UNIVERSITY OF NAIROBI
SCHOOL OF COMPUTING AND INFORMATICS

A FRAMEWORK TO SUPPORT COLLABORATIVE LEARNING USING SOCIAL MEDIA

BY

MWAI JOSPHAT KARANI

P58/61678/2010

SUPERVISOR:

DR. ROBERT OBOKO

A research project report submitted in partial fulfillment of the requirements of the Degree of Master of Science in Computer Science

DECLARATION

This research project, presented on this report is my original work and to my knowledge has not been presented for any other University Award.

Signed: …….. **Date…….**

Student: Mwai Josphat Karani

P58/61678/2010

This project has been submitted as part of fulfillment of the requirements for the award of Master of Science in Computer Science of the School of Computing and Informatics of the University of Nairobi, with my Approval as the University Supervisor.

Signed: …….. **Date…….**

Supervisor: Dr. Robert Oboko

DEDICATION

I dedicate this work to my Parents my Mum; Regina Nyawira and my late Dad; Julius Mwai for their love, endless support and encouragement. Dad you will always be missed and I cherish the memories we had.

ACKNOWLEDGEMENT

First and foremost I would like to thank the Almighty God for providing me with the opportunity and strength to accomplish this task. All the glory unto Him.

I am particularly thankful to my supervisor Dr Robert Oboko, for his patience in guiding me through this research process. I am deeply indebted to you for the help, suggestions and encouragement I received during this period.

To my mum and brother your support during this period was immeasurable and I am grateful.

I owe special thanks to the staff at Kenya Institute of Education for the support they gave me during the evaluation of my prototype.

To all I say thank you.

ABSTRACT

New frameworks for developing interactions between the actors of learning processes require new ways of gaining an understanding of such processes. Today, these frameworks are of a technological kind, and their prime expression is found in social networking sites, which have already been called 'social operating systems'. Students have gone ahead and taken advantage of these platforms by creating virtual learning environments, which are quickly becoming repositories of collective knowledge and teaching resources. Unfortunately, we are not able to fully utilize this collective knowledge being generated through group member's interaction. The focus of this study, was to ensure that the knowledge being generated through social media is not lost. This was achieved by developing a framework that was able to extract, classify and store the classified comments from a social media account to support collaborative leering.

The prototype was then evaluated at the Kenya Institute of Education using a staff induction course that had four modules. Data collection tools used were interviews and questionnaires the interviews were used as part of requirements gathering, this was done to gain background data on the use of social media among the learners. The learners interacted with the prototype for two weeks and evaluated it based on usability and functionality.

Data analysis was presented using descriptive statistics. From the initial survey it was evident that all of the respondents had social media accounts, but very few used them for educational purposes. The study finding indicated that majority of the learners were comfortable interacting with the prototype and that they require minimal training on how to use it. The learners were consistently positive in their assessment regarding functionality of the prototype. From the finding it was evident that web 2.0 tools in particular social media have enormous potential in education and that learners are comfortable using these tools.

The study focused only on developing an enabling environment for learners to interact and be able to capture and store the knowledge generated through these interaction. However, how interactions affects learning and the relationship between the two is a complex pedagogical phenomenon that requires further investigation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	Application programming interface
CCK	Communication Commission of Kenya
FDD	Feature Driven Development
P2P	Peer to Peer
Tf	Term frequency
Idf	Inverse document frequency
CSCL	Computer Supported Collaborative Learning

DEFINITION OF TERMS

Social media- Social media refers to the means of interactions among people in which they create, share, and exchange information and ideas in virtual communities and networks.

Twitter- Twitter is an online social networking service and microblogging service that enables its users to send and read text-based messages of up to 140 characters, known as "tweets"

Microblogging - Microblogging is a broadcast medium in the form of blogging. A microblog differs from a traditional blog in that its content is typically smaller in both actual and aggregate file size.

Tweet – The message you post and send out to your followers is called a 'tweet'

Follower -A follower is a Twitter user who has subscribed to your account so he or she can see all your posts and updates on your own page.

Learning- Learning is acquiring new, or modifying existing, knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information.

Elearning - refers to the use of various kinds of electronic media and information and communication technologies (ICT) in education.

Collaborative learning - is a field of study centrally concerned with meaning and the practices of meaning making in the context of joint activity, and the ways in which these practices are mediated through designed artifacts.

CHAPTER 1.0 INTRODUCTION

This chapter looks at the theoretical background of the study which touches on some issues that are important as a way of introducing the study they include: e-learning, Web 2.0, and the new generation of learners, the chapter also looks at problem statement, objectives of the study, and justification for the study.

1.1 Background of the study

E-learning as we know it has been around for the last decade and it has change from being a radical idea to something that is regarded as mainstream in our education institution especially higher learning institutions. In Most of these institutions, E-learning is in form ofa learning managements system (LMS) which allows tutors to publish course content. This content could be purchased or developed by themselves. Individual learner's then register for courses and the system tracks their activity on the course and provides reports on both course activity like the number of learners taking a course as well as learner activity like what individual leaner have been doing, scores attained etc.

LMS is based on web 1 technology which is "read-only web" that is content is authored and published by the tutor for the learner to consume. However, over the last few years the web has changed and a new breed of technologies by the name Web 2.0 has emerged. Web 2.0 is defined as "read-write web" which enables learners to co-create knowledge, collaborate and share content with others. Time magazine wrote about the Web 2.0 phenomenon "It's a story about community and collaboration on a scale never seen before. It's about the cosmic compendium of knowledge Wikipedia and the million-channel people's network YouTube and the online metropolis MySpace. It's about the many wrestling power from the few and helping one another for nothing and how that will not only change the world, but also change the way the world changes." (Lev Grossman 2006 in Time Magazine3) This change in web technology has come with a new generation of learners the I or Z generation as described by (Jayson S 2006) generation who's technology is part of their DNA,

to them technology is not a tool but a part of life they are constantly online creating and sharing an array of content he puts it that they have no off switch.

This I or Z generation also known as Digital natives are living in collaborative learning communities already. They will expect associations and corporate training departments to accommodate their native preferences for learning. Collaboration among students is a defining feature of constructivist classrooms (Jonassen 1993), and Web 2.0 has wide-ranging potential for social interactivity and the promotion of collaboration and collective learning.

Twitter is an example of a Web 2.0 social networking site, which has enormous potential in the field of education despite the fact that it was not designed as an environment for constructing and managing learning experiences. Interest in collaboration is a natural outgrowth of the trend in education toward active learning, whereby students become involved in constructing their own knowledge through discovery, discussion, and expert guidance. Collaboration affords students the opportunity to share thoughts and interact with peers, facilitators, and experts in a defined area (Francesc 2011).

For this generation of learners going to the LMS to find answers to their questions is not an option they prefer using web 2 tools like Google, Wikipedia or YouTube, or simply posting questions to their networks on Twitter or Facebook in order to get immediate, up-to-date and relevant answers.

In striving to keep up with this new generation of learners, educators need to revisit their conceptualization of teaching and learning. Educators need to engage meaningfully with the world in which students live and strive to integrate technologies and tasks that are meaningful and relevant to the demands of today's networked society (NMC 2007).That is learning professionals need to exploit the tools that their learners are using on a daily basis e.g. Facebook and Twitter. Because it is more appropriate to take the learning to the learners, rather than forcing the learners to come to the learning.

Due to these new trends in the web educationist should try and answer these questions**.** The model of e-learning as we know it where content is produced, published, organized, structured into courses, and consumed by learners, is changing.

Educationist should try and keep up with these changes by envisioning a learning environment that is learner centered. In this new learning environment E-learning ceases to be a medium, and becomes like a platform. It also ceases to be a tool for publishing course content and becomes a content-authoring tool where learning is created. The educational potential of social media is extraordinary, and particularly so if a fresh look is taken at the concepts of education and training, and emphasis is placed on the social nature of knowledge construction that is collaborative learning.

1.2 Collaborative Learning

Computer Supported Collaborative Learning (CSCL) " is a field of study centrally concerned with meaning and the practices of meaning making in the context of joint activity, and the ways in which these practices are mediated through designed artifacts." (Koschmann 2002) The phrase, Professor Stahl a leading proponent of CSCL explains that the phrase "practices of meaning making in the context of joint activity" is a good definition of what he calls "group cognition." (Stahl G, 2006) The wording of the sentence uses the theoretical concepts that help us to understand the behavior of students in CSCL settings. It is about practices, meaning making, joint activity and mediation by technology. He adds that The term "learning" itself carries traditional connotations of an increase in factual knowledge by an individual. So in CSCL, we often talk about "collaborative knowledge building" rather than learning, the goal in a CSCL classroom it to have groups of students develop knowledge artifacts like documents expressing a theory, where the document gradually becomes more and more developed. The nature of collaborative learning is such that a group working together is likely to develop a document that takes into account more issues, uses more abstract conceptualizations and develops more sophisticated arguments than any individual member would have produced on their own.

Through participating in the group process, the individuals may not only learn the theory that the group developed, but also learn to think about the theory from multiple perspectives as well as learning how to work well together with others on this kind of learning task.

The "social nature of collaborative learning is not just some kind of external factor influencing individual learning, but it is the group process itself, created in the interaction of the group and

making the learning at every level possible. When a software designer understands collaborative learning this way, the goal of design is to support productive collaborative knowledge building. This certainly includes providing media for communication within the group across networked computers. But it also involves supporting group processes, like argumentation, seeking information, explaining terms, pointing things out. In addition, it may include making the created knowledge documents easy to modify, persistent for later use and sharable within a larger community.

