio button and drag and drop) , we successfully implemented first two and tested it to know which option will be suitable for the users. We select fifteen volunteers and assigned each of them to tag 50 sentences with solutions as fast as they can to check which method is faster and has less chance of selecting a wrong tag while in a hurry. Results are shown in figure 4.3 and 4.4.

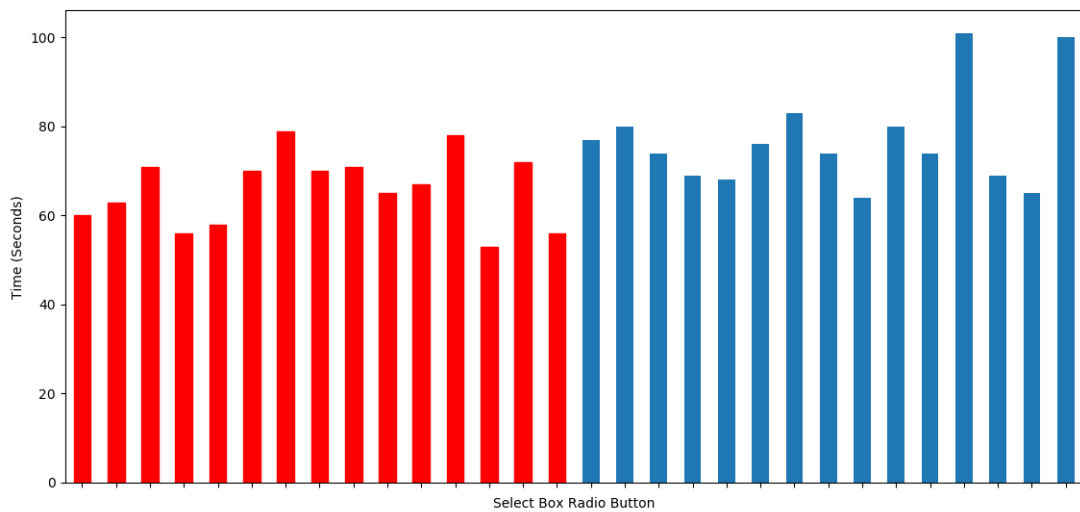


Figure 4.3: Average time to tag a document in HTML select box and radio button

It is seen that using select box takes average of 65 seconds to tag a document, and radio button takes 76 seconds and while using select box, people tags 98.45 percent words correctly while radio button has accuracy of 91.45 percent. So we can have a conclusion that both chance of making silly mistakes and typing speed in radio button is higher than select box. So we've excluded radio button from our site and takes input from user with only HTML Select Box.

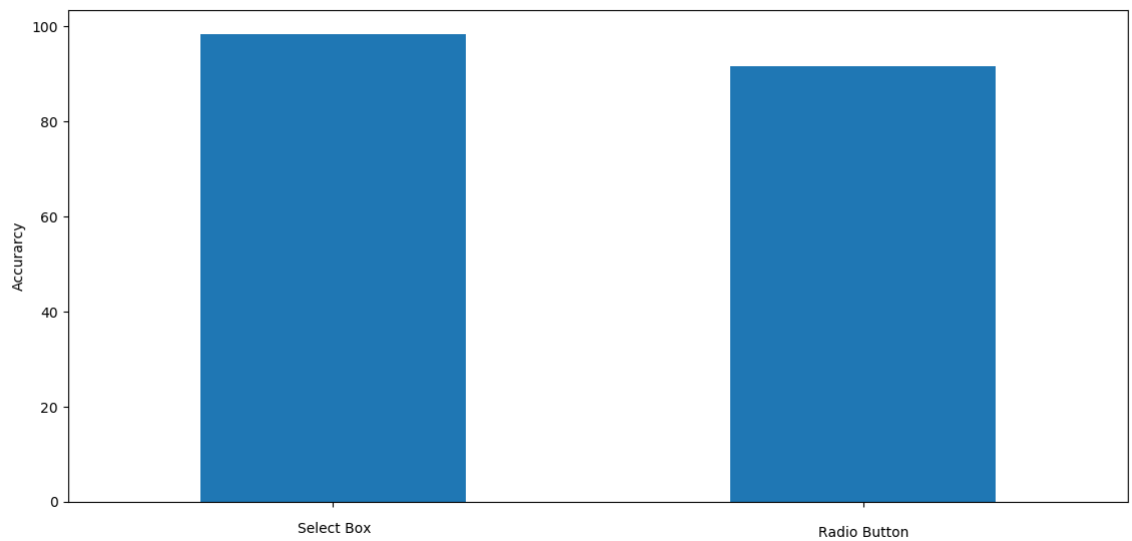


Figure 4.4: Accuracy of tagging a document in HTML select box and radio button

When a user submits a document a confirmation window appears and the user is prompted with a view of the document with the corresponding parts of speech it has been tagged, if user noticed that he/she have tagged a word wrongly , he/she can move back to the previous page and correct it. In this case , user don't have to select all the tags again , because if that happens, nobody wants to do that again. Instead user gets a view where all of his previous selected sentence are tagged and he just have to change the wrong tag and confirm.

