

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

April 23, 2019

Overview

- 1 Introduction
- 2 Problem statement
- 3 Motivations
- 4 Objectives
- 5 Methodology
- 6 Work plan
- 7 Appendices



Le génie pour l'industrie



ITS main purposes:

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

- AutoTutor and its progenies [1] help students learn by holding a conversation in natural language - 2016
- 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 [2] - 2017



[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

Scenario

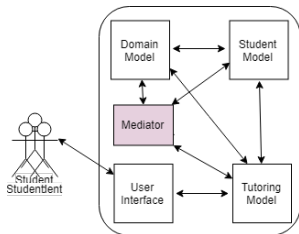


Figure: ITS with Mediator

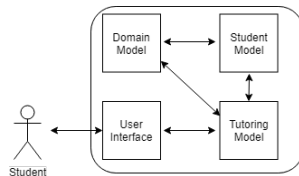
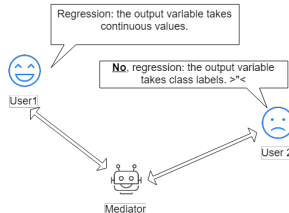
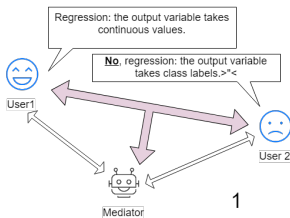


Figure: Original ITS [*]

[*] N. T-Nghe and L. S-Thieme, "Multi-Relational Factorization Models for Student Modeling in Intelligent Tutoring Systems", 17th International Conference on Knowledge and Systems Engineering (KSE) 2015

Scenario



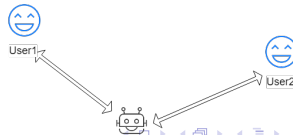
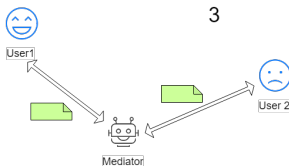
Topic: the definition of regression

Opportunity of intervention: Now

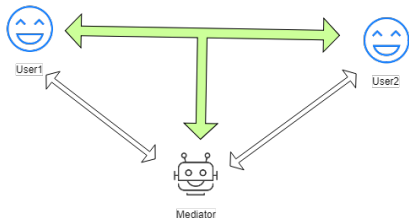
Hints

Regression: given a set of data, find the best relationship that represents the set of data

Classification: given a known relationship, identify the class that the data belongs to



Motivations



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

Objectives

Main objective: Propose a smart Mediator to support constructive discussion based on the Intelligent Tutor System:

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

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

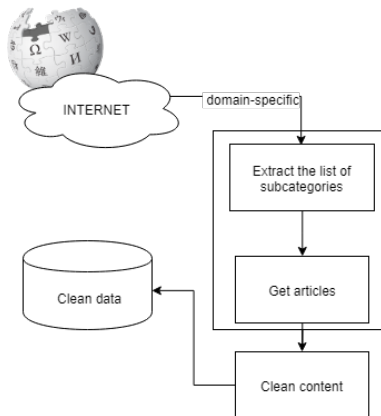
(2) Identify the debated problem

(3) Intervene in the conversation to clarify the problem

19 / 59

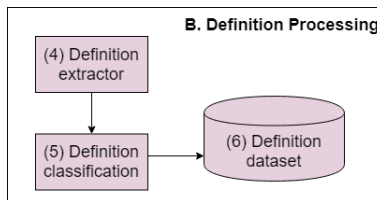
Objective 1 — Methodology — A. Data preprocessing

- (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

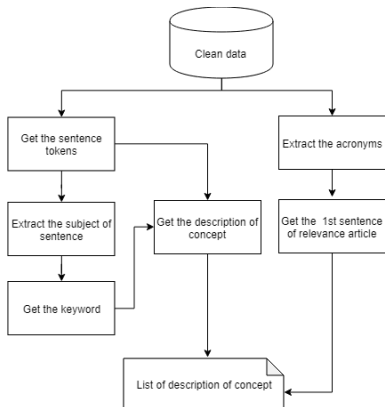


Objective 1 — Methodology — B. Definition processing

- (4) Definition extractor: Extract the description of each concept
- (5) Definition classification: classify type of definition based on supervisor algorithm (Positive/Negative)
 - Using the oversampling methodology to reweight the Positive and Negative samples
- (6) Definition dataset