1.3 Problem Statement

New frameworks for developing interactions between the actors of learning processes require new ways of gaining an understanding of such processes. Today, these frameworks are of a technological kind, and their prime expression is found in social networking sites, which have already been called 'social operating systems' (Francesc 2011).The use of these social networking sites is increasing in all areas of society but, although students have been active in social networking for almost a decade now, during this time, schools and teachers have largely ignored them. Students have gone ahead and taken advantage of these platforms by creating virtual learning environments, which are quickly becoming repositories of collective knowledge and teaching resources. It is common knowledge that when students need to make sense of some information or just needs a quick answer, they turn first to their peers. Hence social media networks like Twitter are providing a platform for this kind of interaction. Through these interaction a lot of collective knowledge being created and shared in this social media sites like Twitter.

Unfortunately, we are not able to fully utilize this collective knowledge being generated through group member's interaction. The focus of this study, is that social media accounts lacks a true system for tagging, filtering, searching and organizing information. The speed at which posts are produced, is the same speed at which it disappears into older posts, they lack a way of 'storing knowledge', whereby knowledge generated can be classified, saved, and easily and quickly retrieved. Trying to locate a message that we know is somewhere within a Twitter account can often be very chaotic. These limitations of social media as a tool for supporting collaborative learning are what this research addresses.

1.4 Goal of the project

Develop a technique for harvesting, classifying and storing for comments from a social media account to support collaborative learning.

1.5 Objectives

a) Develop a way for capturing data from a social media discussion forum.

b) Develop a text classifier that is able to classify the data from social media into groups based various categories.

c) Develop a platform that is able to store the classified data (from b above) for easy and quick retrieval.

1.6 Justification

Social media has an extraordinary level of penetration in society. The digital natives are constantly logged in, sharing information and chatting. It is common knowledge that when students need to make sense of some information or just needs a quick answer, they turn first to their peers and social media networks like Twitter are providing a platform for this kind of interaction. Through this interaction a lot of collective knowledge is being created and shared in this social media site However the speed at which posts are produced, is the same speed at which it disappears into older posts making the process of searching for a message that we know is somewhere within a social media group to be very chaotic. This is the missing link that this research will address the fact that there is a lot of knowledge being generated in social media sites like Twitter which is not being stored.

1.7 Proposed solution overview

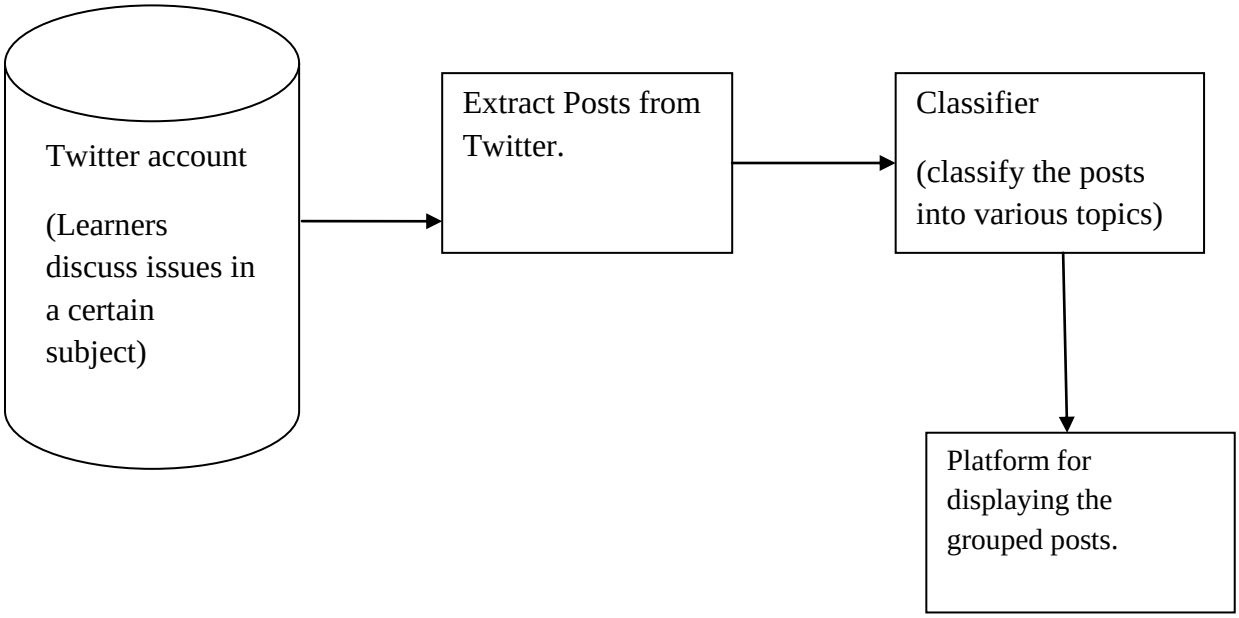

Fig 1: Proposed solution overview

CHAPTER 2.0 LITERATURE REVIEW

2.1 Introduction

This chapter will look at the following aspect: learning community's and their impact on teaching and learning, the evolution of e-leaning, Web 2.0 phenomenon, internet and social media growth in Kenya LMS and social media as tools that support collaborative learning including how and where they are being used to implement collaborative learning.

2.2Learning Communities

Communities may take many forms including political, religion and neighborhoods of the many forms of community, is a learning community "a learning community is characterized by a willingness of members to share resources, accept and encourage new membership, regular communication, systematic problem solving and a preparedness to share success (Moore & Brooks, 2000)"

The concept of learning communities has been discussed for more than two decades (Caverly& McDonald, 2002). Research has clearly shown that functioning in a community can enhance the learning that occurs among community members (Hargis, 2005). When the interactions among community members are directed toward the purpose for which the community was formed, it is considered collaboration. Woods &Ebersole (2003) assert that optimal learning outcomes are "directly tied to the establishment of social networks among participants engaged in a collaborative learning enterprise". Such collaboration has been shown to be very important in the development of a learning community and in achieving the desired learning outcomes for a course (Yuen, 2003).

Collaborative learning is an educational approach to teaching and learning that involves groups of students working together to solve a problem, complete a task, or create a product. According to Gerlach, "Collaborative learning is based on the idea that learning is a naturally social act in which the participants talk among themselves (Gerlach, 1994). It is through the talk that learning occurs."

There are many approaches to collaborative learning. A set of assumptions about the learning process (Smith and MacGregor, 1992) underlies them all:

1. Learning is an active process whereby students assimilate the information and relate this new knowledge to a framework of prior knowledge.
2. Learning requires a challenge that opens the door for the learner to actively engage his/her peers, and to process and synthesize information rather than simply memorize and regurgitate it.
3. Learners benefit when exposed to diverse viewpoints from people with varied backgrounds.
4. Learning flourishes in a social environment where conversation between learners takes place. During this intellectual gymnastics, the learner creates a framework and meaning to the discourse.
5. In the collaborative learning environment, the learners are challenged both socially and emotionally as they listen to different perspectives, and are required to articulate and defend their ideas. In so doing, the learners begin to create their own unique conceptual frameworks and not rely solely on an expert's or a text's framework. Thus, in a collaborative learning setting, learners have the opportunity to converse with peers, present and defend ideas, exchange diverse beliefs, question other conceptual frameworks, and be actively engaged.

2.3 E-Learning History

E-Learning pre-dates the World Wide Web by several decades. The idea of computer-mediated education first emerged in the late 1950s with proposals for building "teaching machines." In the 1960s, Donald L. Bitzer started the PLATO system (Programmed Logic for Automatic Teaching Operation) for mainframe computers. After this university research lab discover that it was possible to transfer data without printing using internet from one computer to another within their network, Email and file transfer were originally used for educational applications.

Then in the early 1980's stand-alone education application came up for individual learners due to the advent of desktops.

In the late 1980's client- server architecture became common and this led to development of applications that would reside in a server and allow clients machines to access content from the server.

The coming of the world wide web in the 1990's moved the concept of e-learning further by allowing one to link via universal resource locator to online resources on any server that was accessible on the internet. However most content developed during this period were "page turners" with lots of text and graphics or were simple imitations of the stand –alone content that was running on desktops previously.

This approach was not popular because it was a bad simulation of a high school or university class ie a presentation of the content followed by a multi choice test.

Gary wood noted in a white paper in 2004 "The principal reason why most people have trouble surfing through an e-Learning course is that there is usually nothing to do but read, look, and take a multiple choice test.

Most often there are no instructional activities that deeply engage the mind of the learners, and 'interactivity' mostly consists of turning from one screen to another. This is especially problematic for the under-40 generation, which has grown up with fast-paced videogames, movies, and television programs. Reading a lot of text on a screen simply doesn't cut it for them."

2.4 Learning Theories

There is no one theory of Computer Supported Collaborative Learning (CSCL). Research in CSCL is guided by and contributes to a diverse collection of theories. For CSCL, theory must take into account interaction in online environments, knowledge building in small groups and cognition at multiple units of analysis.

Although there are many theories that have contributed to CSCL, we can see the major philosophic influences largely through the works of: Mind in Society (Vygotsky, 1930/1978), Situated Learning (Lave & Wenger, 1991),Lectures on Conversation (Sacks, 1962/1995) and Understanding Computers and Cognition (Winograd& Flores, 1986).

