

# Logical Fallacy Detection Using Case-Based Reasoning

## **Team #1**

Anju Vilashni Nandhakumar

Aswatth Ganapathy Subramanian

Navin Prasath Selvakumar

Nikita Vinod Mandal

# INTRODUCTION

---

- Logical fallacies are mistaken beliefs based on unsound arguments.
- They can undermine the validity of an argument.
- **Examples of Fallacies:**
  - Every time I wash my car, it rains. So me washing my car has a definite effect on weather. (**False Causality**)
  - If we teach Tommy how to drive the car, he'll want to learn how to fly helicopters next! (**Slippery Slope**)

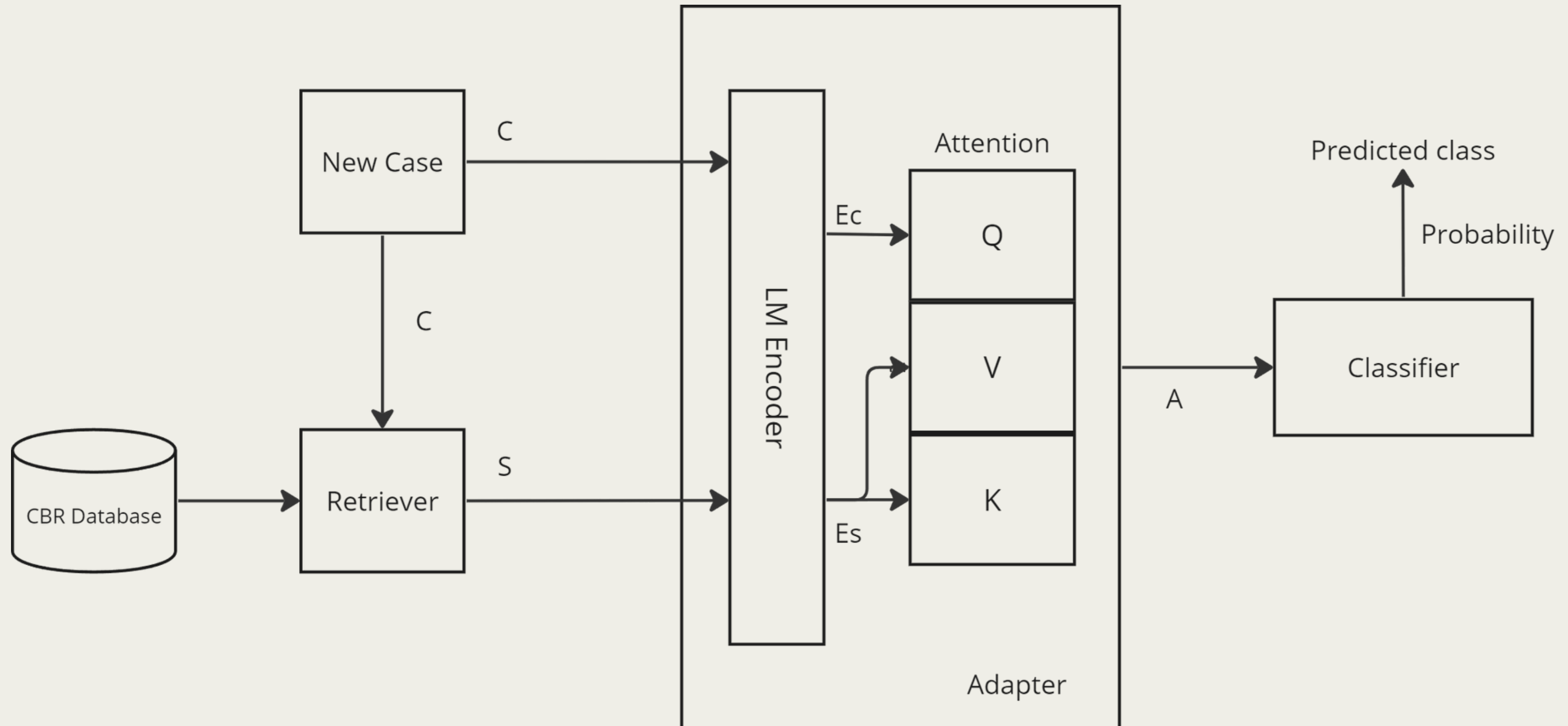
# CASE-BASED REASONING

---

- Case-based reasoning (CBR) solves new problems based on the solutions of similar past problems.
- Enriched case representations: Counterarguments, goal, explanation and structure.
- **Example:** NLP is the best course ever. Everyone is taking it (**Ad populum**)
  - Counterargument: *The speaker may be trying to persuade the listener to study NLP by appealing to their desire to be like everyone else.*
  - Structure: *X is the best Y ever. Everyone is taking Y! .*

# METHODOLOGY

---



# RETRIEVER

---

- Finds  $k$  similar cases( $S_i \rightarrow i$  in  $\{1, 2 \dots k\}$ ) for the new case( $C$ ) from case database using cosine similarity.
- The retriever then picks the  $k$  cases with top cosine similarities from the database.
- The new case is concatenated to its similar cases and is passed as input to the adapter.
- During training, the retriever's weights are frozen, while the adapter and the classifier are trained end-to-end.

