MET CS 777

Scalable Emotion Classification for Affective Computing Using Apache Spark

Group 6: Devin Dyson, Madhur Deep Jain
November 5, 2025

Contents

Building a Scalable Affective Computing Model

Multi-Label Text Classification with Hybrid Lexical-Dimensional Features

What is affective computing?

The study and development of systems and devices that can recognize, interpret, and respond to human emotions and affective states

Affect: relating to moods, feelings, and attitudes

Motivation

- Affective computing systems will improve Human-Computer Interaction (HCI)
- Traditional sentiment analysis (positive/negative/neutral) lacks granularity for affective computing applications

Research Question

- Can we achieve accurate multi-label emotion classification on large-scale social media text using distributed computing framework?

Technical Challenges

- Inter-rater reliability (Cohen's Kappa) consistency of raters classifying items into mutually exclusive categories more difficult with emotions
- Label imbalance: neutral/positive emotions dominate; negative emotions are rare
- Context dependency: lexical semantics shift with domain and syntax

Theoretical Foundations

History

Emergence of Affective Computing

1972 Ekman's Basic Emotions Theory

- Cross-cultural study of facial expressions → 6 universal emotions, lays the foundation for discrete categorical approaches

1980 Plutchik's Psychoevolutionary Model

- Emotion wheel: 8 primary emotions arranged in opposing pairs; Dyads: Adjacent emotions combine (joy + trust = love)

1995 Picard's "Affective Computing" Paradigm

- Seminal work establishing field: systems that recognize and respond to affect (emotion-aware interfaces improve HCI/UX)

2011 Mohammad's NRC Word-Emotion Lexicon

- Crowdsourced emotion annotations via Mechanical Turk, Enables lexicon-based features for ML pipelines

2020 GoEmotions: Fine-Grained Labeled Corpus

- 58K Reddit comments with 27 emotion labels (extends Ekman/Plutchik)
- Multi-label annotations: captures emotion complexity and co-occurrence
- Developed by Google to understand and classify the nuances of human emotion expressed in text

