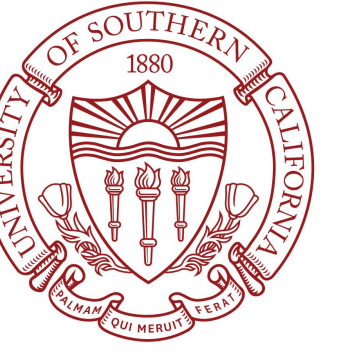


# Swearing **#&%ING** impact on social media engagement

Stefan Niculae, Kendall Crouch, Jeffrey Stone, Santiago Carvajal, Timothy Rhodes



## Abstract

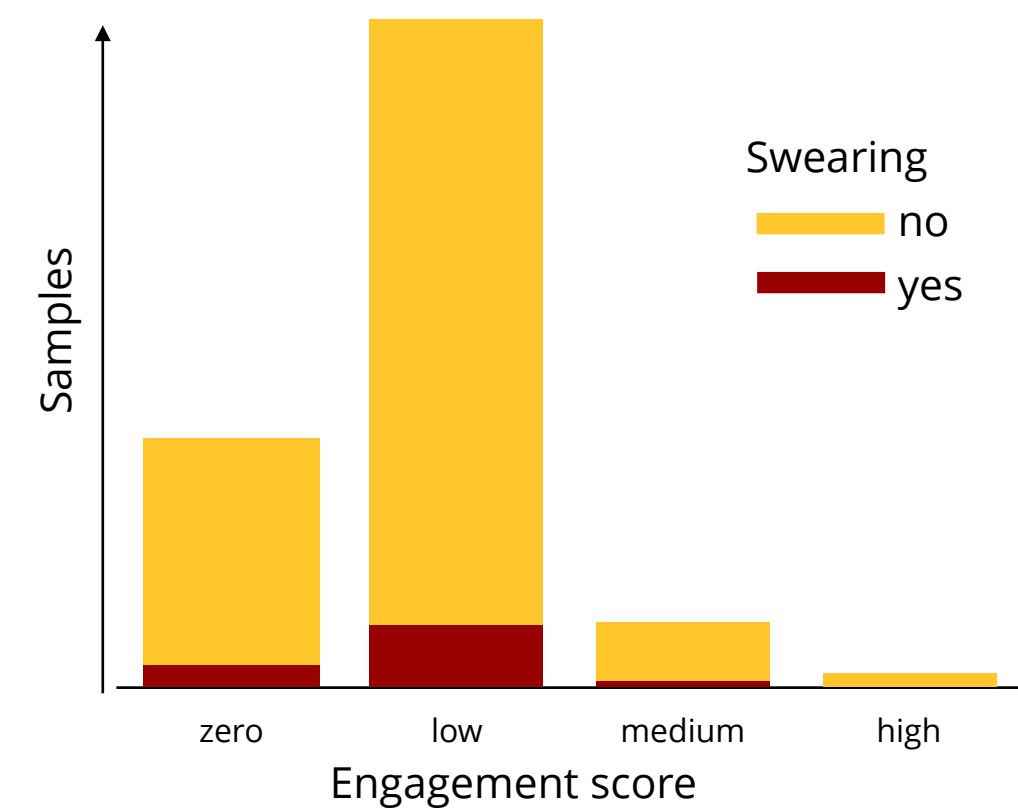
**Question:** How does swearing affect social media engagement?

**Motivation:**

- social media is an integral part of today's life,
- engagement is a core goal sought by individuals and invaluable to companies,
- swearing is an unexplored angle due to its taboo nature,
- understanding is the first step to handling profanities, and other controversial topics.

