MediatorBot: A Mediator bot for supporting collaborative E-learning using an Intelligent Tutor System

Do Dung Vu
Supervisor: Prof. Sylvie Ratté

Département de génie logiciel et des TI

do-dung.vu.1@ens.etsmtl.ca

March 29, 2019

Overview

1. Introduction
2. Problem statement
3. Motivations
4. Objectives
5. Methodology
6. Work plan

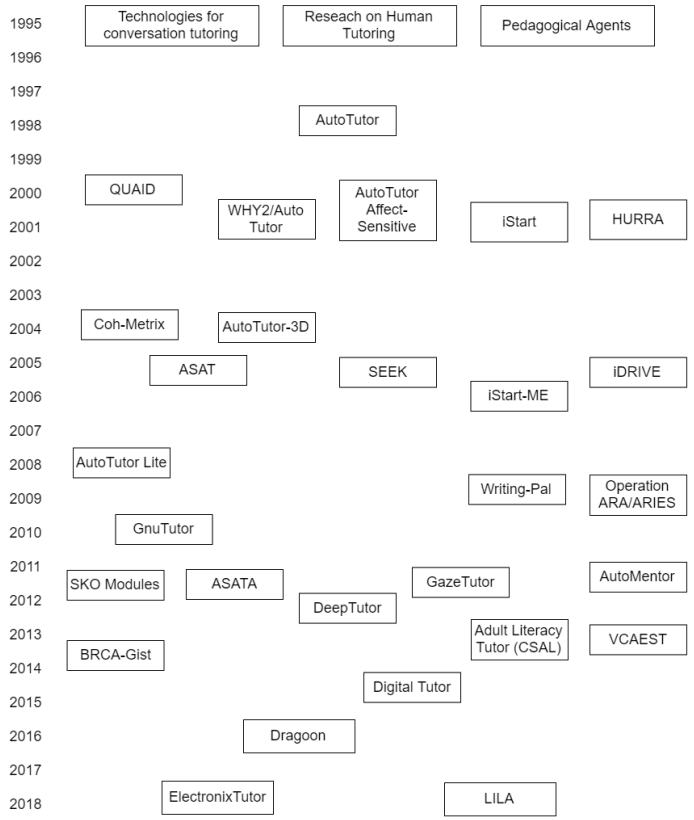
Introduction

Industry Factor	Description
Market size & forecast (Revenue)	USD 402.0 Million (2017) → USD 6,893.4 Million (2024)
Model trend (2017)	Learner Model – 65.33% Pedagogical Model – 22.76% Domain Model – %
Deployment trend (2017)	On-Premise – 84.63% Cloud – 15.37%
Technology trend (2017)	Machine Learning – 22.00% Deep Learning – 5.07% NLP – 68.28% Others – 4.65%
Application trend (2017)	Learning Platform & Virtual Facilitators– 56.37% Intelligent Tutoring System – 21.73% Smart Content – 15.06% Fraud & Risk Management – 2.69% Others – 4.15%
End-Use trend (2017)	Higher Education – 52.37% K-12 Education – 33.98% Corporate Training – 13.66%
Regional trend (2017)	North America – 60.16% Europe – 18.62% Asia Pacific – 16.48% LA – 1.22% MEA– 3.52%

AI in Education industry 3600 synopsis, 2013 – 2024

Source: AAAI, IEEE, WEF, IAAIL, Company Annual Reports, Hoovers, Primary Interviews, Global Market Insights

Introduction



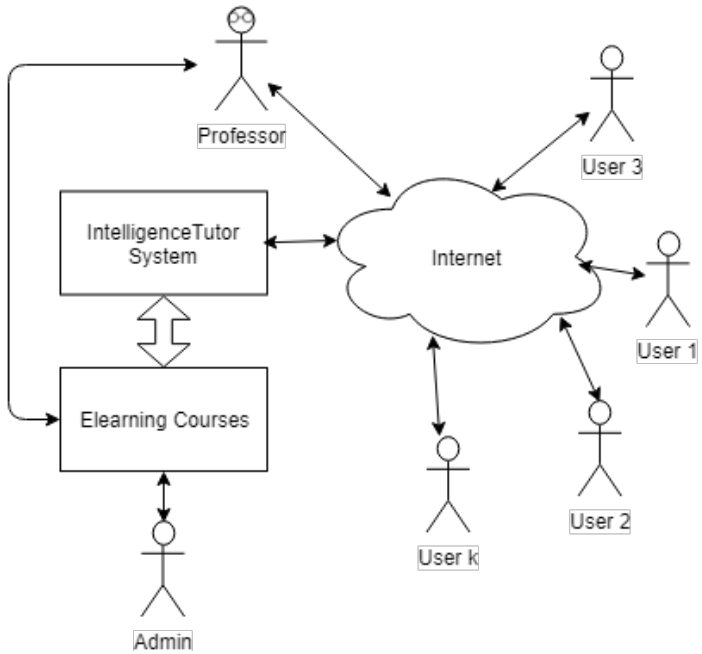
The time life of Intelligent Tutor System (ITS)

ITS main purposes:

- Help students construct expressions of material as answers to questions and solutions to challenging problems
- Ask questions that tap deep levels of reasoning and that involve collaboration
- Solve problems that involve deep argumentation

Context

- Online group learning on a given domain-specific (e.g., statistic)
- The group must discuss about a given topic or assignment (e.g., `https://mydalite.org/en/`)
- The Intelligence Tutor System (ITS) help Professor to monitor the progress of students and Admin to encourage their study



Literature survey

- An ITS monitors students' knowledge, skills, and psychological characteristics and response [1]
- Conversational agents have talking heads that speak, point, gesture, and exhibit facial expressions. [2]
- AutoTutor and its progenies [3] help students learn by holding a conversation in natural language
- Agent intervention aiming to link students' contributions to previously acquired knowledge can improve both individual and group studying when implemented in the context of a collaborative learning activity in higher education [4]

Reference

[1] Sottilare, R, Graesser, AC, Hu, X, Goldberg, B (Eds.) (2014). Design recommendations for intelligent tutoring systems: instructional management, (vol. 2). Orlando: Army Research Laboratory

[2] Johnson, WL, & Lester, JC. (2016). Face-to-face interaction with pedagogical agents, twenty years later. International Journal of Artificial Intelligence in Education, 26(1), 25–36.

[3] Graesser, AC. (2016). Conversations with AutoTutor help students learn. International Journal of Artificial Intelligence in Education, 26, 124–132

[4] Tegos, S., & Demetriadis, S. (2017). Conversational Agents Improve Peer Learning through Building on Prior Knowledge.Educational Technology & Society, 20(1), 99–111

Problem statement

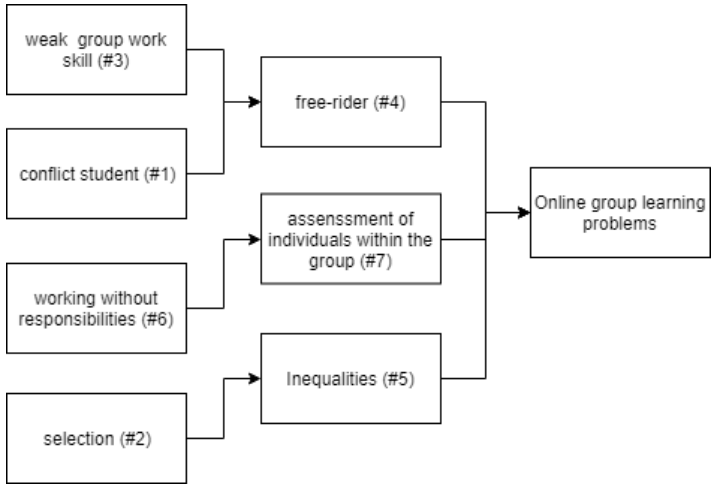
The seven most commonly in Online Learning found in the literature are the following[1]:

- (1): the student has conflicts works in the group
- (2): the selection of the groups is not good
- (3): the students don't have enough group-work skills
- (4): some students want to work alone or become the free-riders
- (5): the possible inequalities of student abilities appears in the group
- (6): some members do not commit to working in the group with their responsibilities
- (7): the assessment of individuals within the groups is not fair

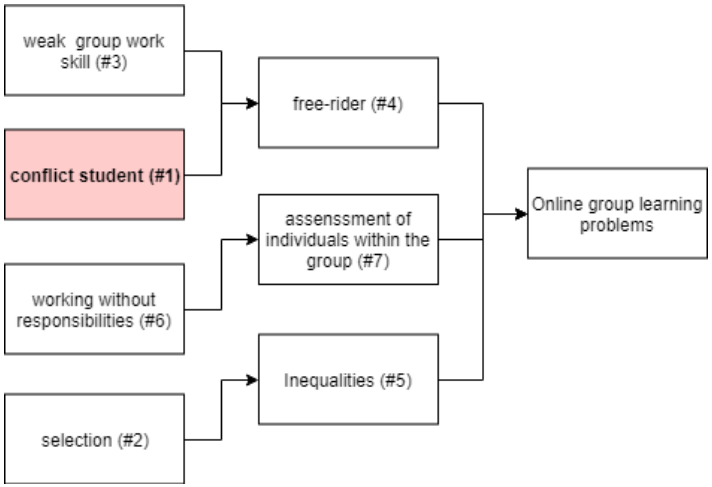
[1] Jianxia Du, Chuang Wang, Mingming Zhou, Jianzhong Xu, Xitao Fan & Saosan Lei (2018) Group trust, communication media, and interactivity: toward an integrated model of online collaborative learning, Interactive Learning Environments, 26:2, 273-286,

Problem statement

Most of these problems above of online group learning are inter-related



Problem statement



#3: solved by orientation training from admin
#6, #2: solved by professor
#1: solved by ITS system

→ in the ITS based on Dialogue System, there are other potential problems in the online group learning that have not been dealt with
→ We want to solve the problem of conflict students with the teamwork.

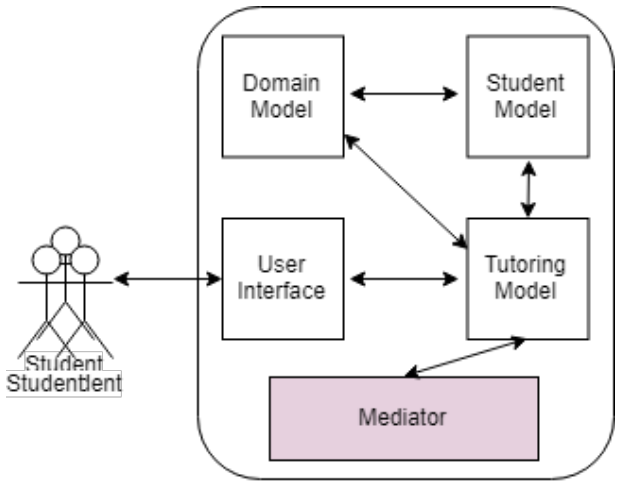


Figure: ITS with Mediator

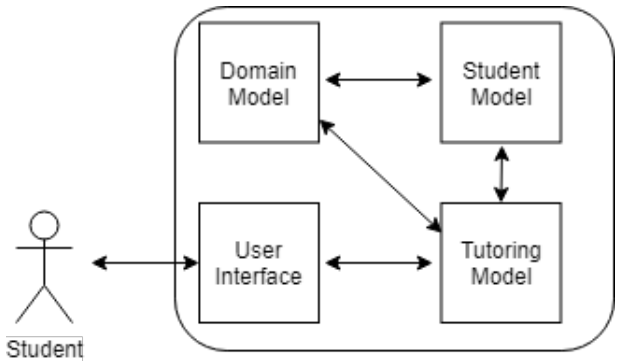
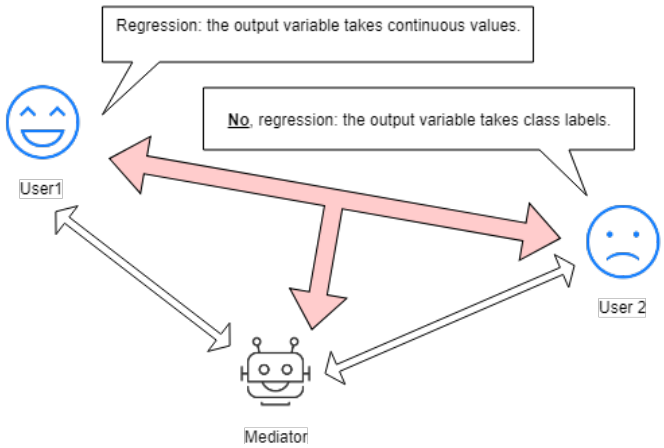


Figure: Original ITS

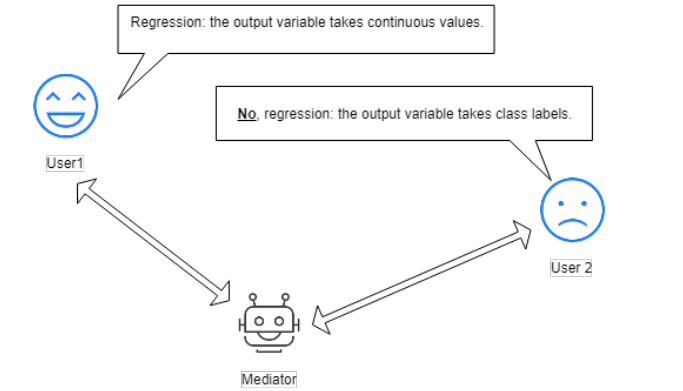
Source: DOI: 10.1109/KSE.2015.9

MediatorBot generates the hints, identifies the debated problem, the opportunities for intervention, and answers the related topic question of students to encourage the users to collaborate more effectively in the online group learning with low price in the specific-domain