Vygotsky's Theory -is the idea that the potential for cognitive development is limited to a certain time span which he calls the "Zone of Proximal Development" (ZPD). Vygotsky defined ZPD as a region of activities that individuals can navigate with the help of more capable peers, adults, or artifacts. In Vygotsky' view, peer interaction, scaffolding, and modeling are important ways to facilitate individual cognitive growth and knowledge acquisition. ZPD can compose of different levels of expertise of individuals (students and teachers), and can also include artifacts such as books, computer tools, and scientific equipments. The purpose of ZPD is to support intentional learning. Vygotsky's sociocultural approach of learning and ZPD can be successfully employed in the study of Computer supported collaborative learning (CSCL) environment.

Constructivist Theory-Basically, constructivism views that knowledge is not 'about' the world, but rather 'constitutive' of the world (Sherman, 1995). Knowledge is not a fixed object, it is constructed by an individual through her own experience of that object. Its goal is to create learning communities that are more closely related to the collaborative practice of the real world. In an authentic environment, learners assume the responsibilities of their own learning, they have to develop metacognitive abilities to monitor and direct their own learning and performance. The contemporary constructivist theory of learning acknowledges that individuals are active agents, they engage in their own knowledge construction by integrating new information into their schema, and by associating and representing it into a meaningful way. Constructivists argue that it is impractical for teachers to make all the current decisions and dump the information to students without involving students in the decision process and assessing students' abilities to construct knowledge.

In other words, guided instruction is suggested that puts students at the center of learning process, and provides guidance and concrete teaching whenever necessary. Perkins (1991) indicates that students may easily get lost in management without any experience to guide them through the information jungle.

Situated learning-It is not possible to separate cognitive tasks from social tasks, because all cognitive tasks have a social component (Perret-Clermont, 1993). Constructivists view cognition as situation-bound and distributed rather than decontextualized tools and product of minds (Lave, 1988 ; Pea, 1994).

Thinking is both physically and socially situated that problem tasks can be significantly shaped and changed by the tools made available and the social interactions that take place during problem solving. Situated cognition, a new paradigm of learning, emphasizes apprenticeship, coaching, collaboration, multiple practice and articulation of learning skills, stories, and technology (Brown, Collins & Duguid, 1989). "Community of practice," a concept emerging from situated cognition, emphasizes sharing and doing, construct meaning in a social unit Situated learning occurs when students work on authentic tasks that take place in real-world setting. As Lave (1991) states that learning is a function of the activity, context and culture in which it occurs, which contrasts with most classroom learning which is abstract and out of context. Education can apply the two basic principles of situated cognition into classroom practice: present in an authentic context, and encourage social interaction and collaboration. It is believed that rich contexts can reflect students' interpretation of the real world and improve their knowledge being transferred in different situations. Collaboration can lead to articulation of strategies that can then be discussed, which, in turn, can enhance generalizing grounded in students' situated understanding. Within CSCL, this approach can be seen in the idea that one learns mathematics by adopting the practices of mathematicians, such as using mathematical symbolisms, making conjectures about mathematical objects and articulating deductive arguments (Sfard, 2008).

2.5 Web 2.0

The term Web 2.0 was coined in 1999 to describe web sites that use technology beyond the static pages of earlier web sites. It is closely associated with Tim O'Reilly because of the O'Reilly Media Web 2.0 conference which was held in late 2004.Although Web 2.0 suggests a new version of the World Wide Web, it does not refer to an update to any technical specification, but rather to cumulative changes in the ways software developers and end-users use the Web. Anderson (2007) argues, the best way to define Web 2.0 is “to make a reference to a group of technologies which have become deeply associated with the term: blogs, wikis, podcasts, RSS feeds etc., which facilitate a more socially connected Web where everyone is able to add to and edit the information space”.

While the short description of Web 2.0 resources shows that they have started to be used, they seem to have little or no impact on the structure and conception of the old learning paradigms on which today's curricula are built Bartolomé, Antonio (2008) argues. He continues to state that most accepted resources are the ones oriented towards document production and distribution such as wikis, blogs and video sharing. E-Learning courses have not yet adopted aspects related to collective intelligence, horizontal relations, dynamic knowledge conceptions and to new information management tools such as tags and social bookmarking. It is too early to speak of a new paradigm, but there are some elements that do not fit easily in the old eLearning paradigms.

One of these elements is what he describes as beyond the level of a single device: PDA, iPod, compute, learning any place any time. E-Learning courses seem to be based on the concept of studying "any place, any time". But in reality, this is not the case. Traditional distance courses do consider studying at any time, but these moments are considered to be "study moments".

That is, there is a time to study and a time for other activities. In Web 2.0, the difference between study times and other times seems to disappear. If the difference between study time and other time disappears, it could mean that we learn as we are actively doing other things, just as children do.

The others element he talks about is user experience richness and learning from peers. A key aspect in eLearning is p2p learning, learning from peers. Of course concepts such as "peer review", "peer tutoring", etc. are not new. But in the traditional paradigm there is a clear difference between professor and students, as in Web 1.0 there is a difference between the expert and the novice.

In Web 2.0 this difference is more subtle. And of course it is not compatible with those eLearning courses that offer pages and pages of pdf documents. In eLearning 2.0 "Rather than being composed, organized and packaged, eLearning content is syndicated, much like a blog post or podcast." (Downes, 2005).

2.6 Growth of internet and social Media in Kenya

One of the key web 2.0 technologies associated with this change in e-learning is social media. According to a survey conducted by Miniwatts Marketing Group in 2012 on the World Internet statistics, Kenya is ranked number four in terms of internet penetration in Africa. Another report by CCK (2010/2011)shows that by the end of the fourth quarter, there were an estimated 12.5 million internet users compared to 11 million reported in the previous quarter, representing an increase of 13.6 percent. Compared to the same period of the previous year this represented 60.1 percent increase.

The summary in the growth of internet users and internet access is as illustrated in Figure below:

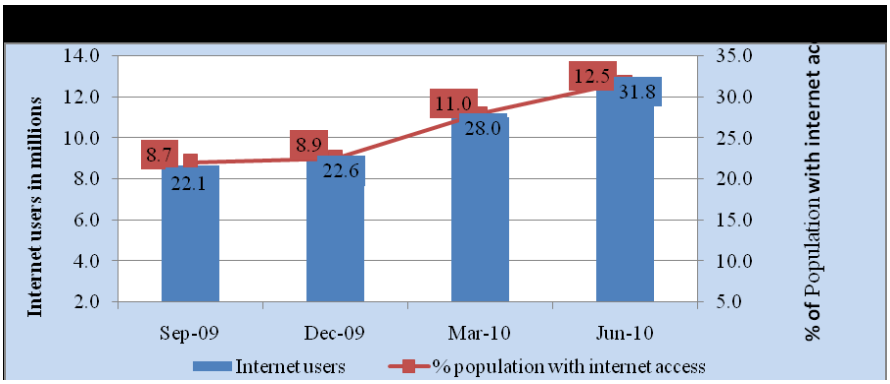

Fig 2: Growth of internet users and internet access

From the above statistics it is clear that internet penetration is on the increase in Kenya and that we know that young people are the ones that are the most active consumers of social media sites. It is because of this that educationists are being forced to look at the educational potential of social media sites so that they can reach the student where they are in an environment that they are already comfortable in.

2.7 Twitter as a collaborative learning tool

Microblogging is a Web2.0 technology, and a new form of blogging that let the users publish online brief text updates, usually less than 140-200 characters, sometimes images too. The posts can be edited and accessed online, or sent as SMS, e-mail or via instant messaging clients. Microblogging enables a real-time interaction between users, using different devices, technologies and applications.

Microblogging is a collaborative technology with a growing interest from users coming from different domains, from eLearning and education in general too. GebrielaGrosseck(2008)

Twitter seems to be the most popular microblogging system. Officially launched in October 2006, Twitter was developed by Obvious Corp. it is a robust, elegant, and simple system has gained a lot of popularity. After creating an account in a few seconds, a user can start twittering. The users of Twitter can send and receive messages via the web, SMS, instant messaging clients, and by third party applications. The notifies can be received in real-time even as SMS.

People use Twitter to communicate, to ask questions, to ask for directions, support, advice, and to validate open-ended interpretations or ideas by discussing with the others. Twitter has mashed up personal publishing and communication, the result being a new type of real-time publishing (GebrielaGrosseck, 2008).

2.8 Educational potential that Twitter offers for collaborative learning

Imagine a 20 minute lecture where all your students back channel about what you're saying. Outside guests or experts are invited in. Someone acts as a "rudder" to keep the conversation on track. The discussion is displayed on a SMARTboard or with a projector. The chatcast is immediately dumped into a wiki. The rest of the class is devoted to reorganizing the wiki clarifying what was said, answering questions (student to student as well as teacher to student; and don't forget the people, students, teachers, mentors or parents beyond the glass walls of the room) summarizing the big ideas, reframing the discussion in terms of what needs to be explained again and where we're going next. Imagine the possibilities Terry Friedman(2007)

As Terry Friedman puts it in the above paragraph the possibilities are endless when it comes to the educational potential of social media. Below are a few areas in which twitter is being used in supporting collaborative learning:

Twitter foresters virtual learning communities: It allows one to extend learning beyond the classroom, Twittering in class or outside of it is about learning.

