HEALTH LEARNING INFORMATION

[1]Mrs.A.K.Jaithunbi ,[2] V.Madhumitha , [3]K.Harika, [4]GanjiPrathyusha

[1]Associate Professor, [2,3,4] Student(B.E),[1,2,3,4]CSE Department,
[1]R.M.D Engineering College, Kavaraipettai, Chennai-601 206, Tamil Nadu, India.

ABSTRACT –Internet has demonstrated itself as a magnificent design over time. Today, billions of people have contact to internet and they promote by making proper use of that technology. This technology can be used in abundant different ways to aid our comfort and ease likewise Tagging is a machine learning technique which provide tags to the information that user can easily identify the related information user searching for but there is flaw in that method most technology uses manual tagging and semi-manual tagging which is time consuming and people must be expert in that domain has the ability to identify the question and tag them it is not possible in real time and high in cost. We propose an automatic tagging method using NLP which automatically tag question where user can get what information he is searching for which overcome earlier methods.

INTRODUCTION-

In this we are using Natural Language Processing (NLP) one of Machine Learning techniques. NLP analyzes understand the human language in a smart and useful way. Machine Learning (ML) is applications of artificial intelligence (AI) that provides systems the ability to automatically learn. The previous systems are build using manual and semi-auto tag which is not suitable in making daily updates of data and is not possible to generate requested information and mines the unwanted answers (information) to the user that are confusing and tags are mismatched wrong it makes difficult to observe the information. The existing problem can be resolved by using NLP. The Medical datasets are collected from the Medinet library for 700 domains.

Medinet library contains 700 medical domain files. if the requested file not found in Medinet it redirects to PDF box it will provide pdf files. related to the question. If the answer is not related to query or not clear, the question will forwarded to the Expert. who can clear the doubts and replay for the query. Then finally if the student want to analyse their knowledge in particular domain quiz will be conducted and providing them scores.

SCOPE OF THE PROJECT

In this project, automatic tagging using NLP provide consistent result to users. It save time of the user. The retrieval of information is easier to understand and cost is low.

RELATED WORKS-

In order to provide the information requested by user they used manual tagging and automatic tagging approach these two approach is not a convenient because manual tagging is time consuming and needs well manual attention likewise in semi automatic tagging the content has to be again processed. This two are time consuming and high cost so we are proposing automatic tagging using Natural Language Processing.

Some of the related works already done is as follows:

- Avigit K. Saha,Ripon K. Saha, Kevin A. Schneider published article on "A Discriminative Model Approach for Suggesting Tags Automatically for Stack Overflow Questions",2013.

  DESCRIPTION:The purpose of this paper is to mine data from millions of questions from Q&A site stack overflow and using a Discriminative Model Approach, we automatically suggest question tags to help a questioner choose appropriate tags for eliciting a response.

- Liqiang Nie, Mohammad Akbari published article on "A Joint Local-Global Approach for Medical Terminology",2014.

  DESCRIPTION:This paper presents a scheme to label question answer(QA) pairs by jointly utilizing local mining and global learning approaches. Local mining attempts to label individual QA pair by independently extracting medical concepts from the QA pair itself and mapping them to authenticated terminologies. Global learning works towards enhancing the local mining via collaboratively discovering missing key terminologies and keeping off the irrelevant terminologies by analyzing the social neighbors.

- Gauri Nalawade, Rekha Ramesh published article on "Design for Semi-Automatic Generation of Question Paper from A Semantically Tagged Distributed Question Repository",2016.

DESCRIPTION: This paper proposing a system design for semi-automatic generation of question paper from semantically tagged distributed question repository.

EXISTING SYSTEM-

- In existing system user search different information in internet to get the knowledge about the content in this environment there is possibility that user can't get the desired information or it's hard to find the required information or can't get the clear knowledge about it and doesn't give any alternative solution to the user these methods are done with manual tagging and semi automatic tagging.

- First manually tagging questions with

  knowledge units needs that the taggers be specialist in that subject a question bank usually contains a huge number of questions, which are updated constantly this makes manual tagging costly in terms of the time taken and related cost and Semi-automatic tagging analyzes content and returns tags that need to be more processed by users, making human help mandatory. . There is high chance that questions can be mismatched.

