INSPAIRED: Cross-cultural Inspiration Detection and Analysis in Real and LLM-generated Social Media Data

Oana Ignat[1][*]   **Gayathri Ganesh Lakshmy**[2][*]   **Rada Mihalcea**[3]

[1]Santa Clara University - Santa Clara, USA
[2]Carnegie Mellon University - Pittsburgh, USA
[3]University of Michigan - Ann Arbor, USA

oignat@scu.edu   gganeshl@andrew.cmu.edu   mihalcea@umich.edu

Abstract

Inspiration is linked to various positive outcomes, such as increased creativity, productivity, and happiness. Although inspiration has great potential, there has been limited effort toward identifying content that is *inspiring*, as opposed to just engaging or positive. Additionally, most research has concentrated on Western data, with little attention paid to other cultures. This work is the first to study cross-cultural inspiration through machine learning methods. We aim to identify and analyze real and AI-generated cross-cultural inspiring posts. To this end, we compile and make publicly available the INSPAIRED dataset, which consists of 2,000 real inspiring posts, 2,000 real non-inspiring posts, and 2,000 generated inspiring posts evenly distributed across India and the UK. The real posts are sourced from Reddit, while the generated posts are created using the GPT-4 model. Using this dataset, we conduct extensive computational linguistic analyses to (1) compare inspiring content across cultures, (2) compare AI-generated inspiring posts to real inspiring posts, and (3) determine if detection models can accurately distinguish between inspiring content across cultures and data sources.

1 Introduction

Inspiration has been a part of our world for millennia, starting with ancient Greece, where Muses were responsible for delivering divine knowledge by whispering in a poet's ear (Leavitt and Leavitt, 1997), all the way to today's creativity domain, where it is still common for artists and scientists to attribute their best ideas to a higher power, independent of their own control.

According to Thrash and Elliot (2003, 2004), inspiration is a general construct that consists of three core characteristics: *evocation*, *transcendence*, and
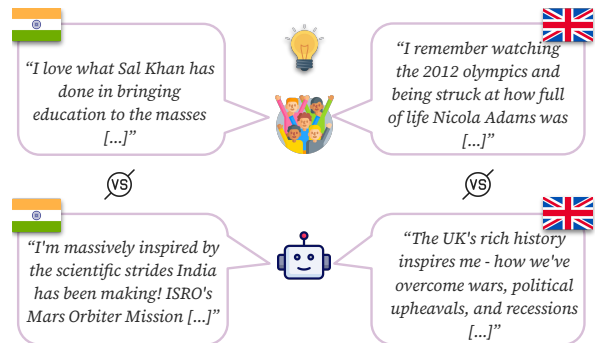

Figure 1: We compare AI-generated and human-written inspiring Reddit content across India and the UK. Although it is challenging for a person to distinguish between them, we find significant linguistic cross-cultural differences between generated and real inspiring posts.

approach motivation. Evocation refers to the process of being triggered by a stimulus, either from within (such as a creative idea that comes from the subconscious) or from outside (such as a person, object, music, or nature). Transcendence allows one to perceive something beyond their usual concerns (Milyavskaya et al., 2012). Finally, an inspired person is motivated to express, transmit, or *act* on their inspiration (Elliot and Thrash, 2002).

Inspiration is an area of study with promising cross-disciplinary applications in creative fields (e.g., advertisement, storytelling), education, therapy, mentorship, coaching, or social media. For instance, social network recommendation systems can mitigate potential harms by showing more positive and inspiring content to users (Ignat et al., 2021). Access to inspiring content can have a positive impact on people's lives by offering them a fresh perspective and motivating them to take action, particularly during periods of uncertainty and concern (Oleynick et al., 2014). Moreover, inspiration facilitates progress towards goals (Milyavskaya et al., 2012) and increases overall well-being (Thrash et al., 2010).