**Approach:** collect and curate data › craft and analyze features › build and interpret models.

## Data & Features



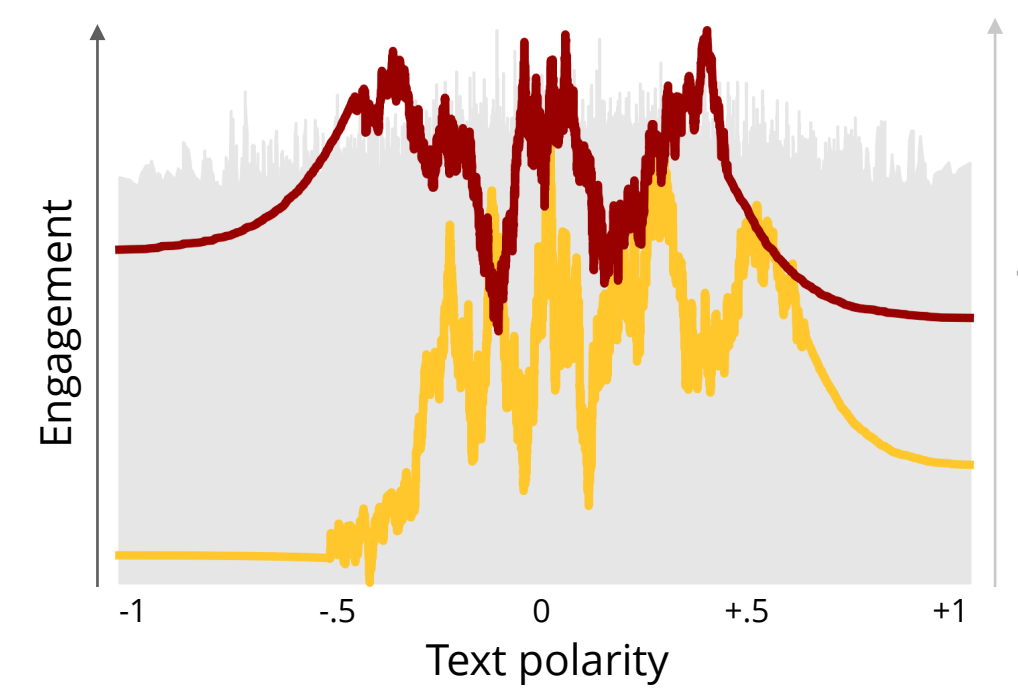
**Data:** 50M tweets, special attention to top 100 most popular users.

**Preprocessing:** general cleaning, lemmatization, emoji transcribing,

Engagement score:

$$\log\left(1 + \frac{\text{likes} + 5 \times \text{retweets}}{\text{followers}} \times 100\right)$$

5.4% of all tweets use swear words.

