DWIT COLLEGE

DEERWALK INSTITUTE OF TECHNOLOGY

Tribhuvan University

Institute of Science and Technology

Bigram Analyzer Model for Devanagari Word Prediction

A PROJECT REPORT

Submitted to

Department of Computer Science and Information Technology

DWIT College

In partial fulfillment of the requirements for the Bachelor's Degree in Computer Science and Information Technology

Submitted by

Sujan Chauhan

Pankaj KC

August 23rd , 2016

DWIT College

DEERWALK INSTITUTE OF TECHNOLOGY

Tribhuvan University

SUPERVISOR'S RECOMENDATION

I hereby recommend that this project prepared under my supervision by NAME OF THE STUDENT [ALL CAPITAL BUT NOT BOLD] entitled **"BIGRAM ANALYZER MODEL FOR DEVANAGARI WORD PREDICTION"** in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Information Technology be processed for the evaluation.

……

Sarbin Sayami

Assistant Professor

IOST, Tribhuvan University

DWIT College

DEERWALK INSTITUTE OF TECHNOLOGY

Tribhuvan University

LETTER OF APPROVAL

This is to certify that this project prepared by SUJAN CHAUHAN and PANKAJ KC entitled **"BIGRAM ANALYZER MODEL FOR DEVANAGARI WORD PREDICTION"** in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Information Technology has been well studied. In our opinion it is satisfactory in the scope and quality as a project for the required degree.

…… Sarbin Sayami [Supervisor] Assistant Professor IOST, Tribhuvan University	…… Hitesh Karki Chief Academic Officer DWIT College
…….. Jagadish Bhatta [External Examiner] IOST, Tribhuvan University	…….. Rituraj Lamsal [Internal Examiner] Lecturer DWIT College

ACKNOWLEDGEMENT

First of all, we would like to express my deepest gratitude to my supervisor associate professor Sarbin Sayami, IOST, TU for his motivation, guidance, advice, and valuable time.

We would like to acknowledge the generous help and support from Mr. Saroj Dhakal, Head of R&D unit, Deerwalk institute of technology for providing the necessary concept and guidance for implementation to undertake this project. We would also like to thank Mr. Hitesh Karki for the whole hearted support.

Sujan Chauhan
TU Exam Roll no: 1815/069

Pankaj KC
TU Exam Roll no: 1804/069

Tribhuvan University

Institute of Science and Technology

STUDENTS'S DECLERATION

We hereby declare that we are the only authors of this work and that no sources other than the listed here have been used in this work.

... ...
Sujan Chauhan

... ...
Pankaj KC

Date: August 23rd , 2016

ABSTRACT

This paper includes the implementation of the Devanagari Keyboard for next word recommendation technique while typing in Devanagari. The prediction for the next word is made using Bigram Analyzer (sequences of words of length 2) for word categorization. For implementation of this project, two techniques have been used. To gather the list of all possible words, Web Crawling has been used and to predict the 'next' word and enable prediction multiple techniques like Content Scrapping, Content Filtering, Split Word have been used. The project successfully implements the predictive algorithm and recommending words that are widely used in very day communication by referring to frequency count of the pair of words and their conditional probability using Bigram Model.

Keywords***:*** Devanagari Keyboard, Web Crawling, Web Scrapping, Bigram Analyzer, Content Filtering.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DK:	Devnagari Keyboard
WCS:	Web Crawling and Scrapping
CF:	Content Filtering
CSS:	Cascading Style Sheet
BA:	Bigram Analyzer
TR:	Transitive Relation
JSP:	JAVA Server Page
HTML:	Hyper Text Markup Language

CHAPTER-1 INTRODUCTION

Text is still the primary medium for all kind of communication. The advent of internet and smart-mobile devices like cell phones, tabs, and etc. have added different dimension to the textual information interchange in the form of SMS(s), e-mails, chats, social media texts: Facebook, Twitter messages and so on. To make these devises user-friendly there are various smart keyboard apps available like SwiftKey, Fleksy and so on and constantly evolving, One of the problem with these apps is they cater to English tying only.. These apps help users by giving multiple suggestions for the next possible word while typing. There are many Nepali typing application such as Hamro keyboard, easy Nepali typing that provides different techniques for word recommendation and consist of existing developed corpus for word recommendation. These recommendations are often limited as the corpus does not include all the words that are used in every day communication.

The developed Devanagari keyboard is capable to predict next possible input word, smart enough to know the dual input words for word recommendation.

1.1 Background

Devanagari keyboard is a Nepali typing interface that has the feature of word recommendation technique. The main aim of this application is to provide easy Nepali typing platform that will contain existing Nepali words crawled from the different Nepali word web sites for recommendation.

1.2 Problem Statement

Currently the applications that are available use the existing database that is prepared by the developer itself and updated whenever the user enter the word on the text editor and recommends only after seeing the user tendency to type in the editor. The Devanagari Keyboard has its database which consists of the collection of the word that are stored

after crawling the web pages with the frequency count of each and every word, Recommends the user even at the initial use of the application.