Despite the compelling motivation, little re-

---

[*]Oana Ignat and Gayathri Ganesh Lakshmy contributed equally to the manuscript

search has been done on the automatic identification of content that is *inspiring*, as opposed to simply positive (Ignat et al., 2021), and no studies have been conducted on how inspiration varies across different cultures, and whether it can be automatically generated. At the same time, the impressive generative ability of LLMs can create new opportunities for automatically generating inspiring content.

In this work, we aim to address these research gaps by focusing on three main research questions. First, *RQ1: How do inspiring posts compare across cultures?* Second, *RQ2: How do AI-generated inspiring posts compare to real inspiring posts across cultures?* Finally, *RQ3: Can detection models effectively differentiate inspiring posts across diverse cultures and data sources?*

We summarize our contributions as follows. First, we share **a novel dataset comprised of cross-cultural real and generated inspiring and non-inspiring posts**, for a total of 2,000 real inspiring posts, 2,000 non-inspiring posts, and 2,000 LLM-generated posts, balanced across India and the UK. Second, we make use of the dataset to conduct **extensive computational linguistic analyses to compare inspiring content across cultures and data sources**, i.e., India vs. the UK and AI-generated vs. human-written. Finally, **we explore the effectiveness of machine learning models for detecting inspiration** across diverse data sources and cultures.

2 Related Work

Automatic Inspiration Detection. There have been a limited number of research studies conducted on inspiration, with the majority of them being carried out by the psychology and sociology communities. These studies, such as the ones by Thrash and Elliot (2003, 2004) and Elliot and Thrash (2002), have established the fundamental characteristics of inspiration. Additionally, they have developed a scale to measure the frequency with which people experience inspiration in their daily lives. These studies found that individuals who are inspired tend to be more open to new experiences and show greater absorption in their tasks. They are also more intrinsically motivated, have a strong drive to master their work, and are less competitive.

The work most similar to ours is from Ignat et al. (2021), who are the first to study inspiration through machine learning methods. To facilitate research in this domain, they release a weakly labeled dataset consisting of inspiring and non-inspiring posts collected from Reddit and annotate these posts with their effect on the reader and the emotions they transmit. They also provide a RoBERTa (Liu et al., 2019) classifier fine-tuned on human labels to provide a strong baseline for determining whether a post is inspiring. Finally, they perform extensive data analyses to gain insights into which topics inspire readers and how they influence them. Our work builds on Ignat et al. (2021) by extending it to other cultures, India and the UK, and by collecting AI-generated inspiring posts in order to compare them to human-written posts.

Human vs. LLM-generated cross-cultural text. With the rapid development of LLMs, these models demonstrate remarkable proficiency in generating human-like text across multiple languages and styles (Wu et al., 2023; Tang et al., 2023).

In particular, LLMs excel at creative writing, such as story generation (Yuan et al., 2022), advertising slogan creation (Murakami et al., 2023), and news composition (Yanagi et al., 2020). Tools like Yuan et al. (2022) can help users in their creative pursuits. By generating inspiring content, our work provides a more indirect approach, with the same final goal of helping the user express, transmit, and act on their insight.

More similar to our work, LLMs have also started to be applied to motivate people (Cox et al., 2023). For example, Karinshak et al. (2023) used GPT-3 to generate messages to persuade people to receive the Covid-19 vaccine. To the best of our knowledge, we are the first to generate inspiring content using LLMs. Our work is also part of the emerging work on modeling cultural factors in LLMs (Huang and Yang, 2023; Fung et al., 2022; Ramezani and Xu, 2023). Inspiration varies across cultures. Therefore, we test the cultural knowledge of LLMs about inspiration in India and the UK, and compare it to inspiring Reddit posts from users in these countries.

Computational Linguistics for Social Media Analysis. The advent of computational linguistics techniques has enabled researchers to analyze vast amounts of social media data for various purposes, including sentiment analysis, topic modeling, and linguistic variation across cultures (Pennebaker et al., 2007; Pang et al., 2008; Imran et al.,

2020). These works have facilitated the extraction of meaningful insights from diverse linguistic contexts, paving the way for studies on cross-cultural communication in online environments.