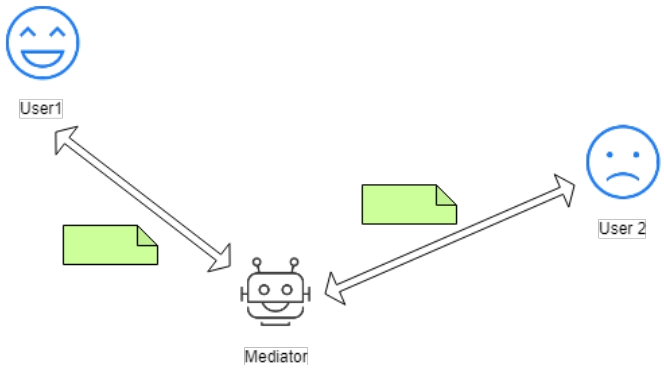
Problem statement



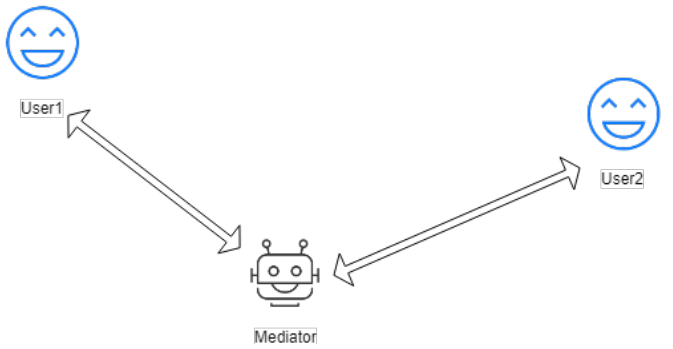
1



2

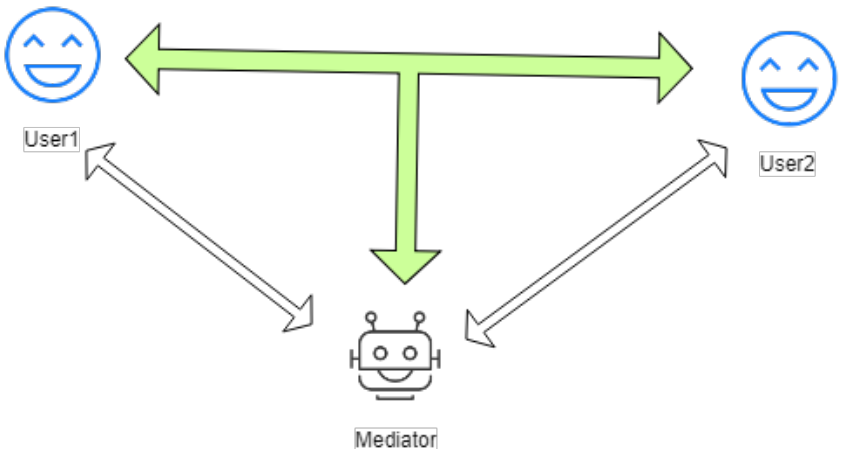


3



4

Motivations



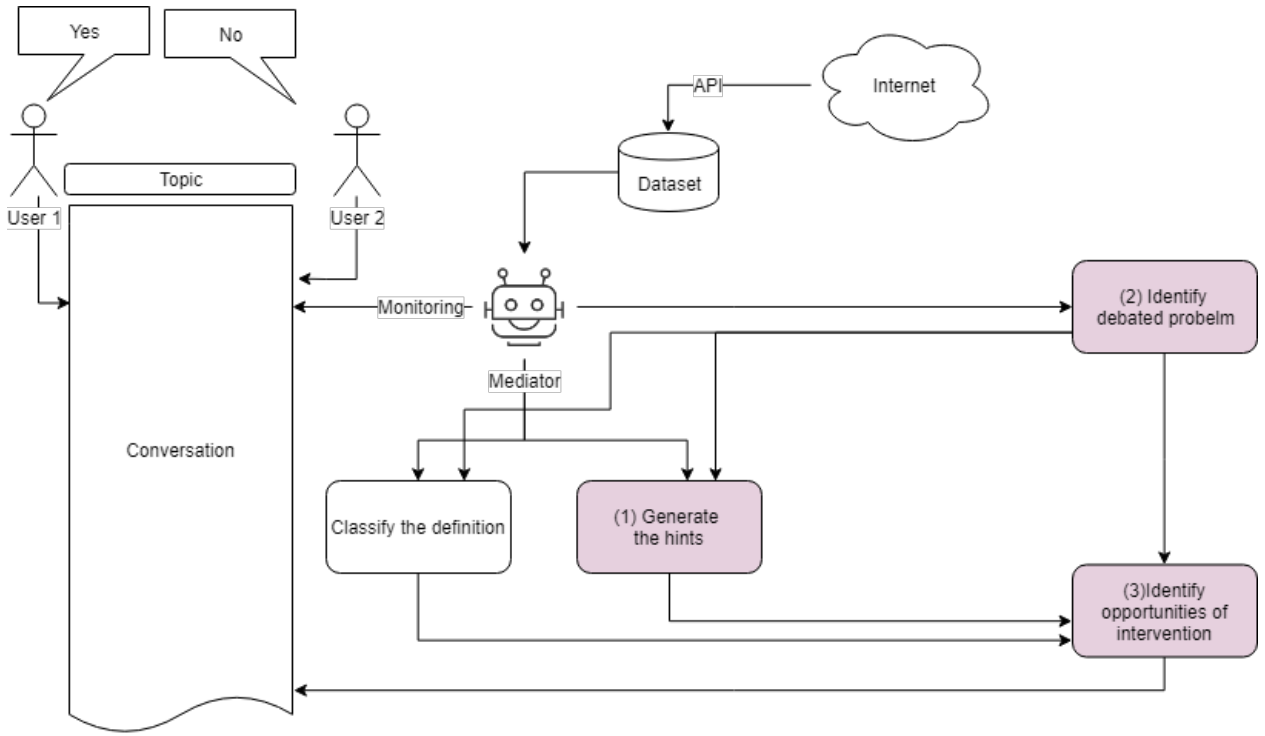
- Future state-of-the-art interventions with low price for intelligent tutor system
- Encourage group student online working
- Easily scalable

Objectives

Main objective: Propose a smart Mediator to support constructive discussion based on the Dialogue system:

- Generate hints to help users solve the topic or problem automatically
- Identify the debated problem
- Intervene in the conversation to resolve the conflict

Mediator system



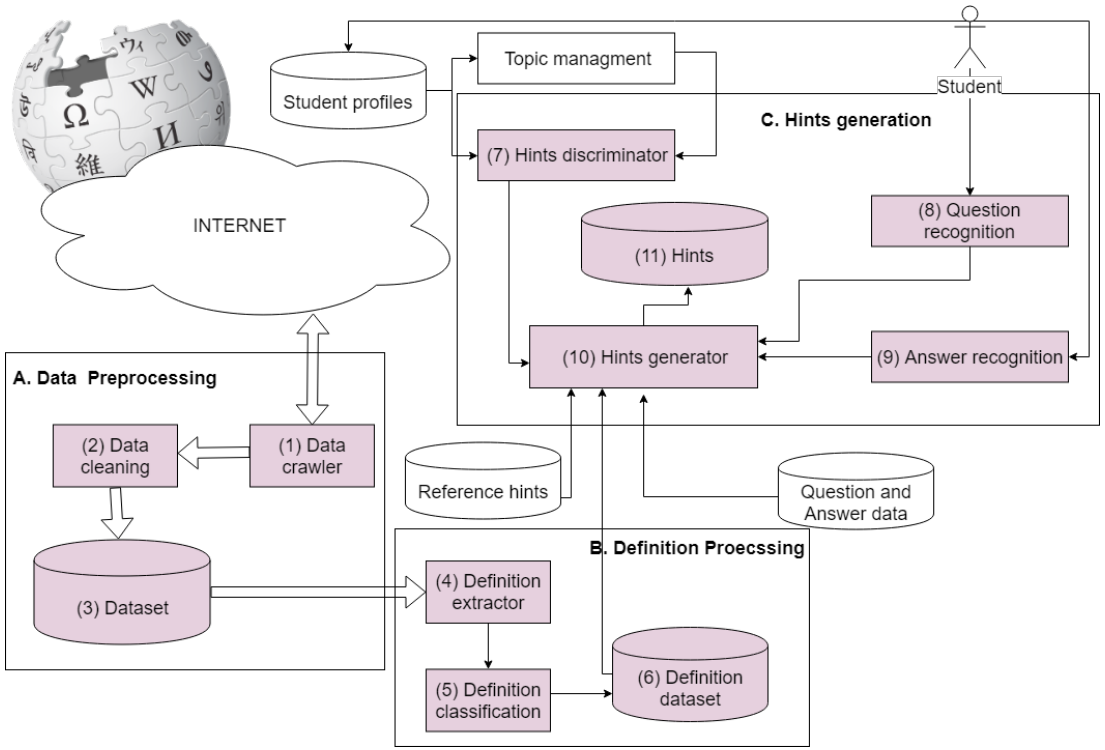
(1) Generate hints to help users solve the topic or problem automatically

