



# **ArguSure: A Gemini-based Chrome Extension for Counterargument Generation**

**Carl Joven M. Marasigan and Concepcion L. Khan**

**University of the Philippines Los Baños  
College of Arts and Sciences  
Institute of Computer Science  
May 31, 2024**

# INTRODUCTION

Filter bubble, a phenomenon where Internet customization effectively isolates individuals from diverse opinions or materials, tends to cause bias and polarization. To alleviate its negative effects, it is important to utilize resources that offer various points of view. With this in mind, this study developed ArguSure. Powered by Google's multimodal LLM called Gemini, it is a counterargument generator Chrome extension designed to provide a way for users to conveniently explore counter viewpoints, especially the ones refuting what they see in digital spaces.

# STATEMENT OF THE PROBLEM

The study seeks to answer the following research questions:

- 1) What is the design and approach of the Chrome extension that can help Internet users conveniently explore counterviewpoints and be introduced to different, contradictory ideas?
- 2) What technologies will be used in developing the Chrome extension and its corresponding web application?
- 3) What prompt designs will be used to elicit counterarguments from Gemini?

# OBJECTIVES

The general objective of the study is to develop a Chrome extension that can be used to conveniently explore arguments against what we see while browsing the Internet. Specifically, the objectives that the researcher set out to accomplish are:

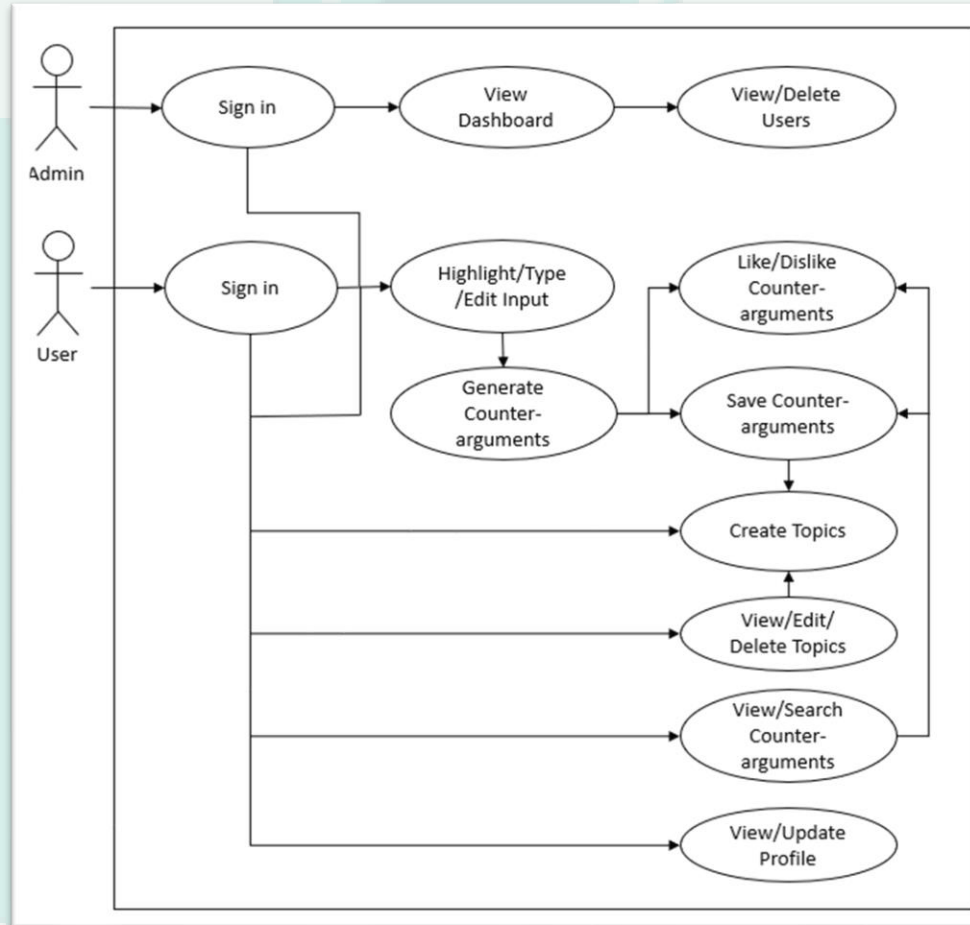
- 1) To develop a counterargument generator Chrome extension that can be used through text selection and the context menu;
- 2) To develop a web application for users to view/organize generated counterarguments and for admins to view/organize application data and users;
- 3) To design prompts and utilize the services provided by Gemini API for input assessment and counterargument generation;
- 4) To evaluate the usability of the Chrome extension and its web application using the System Usability Scale (SUS); and
- 5) To evaluate the quality of the application's generated counterarguments using LLM-based automatic evaluation.