Social media is a key source of inspiring content for younger audiences (Raney et al., 2018). Features such as hope and appreciation of beauty and excellence trigger self-transcendent emotions in videos tagged with "inspiration" on YouTube (Dale et al., 2017), as well as in #inspiring and #meaningful Tumblr memes and Facebook posts (Rieger and Klimmt, 2019; Dale et al., 2020). Similarly, we collect Reddit posts from subreddits related to inspiration across UK and Indian cultures and analyze them using computational linguistic tools such as Pennebaker et al. (2007).

3 The INSPAIRED Dataset

To answer our research questions, we compile a novel dataset, which we refer to as INSPAIRED - AI-generated Inspiring Reddit Content. Our dataset contains inspiring and non-inspiring posts from India and the UK, from two different sources: (1) crawled from Reddit and (2) generated by an LLM.

3.1 👥 Real Inspiring Content

We collect 2,000 weakly labeled inspiring posts and 2,000 weakly labeled non-inspiring posts, balanced across India and the UK. We describe our data collection and annotation process below.

Data Collection. We scrape around 5,300 posts from Reddit, a popular online platform, specifically focusing on culturally inspiring content. Following the data collection process from Ignat et al. (2021), we conduct searches within culture-related or discussion-related flairs of the subreddits using keywords, such as "inspiration" and "motivation", to identify the relevant data in the form of both posts and comments.

More specifically, for Indian data, we primarily target the regions of *Kerala, Karnataka, Maharashtra*, and *Tamil Nadu*. Besides the general "r/india", we also explore subreddits at the *state* level, such as "r/Kerala", "r/Karnataka", "r/Maharashtra", and "r/TamilNadu", which serve as hubs for focused discussions on regional culture, traditions, and social issues. Finally, we also examine subreddits from specific *cities*, including capital cities like *Chennai* and *Bangalore*, to capture more local perspectives and experiences. For the UK data, we follow a similar strategy to collect Reddit posts, targeting both
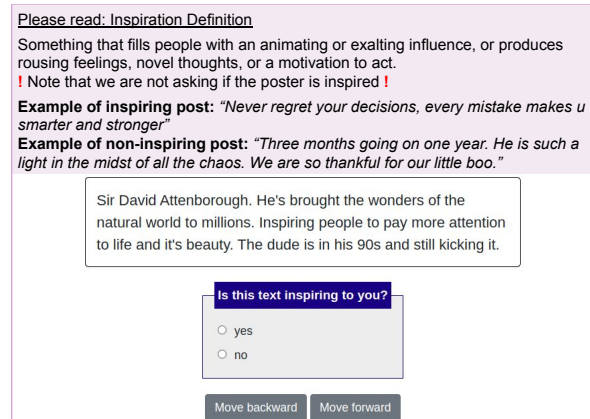

Figure 2: We use the Potato (Pei et al., 2022) annotation platform and provide workers with guidelines to annotate whether a given text is inspiring.

state level subreddits, such as "r/UnitedKingdom", as well as *regional* subreddits representing areas and capital cities, such as "r/London".

We group the collected posts into two main categories: inspiring posts from India and inspiring posts from the UK. Initially, we attempted to create a more fine-grained split at the region or city level, but we faced difficulties in finding annotators from those specific regions. However, we encourage future research to explore this direction and investigate how inspiration varies within a country or within a region, along with exploring other demographic information such as language, age, gender, or income.

Data Filtering. The collected posts are further filtered with an inspiration classifier model. We use the XLM-RoBERTa base (Conneau et al., 2019) from `HuggingFace`.[1] The model is a multilingual version of RoBERTa (Liu et al., 2019), and is pre-trained on 2.5TB of filtered Common-Crawl (Wenzek et al., 2020)[2]. We fine-tune the model for five epochs using the dataset described below, with a learning rate of $2e-5$ and a batch size of 8. More implementation details can be found in the Appendix A.1.