Legend:

NN বিশেষ্য	PR সর্বনাম
DM নির্দেশক	V ক্রিয়া
JJ বিশেষণ	RB ক্রিয়া বিশেষণ
PSP অনুসর্গ	CC সংযোজক
RP অব্যয়	QT পরিমানবাচক
PNC যতিচিহ্ন	RD অন্যান্য

Sentence

গত(RB) ডিসেম্বরে(NN) আন্তর্জাতিক(JJ) ক্রিকেট(NN) থেকে(PSP) অবসর(JJ) নেন(V) পন্ডিং(NN) । সিপিএল(NN) ছাড়াও(PSP) ইন্ডিয়ান(JJ) প্রিমিয়ার(NN) লিগ(NN) মাতাবেন(V) তিনি(PR) । প্রতিযোগিতার(NN) ষষ্ঠ(QT) আসরে(NN) আইপিএলের(NN) দল(NN) মুম্বাই(NN) ইন্ডিয়ান্সের(NN) অধিনায়ক(JJ) হিসেবে(CC) দেখা(V) যাবে(V) তাকে(PR) ।

Back

Confirm

Figure 4.5: Confirmation window

4.4 Adding Leaderboard

As we've mentioned earlier, people only crowdsource when they get a handy reward or got tricked as they are competing or playing a game. So, we've added a leaderboard where everyone can see who is at the top of data tagging. User can see total tag document userwise. Figure x shows the leaderboard interface .

Current Leaderboard

Search:

Show 10 entries

Regi	Tagged Document
2016331108	40
2016331001	27
2015331002	26
2016331087	26
2015331012	25
2016331043	23
2016331101	23
2015331052	22
2015331039	21
2015331040	21

Showing 1 to 10 of 154 entries

Previous

1

2

3

4

5

...

16

Next

Figure 4.6: Interface of the leaderboard

In the leaderboard interface, user can search username, number of tagged documents in the search box.It enhances their compitition.To reach the top place, they started to tag more docu-ments.Fix x shows the search box of the leaderboard.

Current Leaderboard

Show 10 entries

Search: 2015

Regi	Tagged Document
2015331002	26
2015331012	25
2015331052	22
2015331039	21
2015331040	21
2015331031	19
2015331022	18
2015331037	18
2015331058	17
2015331064	16

Showing 1 to 10 of 31 entries (filtered from 154 total entries)

Previous 1 2 3 4 Next

Figure 4.7: search box in the leaderboard

Before adding leaderboard, we've launched the site for two time periods, but we have to work very hard to motivate them properly to crowdsource. We told them that top five from each batch(as our user domain is currently only sust cse department) will get e gift box of books of mzi. But they couldn't compare among themselves and tagging seems a bit boring to them. But as soon as we added the leader-board feature, everyone got crazy. They saw that how many sentences their friends were tagging and tried to reach the leader-board.

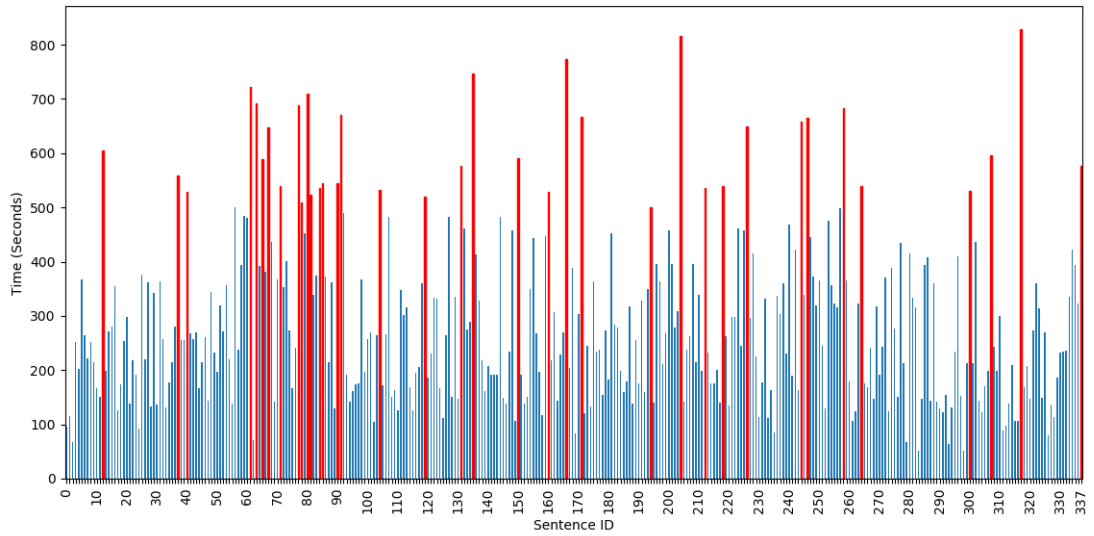


Figure 4.8: Average time to tag a document before adding leader-board.

After adding the leader-board feature, average time to tag a document reduces more than 10 percent than the time takes before. Figure shows the average time it takes to tag the documents between id 150-300 before and after adding leader-board. Snapshot of the leaderboard is given in figure x and y .