# METHODOLOGY

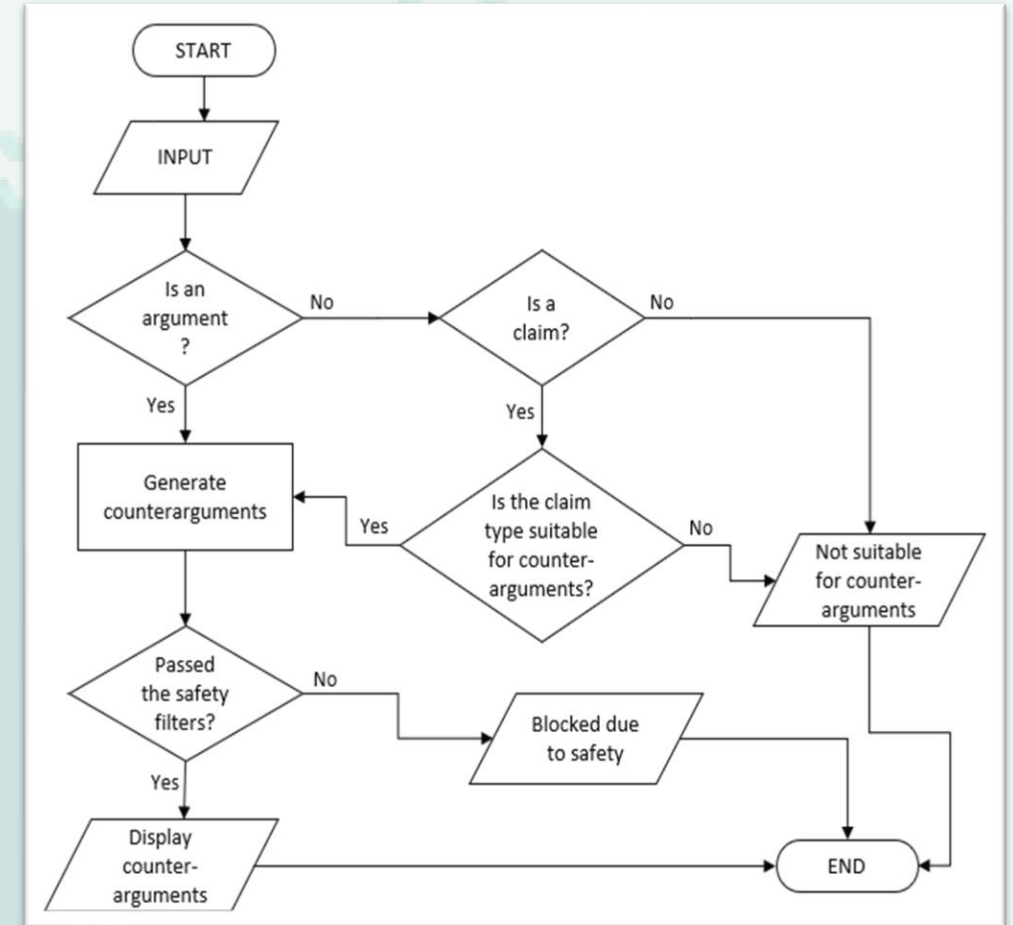
## Technologies

- Node.js
- Express.js
- MongoDB
- React
- Tailwind CSS
- Vite
- Gemini API

## Use Case Diagram



## Counterargument Generation Flowchart



# METHODOLOGY

## Application Testing and Evaluation

1. *System Usability Scale (SUS)*
  - To test and evaluate the usability of the Chrome extension and its web application
  - Human Evaluators (11 UPLB Students)
2. *LLM-based Automatic Evaluation*
  - To test and evaluate the quality of the counterarguments
  - GPT-3.5