(2) Identify the debated or clarify the problem

(3) Intervene in the conversation to clarify the problem: identify the opportunities for intervention, answer the related topic question of users

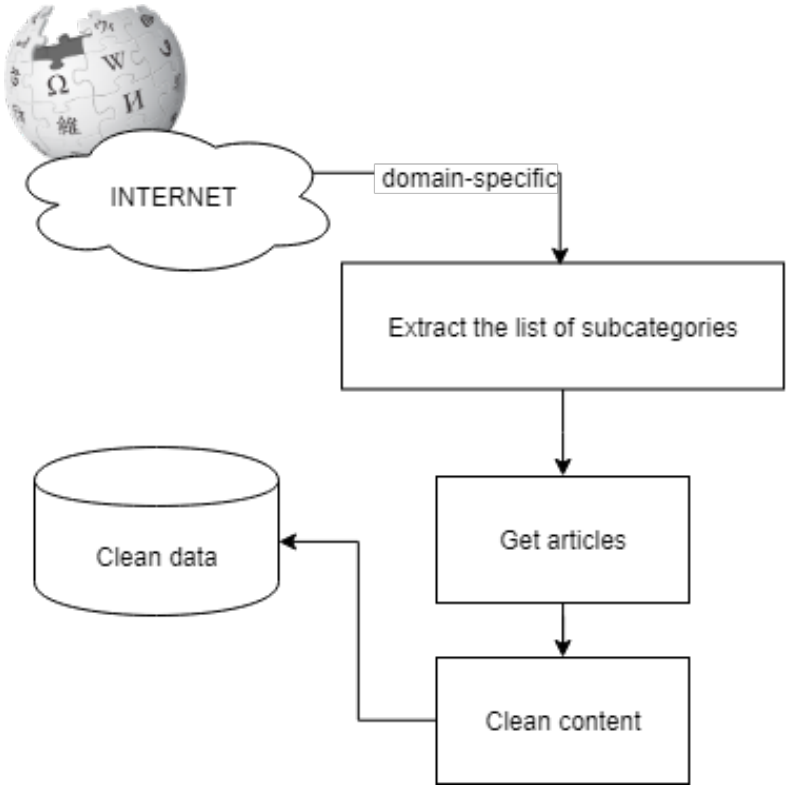
Objective 1— Structure

Generate hints to help users solve the topic or problem automatically



Objective 1 — Methodology — A. Data preprocessing

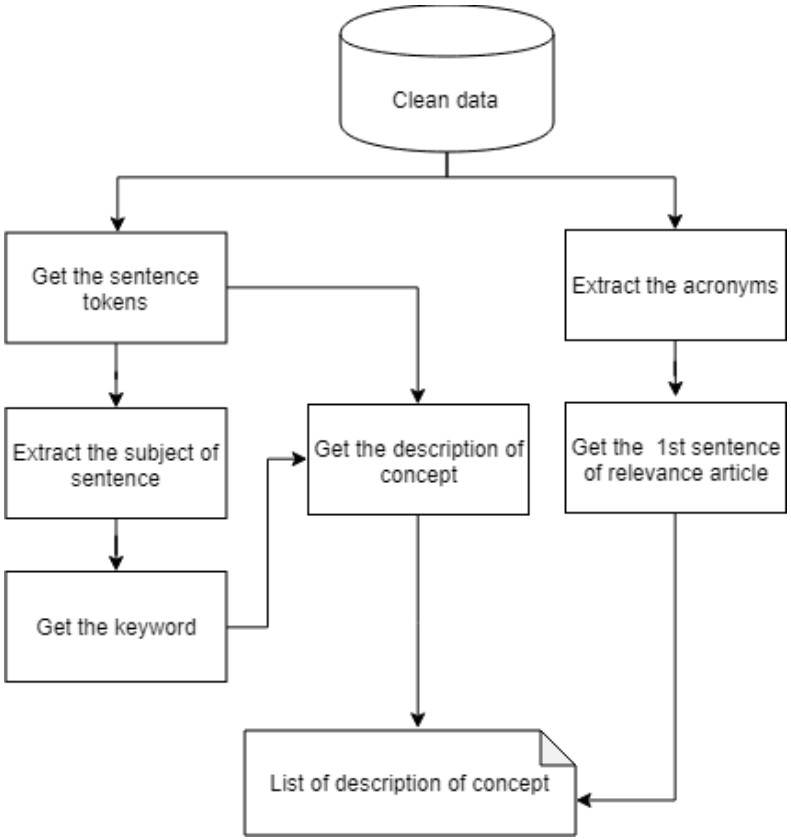
(1) Data crawler: crawling data from wikipedia with a given domain-specific (e.g., statistic)
(2) Data cleaning: clean the unicode, convert xml equation to latex equation, clean punctuation, split raw text to line by line sentence
(3) Dataset: save data to the tsv file with it fields: title, link, content



Objective 1 — Methodology — B. Definition processing

(4) Definition extractor: Extract the description of each concept
(5) Definition classification: classify type of definition based on supervisor algorithm (Good/Not Good)
→ Using the oversampling methodology to reweight the Good and Not Good samples (6) Definition dataset: definition with its' label

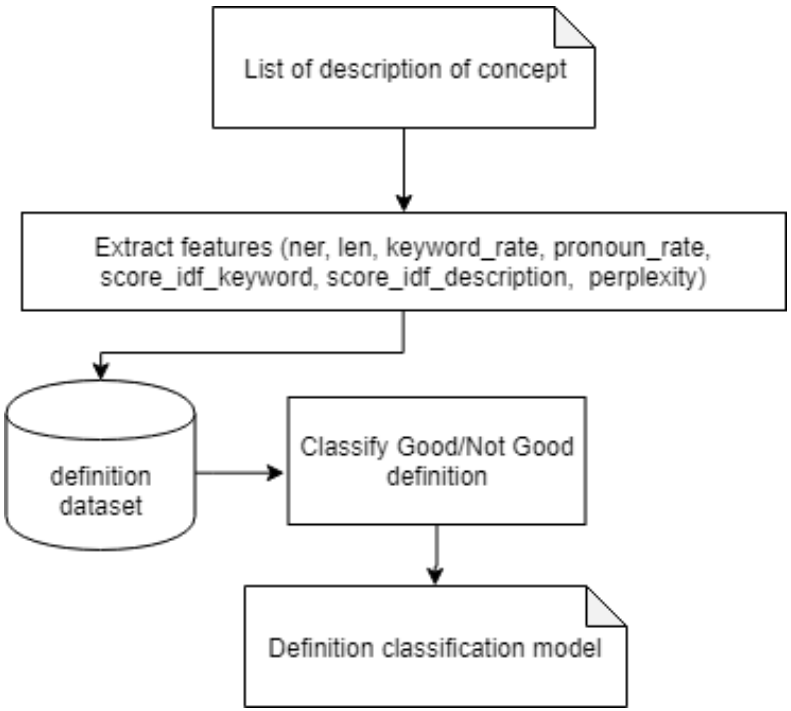
Objective 1 — Methodology — B. Definition processing — Definition extractor