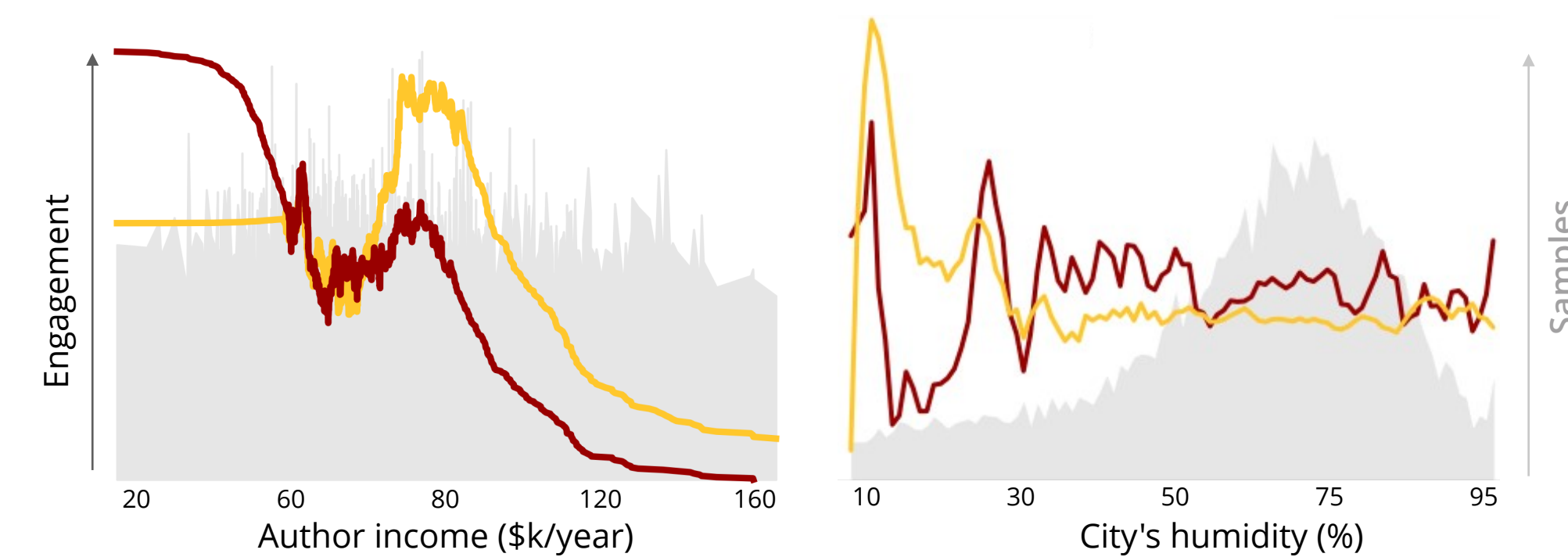


**Textual features:** sentiment polarity, subjectivity, topic modeling, parts of speech, affective, cognitive nuances, punctuation, domain-specific: hashtags, mentions, links.