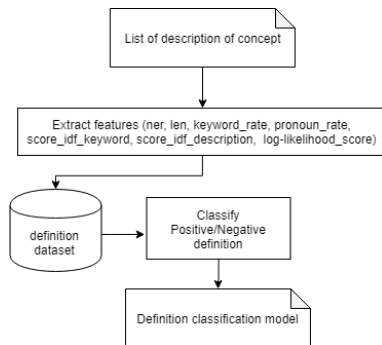
Objective 1 — Methodology — B. Definition processing — Definition extractor



- 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 concept which is called dictionary of definition

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

- Extract the features \leftarrow score table
- Save the score table to the definition dataset
- Classify the Positive/Negative definition based on logistic classification
- Save the classification model



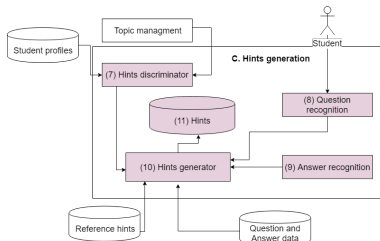
Objective 1 — Methodology — B. Definition processing — Definition classification

Features	Summary
length_of_keyword	the number words in the keyword
length_of_description	the number words in the description
score_keyword	inverse document frequency of keyword
score_description	inverse document frequency of concepts description
ner_in_description	name of entity recognition within the description
coreference_in_description	compute the coreference resolution score
type_of_word	recognize type of word (verb, noun, etc.,)
non_of_word	recognize the none of word (symbol, number, etc.,)
pronouns_rate	the rate of $\frac{\text{pronouns}}{\text{nouns}}$
keyword_rate	the rate of $\frac{\text{keyword_position}}{\text{length_of_description}}$
perplexity	the real value of perplexity of description
likelihood_score	the log-likelihood probability score of description based on sum of probability term by using language model based on RNN

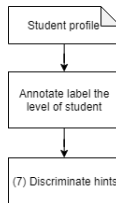
Table: Features of definition

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

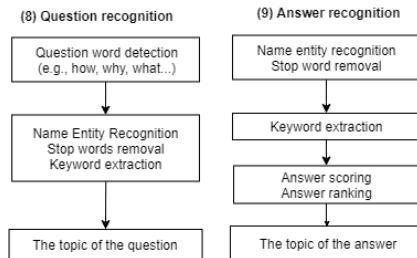


- Classify the students' level based on their profile
- Discriminate hints based on level of student

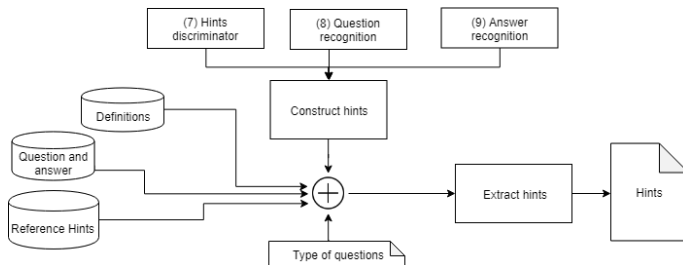


Objective 1 — Methodology — Hint generating — Question & Answer recognition

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



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



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

- * Type of questions: definition, definition & explanation, contrast, explanation, and deeper understanding
- * The hints are phrased in the form of "Think about X" or "Consider X" where X is the part of expectation answer.
- * Using Linear regression model based on the features:

Features	Summary
length_of_hint	the number words in the hint
overlap_question_hint	the rate of overlap between question and hint
score_keyterm	inverse document frequency of keyterm in hint
keyhint_keyquestion_ratio	the ratio of $\frac{\text{number_of_keyhint}}{\text{number_of_keyquestion}}$
topic_overlap	content overlap between the question and hint
pronouns_rate	the rate of $\frac{\text{pronouns}}{\text{nouns}}$ in hint
keyword_rate	the rate of $\frac{\text{keyword_position}}{\text{length_of_hint}}$
perplexity	the real value of perplexity of hint
ner_in_hint	name of entity recognition within the hint
score_of_hint	the log-likelihood probability score of hints based on sum of probability terms by using language model based on RNN

Table: Features of hints

Objective 1 — Methodology — C. Hint generation— Hint generator— Construct hints— Example

Question:

You are given a dataset of images of wildlife in Africa.
You are tasked with building a model which can identify animals in the images.
Is this a regression or classification problem? Explain why?

