Distributed AI-Powered Argument Analysis System

Real-Time Emotion Classification and Multi-Source Fact-Checking

Ifesi Onubogu

AIoT Project

December 7, 2025

Outline

Problem Statement

Challenge

In an era of increasing polarization and misinformation, how do we:

- Analyze emotional dynamics in real-time conversations?
- Verify factual claims against authoritative sources?
- Present crowd-sourced predictions for subjective statements?

Solution

A distributed edge-cloud system that:

- Captures and processes conversations on Raspberry Pi
- Analyzes emotions using fine-tuned transformers on AWS
- Fact-checks claims via RAG, Polymarket, and web search
- Visualizes results in interactive web interface

Key Contributions

1. **Hybrid Fact-Checking Pipeline**
   - First system to combine curated knowledge base (RAG), prediction markets, and web search
2. **Segment-Level Emotion Analysis**
   - Per-utterance emotion classification (73.2% accuracy)
   - 8 emotion classes optimized for argument detection
3. **LRU-Cached Market Discovery**
   - Dynamic cache management for prediction market links
   - API fallback for unseen topics
4. **Privacy-Preserving Edge Architecture**
   - Local processing on Raspberry Pi
   - Cloud inference only for heavy models

System Data Flow

Step	Component	Action
blue!20**Raspberry Pi (Edge) - 4-6 sec**		
1	Microphone	Capture 30-second audio
2	pyannote	Speaker diarization
3	Whisper	Speech-to-text
4	HTTP	POST to AWS
orange!20**AWS EC2 (Cloud) - 2-3 sec**		
5	FastAPI	Receive files
6	Emotion Model	Classify per segment
7	Fact Checker	RAG + Polymarket + Web
8	Database	Store results
green!20**Browser (Client) - ¡1 sec**		
9	Gradio	Load web UI
10	Visualization	Display analysis

Total Latency: 6-10 seconds end-to-end

What Runs Where?

Raspberry Pi

HW: Pi 4 (4GB) + USB Mic

Tasks:

- Audio capture
- Diarization
- Whisper STT

AWS EC2

HW: t2.large (2 vCPU, 8GB)

Tasks:

- Emotion classification
- RAG + Polymarket
- Web search
- Storage

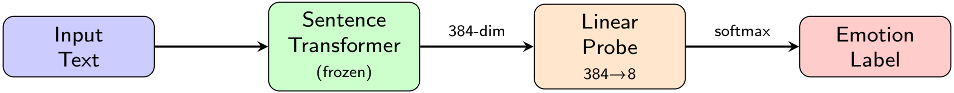
Browser

UI: Gradio web app

Features:

- Chat bubbles
- Emotion badges
- Hover panels
- Source links

Emotion Classifier Architecture



Model:

- Pre-trained embedder (66M params frozen)
- 2-layer MLP (49K trainable)
- Fast inference (100ms)

8 Emotion Classes:

- Calm, Confident, Defensive
- Dismissive, Passionate
- Frustrated, Angry, Sarcastic

Training Data Generation

Synthetic Data via GPT-4

Generated 500 labeled examples using structured prompts for 8 emotion classes

Prompt Structure:

- Specify topic & emotion
- Include emotion patterns
- Generate 2-4 sentences

Topics:

- Remote work, Climate
- EVs, Social media
- Healthcare, Politics

Example Patterns:

- Confident: "Obviously..."
- Defensive: "That's not fair..."
- Frustrated: "This is pointless..."
- Sarcastic: "Oh sure..."

Data Split:

- Train: 400, Val: 100
- Balanced across 8 classes

Training Configuration

Hyperparameters

- Adam optimizer, LR: 0.001
- Batch: 32, Epochs: 20
- Dropout: 0.3
- CrossEntropy loss

Why Linear Probe?

- Only 49K trainable params
- Works with 400 examples
- Fast inference (100ms)
- Low memory footprint

Model: 66M frozen + 49K trainable = 66M total params

Testing Methodology

Test Setup

- 500 synthetic examples (GPT-4)
- Split: 400 train / 100 val
- 100 held-out test (balanced)

Evaluation Process

1. Train 20 epochs, validate, select best
2. Test on 100 held-out examples
3. Real-world validation on recordings

Note

Synthetic data provides strong baseline for transfer learning

Emotion Classifier Results

Metric	Mean	Std	Range
Accuracy	**73.2%**	4.1%	68-79%
Precision	71.8%	5.2%	64-78%
Recall	72.4%	4.8%	66-78%
F1-Score	72.1%	4.5%	66-78%

Best Classes:

- Angry: 82% F1
- Confident: 79% F1
- Calm: 76% F1

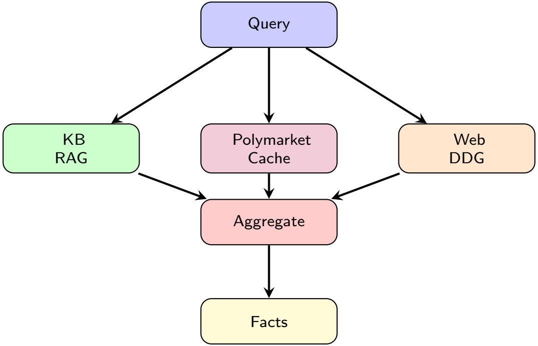
Challenging:

- Sarcastic: 65% F1
- Defensive: 68% F1

Insight

73% for 8-way classification with only 400 examples!