**Swearing features:** count, presence, severity, rarity, censorship.

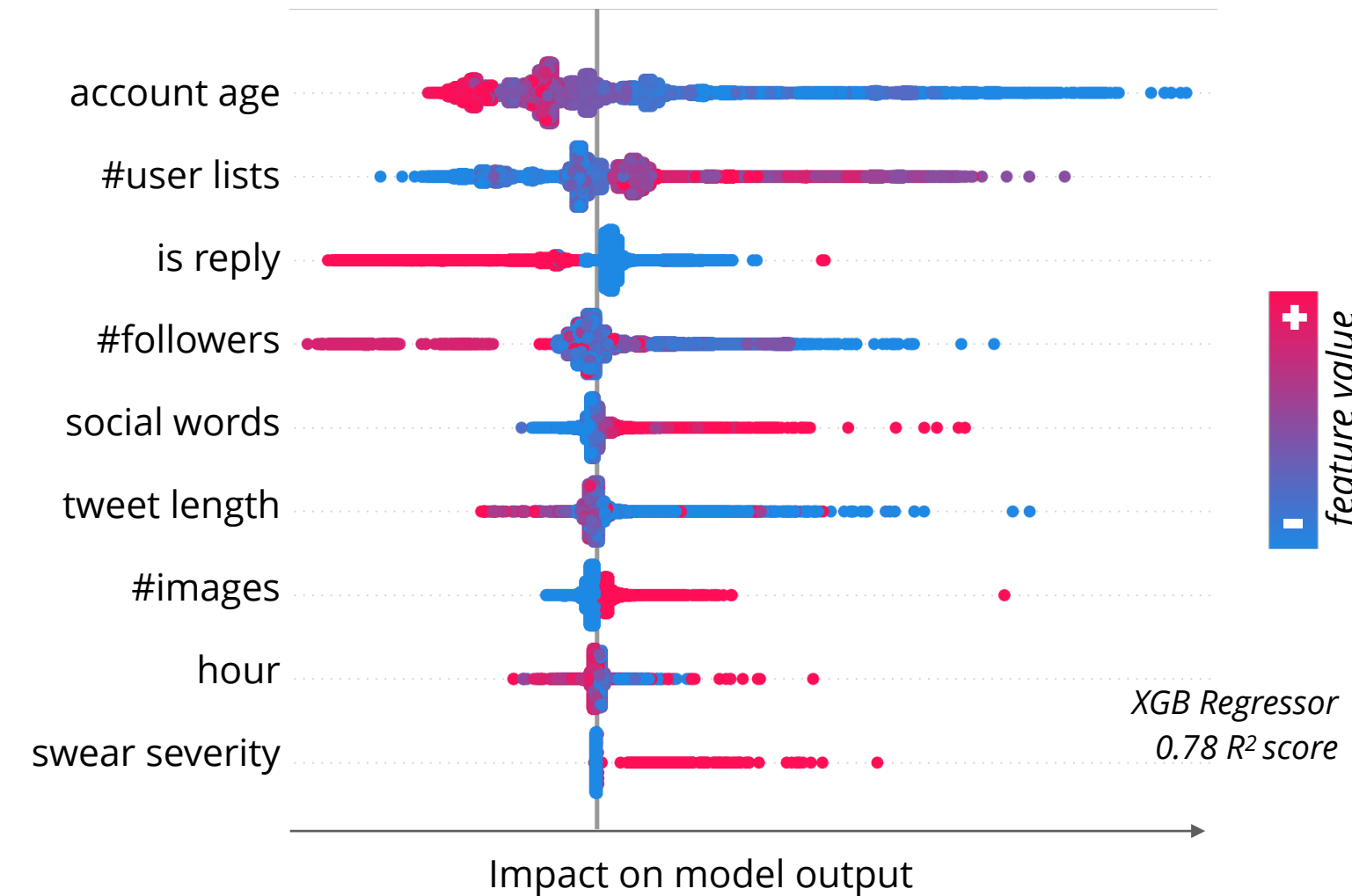
Women get slightly higher engagement, regardless of swearing