1.3 Objectives

1.3.1 General objective

a) To implement Bigram Analyzer algorithm for pair of word recommendation in Nepali language.

1.3.2 Specific objective

a) To type the words user have not think of using which will eventually lead to advancement in Nepali language of user.
b) To save the Nepali words to the database and recommend the word according to the frequency count of a single word.

1.4 Limitations

a) It will not work on other text editor beside its own.
b) No recommendation of words according to transitive relation among words.

1.5 Outline of Document

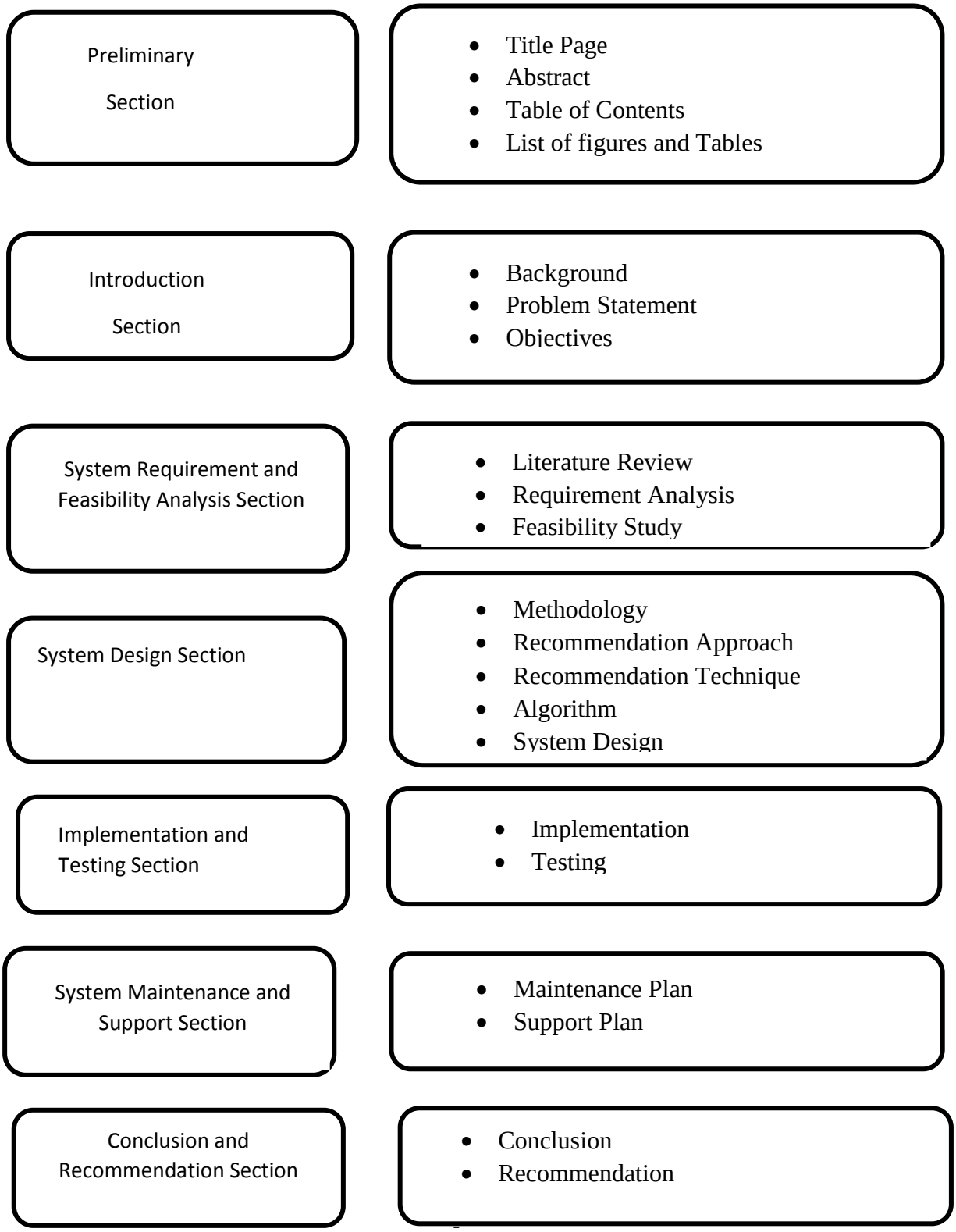

Figure 1 - Outline of document

CHAPTER-2 SYSTEM REQUIREMENT AND FEASIBILITY ANALYSIS

2.1 Literature Review

Research on next word prediction has received researcher's interest recently. A dedicated workshop called Workshop on Advances in Text Input Methods (WTIM) on this issue held twice in 2011 and 2012. In these workshops a few research attempts have been reported on Word prediciton development for Devanagari languages. Those endeavors include languages like Nepali, Hindi, Bengali (Salam et. al., 2012), and Bramhi (Brouillette et. al., 2012). All these works targeted on phonetic typing and back transliteration. The motivation of the present project is entirely different, that is to develop a keyboard for Nepali languages, which can predict next input word.