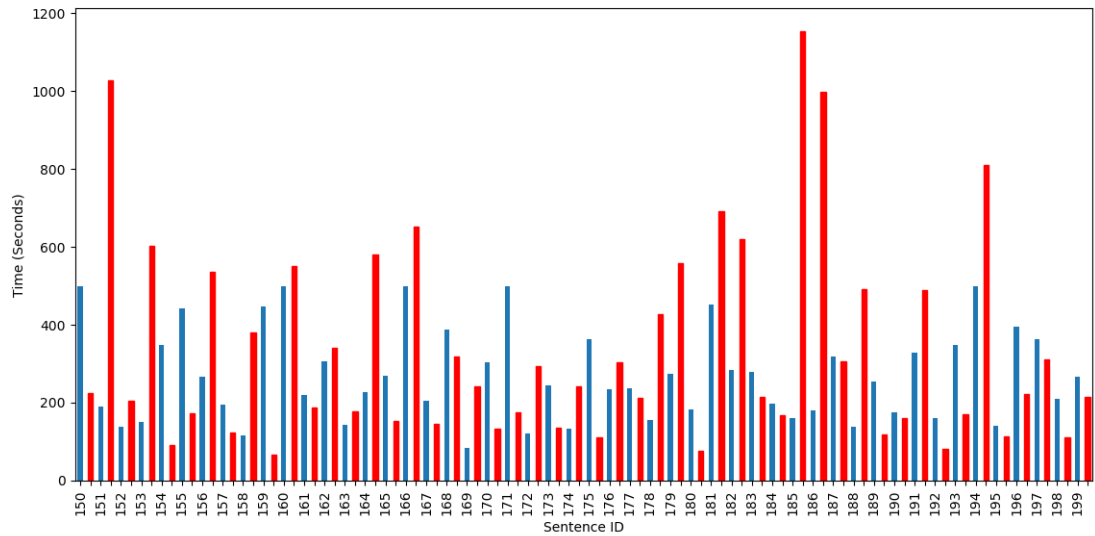


Figure 4.9: Average time to tag a document before and after adding leader-board.

Before adding leader board feature, average time to tag a document by a user is 100 seconds, and after adding leader-board, it reduced to 100 seconds.

4.5 Admin Panel

We've added two admin panel features so that admin can analysis the data from the web site and monitor the crowdsourcing process. From admin view, one can see the all the tagged document in a specific format. Admin can see corresponding profile matrix of the sentences by clicking on view button. Admin can also search username, document id in the search box. Fig 4.10 shows the admin panel view1.

show 200 entries

Search:

username	Document ID	Profile Martix
2014331003	247	View
2014331003	154	View
2014331003	123	View
2014331003	63	View
2014331003	86	View
2014331003	33	View
2014331003	255	View
2014331003	235	View
2014331003	263	View
2014331003	219	View
2014331007	260	View

Figure 4.10: Admin Panel views

In admin panel view2, admins can see each document and how many times they are tagged. They can also see the corresponding profile matrix.

View all tags By users			
Show	200	entries	Search: <input type="text"/>
ID	Tag Count	Profile Martix	Document
0	4	View	এ সময় লোকজনের মধ্যে আতঙ্ক ছড়িয়ে পড়ে। তবে এ আগুনে কেউ হতাহত হয়নি। বৈদ্যুতিক শর্টসার্কিট থেকে আগুনের সূত্রপাত হয়েছে বলে ফায়ার সার্ভিসের সদস্যরা জানিয়েছে।
1	4	View	দ্বিতীয় খেলাটি রুষ্টির জন্য মাঠে গড়াতে পারেনি। তাই তৃতীয় ম্যাচটি ছিল সিরিজ নির্ধারণী। অবশ্য শ্রীলঙ্কা আগেই এক ম্যাচ জিতে সিরিজ হেরে গিয়েছিল।
2	5	View	গড়তে চান নতুন নিয়মও। বিতর্ক তার ছায়াসঙ্গী। তিনি নিজেই তৈরি করেন বিতর্ক।
3	3	View	ভারতের বিখ্যাত নাট্যকার গিরিশ কারনাড রবীন্দ্রনাথকে দ্বিতীয় শ্রেণীর নাট্যকার বললেও বাংলাদেশে রবীন্দ্রনাথই যেন একমাত্র নাট্যকার। এরই দৃষ্টান্ত দেখা গেছে ২০১১ সালের মতো ২০১২। ২০১২ সালেই প্রথম কলকাতার নাট্যোৎসব আমদানি করা হয়েছে ঢাকায়।
4	3	View	গোলের স্থপতি পেদ্রো রদ্রিগুয়েজ। তার পাস মেসিকে খেঁজে পাওয়ার পর বাকি কাজ সহজেই সেরেছেন টানা বারবারের বিশ্ব সেরা ফুটবলার। ভেরানেকে এক ঝটকায় পেছনে ফেলে মেসির বাঁ পায়ের জোরালো ফেরানোর কোনো সুযোগই পান নি লোপেজ।
5	4	View	তবে আইএমএফের শর্তানুযায়ী সংস্থাটির ঋণ ঋকিমুক্ত করতে মুরসিকে কঠোর অর্থনৈতিক সংস্কারের মধ্য দিয়ে যেতে হবে। গত নভেম্বরে আইএমএফের সঙ্গে সমঝোতা হওয়া ৪৮০ কোটি ডলারের ঋণ নিয়ে আলোচনা আবার শুরু করতে চায় বলে জানিয়েছে কায়রো। নভেম্বরের পর থেকে কায়রোজুড়ে শুরু হওয়া সহিংস প্রতিবাদের কারণে ঋণ নিয়ে ওই আলোচনা মিসর সরকারের অনুরোধে স্থগিত করা হয়েছিল।
6	4	View	আগের দিন বিনা উইকেটে ২৩ রান নিয়ে প্রথম ইনিংসে গতকাল দ্বিতীয় দিনের মতো ব্যাট করতে নামে বিসিবি নর্থ জোন। সবকটি উইকেট হারিয়ে দলীয় স্কোরে আর ২৫১ রান জমা পড়তেই দ্বিতীয় দিনের খেলা শেষ হওয়ার আগেই গুঁড়িয়ে যায় দলটি। বোলিং নৈপুণ্যে প্রথম ইনিংসে ৩ রানের লিড পায় সেন্ট্রাল জোন।
7	5	View	দেশের জনগণের আবেগকে পূর্জি করে গণমাধ্যমগুলোর সহযোগিতায় তারা জনগণকে বিভ্রান্ত করার অপচেষ্টা চালায়। দেশপ্রেমী বিপ্লবী তরুণরা এই অপচেষ্টাকে রুখে দিতে ফেসবুক, ব্লগ ও অনলাইন মাধ্যমে ব্যাপক তত্পরতা পরিচালনা করে। এতে ভীত হয়ে পড়ে শাহবাগ আন্দোলনকারীরাও।