**Advanced features:** demographics (education, income, gender) image content and hue, weather (temperature, humidity, conditions), date (hour, weekday, holiday status), semantic category (arts, technology, etc), user type (singer, actor, etc).



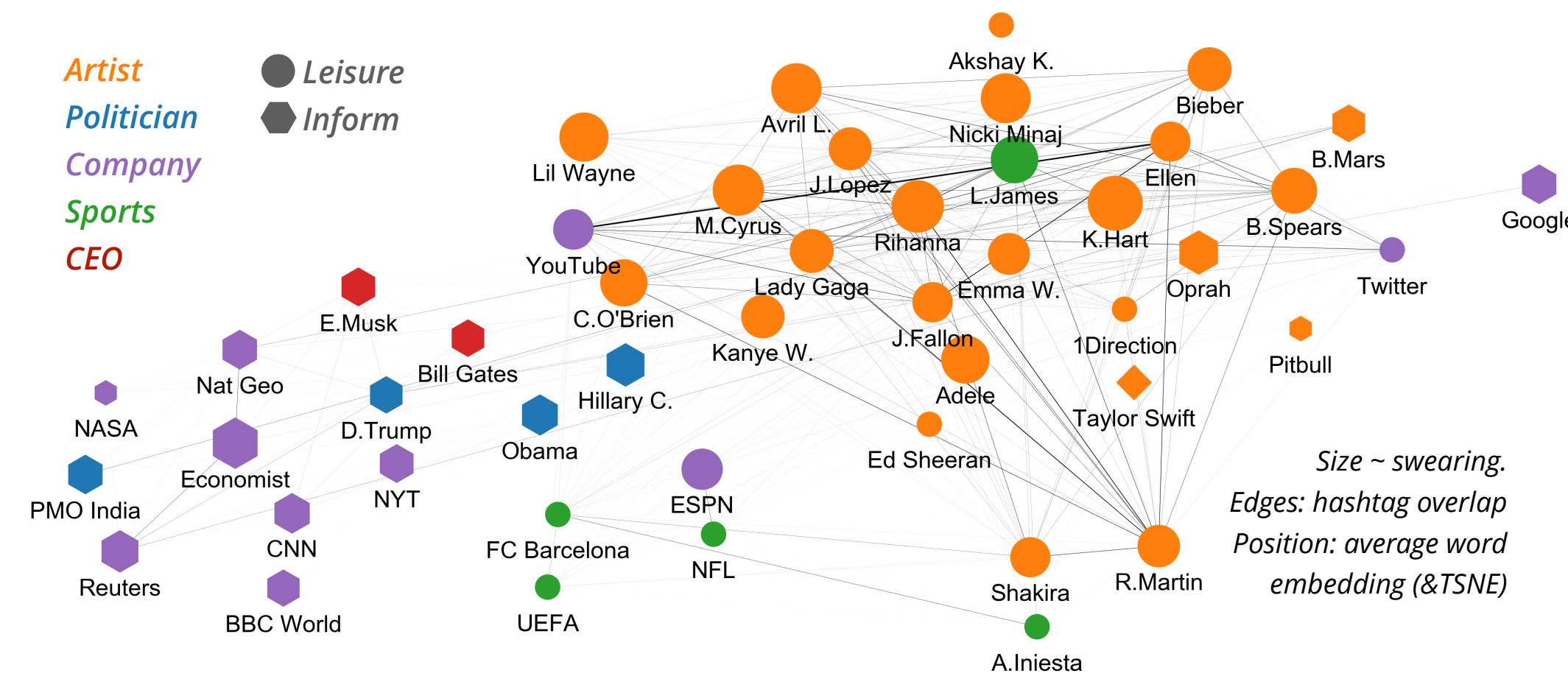
## Analysis

### FEATURE IMPORTANCES



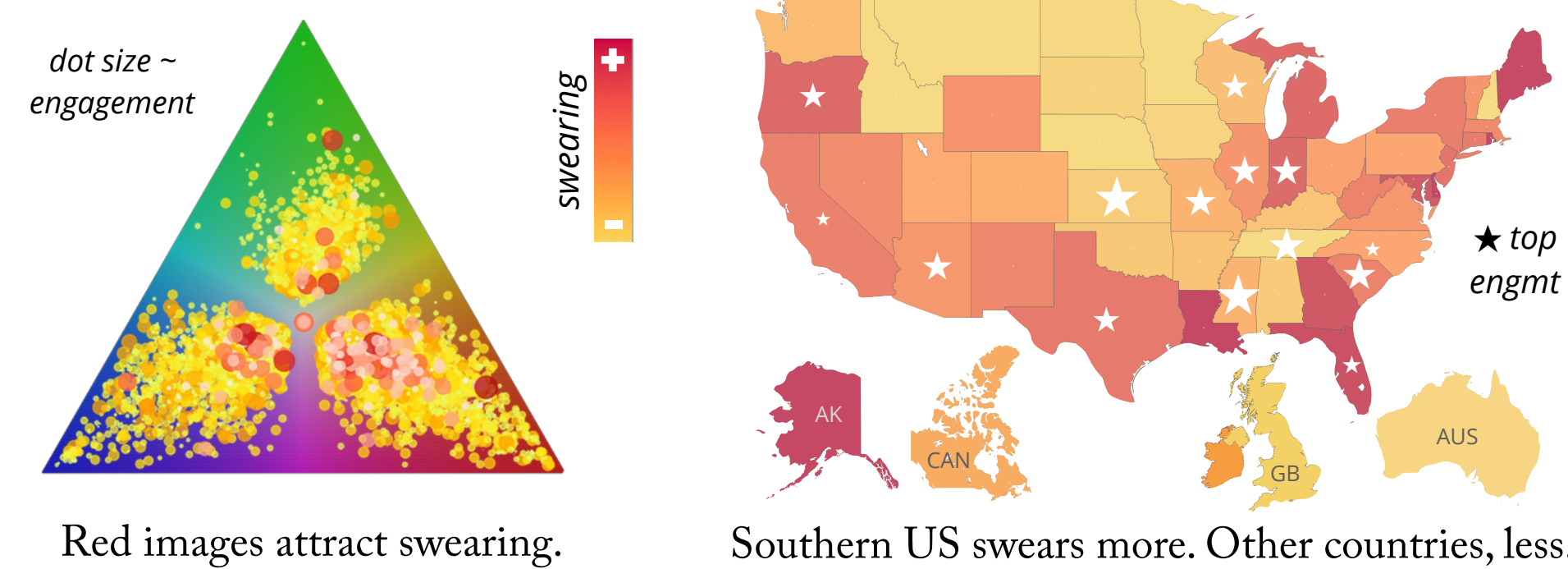
Including images, being part of lists and not mentioning positively impact on engagement. Swearing, when (rarely) used, almost always causes an increase in engagement.