(Amitava, 2014)

Word prediction and word completion are the important phenomena in typing that benefits users who type using keyboard or similar devices. Word completion works so that the user types the first letter or letters of a word and the program provides one or more higher probable words. . If the word user intends to type, is included in the list then the user can select it. If the word that the user wants is not predicted, the user must type the next letter of the predicted word. Word prediction helps disabled people for typing, speed up typing speed by decreasing keystrokes, helps in spelling and error detection. Word prediction is very important and complex task in natural language processing (NLP) to predict the correct word to complete a sentence in a very meaningful way.

(Md. Tarek Habib, 2015)

Word predictors provide the user with a prediction window, i.e. a menu that, at any time, lists the most likely next word candidates, given the input that the user has typed until the current point. If the word that the user intends to type next is in the prediction window, the user can select it from there. Otherwise, the user will keep entering letters, until the target word appears in the prediction window (or, of course, until user finishes typing the

word). Word prediction systems typically base their predictions on various forms of n-gram statistics extracted from one or more training corpora. The (percentage) keystroke savings rate (ksr) is a standard measure used in AAC research to evaluate word prediction systems. The ksr can be thought of as the number of keystrokes, in percent, that a "perfect" user could save by employing the relevant word predictor to type a certain corpus, over the total number of keystrokes that are needed to type the same corpus without using the word predictor.

(Marco Baroni, 2012)

Predictive text systems in place use the frequency-based disambiguation method and predict the most commonly used word above other possible words. In the research done by Sachin Agrawal and Shilpa Arora had the keypad on the phone where multiple texts are mapped to same numeric code and generally used for SMS. That used the bigram language model but after training the Markov model on its email corpus. This shows the 31% of improvement over the traditional frequency based word estimation. But if the error made at t^{th} word propagated in the sequence and hence the error for the current word also reflects the error made on the previous word (on the basis of which it was predicted). So the approach is only suitable for the word prediction on the basis of text but not for the word prediction on the basis of word.

(Sachin Agarwal, 2007)

Word prediction on Bangla Language use N-gram model such as unigram, bigram, trigram, deleted interpolation and backoff model for auto completing the sentence by predicting a correct word in a sentence. That uses the larger data corpus of Bangla Language of different word to predict the correct word with the accuracy as much as possible. The probability of a word depends on the previous word which is called Markov assumption. Unigram looks single item from a given sequence. Bigram is called first-order Markov model which looks one word into the past and trigram is second-order Markov model which looks two words into the past and quadrigram is third-order Markov model which looks three words into the past and similarly an N-gram language model is N-1 Markov model which looks N-1 words into the past. Backoff model and deleted interpolation model are applied to solve the problem of probability of word sequence will very low and zero. In the backoff method, for N = 3 in (1), i.e. for a trigram model, the word sequences will follow trigram probabilities at first; if it could not match then word sequences will follow bigram model; if it also could not match then word sequence will

follow unigram model and predict at least a word. If there are non-zero trigram counts, it rely on trigram counts and don't interpolate bigram and unigram counts at all. In the research, the average accuracy obtained after using the mentioned N-gram model are 21.24%, 45.84%, 63.04%, 63.5% and 62.86% respectively. The N-gram model works well for English Language but its challenging to get the 100% accuracy for Bangla Language. (Md. Tarek Habib, 2015)

Related Existing Application

Hamro Keyboard

Only Nepali keyboard which supports three keyboard layouts, unicode transliteration, MPP Romanized and traditional. Hamro keyboard is a native android keyboard for Nepali language, you can use this keyboard to write Nepali text/Devanagari script in any app. It removes the need to copy and paste. With Hamro Nepali keyboard, you can type directly in any app. It supports three keyboard layouts, Unicode transliteration, MPP based romanized layout and traditional layout. Here, recently added Emoji support and numeric keypad on the latest update. Now type in Nepali and help your friends to type in Nepali.

Figure 2 - Hamro keyboard

In this app, if user type in English and it shows the word translating to it. But initially it does not recommend the word for us. First it saves the word that user initially type in it and it stores the every word in its own database. And if later we type the similar letters in the text area then it suggests user with the words.

2.2 Requirement Analysis

2.2.1 Functional requirement

a) Recommend pair of words with higher frequency.
b) Save the web contents (Nepali words) to database with their frequencies.

2.2.2 Non-Functional requirement

In console, the typed in text should be in Nepali it is done with the help of Google Input tools.

2.3 Feasibility Study

Feasibility study gives an assessment of the practicality of this project. Following feasibility study has been done:

2.3.1 Technical feasibility

Devanagari Keyboard application that uses Java framework. It supports both Windows and Linux platform for its operation. All of the technology required by Devanagari Keyboard are available and can be accessed freely, hence it was determined technically feasible.