Theoretical Frameworks

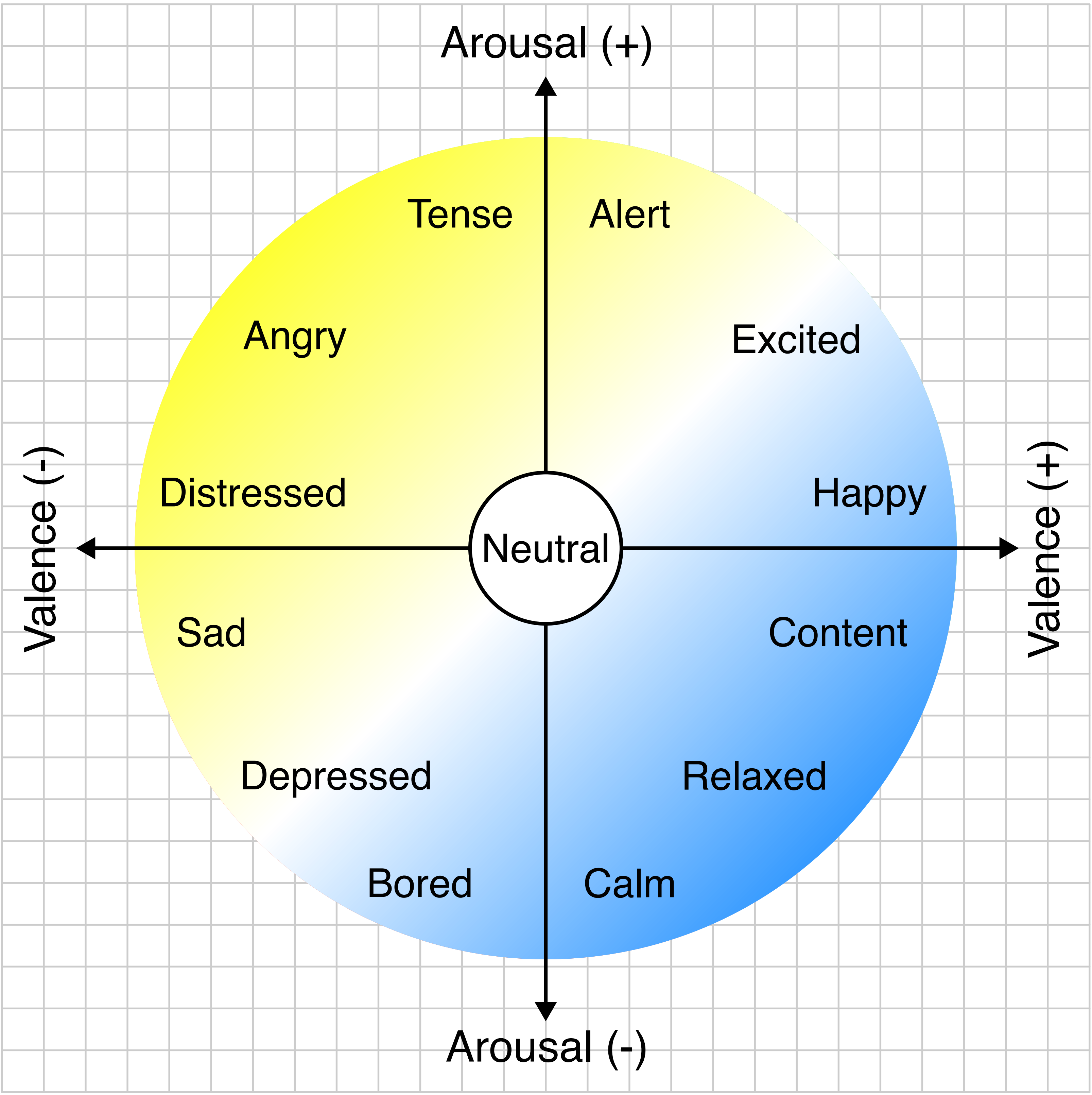
Emotion Taxonomies

Discrete Categorical Models (Ekman, Plutchik)

- Emotions as a finite set of atomic primitives
- Ekman (1972): 6 basic emotions (happiness, sadness, anger, fear, disgust, surprise)
- Plutchik (1980): 8 emotions in opposing dyads (joy-sadness, anger-fear, trust-disgust, anticipation-surprise)

Dimensional Continuous Models (Russell's Circumflex)

- Emotions as coordinates in continuous affective space
- Core dimensions:
  - Valence (V): Pleasantness [-1, +1] (negative ↔ positive)
  - Arousal (A): Activation level [-1, +1] (calm ↔ excited)
  - Dominance (D): Control [-1, +1] (powerless ↔ powerful)
- Example mapping: "Angry" = [-0.6, +0.8, -0.4] (unpleasant, high arousal, low control)



