

Week 2: Discovery Handout

Understanding Emotions in Text

ML/AI for Design Thinking

Pre-Class Exploration

Your Mission

Before we dive into theory, let's discover why understanding emotions in text is harder than it seems!

Work through these exercises. Use the theory hints when stuck. Discuss findings with your partner.

Part A: Sentiment Discovery (10 min)

Exercise 1: Rate these reviews as Positive (+), Negative (-), or Neutral (0):

- | | |
|--|------------------|
| a) "This app is absolutely perfect if you enjoy frustration" | Your rating: ___ |
| b) "Not bad at all" | Your rating: ___ |
| c) "It works." | Your rating: ___ |
| d) "Can't complain about the price, you get what you pay for" | Your rating: ___ |
| e) "I've never been so impressed... by how slow an app can be" | Your rating: ___ |

Question: Which ones were hardest to classify? Why?

Theory Hint 1

Sarcasm uses positive words to express negative sentiment. Look for:

- Exaggerated positive language ("absolutely perfect")
- Unexpected endings ("...if you enjoy frustration")
- Context reversal

Double negatives ("not bad") often express mild positivity.

Part B: Context Changes Everything (10 min)

Exercise 2: The word “fast” appears in these reviews. Determine the sentiment:

- | | |
|---------------------------------------|----------------|
| a) “The delivery was incredibly fast” | Sentiment: ___ |
| b) “The battery drains fast” | Sentiment: ___ |
| c) “Fast customer service response!” | Sentiment: ___ |
| d) “Too fast to be thorough” | Sentiment: ___ |

Discovery Question: Can a single word have fixed sentiment? Explain:

Theory Hint 2

Bidirectional Understanding: BERT reads text in both directions simultaneously:

- Traditional: Read left → right only
- BERT: Read left → right AND right → left
- Result: “fast” + “battery drain” = negative context
- Result: “fast” + “delivery” = positive context

This is why BERT understands “not bad” as positive (sees “not” + “bad” together).

Part C: What Words Matter Most? (10 min)

Exercise 3: In this review, underline the 3 most important words for understanding sentiment:
“The interface is clean and intuitive, but the app crashes constantly, making it completely unusable despite the beautiful design and smooth animations when it actually works.”

Your top 3 words: _____ , _____ , _____

Partner Discussion: Did you choose the same words? Why might they differ?

Theory Hint 3

Attention Mechanism: BERT assigns importance scores to words:

- High attention: “crashes”, “unusable”, “but”
- Lower attention: “the”, “and”, “is”
- Context-dependent: “works” gets high attention near “actually”

Formula preview: Attention = How much should I focus on each word?

Part D: Beyond Positive/Negative (10 min)

Exercise 4: Match these reviews to emotions (Joy, Anger, Fear, Surprise, Sadness):

- | | |
|---|----------------|
| a) "I'm worried my data isn't secure" | Emotion: _____ |
| b) "Can't believe how much this has improved!" | Emotion: _____ |
| c) "Disappointed it doesn't work on older phones" | Emotion: _____ |
| d) "This is absolutely unacceptable!" | Emotion: _____ |
| e) "Exceeded all my expectations!" | Emotion: _____ |

Design Question: How would you redesign a product differently for users feeling fear vs anger?

Theory Hint 4

Emotion Categories: Modern NLP goes beyond binary sentiment:

- | | |
|---|---|
| • Joy: excitement, satisfaction, delight | pected |
| • Anger: frustration, irritation, rage | • Sadness: disappointment, regret, loss |
| • Fear: worry, concern, anxiety | • Disgust: rejection, aversion, distaste |
| • Surprise: amazement, shock, unexpected | |

Each emotion → different design response!

Part E: Words in Space (5 min)

Exercise 5: If words were points in space, which pairs would be closest?

- | | |
|---------------------|-------------------|
| a) good ↔ excellent | d) fast ↔ slow |
| b) good ↔ bad | e) happy ↔ joyful |
| c) fast ↔ quick | f) happy ↔ sad |

Circle the closest pairs. **Bonus:** Can you draw a simple 2D map showing these relationships?

Theory Hint 5

Word Embeddings: BERT converts words to 768-dimensional vectors:

- Similar meaning = closer in space
- "good" is closer to "excellent" than to "bad"
- Context changes position: "bank" (river) vs "bank" (money)
- We use t-SNE to visualize in 2D (like a map projection)

Synthesis: Connect to Design (5 min)

Design Thinking Connection

Final Reflection: Based on your discoveries, answer:

1. **Empathy Challenge:** If you only counted positive/negative keywords, what emotions would you miss?
2. **Scale Problem:** You have 10,000 reviews. How long to read them all? (Assume 30 seconds each)
3. **Design Impact:** Name one feature you'd build differently after understanding hidden emotions:

Prepare for Class

What we'll explore today:

- How BERT achieves 87% sarcasm detection (vs 15% for keywords)
- Why bidirectional reading changes everything
- Attention mechanisms: the key to understanding
- Real examples: Netflix emotions, Airbnb hostile review detection
- Hands-on: Using BERT to analyze product reviews

Key Terms to Remember: BERT, Transformer, Attention, Embedding, Bidirectional, Context, Fine-tuning