2.3.2 Operational feasibility

It replaces the current procedure of Nepali typing on a computer. It can be easily installed on a computer and can be for chatting and writing Nepali words to write documents. Hence Devanagari Keyboard was determined operationally feasible. Hence, it is operationally feasible.

CHAPTER-3 SYSTEM DESIGN

3.1 Methodology

Steps involved:

a) Web Crawling and Web Scrapping
b) Content Filtering
c) Split Word
d) Building Data structure
e) Text Editor for the application

3.1.1 Web crawling and web scrapping

The crawler was designed to crawl the content of web, contains a list of URL's to visit. As the crawler visits these entire URL and identifies all hyperlinks in the page and add them to the list. The crawler pushed the list of pages to visit in stack and popped one by one to visit the pages. The page consists of different link. As the crawler checked all link whether it was already visited or not. If not, then push to the stack. Afterwards, the crawler started to store the content of particular page in the file named with the respective page name. As reading from the pages and writing on the file was the most tedious and time consuming so the crawler wrote the crawled text of the pages only if its size exceeds from 1MB and reset the crawled text and repeat it until all the text of the page were written.

3.1.2 Content filtering

The file containing crawled text may had both English and Nepali words. But Devanagari keyboard was only for the Nepali words. So English word in the file must be filtered. As content filtering read the file line by line and checked the ASCII value for the each and every word. The content filtering wrote only those words holding the ASCII value of

Nepali words excluding the English words and saved under the filter with respective file name.

3.1.3 Split word

The filtered file had only the Nepali word, but it might have lots of word that repeat more than once in the file. The split simply split the bunch of words of whole file individually and counted each and every word reoccurrence in the file and stored under the split with the respective entered file name.

3.1.4 Building data structure

The Devanagari keyboard holds different data structure such as set, list and map. The crawler holds the set of visited page and the list of pages to visit. The reason to have set for visited pages was to determine the unique entries. Simply, no duplication. The reason to have list for pages to visit was to store the bunch of URL. When the crawler visited a pages it collect all the URLs of that page and just append them to the list. While split word holds the map to show the word and their respective counts separated from each other.

3.1.5 Text editor for the application

For the effective user interaction and reliable use of the application, the Devanagari keyboard had the web based user interface. The interface had text field where user can type the word, using that word the application recommends the word just below the text field. The interface takes first word as the input on which recommendation is done based on the recommendation algorithm and the frequency count of each and every word available already on the application. The user interface is designed in a such way that it had encapsulated the “preeti font” so that user can type directly on Nepali or even they type in English, it shows the corresponding Nepali text.

In this example, the word recommendation is done according to the user. The pattern of words is recorded to phone database and later use those word in recommendation purpose.

3.2 Recommendation Technique

Single word recommendation, neighbor word recommendation and transitive word recommendation are the recommendation techniques for the recommending the word. Single word recommendation technique recommends the word even after the user type the single character of the word.

Example: म becomes मलाई, मैले, and so on. Neighbor word recommendation technique recommends the word after the user provide the word in text field with the maximum frequency of that word after the typed word. Example: जान्छु comes after म. That is, म जान्छु. While transitive recommendation technique recommends the bunch of words like phrase after the user type the certain word on the text field. Example: If "डीएर्वाक ईन्स्टिच्युट अफ टेच्नोलोगी" is most repeated then if any of the user type डीएर्वाक then it suggest whole ईन्स्टिच्युट अफ टेच्नोलोगी. That is, डीएर्वाक suggests ईन्स्टिच्युट अफ टेच्नोलोगी.

Analyzing this, the Devanagari keyboard used the neighbor word recommendation technique for the word recommendation after the user type on the interface.

3.3 Algorithm Used

On implementing the neighborhood text recommendation, this application used following techniques

3.3.1 Word frequency count

After word filtration, the application counts the repetition of the each word and saves it. This helps us to find the probability of word that might come after certain word we type. To find the probability we have used Bigram Analyzer that tells what word to predict. It helped us to find the frequency of both words that had been come together.

3.3.2 Bigram analyzer

For implementing Bigram analyzer, first create the probability model for application as: P(w1 w2) denotes the probability of the word pair w1 w2 occurring in sequence (e.g., मेरो

नाम , म घर, तिमि खाउ etc.). If we make one minor simplifying assumption, then the formula is exactly the same as for a single word:

$$P(w1\ w2) = c(w1\ w2)/N$$

In fact, this method extends to n-grams of any size:

$$P(w1 \ldots wn) = c(w1 \ldots wn)/N$$

Below is the Bigram Probabilities of Nepali words from a file, calculated form the above probability model of bigram analyzer algorithm.