Answer:

It is the regression problem because the animal is the independent entity in the africa

Hints:

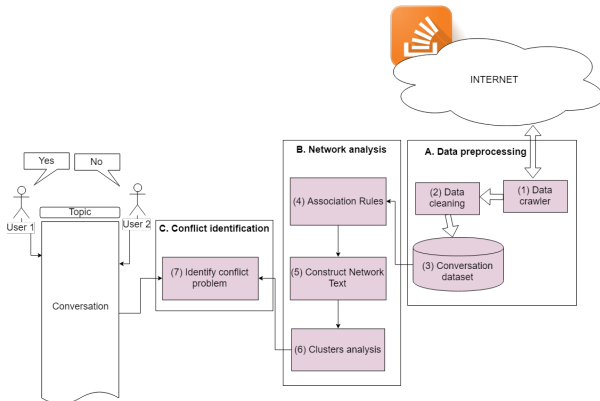
- Recall that each animal is a class.
- Recall that each animal is a discrete class.
- Consider that each animal is a separate class.
- Consider that we are choosing between a set of categories.
- Think about the following: we are choosing between discrete-valued output variables.
- Consider that each image can contain several animals, and therefore the model must predict the existence of each type of animal.

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

(2) Identify the debated problem

(3) Intervene in the conversation to clarify the problem

Objective 2 — Structure

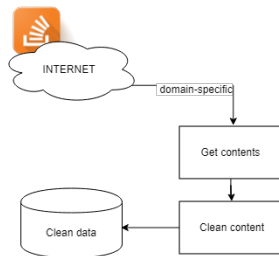


Objective 2 — Methodology — Data preprocessing

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

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

(3) **Conversation dataset:** save the conversation dataset



Objective 2 — Methodology

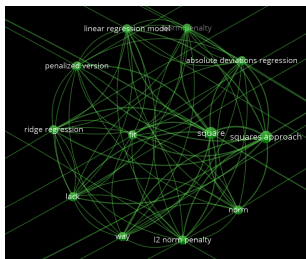
(4) Association rules [*]: find the interesting association or correlation relationship between dominant words

$$\begin{array}{l}
 \text{Rule: } X \Rightarrow Y \\
 \begin{array}{l}
 \nearrow \text{Support} = \frac{\text{freq}(X, Y)}{N} \\
 \rightarrow \text{Confidence} = \frac{\text{freq}(X, Y)}{\text{freq}(X)} \\
 \searrow \text{Lift} = \frac{\text{Support}}{\text{Supp}(X) \times \text{Supp}(Y)}
 \end{array}
 \end{array}$$

- *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

[*] A. Alamsyah, M. Paryasto, F. J. Putra, R. Himmawan,
 "Network text analysis to summarize online conversations for
 marketing intelligence efforts in telecommunication industry",
 in 2016 ICoICT

Objective 2 — Methodology



An example of text network

- (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
 → employ centrality to find the most influential words in the networks and modularity to find words cluster/ groups in the network

Objective 2 — Methodology

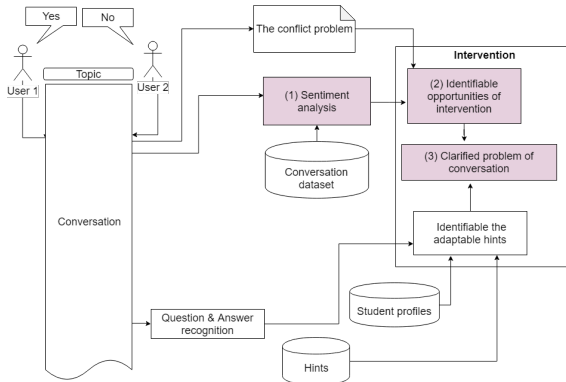
+ Using K-means algorithm to cluster and store Association Rules without favoring or excluding any evaluation measurement of an association rule.[*]

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

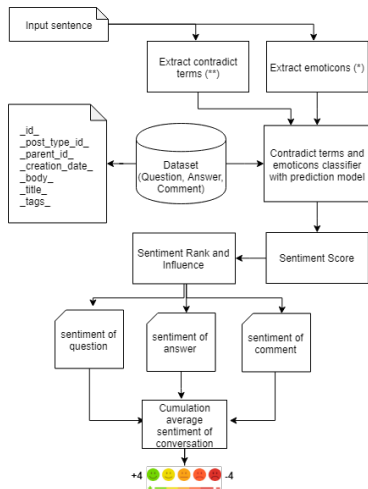
[*]A. Dahbi, M. Mouhir, Y. Balouki and T. Gadi,"Classification of association rules based on K-means algorithm", in 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)