Exploring collaborative writing: it promotes writing as a fun activity, it fosters editing skills, develop literacy skills; it can give our students a chance to record their cognitive trails and then use them to reflect on their work.

Reader response: students can use tweets to send out questions and observations to the group while engaged in classroom activities.

Explore the potential of microblogging in formal and informal settings: Twitter linked to a course/class blog can offer our students opportunities to discuss different kinds of asynchronous online discourse, considering voice, purpose, audience, to organize ideas, reflect, send notes, manage meet-ups, serendipitous discovery etc.

Collaboration across schools, countries (for example, students leave phone-tweets as they come across real-world examples of what they are studying in class or as they conduct fieldwork).

Facilitating virtual classroom discussion by using @username. It directs the 'tweet' at the intended recipient whilst allowing every student to also see it. The virtual aspects of schooling are well represented in Twitter posts. Provides a platform where teachers make themselves available to their students at all times.

2.9 Elimika program

Most Learning Managements System (LMS) out there simply automate the model of a normal classroom. The only difference being that the learners do not have to meet in the same place at the same time. Most of these systems are missing a vital ingredient that is other people. The advent of social media tools has provided us with a way of overcoming this deficiency, although not fully exploited social media has a lot of potential for teaching and learning. It is a tool that allows one to connect with likeminded people with whom they can communicate, share ideas, resources as well as experiences. Hence the reason that we are moving towards integrating LMS with web 2.0 tools. A good example of this is Elimka.

Elimika is a Learning Management System (LMS) developed at the Kenya Institute of curriculum development (KICD). The LMS (elimika) has several functionalities which make the instructional process totally educative, interactive, interesting and above all bringing fun to the participants. There is a real time chat, an interactive discussion forum and an emailing service just to mention a few of the functionalities.

Initially, it was used within the Institute for orientation of new staff on the administrative structure and roles played by each department and their respective divisions. Upon its successful implementation in the staff induction course, the LMS was deemed capable of conducting online orientation of teachers on curriculum and pedagogy issues. The first cohort of Primary teachers was enrolled for the course and they successfully went through and finally graduated in the year 2012. We are living in the era where teachers cannot avoid collaborations and networking among them and of course not forgetting the sharing of knowledge. Collaborations, networking and sharing of knowledge are made possible through the use of social media tools which are embedded on the home page.

The primary teacher orientation course is guides learners on how to interpret the syllabus. Knowledge on the interpretation of the syllabus is very vital because it assists a teacher in determining the appropriate teaching approaches for use while delivering content.

Teachers are also taken through the available teaching and learning resources for different subjects and this includes the available e-resources which teachers can make use of. Assessment is also discussed at length to equip teachers with vital tips which they need to know while doing it.

2.10 Limitations of social media and LMS

From the above discussion on LMS 2.0 (Elimika) and social media (Twitter), it is clear that these web 2.0 enabled platforms are technologically rich, adaptable expandable and capable of supporting collaborative-working experiences in learning communities. However, it is also evident (Francesc 2011) that these platforms have some limitations in the implementation of learning experiences, either because some of the native tools are not well-developed, or because certain tools are quite simply missing. Some of these limitations include:

Both (LMS 2.0 and social media) lack a true system for tagging, filtering, searching and organizing information. The speed at which news is produced, is the same speed at which it disappears into older posts. They do not allow for knowledge generated to be classified, saved, and easily and quickly retrieved.

The drop-down comments system on walls tends to make it hard to see information. If a user is not a habitual user, the presence of new comments may be overlooked. Likewise, the difference between 'read' and 'unread' items is not easy to see, and is occasionally too spurious.

Despite the fact that social media allows photos to be tagged, this function is not very useful for anything other than indicating the one-off presence of elements in them (such as friends' faces), and it does not extend to other types of object.

In both group discussion boards are too basic they do not allow any type of reply nesting, so, despite their name, they simply become lists of replies to the main topic. Trying to locate a message that we know is somewhere can often be very chaotic.

2.11 Conclusion

From the above discussion it is clear social media offers a lot of potential in the area of e-learning through collaboration among peers where learners become active knowledge builders. However, educationists are not taking full advantage of this application. Those who try to shift are simply automating the model of a normal classroom. The only difference being that the learners do not have to meet in the same place at the same time. Most of these systems are missing a vital ingredient that is offered by Web 2.0 tools like social media which is a collaborative learning environment.

CHAPTER 3.0 METHODOLOGY

3.1 Introduction

This chapter will look at the prototype development methodology and the research methodology that will be used, illustrating in detail how they were used in achieving the objectives of the study.

3.2 Prototype Development Methodology

This chapter presents the research methods that were followed to achieve the overall goal of the project. The system development methodology used was Agile development methodology. Agile methodology can be described as iterative and incremental as it allows the assessment the project's direction throughout the development lifecycle. The methodology also supports addition of features to a system incrementally which is important especially when dealing with applications like Web 2.0 which are frequently being updated and changing.

The overall model of the system was created and features were added to it incrementally until all the objectives are met. Since changes were made to different features as the system was being developed the Feature Driven Development (FDD) a variant of agile methodology was used. FDD is a model-driven, short-iteration process. It begins by establishing the overall model shape then it continues with "design by feature, build by feature" iterations. The features are small and useful in the eyes of the client. FDD will allow for features to be defined from the users (learners) point of view.

Agile development methodology was found to be a good option to implement the proposed solution in an incremental approach through an iterative approach. The methodology basically resulted to inspection and adoption approach which greatly reduces both the development cost and time. For each iteration there was an output, a component of the proposed solution that was evaluated against the requirements.

Agile development methodology was followed iteratively through the following steps:

- Requirements analysis
- Architecture and design (feature design)
- Development and evaluate of the feature.

The steps were carried out as a continuous and iterative process as defined in the agile development methodology.

3.3 Methodology Requirement Analysis

Looking at the general domain of collaborative learning and learner support services this research seeks ways to provide and meet the needs of a new generation of learners namely the digital native learner**.** The digital native leaner's has special characteristic that makes them unique and are forcing educationists to change their ways of content delivery. The characteristics of these learners and what this research is about are as follows:

- These are learners who prefer to be at least partially if not fully participative in the learning process.
- To them the web is not a delivery tool but a learning platform.
- To them learning is a continuous activity because they are constantly online.
- They utilize formal, informal and more so social means to learn.

As part of requirements gathering on the use of Web 2.0 tools to support collaborative learning the researcher did interview experts in the area of e-learning on they experience with Web 2.0(chats and online discussion boards) tools.

The experts were three curriculum developers and one system administrator at the Kenya Institute of Education. They have been using a learning Management System (LMS) which has Web 2.0 feature i.e. chats and discussion boards.

The LMS has been running two courses the institute's staff induction course and a Primary Teacher Orientation course for all primary teachers' country wide. All experts strongly agreed that Web 2 tools have enormous potential in the area of teaching and learning.

However they also raised a few challenges the current system faced which informed this study. First the learners who did not go through a face to face training on how to use and interact with the system took a lot of time before they were comfortably interacting with both the system and their peers.

Secondly the system was not able to handle lager numbers of users logged to the chat room at the same time.

Thirdly a significant number, 10% of those who registered did not start the course. This was because they were unable to use the system.

Another major challenge was accessibility. Some of the learners are in remote areas Where they could not access computers with internet connectivity easily.

It was not easy for both the learners and online tutors to go through the comments generated during the chat sessions because of the way they were organized. All these challenges informed the researcher in developing the prototype.

3.4 Architecture and Design

The following activities were carried out:

The first step was extract posts from a twitter account eclass2013.This was achieved through the use of an extraction script that was able to access the data from twitter.

The second step was the preprocessing stage so as to get a set of training and test data to train the classifier. Preprocessing involved cleaning the extracted data, by removing irrelevant features. This was done because some of these irrelevant features might hamper the performance of the classifier.

Third step was to train or learn classifier model a categorization technique was used, this is a systematic way of developing a categorization model from the data extracted from Twitter.

So as to develop the categorization model, a collection of input data sets will be used. The extracted data set will be divided into two i.e. training data and test data. Training data set was a collection of records whose class labels were already be known and used to build the categorization model. This model was then be applied to the test data set. Test data set was a collection of records whose class labels were also known but were used as inputs to build the categorization model, and did return accurate results.

The fourth step was to use the model to classify. The model developed will be used to predict the class label of new posts extracted from Twitter.