Source: Wikipedia

Sentiment vs. Affect

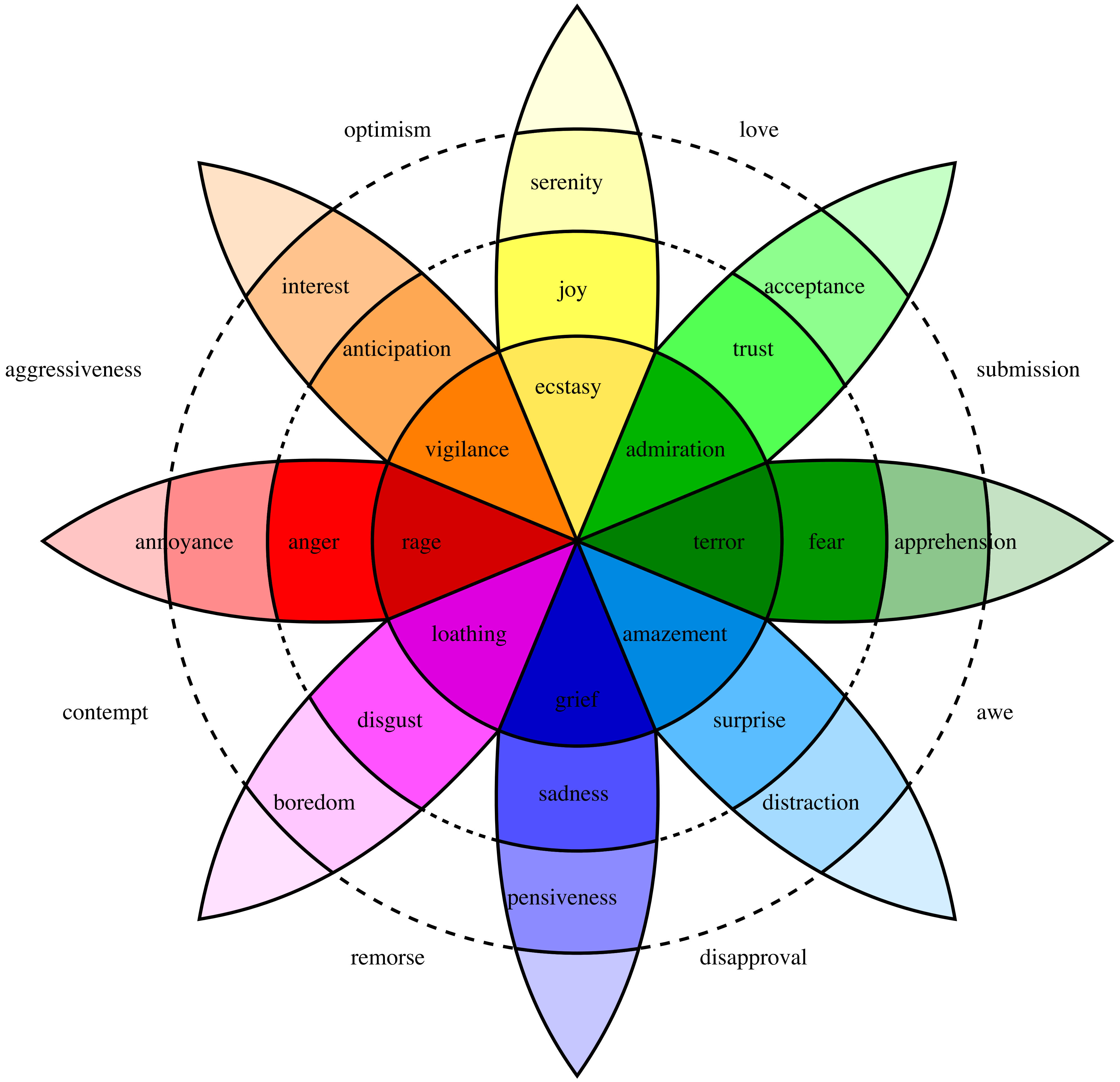
Multi-Label Emotion Classification

Sentiment Analysis

- X → {positive, negative, neutral} single-label classification
- Captures overall evaluative polarity
- Loss of granularity: “I’m furious” and “I’m disappointed” both → negative

Affective Analysis

- X → {0,1}^k where k=8
- Multi-label: each sample can have multiple emotion labels
- Captures emotional complexity: “I’m excited but nervous” → [anticipation:1, fear:1]