- Split raw text dataset to the sentence tokens
- Extract the technical keyword acronyms
- Extract subject (noun phase) ← Get the keyword (concept)
- Filter the right keyword (concept)
- Get the description of concept
- Save the list (dict) description of conecept which is called dictionary of definition

Objective 1 — Methodology — B. Definition processing — Definition classification

- Extract the features ← score table
- Save the score table to the definition dataset
- Classify the p/n definition based on supervised learning algorithm
- Save the classification model

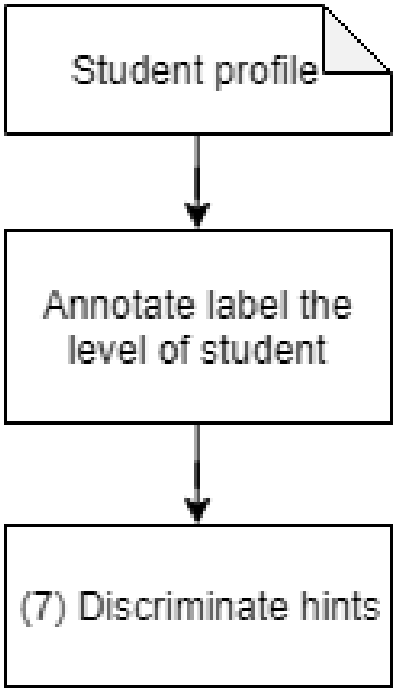


Objective 1 — Methodology — C. Hint generation

(7) Hints discriminator: classify level of hints based on the student profiles
(8) Question recognition: recognize question of student
(9) Answer recognition: recognize answer of student
(10) Hints generator: generate hint based on hint types, level, and language model

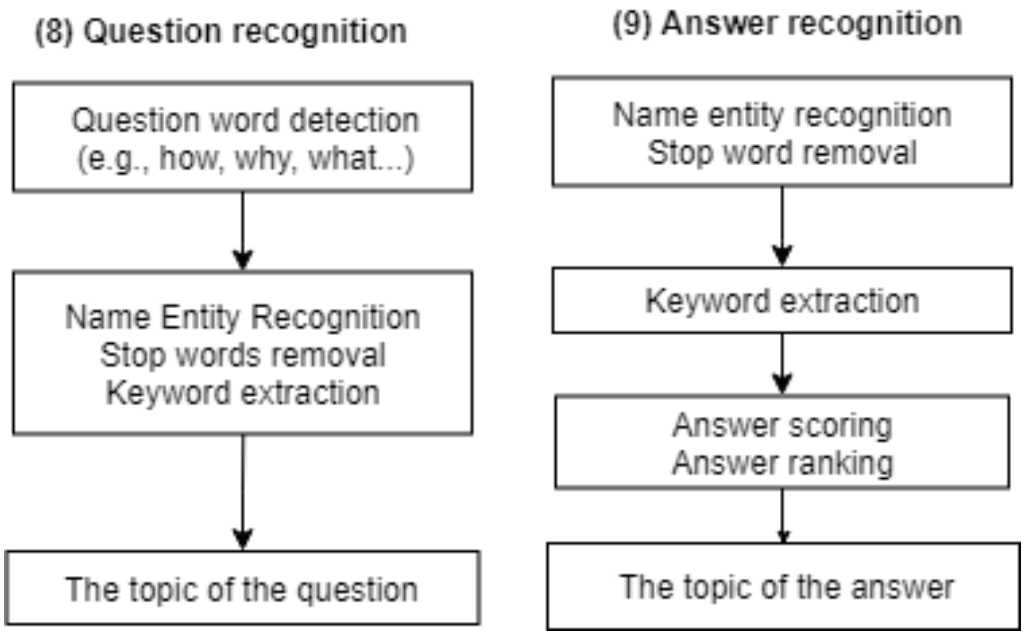
Objective 1 — Methodology — C. Hint generation — Hints discriminator

- Classify the students' level based on their profile
- Discriminate hints based on level of student

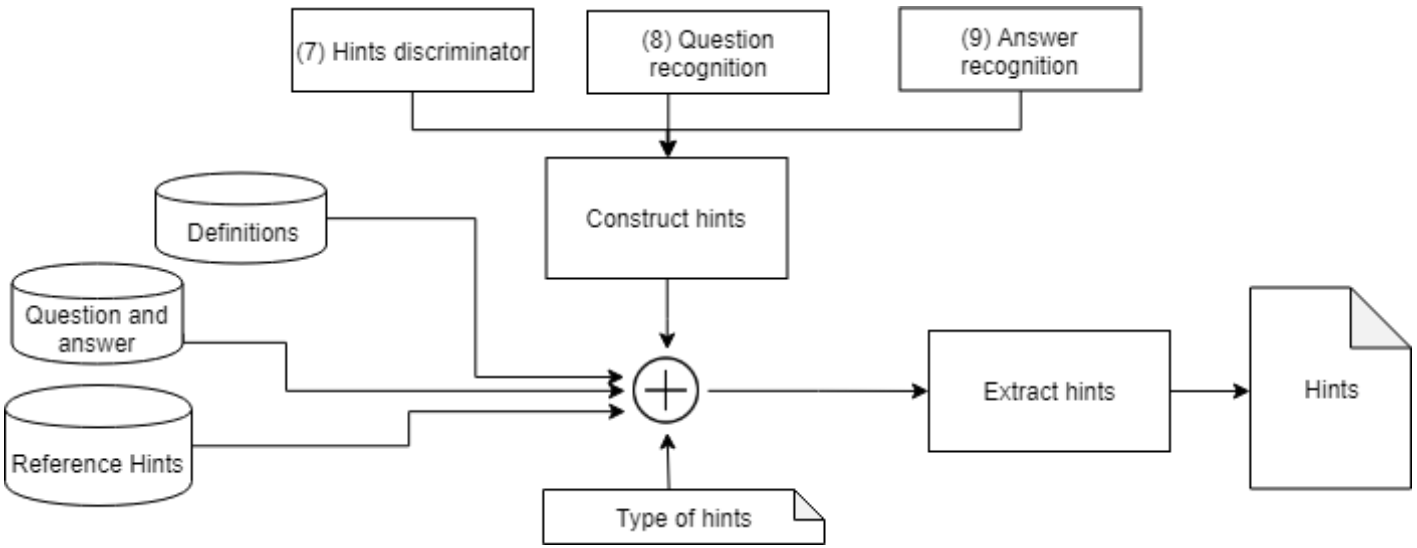


Objective 1 — Methodology — Hint generating — Question & Answer recognition

(8) Question recognition: Recognize the users' questions
(9) Answer recognition: Recognize the users' answers