The last step was to develop a platform for displaying the grouped post to support collaborative learning

3.4.1 High level view of the classifier development process

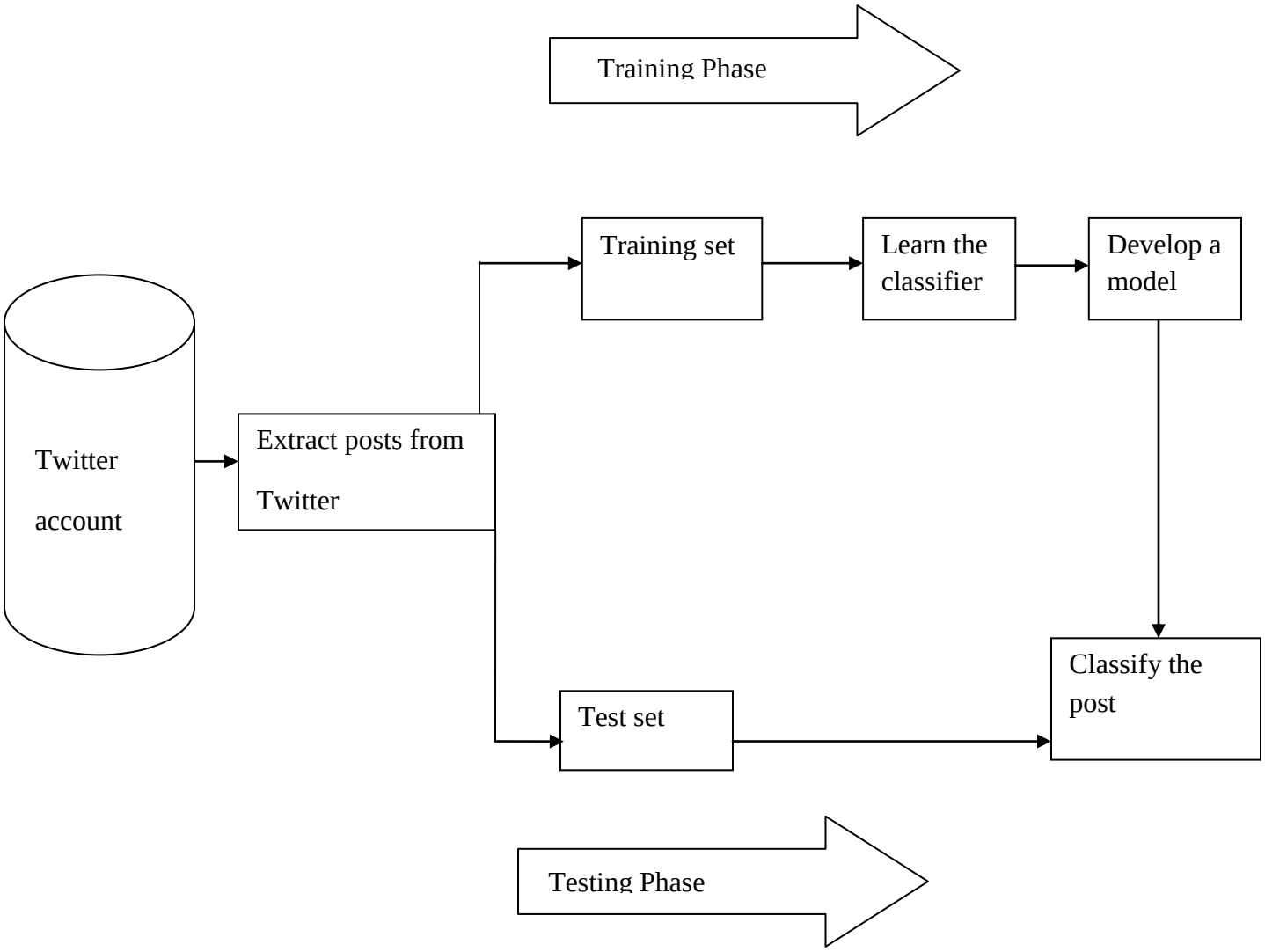

Fig 3 High level view of the classifier development process

3.5 Research Design

The researcher used descriptive analysis design for this study. Descriptive research design was used because it enables one to present data in a more meaningful way, which allows simpler interpretation of the data to be made.

The evaluation of e-learning systems is important for all the actors involved in their development and use. Teachers and students need to evaluate the benefits of using e-learning in comparison with the classical methods of learning" (Posea, Matu and Cristea, 2007).The results of an assessment process should provide information which can be used to determine whether or not intended outcomes are being achieved and how the e-learning material can be improved" (Falco and Soeiro, 2003).

3.6 Evaluation Instrument

To evaluate the impact of the system members of staff of the Kenya institute of Education were involved. The institute had fifteen members of staff who went through an online induction course. The online staff induction course is mandatory for any new member joining the institute. The system was installed on one of the servers in the institute and the members were trained on how to access it. The group used the system for online collaboration for a period of two weeks.

At the end of the two weeks the group was required to evaluate the system based on Usability and functionality.

3.6.1 Questionnaire

The main data collection instrument used was a questionnaire. The questions were based on an inventory of questions developed by the Presences research working group at the Technische University Eindhoven, Netherlands. The questionnaire had three sections; the first section had a set of questions related to participant's use of social media. This section collected background data on whether the participants had social media accounts, how often they used them and if the use them for educational purposes. The second section had a set of questions relating to the usability of the prototype. The third and final part of the questionnaire had questions relating to the functionality of the prototype.

A 5 – point likert type scale was adopted for the second and third sections of the questionnaire the first section of the questionnaire participants were required to respond with “Yes or No”.

3.6.2 Questionnaire Pretesting

This is a very important part of the questionnaire construction. It involves testing the questionnaire in conditions similar as possible to the research but not in order to report results but rather to check for glitches in wording of the questions and lack of clarity of instructions. Pretesting allowed the researcher to identify anything that could have impeded the instrument ability to collect data in an economical and systematic fashion. The researcher carried out a pilot study to pretest the validity of the questionnaire.

The pilot study used descriptive research design and was done with four participants, three learners and one education expert- a curriculum developer. The problems that were identified with the questions included confusion with the overall meaning of a question, as well as misinterpretation of individual terms. This allowed for revision and modification of the questionnaire thereby enhancing validity of the instrument .The data collected during this instrument pre-testing phase was not included as part of the data used in the actual study.

CHAPTER 4.0 SYSTEM DEVELOPMENT AND IMPLEMENTATION

4.1 The system design Process

The prototype development methodology used the iterative feature driven development (FDD) a variant of the agile software development methodology. The design process involved first the development of an overall model for the prototype (refer to Fig 1).The feature development methodology requires that a feature list be developed after the overall model is created hence the next step was identifying a set of features that are required in order to achieve the set objectives of this study. The overall model was broken down to several features based on the three objectives of the study as illustrated below:

NO	Objective	Feature
1	Develop a way for capturing data from a social media discussion forum.	Twitter extractor (extract comments from a twitter account)
2	Develop a text classifier that is able to classify the data from social media into groups based various categories.	A text cleaner A Naïve Bayes classifier
3	Develop a platform that is able to store the classified data for easy and quick retrieval.	A Mysql database to store the extracted and classified tweets A platform for supporting easy and quick retrieval of the classified tweets

4.2 Twitter Extractor

The source of data for this study is a twitter account. I have created a tweeter account "eclass2013" that we be used by learners to post comments or questions on the different subjects they are taking. The tweeter extractor script will then be able to extract these tweets and feed them to the classifier.

4.3 Data Pre-Processing

Many factors affect the success of Machine Learning (ML) on a given task. The representation and quality of the instance data is first and foremost. If there is much irrelevant and redundant information present or noisy and unreliable data, then knowledge discovery during the training phase is more difficult. It is well known that data preparation and filtering steps take considerable amount of processing time in ML problems. Data pre-processing includes data cleaning, normalization, transformation, feature extraction and selection, etc. (S. B. Kotsiantis, et al. 2006). The raw data extracted from twitter is not suitable for use as input for the classifier. Data pre-processing is an important step in the data mining process. The phrase "garbage in, garbage out" is particularly applicable to data mining and machine learning projects. Data-gathering methods are often loosely controlled, resulting in out-of-range values (e.g., Income: −100), impossible data combinations (e.g., Gender: Male, Pregnant: Yes), missing values, etc. Analyzing data that has not been carefully screened for such problems can produce misleading results. Thus, the representation and quality of data is first and foremost before running an analysis.(Pyle, D., 1999.)

If there is much irrelevant and redundant information present or noisy and unreliable data, then knowledge discovery during the training phase is more difficult. Data preparation and filtering steps can take considerable amount of processing time. Data pre-processing includes cleaning, normalization, transformation, feature extraction and selection, etc. The product of data pre-processing is the final training set. Kotsiantis et al. (2006) present a well-known algorithm for each step of data pre-processing.

At this stage of preprocessing of data for this study i have used I script that will able to remove all the elements that are not relevant as input for the classifier these are:

URLs - I don't intend to follow the short urls and determine the content of the site, so we can eliminate all of these URLs via regular expression matching or replace with generic word URL.

Punctuations and additional white spaces - remove punctuation at the start and ending of the tweets. E.g: ' the day is beautiful! ' replaced with 'the day is beautiful'. It is also helpful to replace multiple whitespaces with a single whitespace

Lower Case - Convert the tweets to lower case.

The script is also able to remove the following images and imagemaps, leading whitespace and underscore. Finally the script Removes html comments.

4.4 Classification

The data analysis task classification is where a model or classifier is constructed to predict categories or labels based on a set of attributes. There are two main approaches to classification: supervised learning and unsupervised learning. In supervised learning, the categories that data is assigned to are known before computation. So they are being used in order to 'learn' the parameters that are really significant for those Clusters. However, in unsupervised learning datasets are assigned to segments, without the clusters being known. This research used a classification algorithm Naïve Bayes Classifier to predict the label for a given input sentence. Naïve Bayes classifier is a supervised learning classification technique.

4.5 Python

The task of building the classify was carried out using Python as a programming language The Python programming language is a dynamically-typed, object-oriented interpreted language. Although, its primary strength lies in the ease with which it allows a programmer to rapidly prototype a project, its powerful and mature set of standard libraries make it a great fit for large-scale production-level software engineering projects as well. Python has a very shallow learning curve and is an excellent online learning resource (semanticbible 2008).