Source: Wikipedia

Data Sources

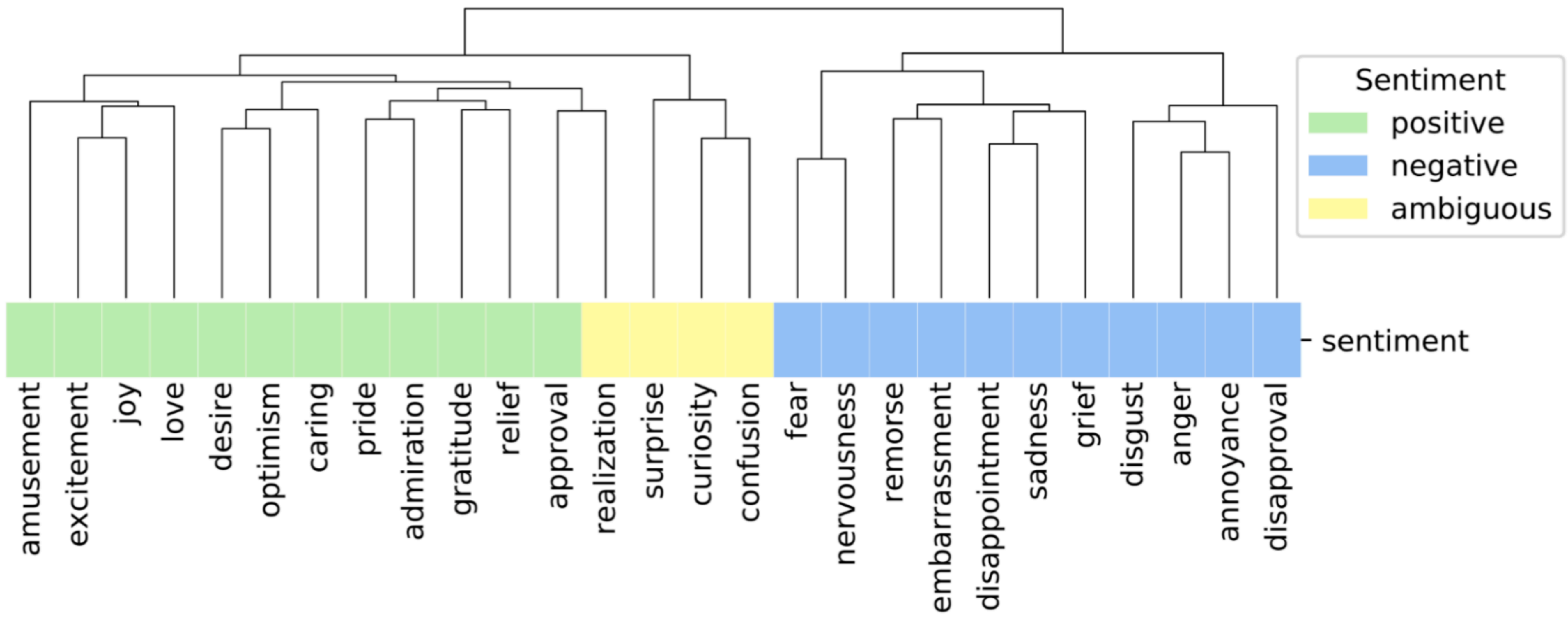
GoEmotions
Dataset Architecture

Corpus Statistics

- Source: Reddit comments (January 2005 - December 2019)
- Size: 58,009 human-annotated instances
- 27 fine-grained emotions + neutral (28 total classes)
- Multi-label: 15.3% of comments have multiple emotion annotations
- Annotators: 3 raters per comment; final label if >= 2 agree (majority vote)

Label Distribution (class imbalance)

- Skewed toward neutral (27.3%) and positive sentiment
- Rare emotions: grief (0.8%), pride (0.9%), relief (1.1%)
- Most frequent: admiration (9.2%), approval (8.1%), neutral (27.3%)



GoEmotions: https://research.google/blog/goemotions-a-dataset-for-fine-grained-emotion-classification/

NRC Word-Emotion Association Lexicon

EmoLex Overview

Lexicon Construction

- Base: 4,454 word-emotion associations
- Annotation: Amazon Mechanical Turk (crowdsourced)
- Schema: binary labels (1/0) for each of 8 Plutchik emotions per word

Feature Extraction for ML Pipeline

- Word-level matching: count words in text matching each emotion category
- Normalized ratios: emotion_count / total_words (controls for length bias)
- Sparse representation: most words map to 0-2 emotions (lexicon coverage ~15% of vocabulary)

Example

- text = “I’m depressed. But happy because I love banana. Please don’t abandon me”
- tokens = [depressed, happy, love, banana, abandon]
- feature vector: [anger: 1, anticipation: 1, disgust: 0, fear: 2, joy: 2, sadness: 2, surprise: 0, trust: 1]
- normalized: [anger: 0.111, anticipation: 0.111, disgust: 0.0, fear: 0.222, joy: 0.222, sadness: 0.222, surprise: 0.0, trust: 0.111]

depressed	anger	1
depressed	anticipation	0
depressed	disgust	0
depressed	fear	1
depressed	joy	0
depressed	negative	1
depressed	positive	0
depressed	sadness	1
depressed	surprise	0
depressed	trust	0

happy	anger	0
happy	anticipation	1
happy	disgust	0
happy	fear	0
happy	joy	1
happy	negative	0
happy	positive	1
happy	sadness	0
happy	surprise	0
happy	trust	1

love	anger	0
love	anticipation	0
love	disgust	0
love	fear	0
love	joy	1
love	negative	0
love	positive	1
love	sadness	0
love	surprise	0
love	trust	0

banana	anger	0
banana	anticipation	0
banana	disgust	0
banana	fear	0
banana	joy	0
banana	negative	0
banana	positive	0
banana	sadness	0
banana	surprise	0
banana	trust	0

abandon	anger	0
abandon	anticipation	0
abandon	disgust	0
abandon	fear	1
abandon	joy	0
abandon	negative	1
abandon	positive	0
abandon	sadness	1
abandon	surprise	0
abandon	trust	0