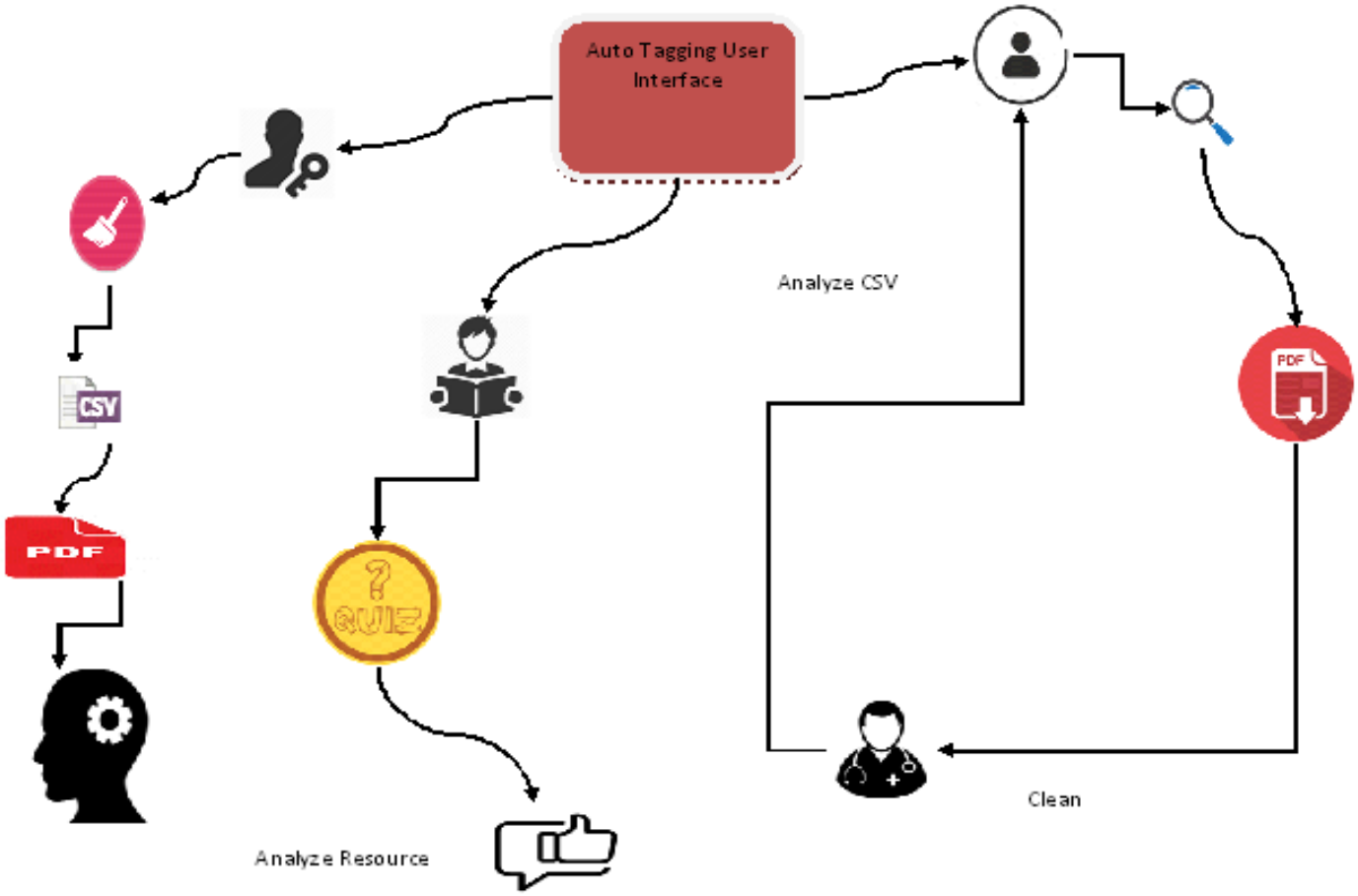
PROPOSED SYSTEM:

In this paper we propose an automatic tagging of questions by using NLP (Natural language processing) one of the machine learning techniques. Machine Learning (ML) is applications of artificial intelligence (AI) that provides systems the ability to automatically learn. The proposed system is designed to eliminate all the drawbacks of the existing system.

- The Medical datasets are collected from the Medinet library. The data's are stored in the CSV file formats for the later use. CSV file contains the questions and answers from the medical health domain and then the automatic tagging takes place of the question and Answer then the user search the query in AWS(Amazon web service) and the query is forwarded to the Medinet library to extract the user queried related answer.

- Medinet library contains 700 medical domain files which is stored in tc2011 api if the requested file not found in Medinet it redirects to PDF box it will provide pdf files related to the question The pdf files are extracted by using lucene indexing which provide fast retrieval of files If the answer is not related to query or not clear, the question will forwarded to the Expert who can clear the doubts and replay for the query. Finally providing the quiz for the student who want know their knowledge in the medical field and providing them domain score and overall score and feedback for the student.

ARCHITECTURE-



ALGORITHM-

A. NATURAL LANGUAGE PROCESSING

Natural language processing analyses understand the human language in smart and useful way.Natural language processing is one of the Machine learning techniques. Machine learning is applications of artificial intelligence (AI) that provides systems the ability to automatically learn. This algorithm automatically tags keyword with the suggestion that is already stored in the format of pdf file.