4.6 Natural Language Toolkit (NLTK)

Although Python already has most of the functionality needed to perform simple NLP tasks, it's still not powerful enough for most standard NLP tasks. This is where the Natural Language Toolkit (NLTK) comes in. NLTK is a collection of modules and corpora, released under an open source license, that allows students to learn and conduct research in NLP. The most important advantage of using NLTK is that it is entirely self-contained. Not only does it provide convenient functions and wrappers that can be used as building blocks for common NLP tasks, but it also provides raw and pre-processed versions of standard corpora used in NLP literature and courses. According to bird et al is of the view that NLTK was designed with four primary goals in mind; Simplicity, Consistency, extensibility, Modularity.

4.7 Naive Bayes Classifier

Naïve Bayes classification is based on Bayes theorem. A simple Bayes classification namely the Naïve classifier is comparable in performance with decision tree and neural network classifiers. Naïve Bayes classifiers have also exhibited high accuracy and speed when applied to large database. Naïve Bayes classifier assumes that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is called *class conditional independence*. It is made to simplify the computations involved and, in this sense, is considered "naïve". While applying Naïve Bayes classifier to classify text, each word position in a document is defined as an attribute and the value of that attribute to be the word found in that position.

Naïve Bayes is formalized as the product of the prior probability which is based on previous experience and the likelihood of a given attribute being in a given class, this forms the posterior probability.

To classify an unlabeled example it is just a matter of using the prior probabilities of a given category and multiplying them together. The category which produced the highest probability would be the label/classification for the unlabeled example. Only the words found in the unlabeled example would be looked up in the feature vector.

The following equation would be used to classify an unlabeled example. Given a document d and a class c, where the goal is to predict the probability that the document d belongs to class c.

$$P(c/d) = \text{argmax } (P(d/c).P(c))$$

4.8 Filtering Data for the classifier

The following was carried out as part of text preprocessing to ensure that we obtain clean data as input for the classifier:

Tokenization: The process of splitting a sentence into its constituent tokens. For segmented languages such as English, the existence of whitespace makes tokenization relatively easier and uninteresting.

Stemmer: Stemming is another common preprocessing step. In order to reduce the size of the initial feature set, we remove misspelled or words with the same stem. A stemmer (an algorithm which performs stemming), removes words with the same stem and keeps the stem or the most common of them as feature. For example, the words “train”, “training”, “trainer” and “trains” can be replaced with “train”.

Removal of Stop words - a, is, the, with etc. The full list of stop words can be found at Stop Word List. These words don't add value to the classifier hence they were removed.

Repeating letters - if you look at the tweets, sometimes people repeat letters to stress the emotion. E.g. hunggrryyy, huuuuuuungry for 'hungry'. We can look for two or more repetitive letters in words and replace them by 2 of the same.

Punctuation - we can remove punctuation such as comma, single/double quote, question marks at the start and end of each word. E.g. beautiful!!!!!! Replaced with beautiful

4.9 Feature selection and extraction

Feature selection is the process of selecting a subset of the terms occurring in the training set and using only this subset as features in text classification. Feature selection serves two main purposes. First, it makes training and applying a classifier more efficient by decreasing the size of the effective vocabulary. This is of particular importance for classifiers that, unlike Naïve Bayes, are expensive to train. Second, feature selection often increases classification accuracy by eliminating noise features. A noise feature is one that, when added to the document representation, increases the classification error on new data.

In order to perform learning it is necessary to extract clues from the text that may lead to correct classification (Yessenov and Misailovic 2009).clues about the original data are usually stored in the form of a feature vector(Fn=fi,f2...fn).Each coordinate of a feature vector represents one clue also called a feature fi of the original text the value of the coordinate may be a binary value indicating the presence or absence of the feature. In most machine learning approaches features in a vector are considered statistically independent from each other. The selection of features, strongly influences the subsequent learning. The goal selecting good features is to capture the desired properties of the training data in numerical form. The study endeavored to select the properties of the training data that are relevant for the classifier but it should be noted that exact algorithm for finding the best feature does not exist.

4.10 Feature Vector

Feature vector is the most important concept in implementing a classifier. A good feature vector directly determines how successful your classifier will be. The feature vector is used to build a model which the classifier learns from the training data and further can be used to classify previously unseen data.

In this study we used the presence of key words that appear in the training data as features. In the training data, consisting of label documents, we split each document into words this are kwon as unigram and added them most defining words to the feature vector which was used to determine the class label of a given tweet. Some of the words did not have any say in indicating the class label of a tweet and hence were filtered out.

4.11 TF-IDF Weighting scheme

The weighting scheme used in classifiers feature vector was tf–idf term frequency–inverse document frequency. This is a numerical statistic which reflects how important a word is to a document in a collection or corpus. It is often used as a weighting factor in information retrieval and text mining. The tf-idf value increases proportionally to the number of times a word appears in the document, but is offset by the frequency of the word in the corpus, which helps to control for the fact that some words are generally more common than others.

This the best known weighting schemes in information retrieval. It combine two concepts of information retrieval that's term frequency (tf) and invers document frequency (idf).Term frequency tf(t,d) refers to the number of times a term t occurs in a document d. Also note that the relevance of tf in classifying a document does not increase proportionally with frequency.ie if a term occurs 30 times in a given document and another word occurs once in the same document this does not necessarily mean that the first term will have more weight when it comes to classification. Invers Document frequency (idf) is premised on the idea that rare terms are more informative than frequent terms, hence they get a high weight. Tf-idf weight of a term is the product of its tf weight and its idf weight.

4.12 Implementation of the text classifier in a web based application

After building training and testing the classifier the next step is to implement it in a web based environment so that it can be used for collaborative learning.

A web based application was develop using python and Django a web development framework.

This application provided an interface for the users/learners to access the classified tweets and be able to respond to the comments or question raised though the twitter account eclass2013. In doing this the platform is supporting the "process of meaning making in the context of a joint activity", which is the definition of collaborative learning.

CHAPTER 5.0 EVALUATION RESULTS AND ACHIEVEMENTS

This chapter provides the results of the evaluation of the prototype as carried out in the Kenya Institute of Education on a staff induction course.

The purpose of this study was to examine how Web 2.0 tools like social media can be used to support the practice of meaning making among learners i.e. collaborative learning. The system was tested against the set overall goal of the study which was to develop a technique that will be able to extract posts from a social media account classifier and store them to support collaborative learning.

Data on actual learner participation in online discussion was collected over a period of two weeks. A satisfaction survey was then conducted at the end of the evaluation period through a questionnaire.

The questionnaire had three sections. The first looked at whether the learners had social media accounts and if they used these accounts for learning purposes. This section of the questionnaire was given to the participants before they started interacting with the prototype. The second section related to usability of the prototype; these were set to evaluate how easy it was to use the prototype with minimal or no training. The third section considered the functionality of the prototype as spelled out by the research objectives that is, was the system able to support collaborative learning through the developed prototype.

5.1 Initial study

This was conducted in order to gain some background data on the intended responses and also to gain more data from experts on what platforms they were currently using to facilitate collaborative learning and the challenges they are facing. The researcher achieved this through the use of two data collection instruments: interviews and a questionnaire.

The questionnaire required participants to provide background data on whether they had social media accounts, and if they did whether they used them for educational purposes. All fifteen participants had social media accounts, and twelve of them were frequent users; they visited their accounts at least once a day.

The participants were also required to indicate if they belonged to any social media group, eleven of them indicated that they belonged to at least one social media group but only 30% of them, which is five learners had used social media for learning. Those who used social media for education purposes said they found it easy and fast to link up with their peers through social media; they were comfortable using social media and did not require training on how it works.

The 70% which is ten learners, that had not used social media for educational purposes indicated that the way comments or post are organized within social media accounts as one of the reasons they are not able to use social media in learning and that social media lacks a true system for organizing information. The graph below illustrates how the learners responded to the first section of the questionnaire.

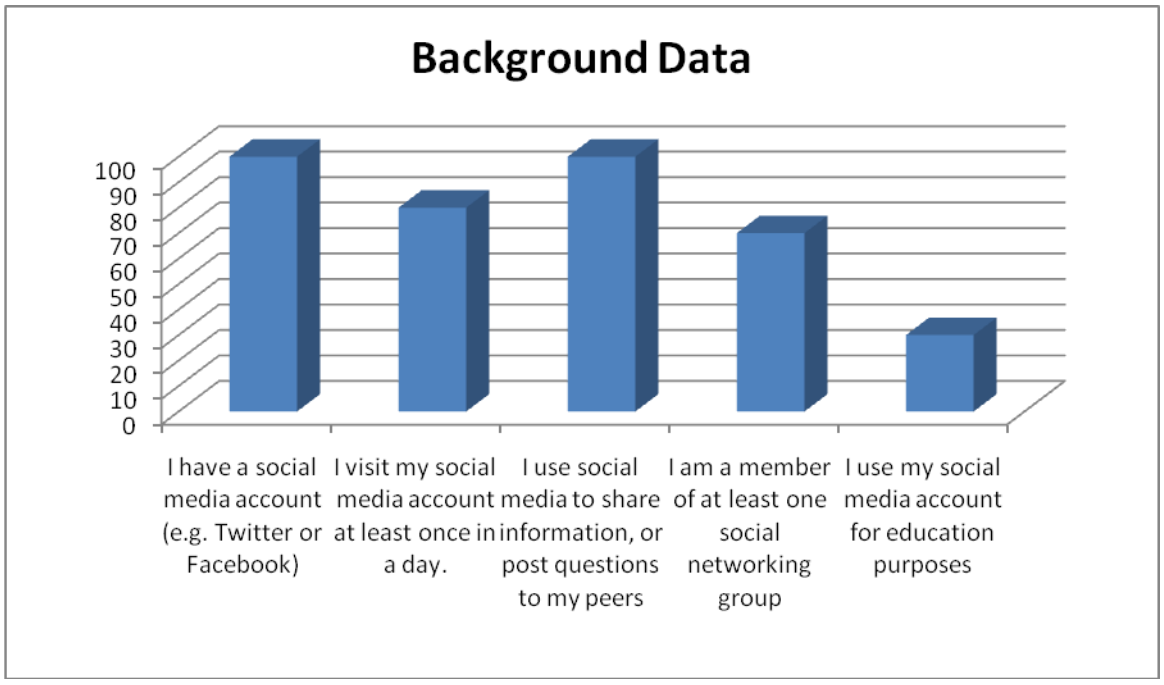

Fig 4, Background data on learners use of social media.