For fine-tuning, we use the data from Ignat et al. (2021). The data contains around 12,000 annotated "general", i.e., no target culture, balanced Reddit inspiring and non-inspiring posts. More details about the fine-tuning data can be found in Ignat et al. (2021). Next, we evaluate the quality of the

---

[1]https://huggingface.co/FacebookAI/xlm-roberta-base
[2]We choose a multilingual model, as the Indian data contains code-mix data.

	# NON-INSPIRING ✗		# INSPIRING 💡	
	Annotated	Weakly-Labeled	Annotated	Weakly-Labeled
🇮🇳	100	900	100	900
🇬🇧	100	900	100	900

Table 1: Final number of inspiring (💡) and non-inspiring (✗) posts across India and the UK.

fine-tuned model predictions by annotating a subset of the posts and comparing them with the model predictions.

Data Annotation. The posts are annotated by crowd-sourced workers from India and the UK. For each country, we select a sample of 200/ 1,000 posts equally sampled from inspiring and non-inspiring predicted posts, to be labeled by three annotators from that country.

We use Potato (Pei et al., 2022) to create the user interface and store the data. Following the guidelines from Ignat et al. (2021), the user interface contains a definition of inspiration and examples of inspiring and non-inspiring posts, as seen in Figure 2. To find and hire UK annotators, we connect Potato to Prolific.[3] We select annotators who have the following qualifications: an approval rate >98%, living in the UK, and a right to vote (as a proxy for age). The annotators independently label 200 posts each, categorizing the posts based on their subjective judgment of whether they found the content inspiring or not. We conduct a similar annotation process with three reliable Indian annotators to label the data from India. To ensure consistency amongst the annotators, they are provided with the same interface as used for the UK annotators and shown in Figure 2.

Finally, we compute the agreement score between the annotators using the Fleiss Kappa measure (Fleiss and Cohen, 1973). We obtain a score of 0.24 for UK data and 0.29 for Indian data, indicating a fair agreement, expected for a subjective task as inspiration detection.

Given that each post is annotated by three annotators, following Ignat et al. (2021), we mark a post as inspiring if at least one annotator labeled it as inspiring.

We find that the predictions made by the fine-tuned model are quite similar to human annotations. For Indian data, the accuracy rate is 72.9% and the F1 score is 75.5%. Similarly, for UK data, the accuracy rate is 73.5%, and the F1 score is 80.1%. Therefore, we decide to not annotate more data and instead use the model predictions. Additionally, we further fine-tune the model on the annotated data and use it to collect more weakly labeled data. This approach aims to leverage the insights gained from the initial fine-tuning process and apply them to the user-annotated subset, thereby refining the model to better capture the nuances present in this specific dataset. We use the fine-tuned model predictions and the annotated subset as our final labeled data.

Quality Assurance. We remove posts that are classified as toxic or hate speech using two fine-tuned XLM-RoBERTa classifiers from `HuggingFace`.[4][5] We also remove profanity from posts using a profanity classifier.[6] We acknowledge that automatic detection of profanity, toxicity, and hate speech are active research areas that have yet to be solved (Dale et al., 2021; Vidgen et al., 2021). Finally, for each country, we manually inspect 100 random posts to ensure syntactic and semantic correctness and find the remaining toxic or hate speech posts. We find that around 99% of the posts are of good quality.

Data Statistics. The final data statistics are shown in Table 1.