Figure 4.11: Admin panel 2.

We've provided users a reference document.Using this document, users get idea of the examples, international labels of the parts of speech. we have used this paper [38] as the reference paper.We will add our own reference document later.

2. Major Language Class

Bangla	: [BNG]	: E.g. lok _[BNG]
English	: [ENG]	: E.g. myan _[ENG]
Foreign	: [FRN]	: E.g. Admi _[FRN]

1. Major Lexical Classes in Bengali

Sl. No.	POS		Label	Example	Comments
1	Noun	বিশেষ্য	\NN\	balak, ghar, shahar, katha, manus	
2	Pronoun	সর্বনাম	\PR\	ami, tumi, se, tara, tui	
3	Demonstrative	নির্দেশক	\DM\	ye, ei, oi, tai,	
4	Verb	ক্রিয়া	\V\	Karchi, kratam, gelo, yabe	
5	Adjective	বিশেষণ	\JJ\	Bhala, manda, sundar, sada	
6	Adverb	ক্রিয়াবিশেষণ	\RB\	Kadachit, babd, karane, sahaje	
7	Postposition	অনুসর্গ	\PSP\	Pare, kache, age, niche	
8	Conjunction	সংযোজক	\CC\	Kintu, athaba, barang	
9	Particle	অব্যয়	\RP\	i, -o, -to na, ne, ni	
10	Quantifier	পরিমাপক	\QT\	ek, dui, tin, pratham, payla, dvitiya	
11	Punctuation	যতিচিহ্ন	\PNC\	,, : ;	
12	Residuals	অবশিষ্ট পদ	\RD\		

POS for Bengali

Sl. No	Category			Label	Annotation Convention**	Examples	Remarks
	Top level	Subtype (level 1)	Subtype (level 2)				
1	Noun (বিশেষ্য)			NN	NN		
1.1		Proper (ব্যক্তিবাচক)		\PPR\	\NN_PPR\	Mohan, ravi,	