5.1 Evaluation of the prototype on Usability

The learners were required to rate the prototype based on ease of use. The learners were consistently positive (Appendix 7.2 -2) in their responses to the questions about the ease of use of the prototype. The majority of the participants evaluated indicated that the system was easy to use. 80% which is twelve learners strongly agreed that the prototype was easy to navigate and that all the important links were visible, since it was similar to social media accounts which they have been using. Thirteen learners, that is 86%, strongly agreed that the comments on the prototype were better organized than on Twitter, since they modules comments were grouped separately hence they were easily accessible. Twelve leaner's indicated that they would recommend others to use the prototype.

5.2 Evaluation of the prototype on Functionality.

To determine if the prototype achieved the set overall goal of the research, the questionnaire contained a set of question on the functionality of the prototype, i.e. did it support the process of meaning making in a joint activity such as collaborative learning (Appendix 7.2- 3). On this 80%, which is twelve learners strongly agreed that the prototype facilitated exchange of information with others and that they were comfortable interacting with their peers. Thirteen learners, which is 86% strongly agreed that they felt motivated to explore content related to the comments posted by others and that the online discussion enabled them to appreciate different view raised by their peers. Nine learners which is 60% of the participants strongly agreed that online interaction were valuable and 80 % that is twelve learners also strongly agreed that they learnt more about the Institute and within a shorter period of time than they would have on their own. 86% which is thirteen learners acknowledged that going through the classified comments on the prototype helped them learn a lot. Twelve learners indicated that they would continue using the prototype.

5.3 Discussion

This section discusses the findings of the research study in relation to objectives.

5.3.1 Objective 1

The first objective was to “Develop a way for capturing data from a social media discussion forum”.

This was achieved through the use of a script. The script was able to access and extract tweets posted by participants who were taking a staff induction course at the Kenya institute of Education. The script was able to pass these tweets to the classifier as input.

5.3.2 Objective 2

The second objective was to “Develop a text classifier that is able to classify the data from a social media account into groups based on various categories”.

The first step in achieving this was to develop a Naïve Bayes classifier and train it using course content for the four modules being offered as part of staff induction course at KIE.

Then tweets extracted from the eclass2013 account were fed as input to the classifier and the output was classified tweets according to the four modules that the classifier was trained on. The last step was developing a MySQL database to store the classified comments.

5.3.3 Objective 3

The last objective was to “Develop a platform that is able to store the classified tweets for easy and quick retrieval”. This was achieved through a web based interface that was able to display classified tweets from the MySQL database based. The platform displayed the tweets based on the four modules of the staff induction course.

This platform achieved the overall goal of the study which was to provide an enabling environment for “meaning making in a joint activity” that is a platform that is able to facilitate collaborative learning.

5.4 Achievements

The results discussed above indicate that the prototype was able to achieve the set goal of the study. The study was able to address the limitation of social media and LMS features when used as tool to support collaborative learning. These limitations were the same challenges that the researcher identified as being faced by institutions offering online courses through LMS and social media which require collaboration among learners through online chats and discussions.

The prototype was able to address the limitations of the social media accounts as tools that support collaborative learning, as well as challenges he identified at the KIE which has been running online courses for the last four years.

These challenges included lack of a system for organizing the post being generated by learners as they chat i.e. they do not allow for the knowledge generated to be classified, this was addressed by the classifier.

According to the experts interviewed by the researcher, chat and discussion boards provided by LMS out there cannot handle large numbers of learners logged in at the same time. The study was able to address this by letting learners use a social media account which was more robust. Letting the learners use social media accounts also solved another challenged raised by the experts, that is the issue of training. No training was required since all the learners already had social media accounts and were comfortable using them to interact with their friends. With regard to the prototype minimal training was required to show the learners how to access the various tweets as classified under the four modules they were taking.

The prototype was also able to improve of learning participation because leaner’s could use their mobile phones to follow the discussion on tweeter as opposed to the use of LMS where learners had to have a computer that has internet connectivity.

CHAPTER 6.0 CONCLUSION AND RECOMMENDATIONS

6.1 Introduction

This chapter covers the conclusion and also points out areas that require further investigation in this field of online collaborative learning.

6.2 Conclusion

The purpose of the study was to develop a framework to support collaborative learning using social media. The research started by pointing out that social media is part of emerging technologies known as Web 2.0. It has enormous potential in the area of teaching and learning. The possibilities are endless when it comes to the educational potential of social media (Terry Friedman,2007)

The study also showed that this potential is not being utilized currently and that the learning institutions that are using these technologies they not able to utilize them fully.It was also pointed out that although Web 2.0 platforms are technologically rich, adaptable and capable of supporting collaborative learning, they have some limitations in the implementation of learning experiences. These limitations of social media as a tool for supporting collaborative learning were the focus of this study.

The researcher developed a prototype that provides an enabling environment for online collaborative learning. The prototype was aimed at addressing the limitations of social media as a tool for supporting collaborative learning, hence taking full advantage of the potential that Web 2.0 has within the field of teaching and learning. The prototype was able to extract tweets from a social media account and then pass them to a classifier. The classifier grouped the tweets into different categories based on various modules that the learners were taking. Finally the prototype was able to store the tweets in a database, which was linked to a web based platform for the learners to interact with.

The prototype was able to address the limitations that social media has as tool for supporting collaborative learning; these included issues of storing the large amounts of knowledge being generated by learners through these sites and providing a way of categorizing comments generated with a social media account.

6.3 Challenges

The researchers encountered the following challenges in the course of the study.

- A lack of standard framework for evaluating online collaborative learning frameworks.
- Upgrading of the twitter system as the study continued.

6.4 Suggestions for further research

This study limited itself to the classification of English statements (tweets) only. Therefore there is need to explore how the classifier can be modified to be able to extract and classify statements tweets in other languages like Kiswahili.

The study focused only on developing an enabling environment for learners to interact and be able to capture and store the knowledge generated through these interaction. However, how interactions affects learning and the relationship between the two is a complex pedagogical phenomenon that requires further investigation.

Reference

Anderson,Paul.2007.What is Web 2.0? Ideas,technologies and implications for education. JISC reports.http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf

Barbara Leigh Smith and Jean T. MacGregor (1992). What is Collaborative Learning? Washington Center for Improving the Quality of Undergraduate Education.

Bartolomé, Antonio (2008). Web 2.0 and New Learning Paradigms.eLearning Papers Nº 8. ISSN: 1887-1542 www.elearningpapers.eu

Berners-Lee, Tim 1999.Weaving the Web. Orion Business Books British Journal of Educational Technology, Vol. 24 No. 3 1993

Brook, C. and Oliver, R. (2003). Online learning communities: Investigating a design framework. Australian Journal of Educational Technology, 19(2), 139-160. http://www.ascilite.org.au/ajet/ajet19/brook.html

Brown, J.S., Collins, A. and Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32-41. HTML

Caverly, D.C., & MacDonald, L. (2002).Online learning communities.Journal of Developmental Education, 25 (3), 36-37.

Downes, Stephen (2005). ELearning 2.0.eLearn Magazine.

http://www.elearnmag.org/subpage.cfm?section=articles&article=29-1

Drs. Erik P.W.M. van Veenendaal CISA (2006) Improve Quality Services BV / Eindhoven University of Technology

Ellison, N. B.; Steinfield, C.; Lampe , C. (2007). "The Benefits of Facebook 'Friends': Social capital and college students' use of online social network sites". Journal of Computer-Mediated Communication.Vol.12, pages 1143-1168.

FrancescLlorensCerdà and NeusCapdeferroPlanasneus (2011) Facebook's Potential for Collaborative e-Learning.

Gerlach, J. M. (1994). "Is this collaboration?" In Bosworth, K. and Hamilton, S. J. (Eds.), Collaborative Learning: Underlying Processes and Effective Techniques, New Directions for Teaching and Learning No. 59.

Hara, N.; Bonk, C. J.; Angeli, C. (2000). "Content analysis of online discussion in an applied educational psychology".InstructionalScience.Vol.28, pages 115-152.

Hargis, J. (2005). Collaboration, community and project-based learning--does it still work online? International Journal of Instructional Media, 32, (2), 157- 162.