3.2 🖥 LLM-Generated Inspiring Content

We generate 2,000 inspiring posts with GPT-4,[7] balanced across India and the UK. Our study can be conducted with any LLM. However, we chose GPT-4 because it is one of the largest LLMs available and has been shown to effectively emulate human texts. (Achiam et al., 2023)

3.2.1 Prompt Design and Robustness

GPT-4 takes as input a list of *message* objects, and returns an inspiring Reddit post. We use *messages*, which are more interactive and dynamic compared to the classical prompt style. Specifically, we use messages with three different roles: *system*, *user* or *assistant*.[8]

---

[3] www.prolific.com

[4] https://huggingface.co/textdetox/xlmr-large-toxicity-classifier
[5] https://huggingface.co/facebook/roberta-hate-speech-dynabench-r4-target
[6] https://github.com/snguyenthanh/better_profanity
[7] https://platform.openai.com/docs/guides/text-generation/chat-completions-api
[8] https://help.openai.com/en/articles/7042661-chatgpt-api-transition-guide

The prompt is first formatted with a *system* role, which sets the behavior of the model. This is followed by a conversation between the *user* and *assistant*, in a **few-shot prompting** fashion. Prior work found that LLMs function better with few-shot prompts (i.e., instructions alongside example output) rather than using zero-shot prompts (with no examples). (Brown et al., 2020)

System Prompt. We find that we can obtain high-quality responses with additional context in our prompts. Therefore, we instruct the model to be a Reddit user from either the UK or India. To ensure that the generated data is diverse and reliable, we collect five versions of *system* prompts with different phrasing, with one example shown below.

> *Imagine you're a person from {location} and use Reddit regularly.*

User-Assistant Prompts. We design two rounds of conversations between a *user* and an *assistant*, where the *user* asks for a Reddit post or comment, and the *assistant* responds to the request. We use **few-shot prompting** by providing the *assistant* in the first round of conversation with an inspiring post, which is randomly extracted from the real posts annotated as inspiring by the majority of annotators. Finally, the answer to the second *user* is automatically generated by the *assistant* and used to collect the GPT-4 generated inspiring post. A *user* message is shown below.

> *Write a Reddit post or comment of maximum 100 tokens about what inspires you.*

Quality Assurance. To ensure the quality of our generated data, we conduct sanity checks to review approximately 200 inspiring posts balanced across the UK and India. The posts are checked for cultural knowledge, factuality, semantic and syntax errors, and style. Based on the feedback, we find that the posts are semantically and syntactically accurate, possess cultural knowledge, and are often more complex than real posts.

4 Cross-Cultural Inspiration Analysis across Real and LLM-generated posts

In line with previous work (Jakesch et al., 2023), we also find during manual inspection that it is challenging to differentiate between LLM-generated texts and those written by humans (see Table 2). Therefore, we perform computational linguistic analyses to compare real and LLM-generated text across cultures. This section addresses our first two research questions: *RQ1: How do inspiring posts compare across cultures?* and *RQ2: How do AI-generated inspiring posts compare to real inspiring posts across cultures?*

4.1 Stylistic and Structural Features

We assess the linguistic style and structure of real and LLM-generated inspiring posts across India and the UK in terms of (1) analytic writing, (2) descriptiveness, and (3) readability.

Analytic Writing index measures the complexity and sophistication of the writing, which can be an indicator of advanced thinking. The formula for analytic writing is $[articles + prepositions - pronouns - auxiliaryverbs - adverb - conjunctions - negations]$ from LIWC scores (Jordan et al., 2019; Pennebaker et al., 2014) More information about LIWC can be found in Section 4.2. We display the scores in Table 3. The low complexity scores of Reddit posts are primarily negative due to their high usage of pronouns and lack of articles. We find that *LLM-generated inspiring posts from the UK and India are more complex than real posts*, which aligns with our initial observations from data quality checks. At the same time, there is *no significant difference in text complexity between real inspiring posts from India and those from the UK*.

Descriptiveness can be measured by the frequency of adjectives used in language patterns. Texts with high rates of adjectives tend to be more elaborate and narrative-like compared to texts with low rates of adjectives. (Chung and Pennebaker, 2008) In Table 3, we find that *LLM-generated inspiring posts from the UK and India are more descriptive than real posts*. Additionally, *real inspiring posts from India are more descriptive than those from the UK*.