Figure 4.12: Reference document

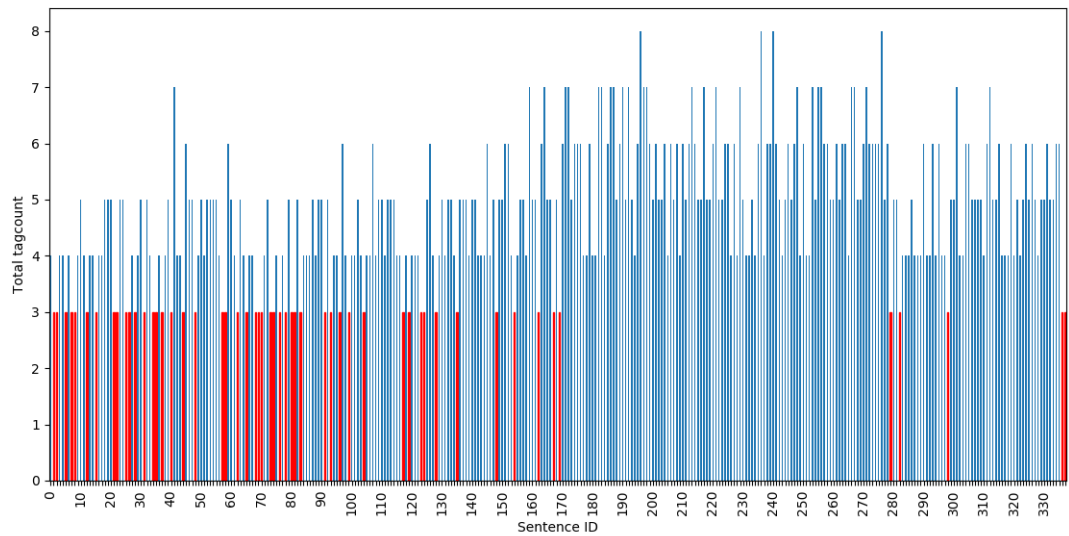
Chapter 5

Results and Analysis

We have collected total of 1591 documents. Each document contains three sentences, so in total 4773 sentences. 159 users have been participated in the crowdsourcing. All of them are the students of department of Computer Science and Engineering, Shahjalal University of Science and Technology. A brief summary is shown in table 5.1 Each document is tagged at least three times, and at most eight times. Our system was build in that manner that all documents will be tagged one after one like a queue, but sometimes crow-sourced worker skips some sentences, but still the counter increases. so this property isn't kept properly. 186 documents are tagged more than or equal to five times, while rest of the 152 words are tagged less than five times. 54 documents have been tagged the least amount of times, which is three. Figure reftgcount shows total tag count of all the documents.

Table 5.1: Summery of the dataset

Total Documents	1591
Total Sentences	4773
Total Words	7657
Unique words	3082
Most frequent word	৩
Crowd worker	159
Average Document tagged	4.72
Average time to tag a document	289 seconds
Maximum number of times a document is tagged	8
Minimum number of times a document is tagged	3



We found that there are 7657 words in our 337 sentences, and there are 3082 unique words. Among them, 3 appeared in 95 documents, which is highest of them all. Table 5.2 shows the most frequent words.

As we haven't used any steamer yet,so করে, করেন, করব, করেছেন all of these words counts as unique words. Table 5.3 shows the effect of not stemming the words before finding unique words,all of the words in the same column can be treated as a single word if we have used a steamer. This is a drawback for us, we will try to overcome it later.

Word	Count	Word	Count
ও	95	এ	90
করে	88	থেকে	71
এই	52	তিনি	52
করতে	47	না	46
এক	44	তার	43
তবে	42	হয়েছে	42
ম্যাচে	39	হবে	36
নিয়ে	35	আর	34
পর	33	প্রথম	33
দুই	33	ম্যাচ	31
শেষ	31	করা	31
জন্য	31	আগে	31
বাংলাদেশ	30	মধ্যে	30
পয়েন্ট	30	ব্যাট	30
হয়ে	29	হয়	29

Table 5.2: Most frequent words

সময়ের	সূচক	হলো	হাসির	শিল্পীরই	সম্ভাব্য	হারাতে	শিল্পের
সময়ে	সূচকও	হলেন	হাসিমুখ	শিল্পীরা	সম্ভাবনাই	হারল	শিল্পখাতে
সময়কে	সূচকটি	হলেও	হাসিকে	শিল্পী	সম্ভবত	হারতে	শিল্পও
সময়	সূচকের	হলে	হাসি	শিল্পীর	সম্ভব	হার	শিল্প

Table 5.3: Effects of not stemming

Before we used profile matrix and consensus score to find the appropriate tag, We run a test to see how many words have been tagged by all the users same every-time. Among the words which have been tagged same by all the users, a pie chart is given to show which parts of speech have been tagged correctly more in figure 5.2. The chart shows that tagging বিশেষ্য and ক্রিয়া are comparatively easier than other documents.