Objective 1 — Methodology — C. Hint generation— Hint generator



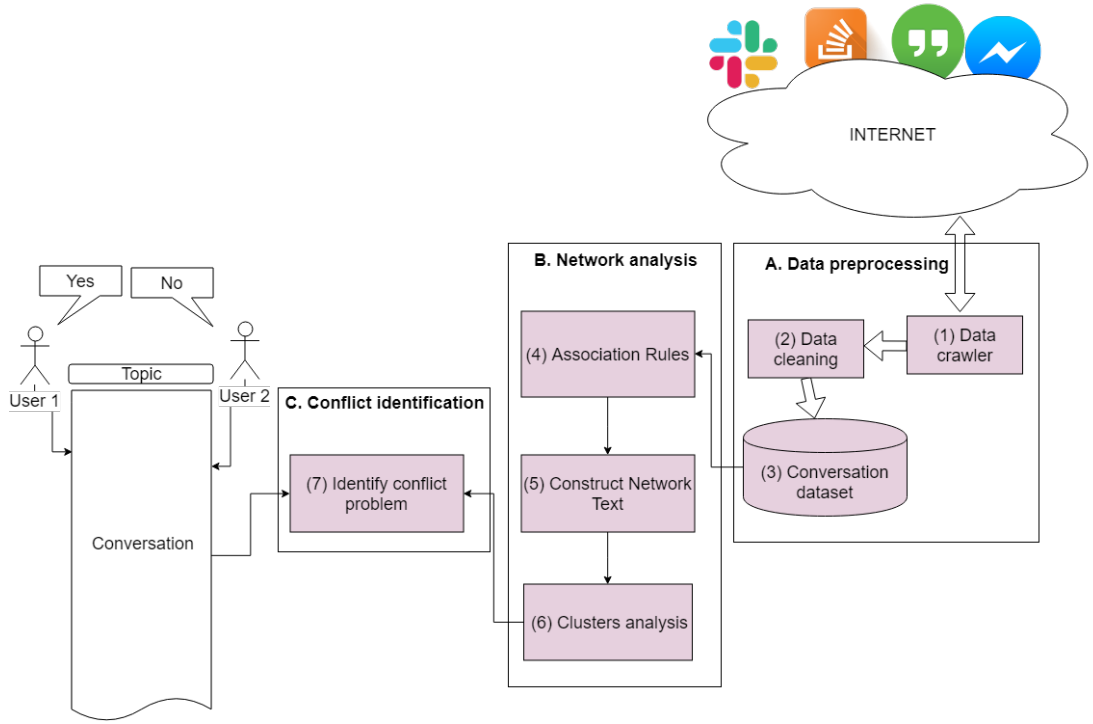
(1) Generate hints to help users solve the topic or problem automatically

(2) Identify the debated or clarify the problem

(3) Intervene in the conversation to clarify the problem: identify the opportunities for intervention, answer the related topic question of users

Objective 2 — Structure

Identify the debated or clarify the problem

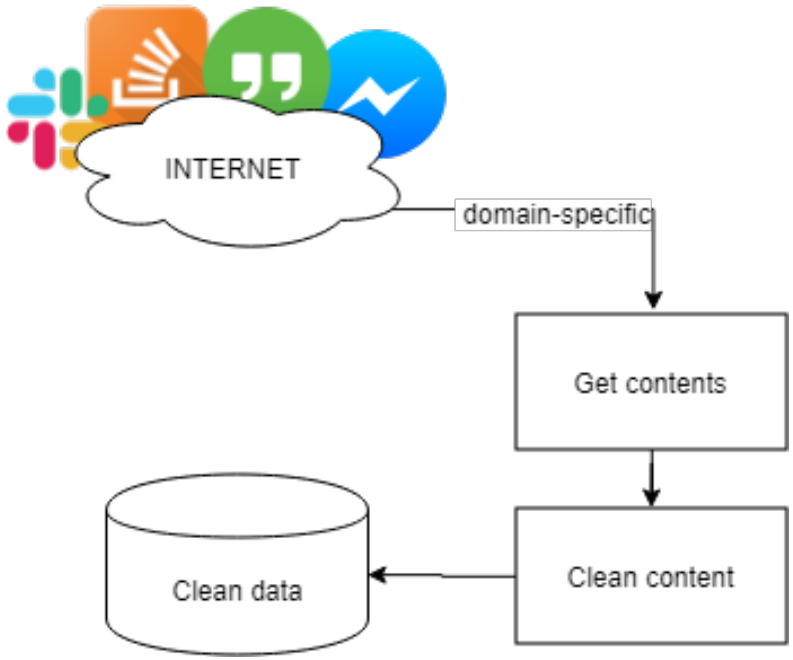


Objective 2 — Methodology — Data preprocessing

(1) Data crawler: crawling data from stackoverflow, hangout, messenger, slackwith a given domain (e.g., statistic)

(2) Data cleaning: Clean content: clean unicode, equation over the conversation

(3) Conversation dataset: save the conversation dataset to the tsv file

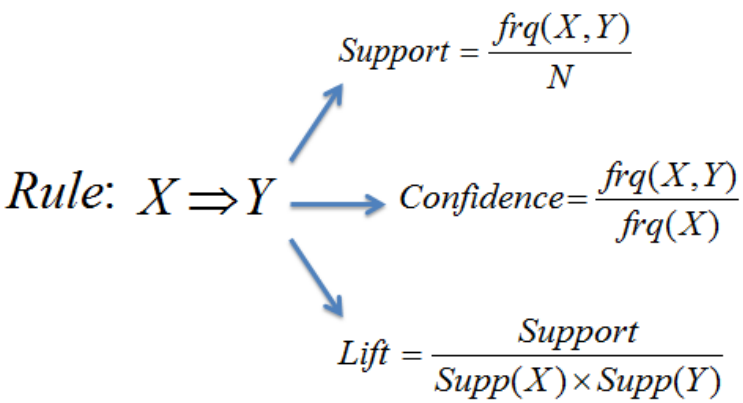


E.g., consider the technical conversation of AI-Educate[2]

[2] https://lilabot.com/

Objective 2 — Methodology

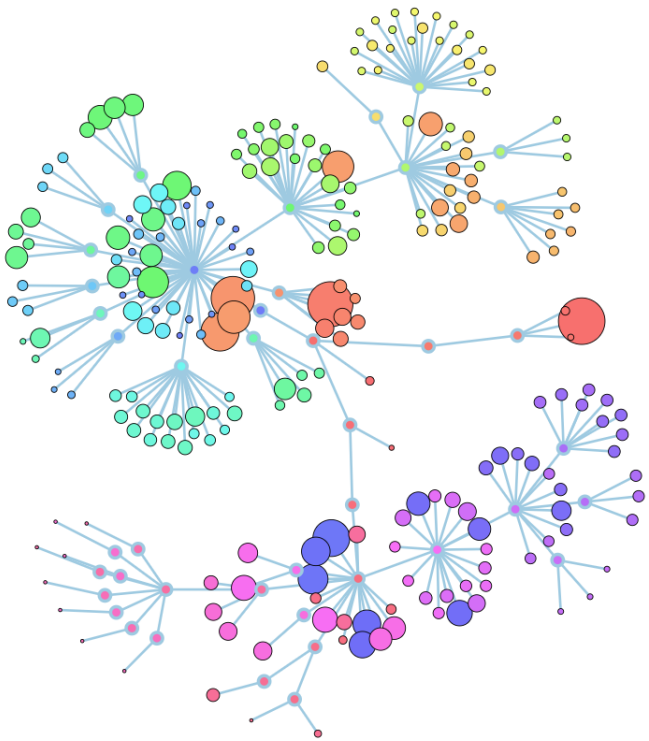
(4) Association rules[3]: find the interesting association or correlation relationship between dominant words



- *Support*: how frequently the itemset appears in the dataset.
- *Confidence*: how often the rule has been found to be true.
- *Lift*: the ratio of the observed support to that expected if X and Y were independent

[3] https://www.saedsayad.com/association_rules.htm

Objective 2 — Methodology



source: http://www.redotheweb.com/

(5) Construct network text of dominant word: include weighted edge result for association rule processes