# OUTPUT

---

```
text = "If you want to be healthy, you need to stop drinking coffee. I read it on a fitness blog. "  
model(text)
```

Executed at 2024.04.10 11:47:46 in 47ms

```
[{'label': 'fallacy of credibility', 'score': 0.5020125508308411}]
```

```
text = "You either like coffee or tea"  
model(text)
```

Executed at 2024.04.10 12:00:13 in 26ms

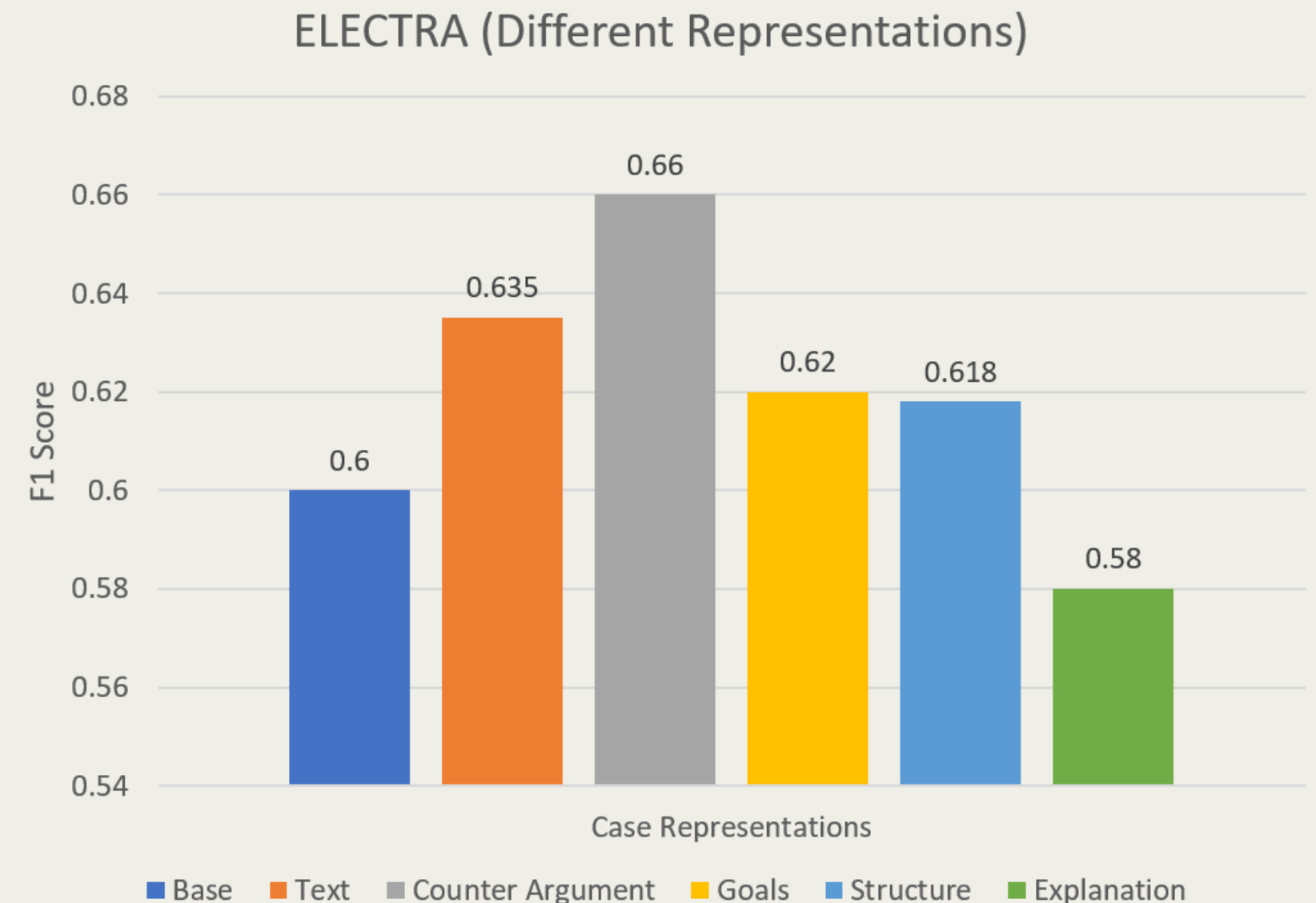
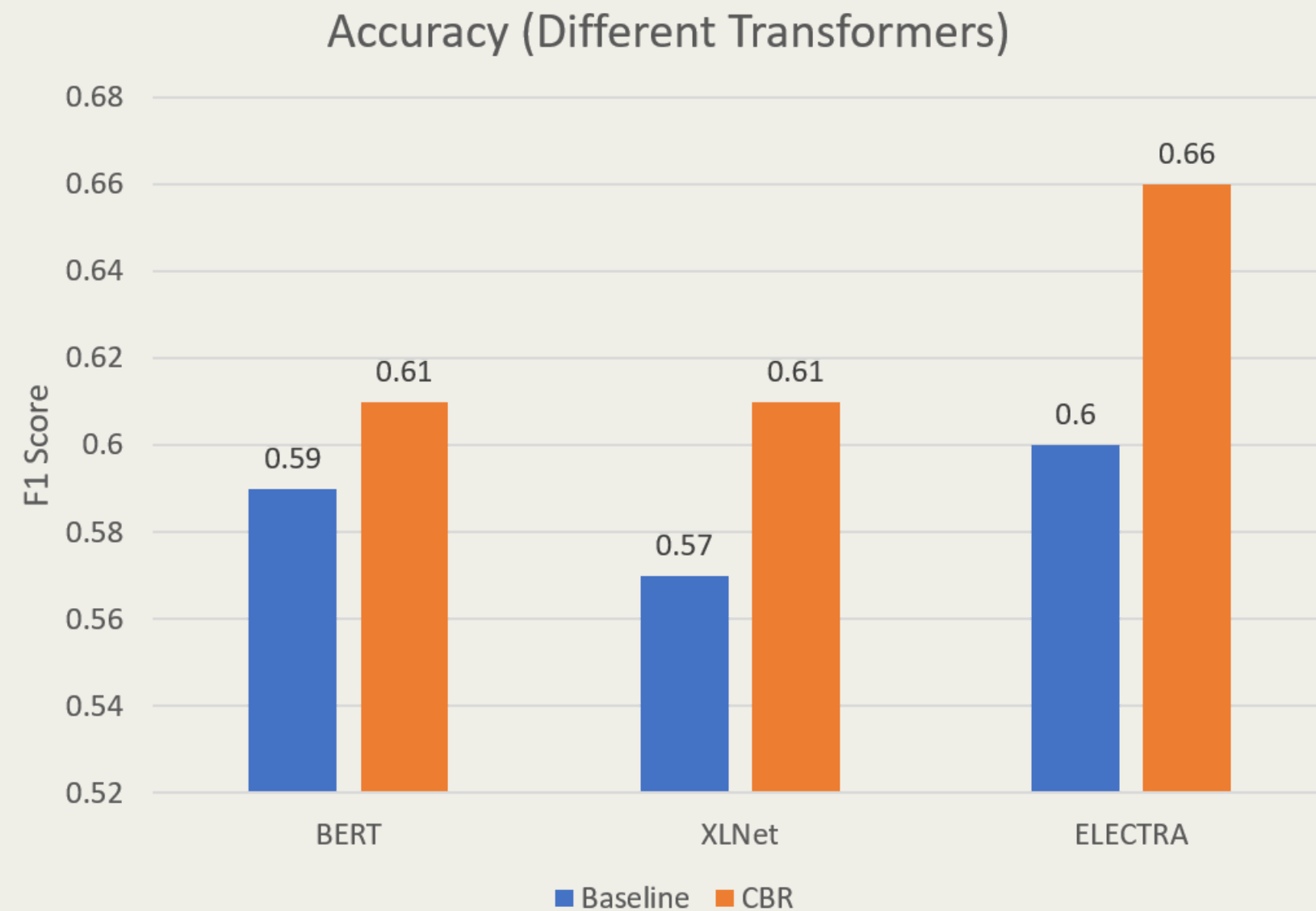
```
[{'label': 'false dilemma', 'score': 0.9633245468139648}]
```

```
text = "'Either you pass this test, or you fail it.' Laughed Sr Malak"  
model(text)
```

Executed at 2024.04.10 11:53:07 in 24ms

```
[{'label': 'false dilemma', 'score': 0.9730586409568787}]
```

# RESULTS



- Structure and just text performs best on some logical fallacy (False Dilemma)
- Retrieving a single similar case gives best accuracy.

## FUTURE WORKS/RECAP

---

- Utilizing case-based reasoning led to a significant 6% increase in model accuracy.
- This tool can be utilized to self-assess work for logical fallacies and to detect propaganda, empowering users with critical analysis capabilities.
- In the future, we envision leveraging hierarchical attention and document pooling for document processing.



# Thank you!

---

Anju Vilashni Nandhakumar

Aswatth Ganapathy Subramanian

Navin Prasath Selvakumar

Nikita Vinod Mandal