# RESULTS AND DISCUSSION

## Development

### 1. *Gemini API*

- Used for input assessment and counterargument generation
- Gemini Pro - inputs/outputs text and can handle zero, one, and few-shot tasks.

### 2. *Prompt Design or Prompting*

- The art and science of figuring out the right wording to get generative models to do what we want.
- Some strategies used were defining the task to perform, specifying constraints, defining the format of the response, breaking down prompts into simple components, and zero-shot prompting.

# RESULTS AND DISCUSSION

## Prompts Used

### To determine if the input is an argument

Strictly yes or no, is “<input>” a claim? Please note that a claim is an assertion open to challenge.

### To determine if the input is a claim

Strictly yes or no, is “<input>” an argument? Please note that an argument is a coherent series of reasons, statements, or facts intended to support or establish a point of view.

### To determine the claim type of the input

Categorize the sentence “<input>” into seven categories:

1. Personal experience (PE): Claims that aren't capable of being checked using publicly-available information, e.g. "I can't save for a deposit."
2. Quantity in the past or present (Q): Current value of something e.g. "1 in 4 wait longer than 6 weeks to be seen by a doctor." Changing quantity, e.g. "The Coalition Government has created 1,000 jobs for every day it's been in office." Comparison, e.g. "Free schools are outperforming state schools.". Ranking, e.g. "The UK's the largest importer from the Eurozone."
3. Correlation or causation (CC): Correlation e.g. "GCSEs are a better predictor than AS if a student will get a good degree." Causation, e.g. "Tetanus vaccine causes infertility." Absence of a link, e.g. "Grammar schools don't aid social mobility."
4. Current laws or rules of operation (CLO): Declarative sentences, which generally have the word "must" or legal terms, e.g. "The UK allows a single adult to care for fewer children than other European countries." Procedures of public institutions, e.g. "Local decisions about commissioning services are now taken by organisations that are led by clinicians." Rules and changes, e.g. "EU residents cannot claim Jobseeker's Allowance if they have been in the country for 6 months and have not been able to find work."
5. Prediction (P): Hypothetical claims about the future e.g. "Indeed, the IFS says that school funding will have fallen by 5% in real terms by 2019 as a result of government policies."
6. Other type of claim (OTC): Voting records e.g. "You voted to leave, didn't you?" Public Opinion e.g. "Public satisfaction with the NHS in Wales is lower than it is in England." Support e.g. "The party promised free childcare" Definitions, e.g. "Illegal killing of people is what's known as murder." Any other sentence that you think is a claim.
7. Not a claim (NAC): These are sentences that don't fall into any categories and aren't claims. e.g. "What do you think?.", "Questions to the Prime Minister!"

Strictly use only one of the 7 labels (PE, Q, CC, CLO, P, OTC, NAC), do not provide any additional explanation.

### For Counterargument Generation

Please provide one argument against “<input>” strictly with summary (in paragraph form labeled as **\*\*Summary:\*\***), body (in paragraph form labeled as **\*\*Body:\*\***), and source (in bullet points labeled as **\*\*Source:\*\***) as the format. The argument should be well-structured and organized in a coherent manner.

Please make sure that the argument will refute “<input>” and not support it.

Please provide another argument against “<input>” strictly with summary (in paragraph form labeled as **\*\*Summary:\*\***), body (in paragraph form labeled as **\*\*Body:\*\***), and source (in bullet points labeled as **\*\*Source:\*\***) as the format. The argument should be well-structured and organized in a coherent manner.

Please make sure that the argument will refute “<input>” and not support it.