Table 1 - Probability count

Words	**Count**	**Probability**
अबको परिवर्तनमा	29	0.00791572
सजाय र	11	0.00635354
गरिबी कती	20	0.002095115
पुग्नु पर्यो	35	002082486
फोन नम्बर	44	0.002717714
पशुपति क्षेत्रमा	22	0.002634364

Word Prediction Mechanism

The application implements the concept of condition probabilities and chain rule for the next word prediction, the probability of seeing word wn given the previous words w1, w2, . . . wn−1
is:

$$P(wn|w1 \ldots wn-1) = \frac{C(w1 \ldots wn)}{C(w1 \ldots wn-1)}$$

In our Bigram model: Maximum likely hood function.

$$P(\text{नम्बर/फोन}) = \frac{C(\text{फोन नम्बर})}{C(\text{फोन})}$$

$$= \frac{12}{70} = 0.171428$$

3.4 System Design

3.4.1 Class diagram

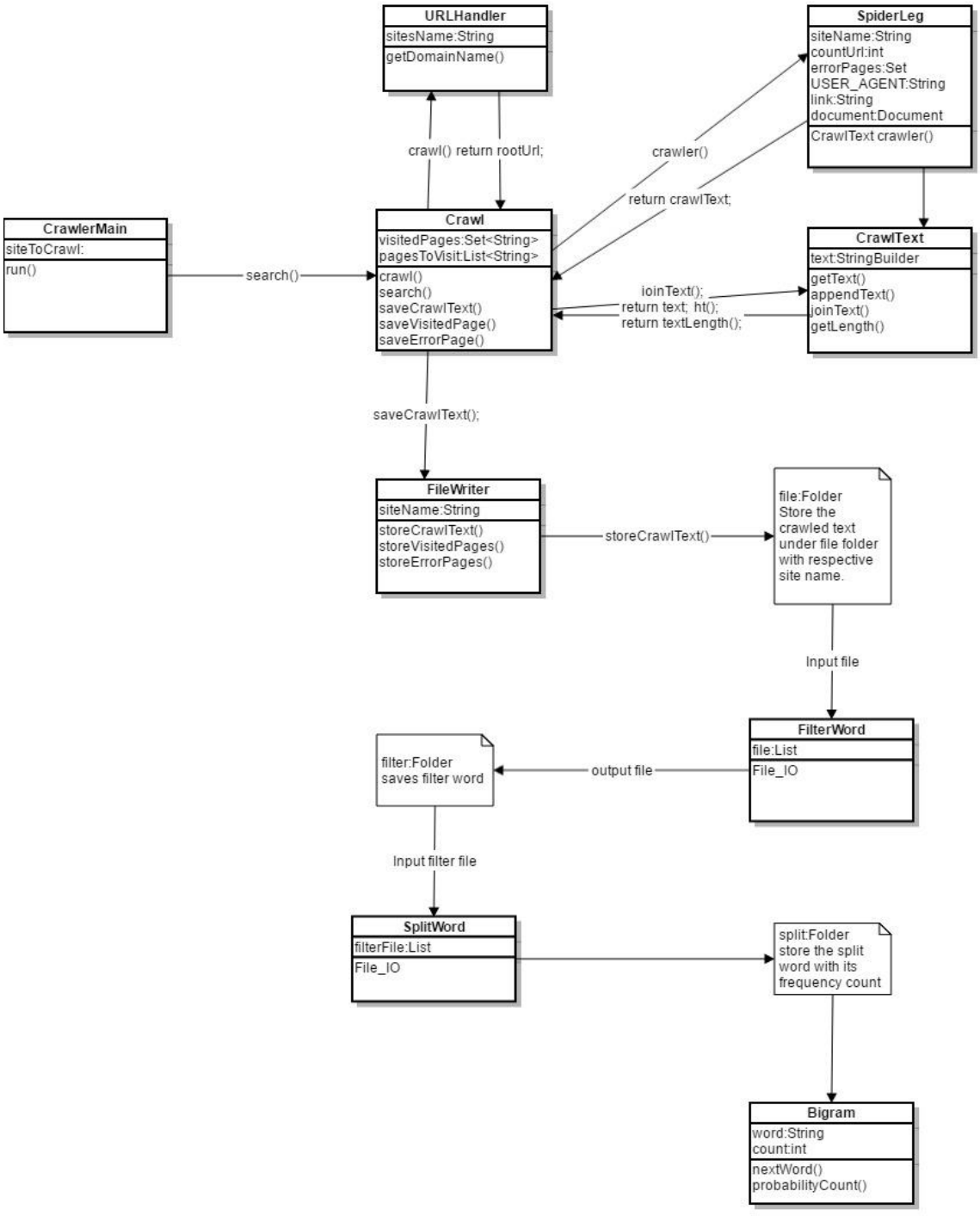

Figure 3 - Class diagram of devanagari keyboard

ClawlerMain above mentioned on the class diagram hold the list of site to crawl and made the thread of which and called the crawl through the search method. Class Crawl hold the proper data structure for the whole system. As it hold the set for the visited site in order to determine the uniqueness of the site as no need to crawl the same site again and again. While it hold the list for the site to visit so that it could store more list of hyperlinks as well.