NRC VAD Lexicon

Valence, Arousal, Dominance

Dimensional Lexicon

- Coverage: 19,970 words
- Annotation: Best-Worst Scaling (BWS) via Mechanical Turk
- Scale: [0, 1] continuous values for each dimension
- Scores in the 0.9 range for Spearman and Pearson correlation with ground truth

Sample Feature Engineering

- text = "I failed the exam. I'm ruined."
- words = [failed, exam, ruined]
- VAD lookups:
  - failed: (V:0.177, A:0.578, D:0.154) - low valence, high stress
  - exam: (V:0.479, A:0.663, D:0.577) - neutral-negative, high arousal
  - ruined: (V:0.135, A:0.564, D:0.194) - very negative, high stress
- Aggregated features:
  - VAD_mean = (0.260, 0.610, 0.308) → negative + stressed + low control
  - VAD_std = (0.156, 0.045, 0.198) → consistent negativity, variable control

fail	0.177	0.578	0.154
exam	0.479	0.663	0.577
ruined	0.135	0.564	0.194

```
text = "I failed the exam. I'm ruined."

def extract_vad_features(text):
    words = re.findall(r'\b[a-z]+\b', text.lower())

valence_sum = 0.177 + 0.479 + 0.125 = 0.781
arousal_sum = 0.578 + 0.663 + 0.588 = 1.829
dominance_sum = 0.154 + 0.577 + 0.192 = 0.923

count = 3 # (failed, exam, ruined)
vad_features = [
    0.781 / 3 = 0.260, # valence (extremely negative)
    1.829 / 3 = 0.610, # arousal (high stress)
    0.923 / 3 = 0.308 # dominance (very low control)
    ]

Text 1: "I'm annoyed by this minor inconvenience"
Text 2: "I'm devastated and utterly destroyed"

# NRC emotion counts might both show:
# - anger: 1
# - sadness: 1
# can't distinguish INTENSITY of negative emotion
```

Technical Approach

Pipeline

Data Preparation

- Load GoEmotions: 58k samples → Spark DataFrame
- Label Mapping: 27 fine-grained → 8 Plutchik emotions (many-to-one hierarchical mapping)
- Stratified Split: 70% train (40,606), 30% test (17,403) preserving label distribution
- Text Cleaning: lowercase, tokenization, URL removal, stop-word removal

Pipeline

Feature Engineering (Multi-Modal Representation)

TF-IDF with n-grams (unigrams + bigrams + trigrams)

- Captures domain-specific vocabulary and rare discriminative terms

Lexicon-Based Emotion Features

- Raw counts and Normalized ratios (length invariance) of NRC emotions
- Captures explicit and interpretable emotion signals

Dimensional Affect Features (VAD)