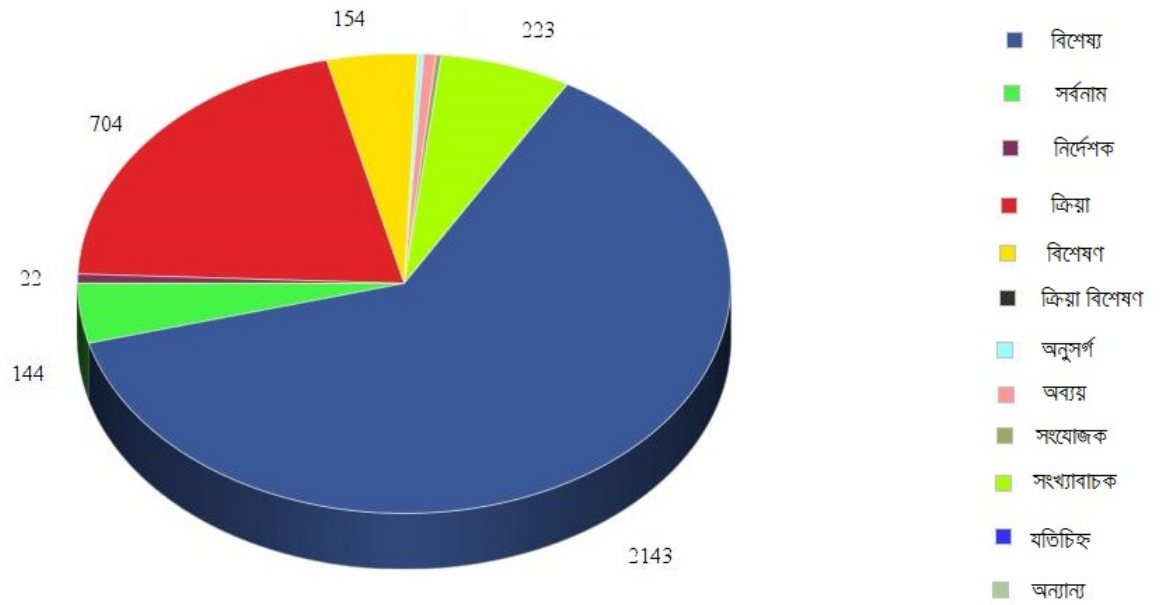


Figure 5.2: Correctness of the parts of speeches

	অস্ট্রেলিয়ার	উদ্বোধনী	জুটি	ভাঙেন	অশ্বিন
বিশেষ্য	7	1	4	0	7
সর্বনাম	0	1	2	0	0
নির্দেশক	0	1	0	0	0
ক্রিয়া	0	0	0	6	0
বিশেষণ	0	3	0	1	0
ক্রিয়া-বিশেষণ	0	0	0	0	0
অনুসর্গ	0	0	0	0	0
সংযোজক	0	0	0	0	0
অব্যয়	0	0	0	0	0
পরিমাণবাচক	0	0	0	0	0
যতি-চিহ্ন	0	0	0	0	0
অন্যান্য	0	1	1	0	0

Table 5.4: Profile Matrix of অস্ট্রেলিয়ার উদ্বোধনী জুটি ভাঙেন অশ্বিন

	গত	অর্থবছর	থেকে	আমদানিতে	স্থবিরতা	দেখা	দেয়
বিশেষ্য	1	5	0	4	2	0	0
সর্বনাম	0	0	0	0	0	0	0
নির্দেশক	2	0	0	0	0	0	0
ক্রিয়া	0	0	0	2	0	7	6
বিশেষণ	1	2	0	1	2	0	0
ক্রিয়া-বিশেষণ	2	0	0	0	3	0	0
অনুসর্গ	0	0	5	0	0	0	1
সংযোজক	0	0	0	0	0	0	0
অব্যয়	1	0	2	0	0	0	0
পরিমাণবাচক	0	0	0	0	0	0	0
যতি-চিহ্ন	0	0	0	0	0	0	0
অন্যান্য	0	0	0	0	0	0	0

Table 5.5: Profile Matrix of গত অর্থবছর থেকে আমদানিতে স্থবিরতা দেখা দেয়

We calculated the profile matrix and consensus score for each of the sentences, some samples and the snapshot of the result set are shown in table 5.4, 5.5. 5.6 . From these tables, we can assume the tags of the words of each of the sentences. But we can see that some words have maximum tag count for different parts of speech, so we cant determine the parts of speech of them. A user tags most of the words in document three as অন্যান্য , and unfortunately we were unable to find a unique maximum tag for the words শর্টসার্কিট and থেকে. Collecting more data from the users will improve our system.

	বৈদ্যুতিক	শর্টসার্কিট	থেকে	আগুনের	সূত্রপাত	হয়েছে	বলে	ফায়ার	সার্ভিসের	সদস্যরা	জানিয়েছে
বিশেষ্য	2	1	0	3	0	0	0	4	2	3	1
সর্বনাম	0	0	0	0	0	0	0	0	0	0	0
নির্দেশক	0	0	0	0	0	0	0	0	0	0	0
ক্রিয়া	0	1	0	0	0	1	4	3	1	0	2
বিশেষণ	1	0	1	0	0	0	0	0	1	1	0
ক্রিয়া-বিশেষণ	0	1	0	2	0	0	0	0	1	0	0
অনুসর্গ	0	0	1	0	0	0	0	0	0	0	1
সংযোজক	0	1	1	0	0	0	0	0	0	0	0
অব্যয়	0	0	1	1	0	0	0	0	0	0	0
পরিমাণবাচক	0	0	0	0	0	0	0	0	0	0	0
যতি-চিহ্ন	0	0	0	0	0	0	0	0	0	0	0
অন্যান্য	1	1	0	1	1	1	0	1	0	0	0

Table 5.6: Profile Matrix of বৈদ্যুতিক শর্টসার্কিট থেকে আগুনের সূত্রপাত হয়েছে বলে ফায়ার সার্ভিসের সদস্যরা জানিয়েছে