Before crawling, Crawl makes sure that all list of site to visit are in proper format or not. If yes then it made the connection to the user agent (browser) through Jsoup. Before connection was made, all the root site are pushed on the stack and popped one by one to crawl the site even with its associated hyperlinks. Afterwards, Crawl started to crawl all the text of the site including the text of all those associated hyperlinks. All those crawled text were need to be save.

But reading and writing at the same time may too long time. So considering that it wrote only after it exceeds the limit included length size of the crawl text and then wrote on the file under its own site name. Class CrawlText was for the purpose of getting the crawled text as it helped on append and join the text through the appendText and hoinText method respectively which was essential on determining the length of the crawled text in order to write on the file on the class writer.

As the crawled text might have other words which are not Nepali, then such words are need to be filter. So the all the file saved after the crawling the site were again made the list and used one by one using the buffer reader which read the file line by line and removed the other words checking the ASCII value of the words and saved those file under the filter with the file excluding the text other than Nepali. After that those file were provided to the class split in order to split the even if the words are separated especially by commas (,) , purnabiram (|) because of which it made easy to implement the bigram analyzer algorithm on each and every pair of the words available on the file which was saved after the words had been separated by the Split.

3.4.2 Process View

Each loop represents a thread of control.

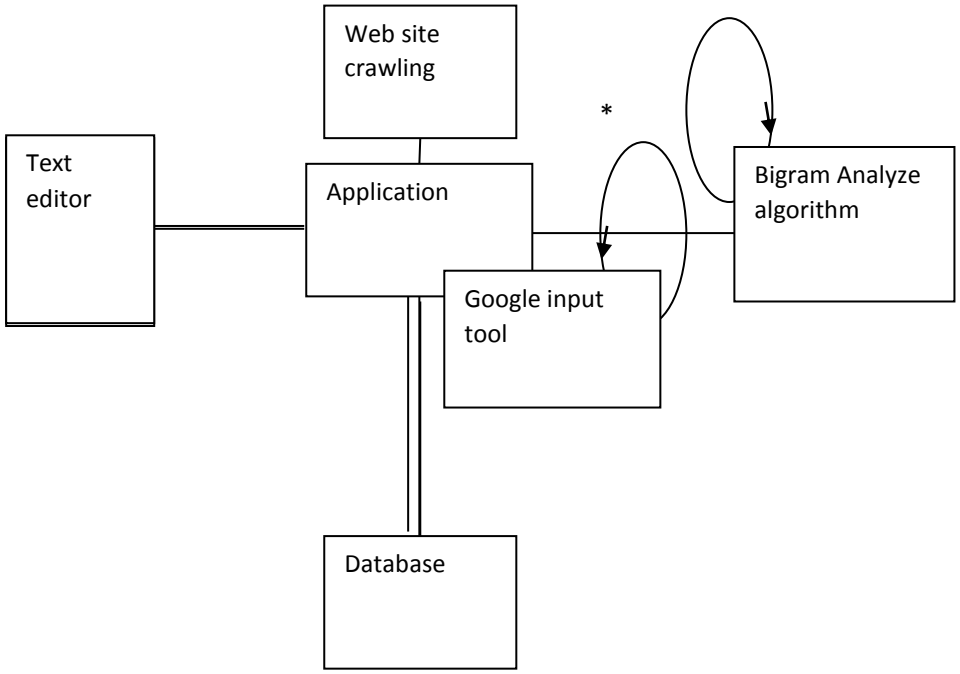

Figure 4 - Process model of devanagari keyboard

Step1: Use Google input tools for input.

Step2: Type word on text editor.

Step3: Crawl the Nepali website

Step4: Send the word to Application

Step5: Use Bigram Analyzer for word pair of word recommendation.

Step6: Display the Recommendations.

CHAPTER-4 IMPLEMENTATION AND TESTING

4.1 Implementation

4.1.1 Tools used

Java

Java is the object oriented programming language and is the cross platform. Java programming was used as the tool as it was easy to crawl the website in java which could be done using the JSOUP library provided by the java.

Jsoup library

JSOUP is the java library for working with real world HTML. As it is very convenient API for extracting and manipulating data using the DOM and CSS. This library was used to connect with the URL and extract the data or information from that page.

Gliffy

Gliffy is the software for diagramming different diagram like UML, flowchart, etc. Here it was used to create the class diagram of the application.

Google input tools

Used for word input in a personal computer. It supports multiple language typing. For our application we have only used its nepali tool for typing.

4.2 Testing

Black-Box Testing

The technique of testing without having any knowledge of the interior workings of the application is called black-box testing.

In this application the tester is oblivious to the system architecture and does not have access to the source code. Typically, while performing a black-box test, a user interacted with the system's user interface by providing inputs and examining outputs without knowing how and where the inputs are worked upon.

4.3 User acceptance testing

Black box testing was conducted. A random user used the application which provided the input and the output was examined.

4.3.1 Testing module

Module A:	Module B:
Single Word Prediction Display	Dual word Prediction Display

Test case 1: Module A

Title: Single Word Prediction Display

Precondition: Word pair count is saved in database using Bigram analyzer.