- (1) Generate hints to help users solve the topic or problem automatically
- (2) Identify the debated 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
- Cumulate the setiment of question, answer, and comment for evaluating the sentiment of conversation

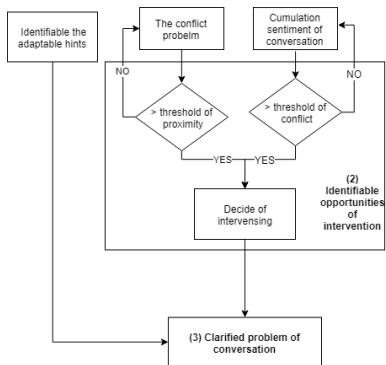
(*) L. Ling, S. Larsen, "Sentiment Analysis on Stack Overflow with Respect to Document Type and Programming Language", KTH ROYAL INSTITUTE OF TECHNOLOGY, 2018

(**) M. Marneffe, A. N. Rafferty, and C. D. Manning. 2008. Finding contradictions in text. In Proc. ACL

Objective 3 — Methodology — Intervention

(2) Identifiable opportunities of intervention: analysis the serious of conversation and conflict problem

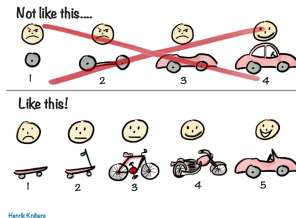
(3) Clarified problem of conversation: give the right intervention



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.

Evaluation — Approach



Source: <https://www.jpatttonassociates.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 LILA⁴, friends (if REB is valid) or Amazon Mechanical Turk⁵

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

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														
Experiment														
Research ethics board certificate														
Publication														
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

Appendices

ASAT: AutoTutor Script Authoring Tool is the primary authoring tool for AutoTutor. Can direct multiple agents and external events/controls.

ASATA: AutoTutor Script Authoring Tool for Assessment is a specialized authoring tool developed with the Educational Testing Service for developing for building dialog-based high stakes assessments.

AutoMentor (STEM Thinking): Uses epistemic analysis of discourse in student group chats to help students learn how to think and act like STEM (science, technology, engineering, and mathematics) professionals in a multi-party serious game simulation of urban planning.

AutoTutor (Computer Literacy): Core AutoTutor natural language tutoring system, which uses expectation-misconception dialog and deep questions, latent semantic analysis & regular expressions, and talks with user through the animated agent(s).

AutoTutor-3D (Physics): An extension of AutoTutor for physics, AutoTutor-3D added interactive three dimensional simulations of physics problems designed in 3D Studio Max.

AutoTutor Affect-Sensitive (Computer Literacy): AutoTutor-AS detected affect using natural language and discourse, facial expressions, body posture, and speech. Feedback considered student emotions and cognitive states. Sometimes called AutoTutor-ES (Emotion Sensitive).

AutoTutor Lite (General): AutoTutor Lite (ATL) is a web-based variant of AutoTutor designed for simpler authoring, rapid deployment, and integration into thirdparty systems.

BRCA-Gist (Breast Cancer Risk): An AutoTutor Lite tutor led by the Miami University, intended to tutor understanding of risk probabilities and personal breast cancer risk.

Coh-Metrix: A linguistic analysis toolkit with over 200 metrics. The “Coh” stands for cohesion and coherence.

CSAL Adult Literacy Tutor (Reading): This tutoring system project for the Center for the Study of Adult Literacy (CSAL) is intended to help learners who 460 Int J Artif Intell Educ (2014) 24:427–469 struggle with print media, through closer integration of trialogs, web pages, and multimedia.

DeepTutor (Physics): Tutor that uses learning progressions to foster deep learning of physics concepts, as well as enhanced semantic analysis, such as entailment.

Appendices

GazeTutor (Biology): Enhanced version of Guru Tutor that monitors and reacts to student gaze.

Gnu Tutor (General): An open source Java release of an early version AutoTutor Lite.

Guru Tutor (Biology): Tutoring system for biology designed based on observation of expert tutors. Uses collaborative lecturing and concept maps to support learning.