File	Edit	Format	View	Help
এ 0 1 2 0 0 0 0 0 1 0 0 0	score 2			
সময় 3 0 0 0 1 0 0 0 0 0 0 0	score 1			
লোকজনের 4 0 0 0 0 0 0 0 0 0 0 0	score 0			
মাঝে 0 0 0 0 2 0 0 1 1 0 0 0	score 2			
আতঙ্ক 1 0 0 0 3 0 0 0 0 0 0 0	score 1			
ছড়িয়ে 0 0 0 1 0 3 0 0 0 0 0 0	score 1			
পড়ে 0 0 0 4 0 0 0 0 0 0 0 0	score 0			
consensus	score 7			
তবে 0 0 0 0 0 0 2 2 0 0 0 0	score 2			
এ 0 1 2 0 0 0 0 0 1 0 0 0	score 2			
আগুনে 3 0 0 0 1 0 0 0 0 0 0 0	score 1			
কেউ 0 4 0 0 0 0 0 0 0 0 0 0	score 0			
হতাহত 1 0 0 0 2 1 0 0 0 0 0 0	score 2			
হয়নি 0 0 0 2 1 0 0 0 1 0 0 0	score 2			
consensus	score 9			
বৈদ্যুতিক 2 0 0 0 1 0 0 0 0 0 0 1	score 2			
শর্টসার্কিট 1 0 0 1 0 1 0 0 0 0 0 1	score 2			
থেকে 0 0 0 0 0 0 1 2 1 0 0 0	score 2			
আগুনের 3 0 0 0 0 0 0 0 0 0 0 1	score 1			
সূত্রপাত 0 0 0 1 0 2 0 0 0 0 0 1	score 2			
হয়েছে 0 0 0 4 0 0 0 0 0 0 0 0	score 0			
বলে 0 0 0 3 0 0 0 0 0 0 0 1	score 1			
ফায়ার 4 0 0 0 0 0 0 0 0 0 0 0	score 0			
সার্ভিসের 2 0 0 1 1 0 0 0 0 0 0 0	score 2			
সদস্যরা 3 0 0 0 1 0 0 0 0 0 0 0	score 1			
জানিয়েছে 1 0 0 2 0 0 1 0 0 0 0 0	score 2			
consensus	score 14			
দ্বিতীয় 0 0 0 0 1 0 0 0 0 2 0 0	score 1			
খেলাটি 2 0 0 1 0 0 0 0 0 0 0 0	score 1			
বৃষ্টির 2 0 0 1 0 0 0 0 0 0 0 0	score 1			
জন্য 0 0 0 0 0 0 2 0 0 0 0 1	score 1			
মাঠে 2 0 1 0 0 0 0 0 0 0 0 0	score 1			
গড়াতে 0 0 1 2 0 0 0 0 0 0 0 0	score 1			
পারেনি 0 0 0 2 0 1 0 0 0 0 0 0	score 1			
consensus	score 7			
তাই 0 0 1 0 0 0 0 1 1 0 0 0	score 0			

Figure 5.3: Profile matrix and consensus score of dataset

Chapter 6

Future Work

Now we have studied some previous works on crowdsourcing systems. We have built a webApp for collecting data by crowdsourcing, collected some data, do some analysis on the data and find consensus score and profile matrix of each sentence of our dataset. Now we are going to propose some works that we are going to do in our next semester.

- Now we've collected data from a small amount of people, just from department of CSE. In future We'll try to collect more data from whole university and with association with pipilica team, we'll try to collect data from around the country.
- We'll use HMM and SVM or GA once we get enough data, we are planning to implement hidden markov model.
- As deep learning is trending now-a-days, we'll try to use deep learning. We'll integrate our research part in the interface so that the system can tag document when a user gives input sentence.

References

- [1] J. Howe., “Crowdsourcing: Why the power of the crowd is driving the future of business.” Crown Business, 2008.
- [2] L.-J. C. M.-C. Yuen and I. King, “A survey of human computation systems. in cse ’09: Proceedings of ieee international conference on computational science and engineering, pages 723–728.” .IEEE Computer Society, 2009.
- [3] I. K. M.-C. Yuen and K.-S. Leung., “Task matching in crowdsourcing.” InCPSCom ’11: Proceedings of The 4th IEEE International Conferences on Cyber, Physical and Social Computing. IEEE Computer Society, 2011.
- [4] A. J. Quinn and B. B. Bederson., “Human computation: a survey and taxonomy of a growing field.” In Proceedings of the 2011 annual conference on Human factors in computing systems, CHI ’11, pages 1403–1412, New York, NY, USA, 2011. ACM.
- [5] A. mechanical turk., <https://www.mturk.com/>.
- [6] V. D. B. J. Gordon and L. K. Schubert., “Evaluation of commonsense knowledge with mechanical turk.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 159–162, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [7] Q. Gao and S. Vogel, “Consensus versus expertise: a case study of word alignment with mechanical turk.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 30–34, Morristown, NJ, USA, 2010. Association for Computational Linguistics.

- [8] K. T. S. R. . M. Jha, J. Andreas and K. McKeown, “Corpus creation for new genres: A crowdsourced approach to pp attachment.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 13–20, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [9] G. Parent and M. Eskenazi., “Clustering dictionary definitions using amazon mechanical turk.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 21–29, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [10] A. Skory and M. Eskenazi., “Predicting cloze task quality for vocabulary training.” In Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications, IUNLPBEA ’10, pages 49–56, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [11] J. W. C. Akkaya, A. Conrad and R. Mihalcea., “Amazon mechanical turk for subjectivity word sense disambiguation.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 195–203, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [12] C. Callison-Burch and M. Dredze., “Creating speech and language data with amazon’s mechanical turk.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 1–12, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [13] E. M. C. Higgins and L. Moretto., “Mturk crowdsourcing: a viable method for rapid discovery of arabic nicknames?” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 89–92, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [14] A. K. N. K. J. M. T. Finin, W. Murnane and M. Dredze., “Annotating named entities in twitter data with crowdsourcing.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 80–88, Morristown, NJ, USA, 2010. Association for Computational Linguistics.