### TOP USERS SEMANTIC CORRELATIONS



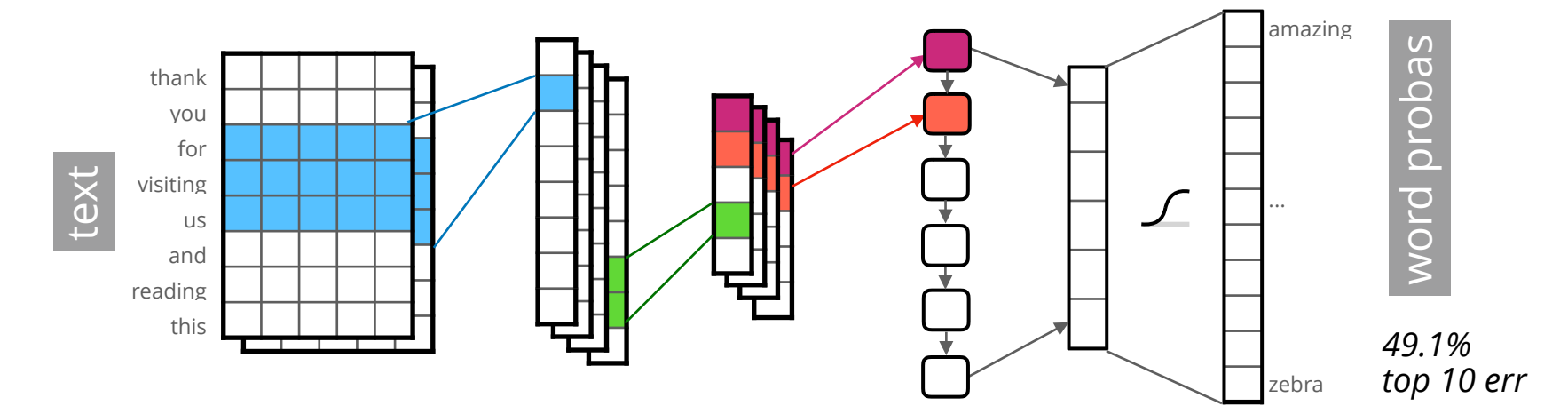
Tweet accounts conform to their category, semantic content average and intent trends.

### SWEARING & ENGAGEMENT



## ML Models

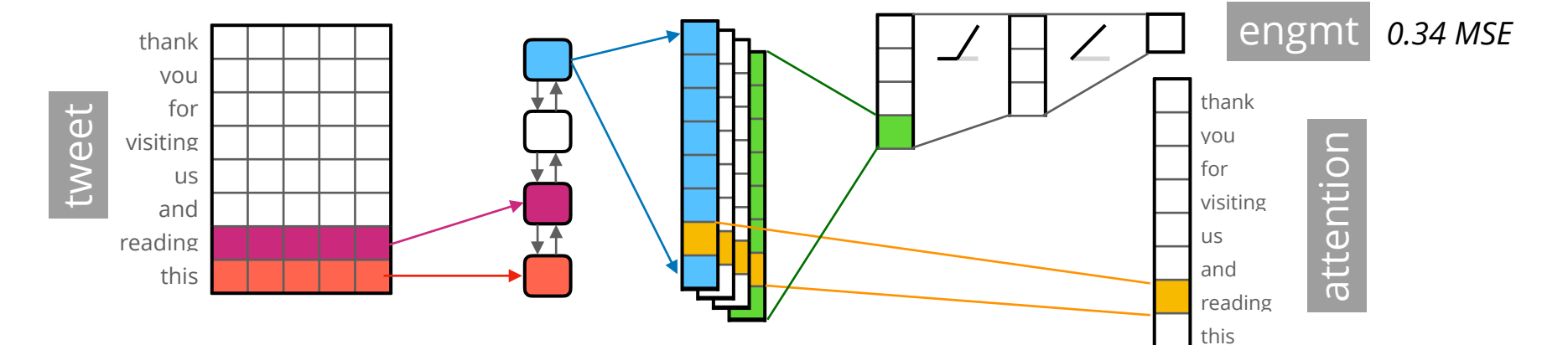
**Placement:** where and which swear word fits into this text?