(6) Network analysis: create context, keyword, and sense from network text

$\rightarrow$ employ centrality to find the most influential words in the networks and modularity to find words cluster/ groups in the network

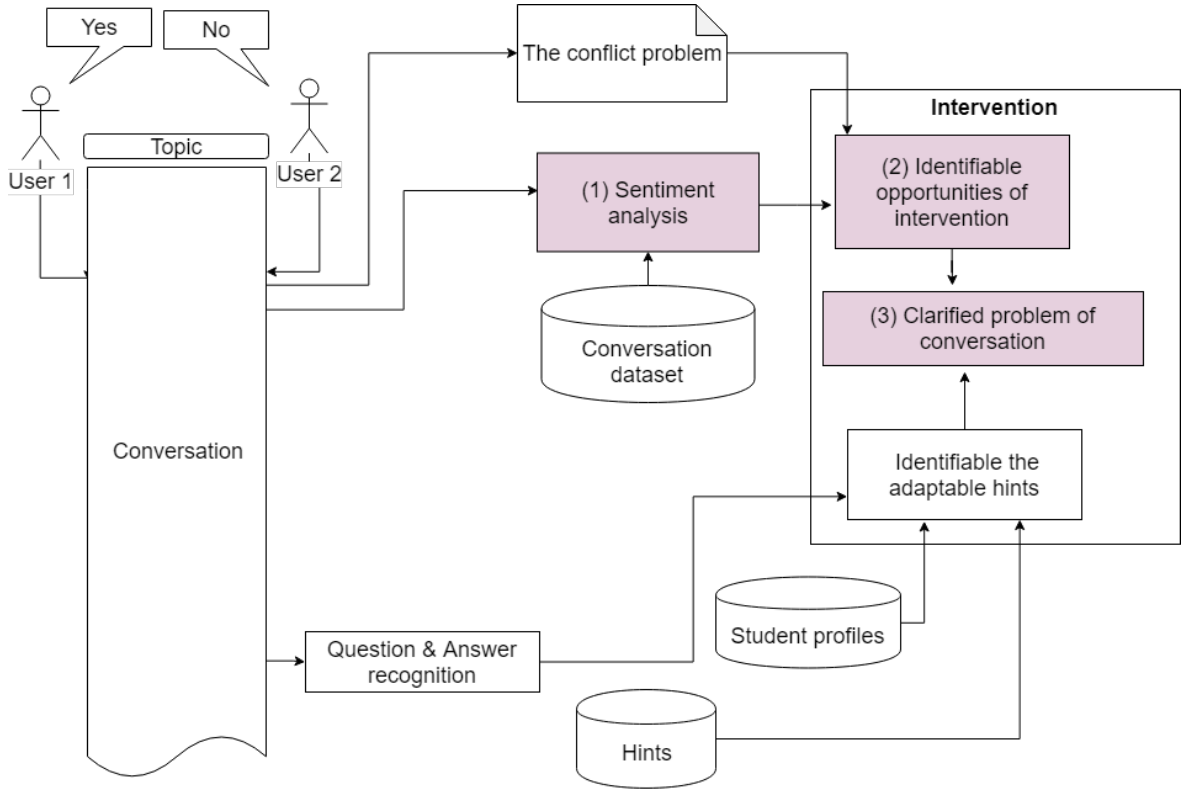
(7) Identify conflict problem: get the conflict problem related to the topic by mapping conversation to clusters analysis

(1) Generate hints to help users solve the topic or problem automatically

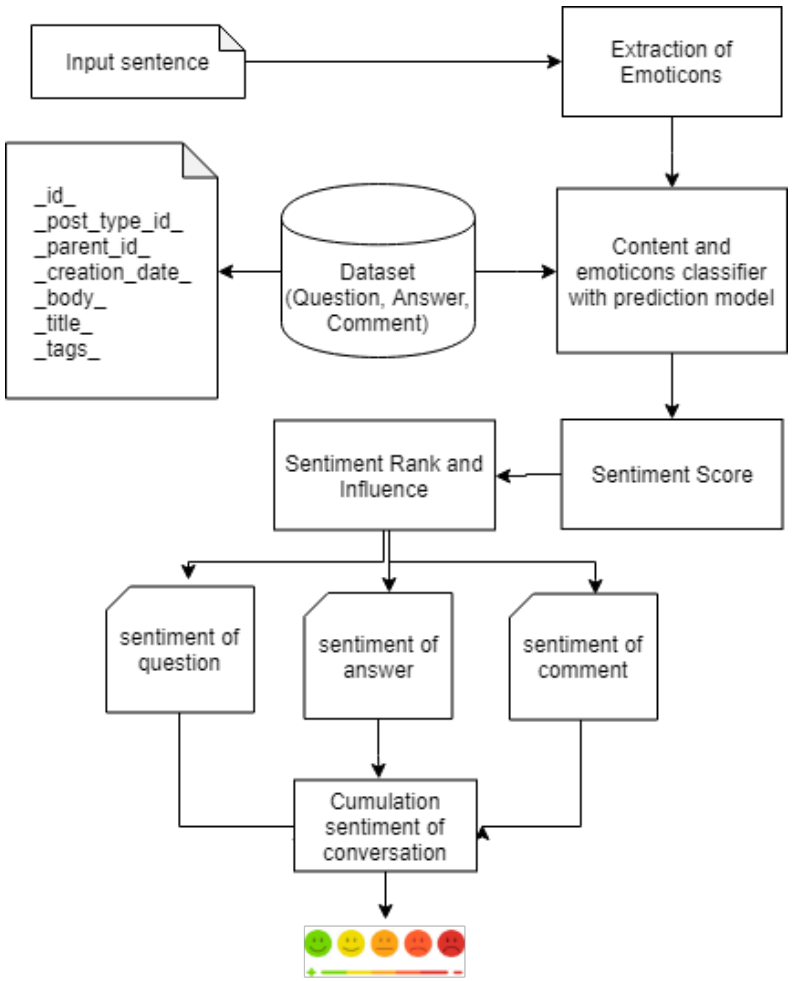
Identify the debated or clarify the problem

(3) **Intervene in the conversation to clarify the problem**

Objective 3— Structure



Objective 3 — Methodology — Sentiment analysis

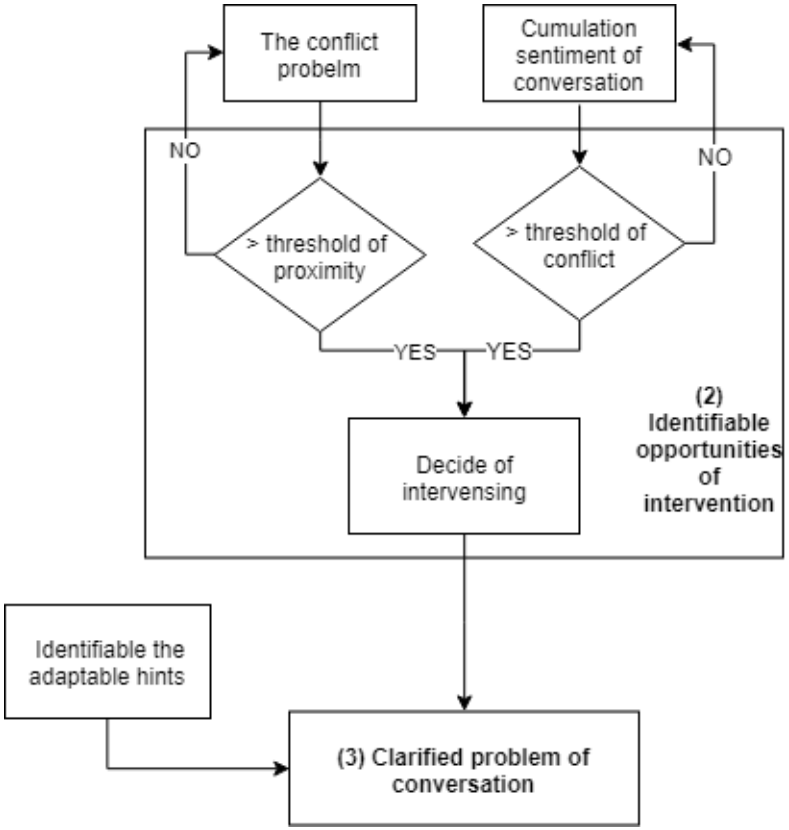


- Listening the conversation
- Using SVM in classifying the Emoticons of content
- Cummulate the setiment of question, answer, and comment for evaluating the sentiment of conversation