- [15] M. P. N. Lawson, K. Eustice and M. Yetisgen-Yildiz., “Annotating large email datasets for named entity recognition with mechanical turk.” InProceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 71–79, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [16] F. X. M. Yetisgen-Yildiz, I. Solti and S. R. Halgrim., “. preliminary experience with amazon’s mechanical turk for annotating medical named entities.” InProceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 180–183, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [17] J. G. J. C. M. R. C.-j. a. B. Mellebeek, F. Benavent and R. Banchs., “Opinion mining of spanish customer comments with nonexpert annotations on mechanical turk.” InProceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 114–121, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [18] R. Wang and C. Callison-Burch., “Cheap facts and counter-facts.” In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 163–167, Morristown, NJ, USA, 2010. Association for Computational Linguistics.
- [19] D. E. R. O. Alonso and B. Stewart., “Crowdsourcing for relevance evaluation.” SIGIR Forum, 42:9–15, November 2008.
- [20] F. E. B. D. R. Choffnes and Z. Ge, “F. e. bustamante, and z. ge. crowdsourcing servicelevel network event monitoring. sigcomm comput. commun. rev.” 40:387–398, August 2010.
- [21] D. C. Brabham., “Crowdsourcing as a model for problem solving: An introduction and cases.” Convergence: The International Journal of Research into New Media Technologies, 14(1):75–90, 2008.

- [22] U. B. J. Leimeister, M. Huber and H. Kremer, “Leveraging crowdsourcing: Activation-supporting components for it-based ideas competition.” *J. Manage. Inf. Syst.*, 26:197–224, July 2009.
- [23] M. G. G. Little, L. B. Chilton and R. C. Miller, “Turkit: human computation algorithms on mechanical turk.” In *Proceedings of the 23rd annual ACM symposium on User interface software and technology, UIST ’10*, pages 57–66, New York, NY, USA, 2010. ACM.
- [24] B. S. A. Kittur and R. Kraut, “Crowdforge: crowdsourcing complex work.” In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems, CHI EA ’11*, pages 1801–1806, New York, NY, USA, 2011. ACM.
- [25] M. S. S. A. Z. J. Ross, L. Irani and B. Tomlinson, “Who are the crowdworkers?: shifting demographics in mechanical turk.” In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems, CHI EA ’10*, pages 2863–2872, New York, NY, USA, 2010. ACM.
- [26] S. S. J. S. Downs, M. B. Holbrook and L. F. Cranor, “Are your participants gaming the system? screening mechanical turk workers.” In *Proceedings of the 28th international conference on Human factors in computing systems, CHI ’10*, pages 2399–2402, New York, NY, USA, 2010. ACM.
- [27] C. G. Harris, “You’re hired! an examination of crowdsourcing incentive models in human resource tasks.” In *Proceedings of the Workshop on Crowdsourcing for Search and Data Mining (CSDM 2011), WSDM 2011*, pages 15–18, New York, NY, USA, 2011. ACM.
- [28] G. Kazai, “An exploration of the influence that task parameters have on the performance of crowds.” In *Proceedings of the CrowdConf 2010, CrowdConf 2010*, 2010.
- [29] B. K. S. A. Moreno, J. L. de la Rosa and J. M. Barcenas, “Reward system for completing faqs.” In *Proceeding of the 2009 conference on Artificial Intelligence Research and Development: Proceedings of the 12th International Conference of the Catalan Association for Artificial Intelligence*, pages 361–370, Amsterdam, The Netherlands, The Netherlands, 2009. IOS Press.

- [30] J. J. Horton and L. B. Chilton., “The labor economics of paid crowdsourcing.” . In Proceedings of the 11th ACM conference on Electronic commerce, EC ’10, pages 209–218, New York, NY, USA, 2010. ACM.
- [31] D. M. R. B. A. Huberman and F. Wu., “Crowdsourcing, attention and productivity.” J. Inf. Sci., 35:758–765, December 2009.
- [32] D. M. W. F. Wu and B. A. Huberman., “Feedback loops of attention in peer production.” In- Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04, pages 409–415, Washington, DC, USA, 2009. IEEE Computer Society.
- [33] P. G. Ipeirotis., “Analyzing the amazon mechanical turk marketplace.” XRDS, 17:16–21, December 2010.
- [34] C. Grady and M. Lease., “Crowdsourcing document relevance assessment with mechanical turk.” InProceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk, CSLDAMT ’10, pages 172–179, Morristown, NJ, USA, 2010.
- [35] L. A. A. J. Yang and M. S. Ackerman., “Crowdsourcing and knowledge sharing: strategic user behavior on taskcn.” InProceedings of the 9th ACM conference on Electronic commerce, EC ’08, pages 246– 255, New York, NY, USA, 2008. ACM.
- [36] W. Mason and D. J. Watts., “Financial incentives and the performance of crowds. sigkdd explor.” Newsl., 11:100–108, May 2010.
- [37] C. Eickhoff and A. P. de Vries., “How crowdsourcable is your task?” In Proceedings of the Workshop on Crowdsourcing for Search and Data Mining (CSDM 2011), WSDM 2011, pages 11–14, New York, NY, USA, 2011. ACM.
- [38] N. S. Dash., “Pos tagset for bangla document.” Microsoft Research India, Aug, 2010.