Reuters, NBA, Google | B.Mars | K.Hart, C.O'Brien  
least | swearing fitting | most

- Google: Hi there, **dummy**, try these tips if you're unable to connect [...]
- Barack Obama: Happy 4th of July. God bless you, God f\*\*\* bless America!
- Emma Watson: Don't forget to vote on Thursday, no **damn** excuses!

**Attention:** what'd be a tweet's engagement, and what words matter the most?

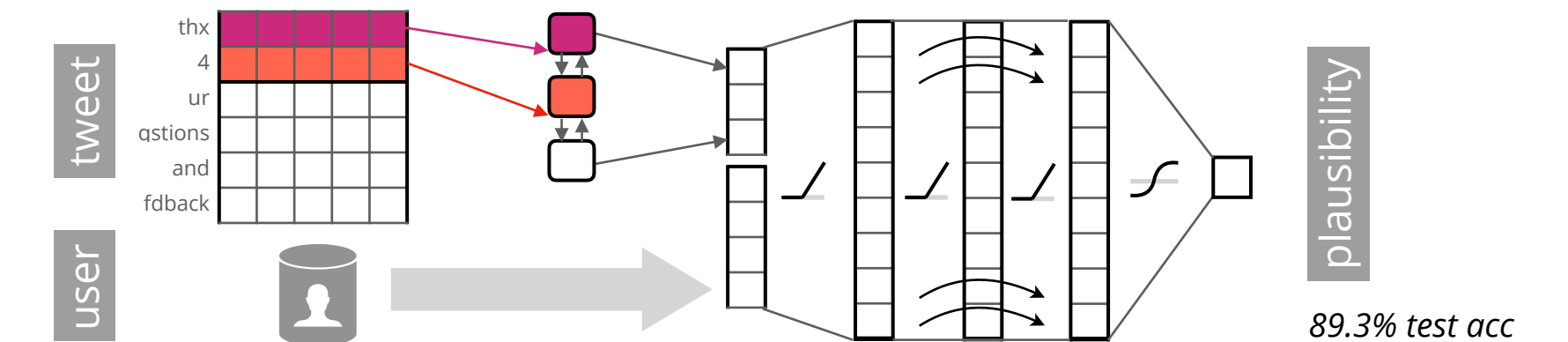


B.Gates, M.Carey, L.James | B.Mars | Rihanna, K.Hart, Z.Malik  
decrease | artificial tweets engagement | increase

When a swear word was most focused (2.23x more likely), actual engagement was +28%.

- You will get there... But it's so **f\*\*\* hard** and frustrating at first.
- Screaming "**congrats**" to **amp** my brother! Love the s\*\*\* out of you!

**Discriminator:** how characteristic would this tweet be for this user?



On artificial tweets: 7.14% classified as plausible (6.8% average plausibility).

## Conclusions

1. Swearing, while not very prevalent, increases engagement in certain contexts.
2. Swearing roles: aggression / emphasis /affection /comedic effect.
3. Tweet- and user-features are predictive of engagement.

**Next steps:**

- Study long term ramifications of swearing, and effects of incorporating swearing
- Personalized real-time feedback to optimize social media engagement
- Other platforms: reviews, longer text, anonymity, video, correlation to current events