HURAA (Research Ethics): The Human Use Regulatory Affairs Advisor for training ethics in human experiments. AutoTutor agents helped navigate hypertext multimedia containing case-based reasoning and multiple information retrieval mechanisms.

iDRIVE (Computer Literacy, Physics, Biology): Instruction with Deep-Level Reasoning Questions in Vicarious Environments where the learner observes two pedagogical agents demonstrate deep explanations and model effective learning behavior (e.g. question-asking).

iSTART (Reading): Interactive Strategy Training for Active Reading and Thinking is a tutoring system for improving reading comprehension by training reading strategies. Uses multi-agent conversations and specialized semantic analysis to tutor reading strategies.

iSTART-ME (Reading): The Motivationally-Enhanced (ME) version of iSTART provides tutoring using an interactive game environment.

MetaTutor (Biology): Tutors self-regulated learning (SRL) skills inside a hypermedia setting.

Operation ARA (Scientific Reasoning): Operation Acquiring Research Acumen is an extension of the Operation ARIES project that adds additional features and game content.

Operation ARIES (Scientific Reasoning): Operation Acquiring Research, Investigative, and Evaluative Skills is a trialog-based tutoring system and serious game for teaching critical thinking. Learners resolve inconsistent information about scientific methods inside a serious game narrative.

Appendices

QUAID: Question Understanding Aid was a tool to evaluate the comprehensibility of questions.

SEEK Web Tutor (Critical Thinking): The Source, Evidence, Explanation, and Knowledge Tutor was designed to help learners evaluate the credibility and relevance of information using tutoring-enhanced web search, with spoken hints, pop-up ratings and metacognitive journaling.

SKO Modules (General): Sharable Knowledge Object Modules are encapsulated, cloud-hosted modules that compose web services to provide tutoring. Currently being applied to Algebra.

VCAEST (Medical): Virtual Civilian Aeromedical Evacuation Sustainment Training is designed to train civilian medical personnel on federal guidelines for emergency situations and triage.

WHY2/AutoTutor (Physics): Extension of AutoTutor that approached tutoring conceptual physics. This was part of a larger WHY2 project that included Int J Artif Intell Educ (2014) 24:427–469 461

WHY2/Atlas. WHY2 was a reference to an old tutoring system called WHY and the year 2000 (e.g., Y2K).

Writing-Pal (Writing): This tutor attempts to improve essay and academic writing

skills and provides automated evaluation and feedback on essays. It is related to the iSTART system.

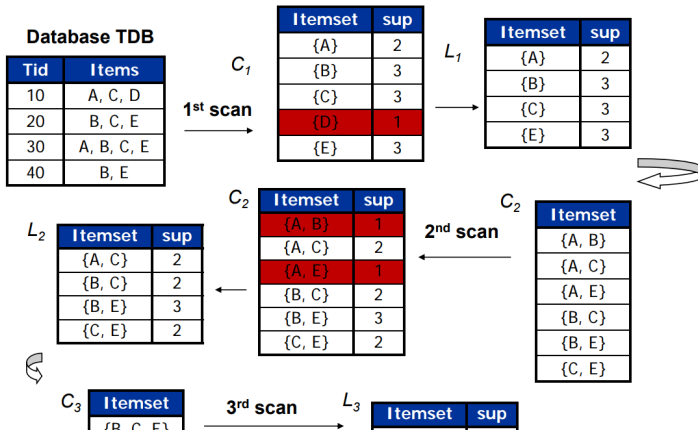
Appendices

TID	s	c
bread \Rightarrow peanut-butter	0.60	0.75
peanut-butter \Rightarrow bread	0.60	1.00
beer \Rightarrow bread	0.20	0.50
peanut-butter \Rightarrow jelly	0.20	0.33
jelly \Rightarrow peanut-butter	0.20	1.00
jelly \Rightarrow milk	0.00	0.00

TID	Items
T1	bread, jelly, peanut-butter
T2	bread, peanut-butter
T3	bread, milk, peanut-butter
T4	beer, bread
T5	beer, milk

Appendices

Example of Apriori Run



* We extract contradiction features on which we apply logistic Features for contradiction detection

Features	Summary
polarity	The polarity features capture the presence (or absence) of linguistic markers of negative polarity contexts
numeric	The numeric features recognize (mis-)matches between numbers, dates, and times
antonymy	list of antonyms and contrasting words comes from WordNet, from which we extract words with direct antonymy links and expand the list by adding words from the same synset as the antonyms
structural	determine whether the syntactic structures of the text and hypothesis create contradictory statements.
factivity	The context in which a verb phrase is embedded may give rise to contradiction
modality	Simple patterns of modal reasoning are captured by mapping the text and hypothesis to one of six modalities ((not)possible, (not)actual, (not)necessary), according to the presence of predefined modality markers such as can or maybe