B. CLUSTERINGANDCLASSIFICATION

Classification is used to map the dataset and the query of the user. It is a supervised learning algorithm analyses training dataset. The dataset are trained and preprocessed in order to get the desired result .In clustering is used to cluster the dataset and user query.

USECASE DIAGRAM

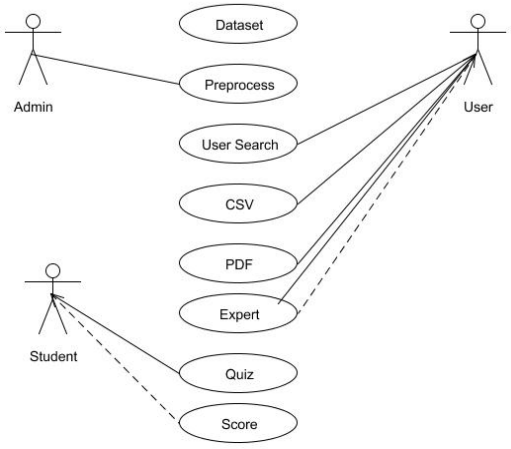


Fig-2: Flow of the process.

MODULES

- **ADMIN**:
  - In this module Admin has to register first and then have to provide admin details (Name and password), valid admin enters the page.
  - Admin is responsible for the whole operation admin work includes cleaning, adding the CSV files, analyzing resources, NLP (Natural language processing) and cleaning NLP.
- **AUTO TAGGING QUESTIONS:**
  - In this module the user has to register their details first after that user login using (Name and Password).
  - If user is valid then they can enter the page, then user chooses the categories of domain present in the page user clicks the categories it shows the list available domains, then user select any one domain based on that related information or question and answers to that domain will be displayed.
  - If the requested domain information or question answer is not available it provides PDFs related to that domain the provided answers of the question are not satisfied when user feels. User can ask question to the Expert (doctor) who can clearly answer the user query.
- **EXPERT ANSWERING PROCESS:**
  - In this module Expert (doctor) has to register first like (Name, Expert id, category of doctor) after that login using valid credentials.
  - Here expert id is unique id for the expert and category is to identify the expert domain.
  - Expert enters the page and he gets the notification from the user who asked the question after that Expert analyze

    the question and provides the related answer to the question the answer is forwarded to the user profile. User can view the answer any time.
- **STUDENT ASSESSMENT:**
  - student registers first after the login page is forwarded to the quiz page the student is asked to write quiz in medical domain.
  - Questions from medical domain is provided student has to answer it finally student gets the overall score of the test, domain score, and gets the feedback for the test which the student can improve knowledge in that domain.

CONCLUSION-

Thus the aim of the project is to give the health or disease related information to the user by receiving user query and also providing score for the student by conducting quiz to understand knowledge in particular domain. To achieve this we are implementing automatic tagging using natural language processing was implemented in this project.

REFERENCES:

[1] Zi-Ke Zhang, Tao Zhou, Yi-Cheng Zhang published article on"Tag Aware Recommender Systems: A State-of-the-art Survey"on February 2012.

[2] Avigit K. Saha, Ripon K. Saha," A Discriminative Model Approach for Suggesting Tags Automatically for Stack Overflow Questions",in proc. 10th work. Conf. Mining software Repositories. IEEE,2013.

[3] Andrzej Janusz, Sebastian Stawicki," Adaptive Learning for Improving Semantic Tagging of ScientificArticles",vol.2,pp.27-34,IEEE on 2014.

[4] Liqiang NIE, Yiliang ZHAO," Bridging the Vocabulary Gap between Health Seekers and HealthcareKnowledge",vol.27,pp.396-409,2015.

[5] Gauri Nalawade, Rekha Ramesh," Design for Semi-Automatic Generation of Question Paper from A Semantically Tagged Distributed Question Repository",vol.4,2016.

[6] Liqiang Nie, Mohammad Akbari," A Joint Local-Global Approach for Medical TerminologyAssignment",IEEE,2013,pp.73-76.

[7] Rekha Ramesh, Shitanshu Mishra," Semi-Automatic Generation of Metadata for Items in a Question Repository",IEEE,2014.

[8] Alex Krizhevsky, Ilya Sutskever,"ImageNet Classification with Deep Convolutional Neural Networks",pp.1097-1105,2012.

[9] Nitish Srivastava, Geoffrey Hinton," Dropout: A Simple Way to Prevent Neural Networks from Overfitting", 2014.

AUTHORS PROFILE

1)MADHUMITHA.V is a student currently studying final year B.E (CSE) in R.M.D Engineering College, Chennai, Tamil Nadu, India.

2)K.HARIKA is a student currently studying final year B.E (CSE) in R.M.D Engineering College, Chennai, Tamil Nadu, India.

3)GANJI PRATHYUSHA is a student currently studying final year B.E (CSE) in R.M.D Engineering College, Chennai, Tamil Nadu, India.