Test Steps:

1. Single word Input was successfully given to the program.
2. List of recommended word was displayed.

Outcome:

Hence, According to the input word its neighbor words according to its frequencies were successful displayed.

Example:

Typed Input: मेरो Recommended words

घर

नाम

पनि

Test case 2: Module B

Title: Dual word Prediction Display

Precondition: Single word Prediction Successfully displayed.

Test Steps:

1. Input a Word pair.
2. Display the neighbor words of the second word.

Outcome:

Hence, According to the input of a pair of word, neighbor words of second word according to its frequencies were successful displayed.

Example:

Input word: मेरो देश Recommended words

नेपाल

प्यारो

राम्रो

CHAPTER-5 SYSTEM MAINTENANCE AND SUPPORT

If any maintenance, changes, or additional functionality needs to be made to the Devanagari Keyboard application, it is helpful to know some of the ins and outs of the code and to be made aware of some of the specifics of the code. This section of the document will go into a bit of detail on some parts of the project to help the maintenance person understand the project. Because this project is a 2-tier project, this discussion of the code will be grouped into the three different tiers of the architecture.

Front-End

The visual interface that the user sees in the project is made up of a web text editor. The code for each form is grouped into two files: a .java file and a Designer.java file.

Database

The physical database is created within the code and is implemented using the integrated phpmyadmin database.

If any failure occurred, knowing above architecture it will be easy to maintain the application. The documentation is done of each and every part implementation phases.

CHAPTER-6 CONCLUSION AND FURTHER RECOMMENDATIONS

6.1 Conclusion

Hence, The **Devanagari keyboard** allows the user to type on the text editor and recommends the words, after word that had been typed in by the user using the neighbor word recommendation technique. Simply the **Devanagari keyboard** recommends the user with the **Bigram analyzer algorithm** implementing frequency count of each pair of word. This application was specific to next word recommendation from which it has produced a significant amount of accuracy for predicting the next Nepali word. Hence, with all available technical knowledge the application was successfully developed.

6.2 Recommendation

Other specified recommendation techniques can be implemented as below:

Transitive recommendation technique recommends the bunch of words like phrase after the user type the certain word on the text field. Example: If "डीएर्वाक ईन्स्टिच्युट अफ टेच्नोलोगी" is most repeated then if any of the user type डीएर्वाक then it suggest whole ईन्स्टिच्युट अफ टेच्नोलोगी. That is, डीएर्वाक suggests ईन्स्टिच्युट अफ टेच्नोलोगी.

REFERENCES

Amitava, D. (2014). Antaryāmī: The Smart Keyboard for Indian Languages. *The Smart Keyboard for Indian Languages*.

Fry, J. (2013). *BIGRAMS AND TRIGRAMS.*

James Downey (October 2011). Dynamic Approach to web crawling.

Marco Baroni, J. M. (2012). *Predicting the Components of German Nominal.*

Md. Tarek Habib, M. M. (2015). AUTOMATED WORD PREDICTION IN BANGLA LANGUAGE USING STOCHASTIC LANGUAGE MODELS. *International Journal in Foundations of Computer Science & Technology (IJFCST).*

Moumie Soulemane, M. R. (2012). At all web crwaling. *Crawling the Hidden Web: An Approach to*.

Petersen, Dustin Hillard and Sarah. (2009). *N-Gram Language Model.*

Riloff. (1995). *Little words can make a big difference for text classification.*

Sachin Agarwal and Shilpa Arora, S. A. (2007). Context Based Word Prediction for Texting Language.