Table: Features of contradict detection

Table of Emoticon

Advantages of Logistic Regression:

- Lots of ways to regularize your model, and you don't have to worry as much about your features being correlated, like you do in Naive Bayes.
- You also have a nice probabilistic interpretation, unlike decision trees or SVMs, and you can easily update your model to take in new data (using an online gradient descent method), again unlike decision trees or SVMs.
- Use it if you want a probabilistic framework (e.g., to easily adjust classification thresholds, to say when you're unsure, or to get confidence intervals) or if you expect to receive more training data in the future that you want to be able to quickly incorporate into your model.

Algorithm	Problem Type	Results interpretable by you?	Easy to explain algorithm to others?	Average predictive accuracy	Training speed	Prediction speed	Amount of parameter tuning needed (excluding feature selection)	Performs well with small number of observations?	Handles lots of irrelevant features well (separates signal from noise)?	Automatically learns feature interactions?	Gives calibrated probabilities of class membership?	Parametric?	Features might need scaling?
KNN	Either	Yes	Yes	Lower	Fast	Depends on n	Minimal	No	No	No	Yes	No	Yes
Linear regression	Regression	Yes	Yes	Lower	Fast	Fast	None (excluding regularization)	Yes	No	No	N/A	Yes	No (unless regularized)
Logistic regression	Classification	Somewhat	Somewhat	Lower	Fast	Fast	None (excluding regularization)	Yes	No	No	Yes	Yes	No (unless regularized)
Naive Bayes	Classification	Somewhat	Somewhat	Lower	Fast (excluding feature extraction)	Fast	Some for feature extraction	Yes	Yes	No	No	Yes	No
Decision trees	Either	Somewhat	Somewhat	Lower	Fast	Fast	Some	No	No	Yes	Possibly	No	No
Random Forests	Either	A little	No	Higher	Slow	Moderate	Some	No	Yes (unless noise ratio is very	Yes	Possibly	No	No
AdaBoost	Either	A little	No	Higher	Slow	Fast	Some	No	Yes	Yes	Possibly	No	No
Neural networks	Either	No	No	Higher	Slow	Fast	Lots	No	Yes	Yes	Possibly	No	Yes

Table of supervisor learning algorithm comparison

Appendices

Using wikiSynonyms API to get the synonyms keywords

E.g., Linear regression model 'Linear modeling', 'Regression coefficient', 'Linear Regression', 'Regression coefficients', 'Regression line', 'Linear weights', 'Multiple linear regression', 'Line regression', 'Linear regression model', 'Linear trend', 'Multi-linear regression', 'Linear fit', 'Line of regression', 'Linear regression'

Appendices

TABLE 1 : SOME INTERESTINGNESS MEASURES

Measures	Formula
Lift(Lift)	$Lift(X \rightarrow Y) = \frac{P(XY)}{P(X)P(Y)}$
Information Gain(GI)	$GI(X \rightarrow Y) = \log_2 \frac{P(XY)}{P(X)P(Y)}$
Example & Counter Example Rate(ECR)	$ECR(X \rightarrow Y) = 2 - \frac{1}{conf(X \rightarrow Y)}$
Jaccard(JRD)	$JRD(X \rightarrow Y) = \frac{P(XY)}{P(XY) + P(Y)}$
Cosinus(COS)	$COS(X \rightarrow Y) = \frac{P(XY)}{\sqrt{P(X)P(Y)}}$
Pearl(PRL)	$PRL(X \rightarrow Y) = P(X) P(\frac{Y}{X}) - P(Y) $
Loevinger(LVG)	$LVG(X \rightarrow Y) = \frac{P(\frac{Y}{X}) - P(Y)}{1 - P(Y)}$
Conviction(CNV)	$CNV(X \rightarrow Y) = \frac{P(X)P(\bar{Y})}{P(XY)}$
Zhang(ZHN)	$ZHN(X \rightarrow Y) = \frac{P(XY) - P(X)P(Y)}{\max\{P(XY)P(Y), P(Y)P(XY)\}}$
Piatetsky Shapiro(PS)	$PS(X \rightarrow Y) = P(XY) - P(X)P(Y)$
Sebag-Schoenauer(SBG)	$SBG(X \rightarrow Y) = \frac{P(XY)}{P(XY)}$