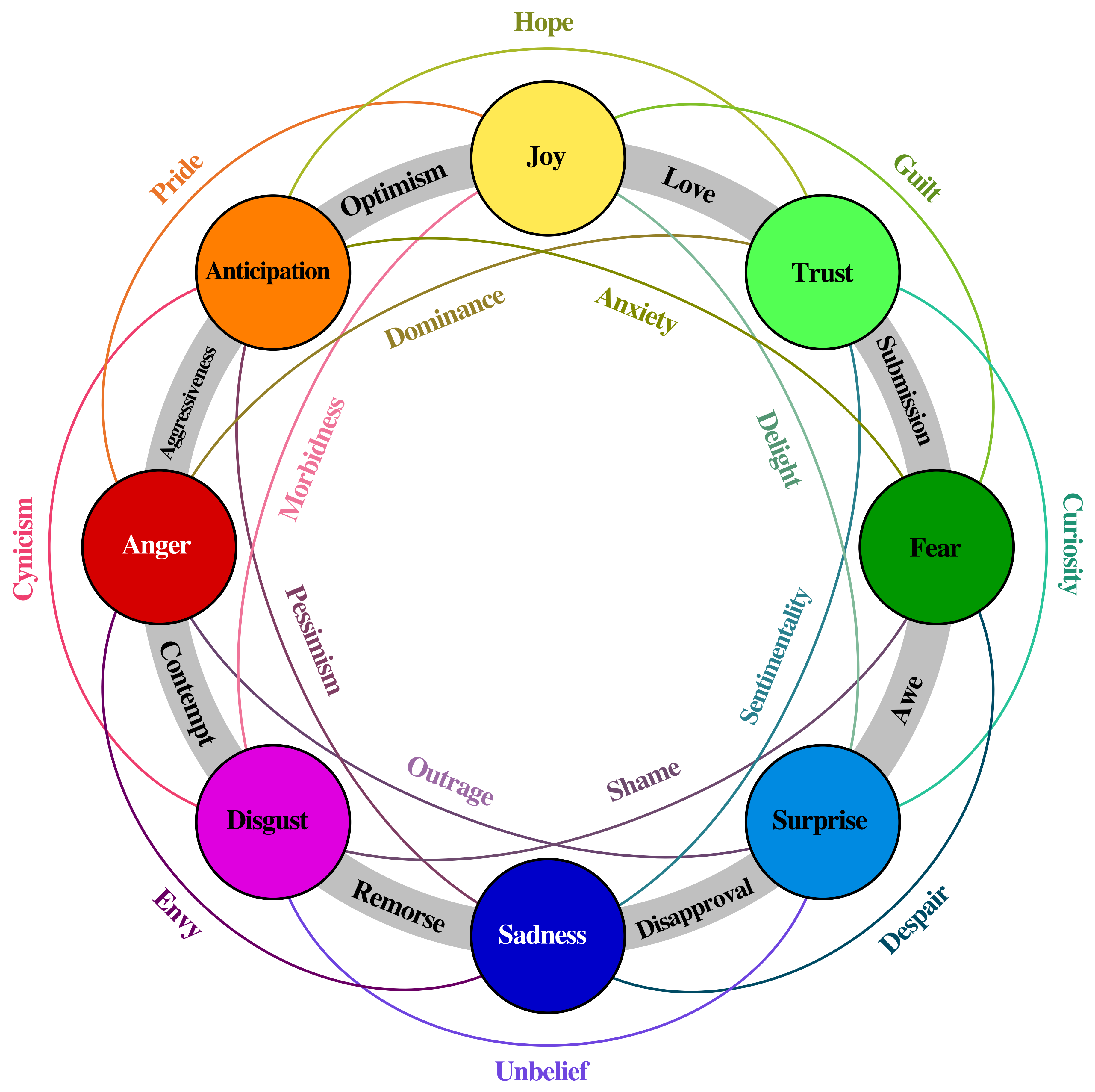
- Aggregates (mean, std and range) of VAD for each word
- Captures emotion intensity and consistency across dimensions

Linguistic Signals

- Length: word_count, char_count
- Punctuation density + emphasis markers: exclamation_ratio, question_ratio, ellipsis_count, CAPS_ratio, repeated_chars (!!!, ???)

Vector Assembly

- Concatenation: x = [TF-IDF ⊕ NRC ⊕ VAD ⊕ Linguistic]



Source: Wikipedia

Pipeline

Model Zoo

Model Training Strategy

- For each emotion $e_i \in$ {joy, trust, fear, surprise, sadness, disgust, anger, anticipation}:
  - Binary labels: $y_i \in \{0, 1\}$
  - Train classifier: $h_i(x) \to [0, 1]$ (probability)
  - Algorithms: Logistic Regression, Linear SVM, Naive Bayes, Random Forest
- Parallelization: 8 independent classifiers trained concurrently

Optimization Techniques

- Broadcast Variables: share NRC/VAD lexicons across executors (avoid data transfer)
- Threshold Tuning: per-emotion optimal thresholds via ROC analysis (maximize F1)
- Ensemble Aggregation: majority voting across 4 algorithms with confidence weighting

Results & Demo

Sample Output and Evaluation Metrics

Key Findings

- Metrics (for the majority vote ensemble)
  - F-1 score: 0.51
  - Precision: 0.51
  - Recall: 0.55
- Positive emotions (joy, trust) achieve 15-20% higher F1
- Logistic Regression - best individual classifier

