

# Module 1: Reasoning and Proof Models

# Models of Argument

- **Definition:** Structured frameworks for reasoning, showing how claims are supported by evidence and logic.
- In NLP, it is used in argument mining, discourse analysis, and AI reasoning.
- **Example:**
  - **Claim:** “The model is biased.”
  - **Evidence:** “It misclassifies 70% of sentences with gendered pronouns.”
  - **Warrant:** “Consistent misclassification indicates bias.”

# Proof by Demonstration

- **Definition:** Showing correctness by **working example** or execution, not formal math.
- Demonstrating an algorithm works on a real input.
- **Example:**
  - Running a sentiment analysis model on the text:
    - **Input:** "The movie was fantastic!".
    - **Output:** Positive.
- Demonstrates that the sentiment classifier functions.

# Proof by Empirical Methods

- **Definition:** Validating a claim by repeated experiments and data-driven evidence.
- In NLP, it is used for model evaluation through benchmarks, test sets, or metrics.
- **Example:**
  - An NLP classifier tested on 10 000 emails achieves 92% accuracy, 0.89 F1-score.
  - This is empirical proof that the model performs well on average.

# Mathematical Proofs

- **Definition:** Rigorous derivations using logic, theorems, and formal definitions.
- Used in algorithm analysis, complexity proofs, or correctness proofs.
- **Example:**
  - **Claim:** “The edit distance dynamic programming algorithm runs in  $\mathcal{O}(nm)$ ”.
  - **Proof:** By analyzing the DP (dynamic programming) table size ( $n \times m$ ) and constant-time cell updates.

# Proof by Hermeneutics

- **Definition:** Interpretation-based justification, often used in qualitative, philosophical, or human-centered CS/NLP research.
- In NLP, it is used for explaining results by interpreting meaning rather than strict numerical or formal reasoning.
- **Example:**
  - Analyzing why a chatbot's response "I understand your feelings" is more empathetic:
    - **Interpretation:** It mirrors human conversational norms.
  - Proof through interpretive understanding of language use.

# Lab 1

To get familiar with the basics of the aforementioned methods of proof in Python.