Multi-Source Fact-Checking Pipeline



Three Sources: Curated KB (semantic) — Polymarket (crowd opinion) — Web (current news)

Vector Database for Semantic Search

What is a Vector Database?

Stores text as numerical vectors enabling similarity-based retrieval instead of keyword matching

Keyword Search:

- Matches exact words only
- Misses synonyms

× Limited by phrasing

Semantic Search:

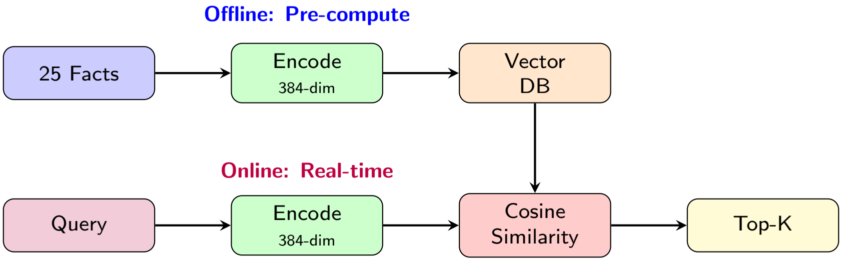
- Matches similar **meaning**
- Finds paraphrases

✓ Understands context

Our Implementation

25 facts × 384 dimensions stored in memory for fast cosine similarity search

RAG: Vector Database Architecture



- **Model:** all-MiniLM-L6-v2, **Dimension:** 384
- **Metric:** Cosine similarity, **Threshold:** 0.3 (30% match)

RAG: Knowledge Base Search

Semantic Search Algorithm

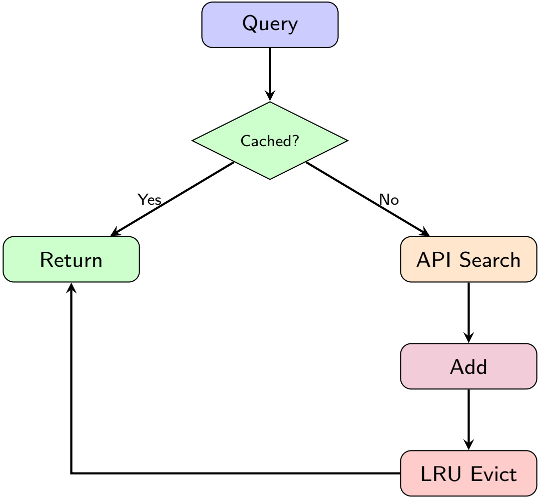
```
def search(query, top_k=3, threshold=0.3):
    # Encode query to 384-dim vector
    query_emb = model.encode(query)

    # Cosine similarity with all facts
    sims = cosine_similarity(query_emb, fact_embeddings)

    # Return top-k above threshold
    top_idx = argsort(sims)[-top_k:]
    return [facts[i] for i in top_idx if sims[i] >= threshold]
```

Key Steps: Encode query → Compute similarity → Filter by threshold → Return top matches

Polymarket Integration with LRU Caching



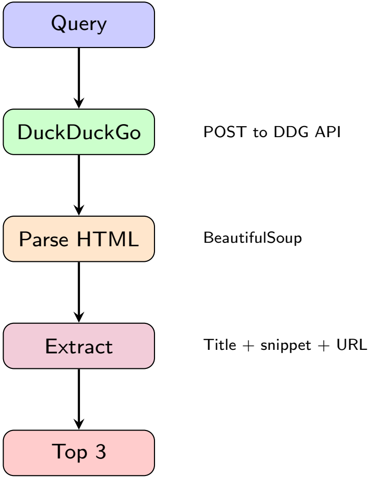
Strategy:

- 22 preloaded markets

Topics:

- Sports, Politics

Web Search Integration



- **Fallback** when KB/Polymarket lack info
- **Speed:** 1-2 sec, **Purpose:** Latest news

Use Case 1: Argument De-escalation in Microclimates

Scenario

Two roommates arguing about who should do dishes, household temperature settings, or shared space usage

How System Helps:

- **Emotion Tracking:** Detects when conversation becomes defensive or frustrated
- **Fact-Checking:** Provides objective data (e.g., "Studies show fair chore distribution reduces conflict by 40%")
- **De-escalation:** Visual emotion timeline shows patterns, encourages cooling down

Example

Speaker A: "You never do the dishes!" (Emotion: Angry 85%)
Speaker B: "That's not true, I did them yesterday!" (Emotion: Defensive 78%)
System shows: Polymarket on household chore distribution fairness
Psychology research on effective conflict resolution

Use Case 2: Corporate Conflict Resolution

Scenario

Team members disagreeing on project direction, resource allocation, or deadlines

How System Helps:

- **Meeting Analysis:** Records team discussions, tracks emotional dynamics
- **Evidence-Based:** Provides market data, industry benchmarks, case studies
- **Post-Meeting Review:** Managers review conversation for missed concerns

Example

Engineer: "Everyone in the industry is moving to microservices" (Confident)
System shows: Polymarket: 55% of companies still use monoliths
Recent survey: 40% of microservice migrations failed
Case study: When to choose monolith vs microservices

Use Case 3: Educational Debate Analysis

Scenario

Students debating controversial topics in classroom settings

How System Helps:

- **Real-Time Feedback:** Students see emotion labels, learn self-awareness
- **Source Quality:** Distinguishes opinion claims from factual claims
- **Critical Thinking:** Teaches students to verify assertions with evidence

Example

Student A: "Most people think climate change is the biggest threat" (Passionate)
System shows: Polymarket: 68% rate climate in top 3 global risks
Pew Research: Public opinion varies by age and region
Latest IPCC report on climate consensus

Use Case 4: Political Debate Fact-Checking

Scenario

Live political debates, town halls, or campaign events

How System Helps:

- **Real-Time Verification:** Fact-checks claims as candidates speak
- **Public Opinion:** Shows what voters actually think via prediction markets
- **Emotional Analysis:** Detects when candidates become defensive or evasive

Example

Candidate: "Everyone agrees my healthcare plan is the best solution" (Confident)
System shows: Polymarket: Plan has 45% approval, 38% disapproval
CBO analysis: Plan would cost $2.1T over 10 years
Kaiser poll: 52% prefer alternative approach

System Performance Metrics

Component	Metric	Value
Raspberry Pi		
Speaker Diarization	Processing Time	2-3 sec
Whisper Transcription	Processing Time	2-3 sec
Total Pi Processing	Latency	4-6 sec
AWS EC2		
Emotion Classifier	Accuracy	73.2%
Emotion Classifier	Inference Time	100 ms
RAG Search	Query Time	50 ms
Polymarket Lookup	Query Time	80 ms
Web Search	Query Time	1-2 sec
Total AWS Processing	Latency	2-3 sec
End-to-End		
Total System Latency	Pi + AWS + Network	6-10 sec

Interactive Visualization

Features:

- Color-coded speaker bubbles
- Emotion badges per segment
- Hover to see fact-checks
- KB + Polymarket + Web sources

Live Demo

`54.209.249.85:7863`

Example Output:

- Speaker emotions
- Supporting facts
- Contradicting claims

Future Work: Technical Improvements

1. **Real-Time Processing**
   - Stream audio continuously instead of 30-second windows
   - WebSocket connections for live emotion updates
   - Reduce end-to-end latency to ¡3 seconds
2. **Improved Emotion Classification**
   - Fine-tune on real conversation data (not synthetic)
   - Add audio features (pitch, volume, speaking rate)
   - Multi-modal model combining text + audio
   - Target 85%+ accuracy
3. **Enhanced Fact-Checking**
   - Integrate scholarly databases (arXiv, PubMed)
   - Add fact verification model (check claim truthfulness)
   - Expand knowledge base to 1000+ curated facts
   - Support multilingual fact-checking

Future Work: New Features

1. **Argument Summarization**
   - Generate concise summaries of key points
   - Identify areas of agreement vs disagreement
   - Extract action items and next steps
2. **Bias Detection**
   - Detect logical fallacies (ad hominem, straw man)
   - Identify cognitive biases (confirmation bias, anchoring)
   - Flag misleading statistics or cherry-picked data
3. **Multi-Speaker Scaling**
   - Support 3+ speakers in group discussions
   - Track interruptions and speaking time balance
   - Identify dominant speakers and quiet participants

Future Work: Deployment & Applications

Deployment:

- Mobile app for iOS/Android
- Zoom/Teams plugin for virtual meetings
- Smart speaker integration (Alexa, Google Home)
- Offline mode for privacy-sensitive contexts

Privacy Enhancements:

- On-device emotion classification
- Differential privacy for stored data
- User consent and data deletion
- End-to-end encryption

New Applications:

- Mental health therapy sessions
- Customer service quality monitoring
- Podcast/media content analysis
- Diplomatic negotiations
- Legal mediation

Research Directions:

- Causal analysis of de-escalation
- Longitudinal studies of conflict patterns
- Cross-cultural emotion detection
- Ethical guidelines for AI mediation

Summary

What We Built

A distributed edge-cloud system for real-time conversation analysis:

- **73.2%** emotion accuracy (8 classes)
- **3-source** fact-checking: RAG + Polymarket + Web
- **6-10 sec** end-to-end latency

Key Innovations

1. Hybrid fact-checking with prediction markets
2. LRU-cached market discovery
3. Segment-level emotion analysis
4. Privacy-preserving edge architecture

Impact

Reduces bias by providing objective facts and public opinion data

Thank You!

Questions?

Project Links:

GitHub: `github.com/ifesionubogu/aiot_project`
Live Demo: `http://54.209.249.85:7863/`
Technical Report: See `TECHNICAL_REPORT.pdf`