Ref: L. Ling, S. Larsen, "Sentiment Analysis on Stack Overflow with Respect to Document Type and Programming Language", KTH ROYAL INSTITUTE OF TECHNOLOGY

Objective 3 — Methodology — Intervention

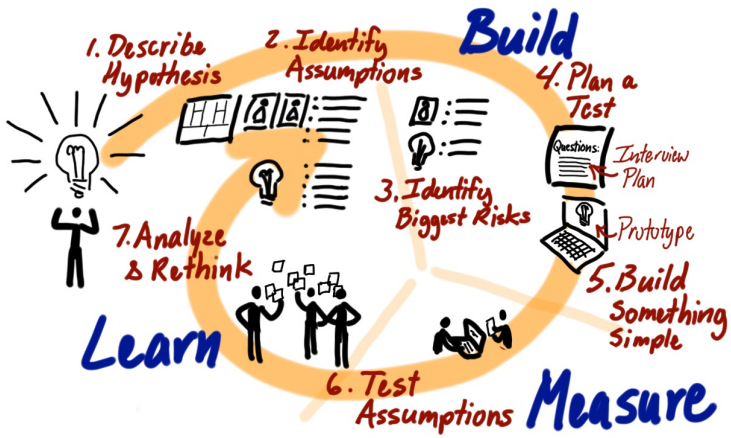
(2) Identifiable opportunities of intervention: analysis the serious of conversation and conflict problem
(3) Clarified problem of conversation: give the right intervention



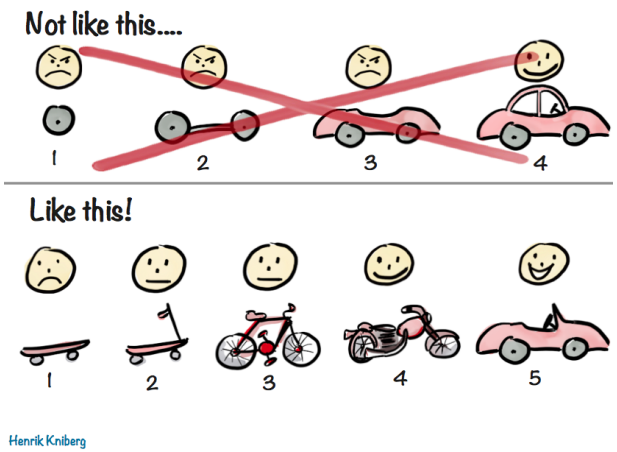
Evaluation measurement

Because this is the conversation between Human and machine, so we prefer to use the users' experiment test to get feedback score in range (1,5) and expert recommendations.

User experience — Approach



Source: https://www.jpattonassociates.com/



Source: https://quickleft.com

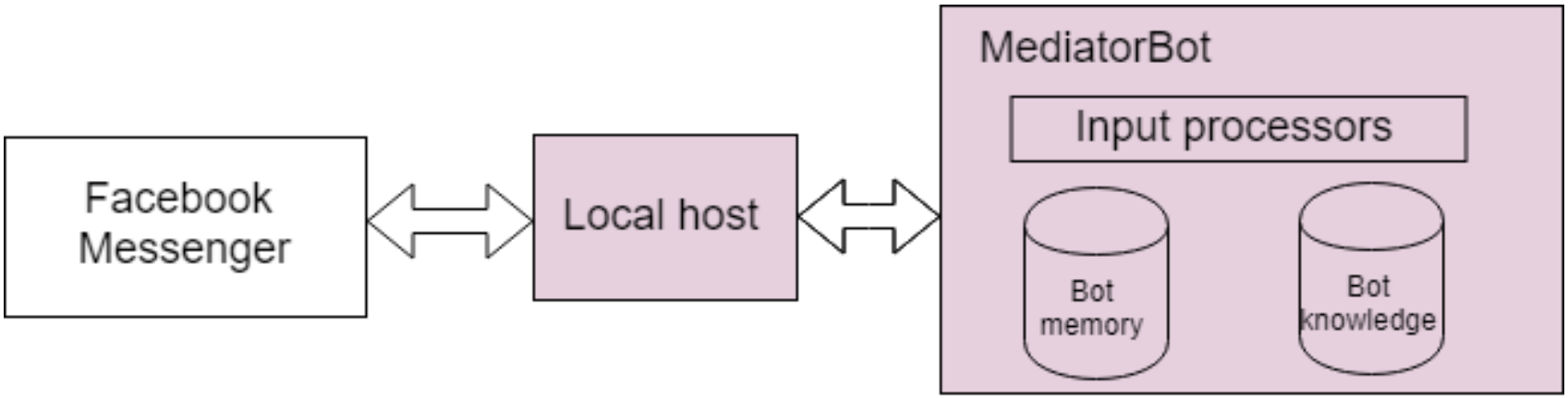
→ We evaluate our system by using the user experiments testing.

- students' experiments
- professor recommendations

→ Users: students at the class offline, students on LILA[4], friends (if REB is valid) or Amazon Mechanical Turk[5]

4.https://lilabot.com 5.https://www.mturk.com/

User experience — Environment



- (1) Use the Facebook messenger API to set up the conversation environment
- (2) Set up the flask server for local host
- (3) Process the conversation with the given bot memory and knowledge
- (4) Make the report feedback statistic evaluation (`https://docs.gogle.com/forms/u/0/`)
- (5) Using Cohen's kappa for evaluating the agreement of human and machine experiment

Achievements

(1) 3 years Mitacs accelerate grant for Natural Language Generation for Intelligent Tutoring Systems

(2) Directly apply the results to LILA and Korbit systems at Ai-educate Inc
`https://lilabot.com/`
`https://www.youtube.com/watch?v=MP0et2zE87I&feature=youtu.be`

(3) Get the good feedback from the students though LILA system (Ai-educate has the REB for this experiement)

Work plan

Activity	2017		2018			2019			2020			2021		
	S	F	W	S	F	W	S	F	W	S	F	W	S	F
DGA1005														
DGA1031														
MTI830														
DGA1032														
DGA1033														
Conference														
Mitacs proposal for PhD grant														
Literature review														
Experiement														
Research ethics board certificate														
Publishcation								[1]		[2]			[3]	
Writing thesis													(*)	(*)

Figure: Work schedule

Journals:

[1] Journal of Artificial Intelligence Research
[2] Technology, Knowledge and Learning
[3] Education and Information Technologies

Finished courses:

(1) DGA1005
(2) MTI830

Thank You