Again, please provide another argument against “<input>” strictly with summary (in paragraph form labeled as **\*\*Summary:\*\***), body (in paragraph form labeled as **\*\*Body:\*\***), and source (in bullet points labeled as **\*\*Source:\*\***) as the format. The argument should be well-structured and organized in a coherent manner.

Please make sure that the argument will refute “<input>” and not support it.

### For LLM-based Automatic Evaluation

Proposition: "<input>"

Counterargument: "<output body> \n\nTo summarize, <output summary>"

You are a lecturer of the writing class. You are given the above proposition and the counterargument that tries to refute it. You need to carefully read the proposition and the counterargument, and evaluate the counterargument based on the criteria:

1. Clarity: The counterargument should be expressed clearly, with a well-defined structure that is easy to follow. Ambiguity or vagueness can detract from the argument's coherence;
2. Relevance: The counterargument should directly address the proposition and stay focused on the topic. Irrelevant points or anecdotal evidence can detract from its coherence;
3. Validity of reasoning: Evaluate the clarity and coherence of the counterargument's reasoning. Is the line of reasoning easy to follow? Does it present a clear cause-and-effect relationship or logical progression? A well-structured and coherent counterargument should present a logical flow of ideas;
4. Logical consistency: Assess the counterargument for internal consistency. It should not contain any contradictory statements or logical fallacies that undermine its coherence. Look for logical connections and coherence between the counterargument's claims, evidence, and reasoning;
5. Overall effectiveness: The counterargument should effectively challenge the initial proposition and provide a convincing alternative viewpoint, and is likely to persuade the reader to reconsider their initial position.

Now you need to assign a score for each criterion on a scale of 1 to 5 where:

- 1 means you strongly disagree that the counterargument satisfies the criterion;
- 2 means you disagree that the counterargument satisfies the criterion;
- 3 means you prefer to be neutral on whether the counterargument satisfies the criterion;
- 4 means you agree that the counterargument satisfies the criterion; and
- 5 means you strongly agree that the counterargument satisfies the criterion.

Note that you should be very strict when giving the score.



# RESULTS AND DISCUSSION

**\* Chrome Extension and its Web Application Video Demo**

# RESULTS AND DISCUSSION

## System Usability Testing and Counterargument Quality Evaluation

Respondent	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Score
1	4	1	5	1	4	2	2	2	4	1	80
2	4	4	5	2	5	3	5	1	4	2	77.5
3	5	1	4	1	5	1	5	1	5	1	97.5
4	5	4	5	3	4	1	5	1	4	1	82.5
5	4	2	5	1	4	4	5	1	5	1	85
6	5	1	4	1	5	1	5	1	4	1	95
7	5	1	5	1	5	1	5	1	5	1	100
8	4	1	5	1	5	2	5	1	5	1	95
9	4	2	4	1	4	2	4	2	4	2	77.5
10	5	1	5	1	5	1	5	1	5	1	100
11	5	1	3	5	5	3	4	2	3	3	65
Mean Score:											86.82

System Usability Scale Scores

Criteria	Mean Score
Clarity	4.65
Relevance	4.98
Validity of reasoning	4.55
Logical consistency	4.79
Overall Effectiveness	4.67

Counterargument Quality Evaluation Scores

# CONCLUSION

- The study was able to develop ArguSure. ArguSure is a counterargument generator Chrome extension (with web application) designed to provide a way for users to conveniently seek and explore counter viewpoints, especially the ones refuting what they see in digital spaces.
- Its counterargument generation is powered by Google's multimodal LLM called Gemini. By feeding it the carefully crafted prompts, Gemini, through its API, assesses the input and tries to generate counterarguments for it.
- With a SUS score of **86.82** obtained from 11 respondents, ArguSure can be classified as an above-average in terms of user experience
- With mean scores of **4.55** to **4.98** obtained from automatic evaluation, the counterarguments are perceived to be clear, relevant, valid reasoning-wise, logically consistent, and effective overall.
- The findings also show that, compared to each criterion, the counterarguments are strongest in relevance and weakest in validity of reasoning.



**THANK YOU!**