ISO 9241-11 (1995), Ergonomic Requirements for office work with visual display terminals (VDT's) -Part 11 : Guidance on usability, International Organization of Standardization

Jonasen, D. H.; Mayes , J. T.; McAlese, R. (1993). "A manifesto for a constructivist approach to technology in higher education". In: t. duffy, d. jonassen, j. lowyck (eds.). Designing constructivist learning environments. Heidelberg, FRG: Springer-Verlag.

Kenya Facebook statistics (2012) Socialbakers.www.socialbakers.com/facebook-statistics/kenya

Kenya institute of curriculum development web site www.kicd.ac.ke

Kirakowsky, J. (1996). The software usability measurement inventory: Background and usage. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (eds.). Usability evaluation in industry. Taylor & Francis, London, 169-178.

Koschmann, T. (2002a). Dewey's contribution to the foundations of CSCL research. Paper presented at the Computer Support for Collaborative Learning (CSCL 2002), Boulder, CO.

LAVE, J.; WENGER, E. (1990). Situated Learning: Legitimate Peripheral Participation. Cambridge (United Kingdom): Cambridge University Press.

Prof. Dr. MatjažDebevc (2000),Usability testing of e-learning content as used in two learning management systems.

Lev Grossman 2006 Web 2.0 phenomenal in Time Magazine3

Jayson s. (2006)I Generation has no Off switch USA Today February 2010.

Measuring Impact of e-Learning on Quality and Learning Outcomes.(2003) Rita Falcão, Alfredo Soeiro.

Mashups, SOAP, and Services: welcome to Web Hybrid e-learning Application (2006) Gray and Carlos Oliveira. The e-learning Guild learning solution e-Magazine.

NMC. 2007. The Horizon report: 2007 edition. Austin, TX: The New Media Consortium. http://www.nmc.org/pdf/2007_Horizon_Report.pdf (May31,2008). Archived at http://www.webcitation.org/5YEmETRCx.

Omondi Oscar Odhiamb , 2009, Comparative study of the e-learning platforms used in Kenyan universities: case study of Jomo Kenyatta University of agriculture and technology and United States International University.

O'Reilly, Tim (2005). What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software. O'Reilly Network. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

O'Reilly, Tim (2006). Web 2.0 Compact Definition: Trying Again. O'Reilly Network. http://radar.oreilly.com/archives/2006/12/web-20-compact-definition-tryi.html.

Perret-Clermont, A. N. (1993). What is it that develops? Cognition and Instruction, 11, 197-205.

Porteous, M.; Kirakowsky, J. and Corbett, M. (1993).SUMI user handbook. Human Factors Research Group, University College Cork.

Quarterly sector statistics report 4th quarter (April to June 2010/2011) communication commission of Kenya.

University of South Florida (USF) http://www.usforacle.com/usf-teacher-uses-facebook-in-class-1.2191262

Voithofer, R. J.; Foley, A. (2007). "Digital dissonances: Structuring absences in national discourses on equity and educational technologies". Equity and Excellence in Education. Vol. 40, No 1, pages 14-25.

Vygotsky, L. (1930/1978). *Mind in society*. Cambridge, MA: Harvard University Press.

VladPosea, Stefan Trausan-Matu, ValentinCristea(2007) Online Evaluation of Collaborative Learning Platforms.

WENGER, E. (1998). "Communities of Practice.Learning as a social system". In: Systems Thinker. http://www.co-i-l.com/coil/knowledge-garden/cop/lss.shtml

World Internet statistics (2012) Miniwatts marketing group.www.internetworldstats.com/africa.html

Woods, R. &Ebersole, S. (2003) Social networking in the online classroom: Foundations of effective online learning. EJournal, 12-13 (1). Available at: http://www.ucalgary.ca/ejournal/archive/v12-13/v12-13n1Woods-browse.html

Sherman, L. W. (1995). A Postmodern, constructivist and cooperative pedagogy for teaching educational psychology, assisted by computer mediated communications. In Proceedings of CSCL 95' Conference.

Social learning handbook (2010) Jane Hart Center for learning and performance technologies

Sfard, A. (2008). Thinking as communicating: Human development, the growth of discourses and mathematizing. Cambridge, UK: Cambridge University Press.

Smith, B. L., and MacGregor, J. T. (1992). "What is collaborative learning?" In Goodsell, A. S.,

Stahl, G. (2006). *Group cognition: Computer support for buildingcollaborative knowledge*. Cambridge, MA: MIT Press. Available athttp://GerryStahl.net/mit/.

Maher, M. R., and Tinto, V. (Eds.), Collaborative Learning: A Sourcebook for Higher Education. National Center on Postsecondary Teaching, Learning, & Assessment, Syracuse University.

Terry Friedman,2007,Twittering in the classroom.

Yuen, A.H. (2003). Fostering learning communities in classrooms: A survey research of Hong Kong schools. Education Media International, 40, 153-162.

Appendix

Questionnaire

(1) Please indicate Yes or No for the following questions:

Give three reasons to support your answer for the last question.

(2) Please indicate your level of agreement with the following statements regarding the platform you have been using

1-Strongly Agree, 2 -Agree, 3-No Opinion, 4-Disagree, 5-Strongly Disagree

	1	**2**	**3**	**4**	**5**
The site was easy to navigate around when viewing the various categories of comments.	12	2	0	1	0
The important links were visible on the platform.	12	2	0	1	0
I manage to easily interact with the interface of the platform.	11	2	1	1	0
The comment displayed on the platform were better organized than on twitter	13	1	1	0	0
Would you recommend others to use it.	12	2	0	0	0

(3) Please indicate your level of agreement with the following statements regarding the platform you have been using

1-Strongly Agree, 2 -Agree, 3-No Opinion, 4-Disagree, 5-Strongly Disagree

	1	**2**	**3**	**4**	**5**
The system enabled me to exchange information with other participants.	11	3	0	1	0
I felt comfortable interacting with other course participants.	11	2	1	1	0
I felt motivated to explore content related to the questions raised by others.	13	1	1	0	0
Online discussions were valuable in helping me appreciate different perspectives.	9	3	1	2	0
I learnt a lot by going through the classified comments on the platform, than I would have learned working on my own	11	3	0	1	0
The others learn something from me.	13	1	1	0	0
I will continue using the application	12	2	1	0	0

Sample code

```
fromnltk.corpus import stopwords

fromnltk.stem import PorterStemmer

fromnltk.tokenize import WordPunctTokenizer

fromnltk.collocations import BigramCollocationFinder

fromnltk.metrics import BigramAssocMeasures

fromnltk.classify import NaiveBayesClassifier

fromnltk.classify.util import accuracy


fromclassify.training_files_db import get_categories

defextract_words(text):

   '''

   Here we are extracting features to use in our classifier.

   '''

stemmer = PorterStemmer()

tokenizer = WordPunctTokenizer()

tokens = tokenizer.tokenize(text)

bigram_finder = BigramCollocationFinder.from_words(tokens)

bigrams = bigram_finder.nbest(BigramAssocMeasures.chi_sq, 500)


forbigram_tuple in bigrams:

      x = "%s %s" % bigram_tuple

tokens.append(x)

result =  [stemmer.stem(x.lower()) for x in tokens if x not in stopwords.words('english') and len(x) > 1]
```

```
return result

defget_feature(word):
returndict([(word, True)])

defbag_of_words(words):
returndict([(word, True) for word in words])

defcreate_training_dict(text, sense):
    ''' returns a dict ready for a classifier's test method '''
tokens = extract_words(text)
return [(bag_of_words(tokens), sense)]

defmy_classifier(questions_list):
output_list = []
features_list = []
    # create our dict of training data
texts = {}
categories_from_db = get_categories()
ifcategories_from_db:
for cat in categories_from_db:
texts[cat[0]] = 'documents/' + cat[1]

else:
```

```
raiseNotImplementedError('No training data in database')

train_set = []

    # loop through each item, grab the text, tokenize it and create a training feature with it
for sense, file in texts.iteritems():
print "training %s " % sense
text = open(file, 'r').read()
features = extract_words(text)
train_set = train_set + [(get_feature(word), sense) for word in features]

classifier = NaiveBayesClassifier.train(train_set)
features_list[:] = classifier.most_informative_features(30)
for question in questions_list:

tokens = bag_of_words(extract_words(question))
decision = classifier.classify(tokens)
result = "%s --- %s" % (decision, question)
iflen(result) > 30:
output_list.append(result)

return (output_list, features_list)
```

PROJECT SCHEDULE A FRAMEWORK TO SUPPORT COLLABORATIVE LEARNING USING SOCIAL MEDIA			
	Duration in Weeks	**Time Line**	**Remarks**
Initial study	4	**01/02/2013 To 28/02/2013**	**Finished**
Identifying the gaps	2		
Defining the Problem	1		
Set the goal and objectives of the study	1		
Requirements specification	4	**1/03/2013 To 30/03/2013**	**Finished**
Proposal presentation	**1**		**Finished**
System Design and Development	4	**31/03/2013 To 28/5/2013**	**Finished**
Develop a tool for cleaning Data	1		**Finished**
Developing a classification model			
Training the model	1		**Finished**
Test the model			
Develop a tool for Extracting tweets from twitter	1		**Finished**
Develop a web based application to facilitate collaborative learning	1		**Finished**
Progress presentation	**1**	**4/07/2013**	**Finished**
Implementation and Evaluation	4		
Learner interacting with the system	3	**1/06/2013 To 1/07/2013**	**Finished**
Analysis of the results using student's t-Test	1		
Final Documentation	2		
Final presentation	1	**22/07/2013**	**Finished**