WHAT MAKES A POST FUNNY?

Humor in Headlines

Basmah Alrashid, Noura Alamro, Reema Alraqibah, Shooq Alawadah
Supervised by: Dr. Abeer Aldayeel

Introduction

Decoding What Makes Us Laugh Online

What makes a social media post funny? Is it the words, the length, or the emojis? Our project, "Humor in Headlines," uses Data Science to decode the patterns of online humor. By analyzing thousands of Reddit posts, we explored how language, community, and context influence what we find funny. This research helps us understand digital culture and teaches AI to better understand human communication.

Objectives

We aimed to answer one central question with two key tasks:

How does language and social context influence the structure and perception of humor in short texts?

Key Tasks:

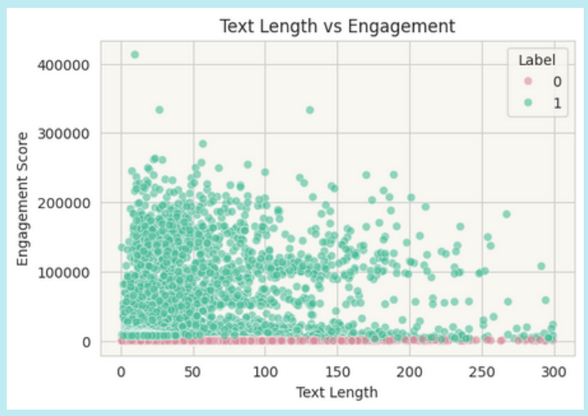
1. Humor Detection: Build a model to automatically classify a post as "Humorous" or "Not Humorous."
2. Sociolinguistic Analysis: Uncover how humor styles vary across different online communities.

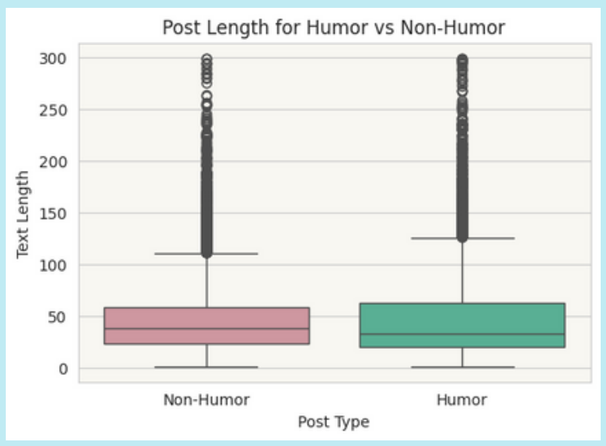
Data Collection

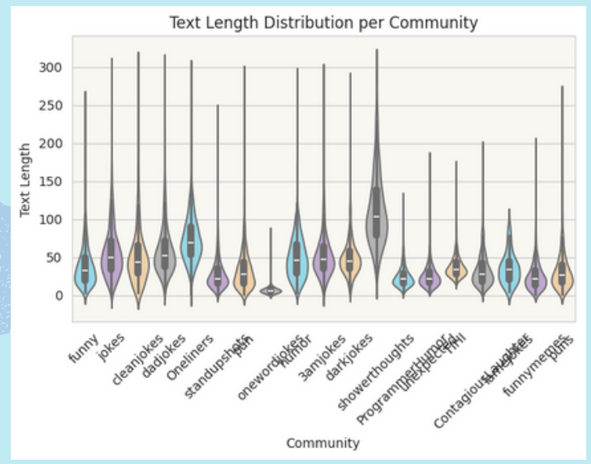
1. Primary Source: Reddit
- 19 different communities (subreddits) analyzed, including r/funny, r/memes, r/unexpected, and r/ProgrammerHumor.
- Collected 20,804 posts using the Reddit API (PRAW).
- Data included post titles, upvotes, number of comments, and community name

2. Secondary Source: Short Jokes
- Augmented with 231,657 traditional one-liner jokes from Hugging Face.
- Used to teach our AI a wider vocabulary of humor.

Data Analysis







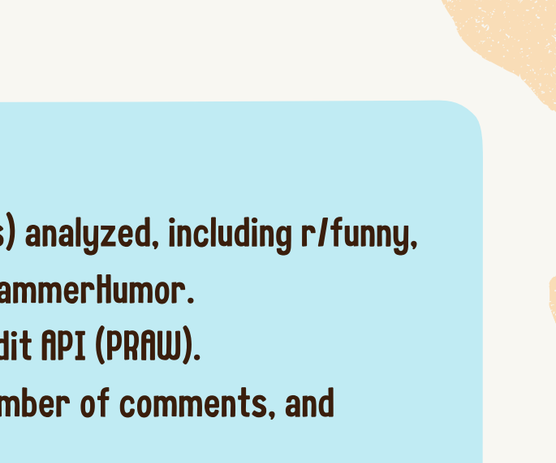
Models and Findings

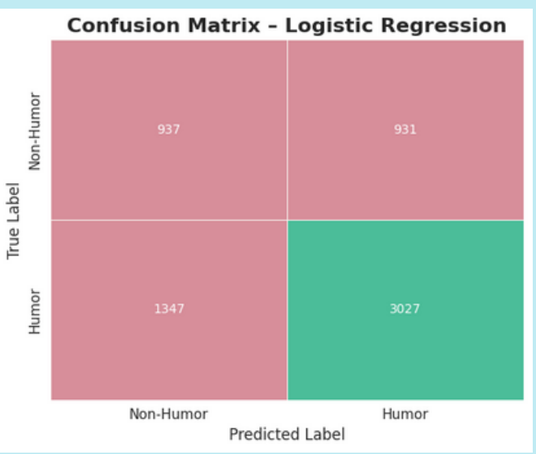
We built two AI models to detect humor automatically. A simpler, faster Logistic Regression model analyzed word patterns, while a more complex Recurrent Neural Network (RNN) learned from sentence structure.

The Results:

- Logistic Regression outperformed the RNN (63.5% vs 62.7% accuracy)
- Simple proved better than complex for this task
- Both models struggled with non-humor - showing how subjective comedy can be
- Informal language was a strong humor signal - slang and casual wording consistently predicted funny content

Key Insight: Even advanced neural networks couldn't beat traditional methods, proving that basic word patterns are surprisingly effective for humor detection.





Conclusions

Main Insights:

Brevity is key - Successful humor posts average only 46 characters

Context rules - Humor varies dramatically by community (99% success in r/unexpected vs 7% in r/pun)

Emojis don't drive humor - Only 1.7% of posts use emojis with no significant impact

Simple AI works best - Logistic Regression (63.5% accuracy) beat complex neural networks

Implications:

For Content Creation: Focus on short, community-aware posts rather than visual embellishments

For AI Development: Simple, interpretable models can effectively detect humor patterns

For Digital Communication: Understanding community-specific humor patterns enhances online engagement