APPENDIX I

Figure: Web Crawling

र्ति बढेको छ । कलेजहरु विद्यार्थीलाई आकर्षित गर्न भवन सजाउने देखि अनेक तिकडम अपनाउन तिर लागि परेका छन् । गुणस्तरिय शिक्षा दिने नाममा धेरैजसो कलेज
। थालेपछि उनका विश्वास पत्राहरूले छाड्न थालेका हुन्...काठमाडौं । प्रधानमन्त्री केपी ओलीले दुई वर्षभित्र खाद्यान्नमा देशलाई आत्मनिर्भर बनाउने घोषणा गरेको चार म
को करोडौं रकममा मस्ती गर्नेहरुलाई भने कुनै कारबाही हुनसकेको छैन । यो करोडौं रकम सर्वसाधरणको हो । यूनिटीले विभिन्न कम्पनीहरु र संस्थामा शेयर लगानी गरे
येका निशा अधिकारीको कनेक्सन देश टुक्रयाउने अभियानमा जुटेका सीके राउतसँग रहेको खुलासा भएको छ । निशाले मंगलबार सीके...कञ्चनपुर, असार १५, न्यूज अ
हरु विद्यार्थीलाई आकर्षित गर्न भवन सजाउने देखि अनेक तिकडम अपनाउन तिर लागि परेका छन् । गुणस्तरिय शिक्षा दिने नाममा धेरैजसो कलेजहरुले ब्रहत्ुट नै मच्चाए
विश्वास पत्राहरूले छाड्न थालेका हुन्...काठमाडौं । प्रधानमन्त्री केपी ओलीले दुई वर्षभित्र खाद्यान्नमा देशलाई आत्मनिर्भर बनाउने घोषणा गरेको चार महिनापछि कृषिमन्त्री
मस्ती गर्नेहरुलाई भने कुनै कारबाही हुनसकेको छैन । यो करोडौं रकम सर्वसाधरणको हो । यूनिटीले विभिन्न कम्पनीहरु र संस्थामा शेयर लगानी गरेको थियो । तर उक्त
रीको कनेक्सन देश टुक्रयाउने अभियानमा जुटेका सीके राउतसँग रहेको खुलासा भएको छ । निशाले मंगलबार सीके...कञ्चनपुर, असार १५, न्यूज अभियान । कञ्चनपुरत
नलाईन पत्रीका,रेडियो तथा टेलिभिजन मार्फत विभिन्न ज्ञानमूलक सामाग्रीका साथै राजनिती, आर्थिक, सामाजिक, स्वास्थ्य, खेलकुद,साहित्य, मनोरंजनका बिबिध सामाग्री
आइरहेको छ । समाजमा हरेक व्यक्तिलाइ सूसूचित गराउने प्रयास अन्तर्गत बेला बेलामा बिभिन्न क्रियाकलाप समेत संचालन गर्दै आई रहेको सगौरव जानकारी गराउदछ
र्क गर्नुहोस, समाज परिवर्तनको लागि हामी केही गर्दै छौं...। तसर्थ आउनुहोस हामी सबै मिली समाज परिवर्तनको लागि केही गरौं, हामी केही गर्दै छौं, गर्दै जानेछौं
, कठमाण्डौं । फोन नम्बर : ०१-४२६९९४१ Copyright © 2013 - 2016 Newsabhiyan.com.np All
मलद्वार (बेर्नो) बाहिर निस्कने, दिशागर्न गाह्रो हुने, दिशागर्दा रगतआउने, फिस्टुला आदि । कब्जियत (दिशा कडा हुने, खुलेर नहुने, धेरैपटक टाइलेट जान पर्ने–धेरै सम
ाँदा पनि ठिक नभएका पाठेघर संबन्धिअन्य समस्याहरू ।मोटो तथाव्ब्लोपन ः– पेट तथाहिप बढेको, जस्तोसूकै उपायअपनाउँदापनि नघटेको मोटोपन तथाकुनै पनि
राश्यता, मर्ने चिन्ताहुने, तनाब, भविष्यको बारेमा सोँचेर दिक्कमान्ने–निश्चित दिशा अवलम्वन गर्न नसक्ने आदि ।मधुमेह ः– मधुमेह(चिनी रोग), यौन शिथिलता, शारीरिक
कठमाण्डौं । फोन नम्बर : ०१-४२६९९४१नेशनल भिजन मिडिया ग्रुप प्रा.ति. द्वारा संचालित सम्पर्क कार्यलय : – गौतमबुद्ध मार्ग, अनामनगर, कठमाण्डौं । फोन नम्ब
, साइटिका, अण्डकोष सूनिने–दुख्ने, नशाको कमजोरी, नशा च्यापिएको, अपरेशन गरेर पनि ठिक नभएको जोर्नीहरूको समस्याआदिको विशेष उपचार ।धातु÷यौन रोग
– नाकबन्दहुने, गन्धथाहानपाईने, नाकमामास् बढेको, निधार दुख्ने, आधा–आधा टाउको दुख्ने, रूघा लागिरहने(क्रयमि बिि भचनधरक्रयmmयत अयमि) पटक पटकनि
शाला पुनर्जीवन प्राप्तिको केन्द्र भईरहेको छ ।संपर्क स्थान ः ॐ आयुर्वेद आरोग्यशाला कमलमार्ग घर नं. १७२ केसीटोल कमलपोखरी काठमाडौं ।फोन ०१–४४३१७४७,
गाल्छ ।उनी मैदानमै रुन थाले ।उनलाई मैदानबाट सम्हालेर फर्काउन पोर्चुगलका प्रशिक्षक नै मैदानमा पुग्नु पर्यो । एजेन्सीएनभिएम ग्रुपद्वारा संचालित यस् न्यूज अभियान

APPENDIX II

Figure: Content Filtering

APPENDIX III

Figure: Database Structure

wordOne	wordTwo	count
जय	नेपाल	33
जय	जनता	30
नेपाल	कांग्रेस	35
नेपाल	सरकार	37
मेरो	देश	25
मेरो	नाम	18
मामा	माइजु	18
गरीब	जनता	18
गरीब	देश	25
केटा	केटी	30
नेपाल	सुन्दर	22
मेरो	घर	23
माया	नमार	23
नेपाली	शब्द	32

APPENDIX IV

Figure: Word Suggestion