Demo

- Combines top emotions into a narrative
- E.g., “This demo is awesome” -> “Steady notes of trust lead the narrative, with a uplifting undertone from joy.”


```
{
  "text": "Wow this is amazing. LOVE THIS!!!",
  "predictions": [
    {
      "label": "joy",
      "predicted": true,
      "score": 0.8007311486763947,
      "raw": null,
      "threshold": 0.4
    },
    {
      "label": "trust",
      "predicted": true,
      "score": 0.9170571444419771,
      "raw": null,
      "threshold": 0.3
    },
    {
      "label": "fear",
      "predicted": false,
      "score": 0.008815574638560714,
      "raw": null,
      "threshold": 0.3
    },
    {
      "label": "surprise",
      "predicted": true,
      "score": 0.5825290608402673,
      "raw": null,
      "threshold": 0.3
    },
    {
      "label": "sadness",
      "predicted": false,
      "score": 0.023865459561988243,
      "raw": null,
      "threshold": 0.3
    },
    {
      "label": "disgust",
      "predicted": false,
      "score": 0.015072881541916815,
      "raw": null,
      "threshold": 0.3
    },
    {
      "label": "anger",
      "predicted": false,
      "score": 0.04585021930400218,
      "raw": null,
      "threshold": 0.3
    },
    {
      "label": "anticipation",
      "predicted": false,
      "score": 0.05138461374788547,
      "raw": null,
      "threshold": 0.4
    }
  ],
  "positive_labels": [
    "joy",
    "trust",
    "surprise"
  ],
  "top_labels": [
    "trust",
    "joy",
    "surprise",
    "anticipation",
    "anger"
  ],
  "story": "Steady notes of trust lead the narrative (92%), with a uplifting undertone from joy (80%)."
},
```

Applications and Future Directions

Real-World Impact in HCI, Mental Health, and Content Moderation

Emotion-Aware Human-Computer Interaction

- Affectiva (affectiva.com) emotion AI for automotive, media, and market research
  - In-cabin monitoring: Detects driver drowsiness, distraction, and emotional state, used by BMW, Honda, Kia…
- Amazon Alexa Emotion Detection (Patent US10573312B1)
  - Detects frustration in voice to adjust response strategy

Mental Health Monitoring & Crisis Detection

- Crisis Text Line (crisistextline.org) emotion analysis for suicide prevention
  - ML models triage incoming messages by emotional distress level, partners with: Facebook, Instagram for at-risk user detection
- OpenAI Safety Systems (openai.com/safety)
  - ChatGPT detects sensitive conversations (self-harm, crisis), provides crisis resources (988 Suicide & Crisis Lifeline)

Content Moderation at Scale

- Discord's AutoMod (discord.com/safety)
  - Real-time emotion analysis to flag hostile language, includes customizable thresholds for different community standards

Future Research Directions

- Zoom's Emotion AI (Patent 2021) meeting sentiment analysis through multimodal fusion of text + voice + facial expressions
- USC's SimSensei virtual therapist funded by DARPA used microphone, webcam, and depth sensors to track 60 non-verbal behaviors per second, including facial expressions, eye gaze, body posture, and voice intonation

References

[1] Rosalind W. Picard. 1995. Affective Computing. Technical Report 321. MIT Media Laboratory, Cambridge, MA.

[2] Paul Ekman. 1992. An argument for basic emotions. Cognition & Emotion 6, 3-4 (1992), 169–200.

[3] Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In Theories of Emotion, Robert Plutchik and Henry Kellerman (Eds.). Academic Press, New York, NY, 3–33.

[4] James A. Russell. 1980. A circumplex model of affect. Journal of Personality and Social Psychology 39, 6 (1980), 1161–1178.

[5] Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a Word-Emotion Association Lexicon. Computational Intelligence 29, 3 (2013), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

[6] Saif M. Mohammad. 2018. Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, Melbourne, Australia, 174–184. https://doi.org/10.18653/v1/P18-1017

[7] Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. GoEmotions: A Dataset of Fine-Grained Emotions. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, 4040–4054. https://doi.org/10.18653/v1/2020.acl-main.372

[8] Daniel Jurafsky and James H. Martin. 2024. Speech and Language Processing (3rd ed. draft). Chapter 20: Lexicons for Sentiment and Affect. Retrieved from https://web.stanford.edu/~jurafsky/slp3/

Q&A