Readability considers not only word count, but also word complexity. For instance, longer words are more complex than shorter ones. We use the Flesch Reading Ease metric (Flesch, 1948), which counts the number of words per sentence and syllables per word. In Table 3, we find that *LLM-generated inspiring posts from the UK and India are less readable than real posts*. Additionally, *real inspiring posts from India are less readable than those from the UK*. Furthermore, when measuring post length, *LLM-generated UK posts are*

	Real Post 👥	GPT4-generated Post 🖥
🇮🇳	💡 *The youngest freedom fighter martyr of the country, the brave boy who refused to take the British cops across the river and was shot to death. His name was Baji Rout and he was 12 at the time of his death.*	💡 *Dr. APJ Abdul Kalam - His humble beginnings, insatiable thirst for knowledge, and absolute dedication to his country have been my greatest inspiration. It pushed me to work harder, dream big, and contribute to society.*
	✗ *What is a common meal there? I've always thought Indian palates to be quite spicy and rich.*	✗ N/A
🇬🇧	💡 *Dr. Helen Sharman. I'm very pleased to come from a country whose first astronaut isn't a man by default. Equality shouldn't be about women catching up, it should be about women being first 50% of the time.*	💡 *Absolutely love Sir David Attenborough's documentaries. His passion and commitment to preserving the environment is truly inspiring in these challenging times.*
	✗ *Mate all he does is play football and misses pens. Not that special*	✗ N/A

Table 2: Random samples of inspiring (💡) and non-inspiring (✗) posts from India and UK.

	🇮🇳		🇬🇧	
	💡👥	💡🖥	💡👥	💡🖥
Analytic	-16.7 ± 18.0	-6.9 ± 10.0	-17.1 ± 19.9	-7.0 ± 11.4
Descriptive	8.2 ± 7.8	8.9 ± 3.6	6.6 ± 5.9	9.1 ± 3.9
Readable	36.3 ± 63.1	12.1 ± 20.6	53.2 ± 52.2	29.6 ± 18.0
Word Count	61.1 ± 59.8	66.9 ± 20.3	43.9 ± 49.3	49.9 ± 16.8

Table 3: To what degree is the LLM-generated text (🖥) stylistically and structurally different from the real text (👥)? We compute the mean and standard deviation for the inspiring posts, across cultures. The differences are statistically significant, based on the Student t-test (Student, 1908), p-value < 0.05.

longer than real posts and *UK posts are shorter than Indian posts*.

4.2 Semantic and Psycholinguistic Features

We assess the semantic and psycholinguistic differences between LLM-generated and real inspiring posts across India and the UK, using topic modeling and LIWC psycholinguistic markers.

Data Pre-processing. We use `spaCy` library[9] to pre-process the data: tokenize each post, lowercase tokens, remove stop words, remove numbers, symbols, emojis, links, and lemmatize tokens.

Topic Modeling. We use Scattertext (Kessler, 2017), a tool used to create interactive visualizations of linguistic patterns. After initial preprocessing, we structure the text data for a better representation of information. Next, we use the n-gram representation of the text data to conduct topic extraction through sentence-level clustering. We utilize methods like Non-negative Matrix Factorization (NMF) to decompose the TF-IDF matrix

to identify the latent thematic structures within the corpus (Ramos et al., 2003).

Results. We analyze the n-gram and topic distributions by various dimensions: inspiring vs. non-inspiring, Indian vs. UK, and real vs. generated. We display the topic distribution across real and generated UK inspiring posts in Figure 3. Further topic and n-gram distributions are displayed in Appendix A.2.

💡 vs. ✗: Comparing inspiring to non-inspiring real posts, we find that, in the Indian data (Fig. 5, 7), the topic of *people* is frequently discussed in an inspiring context. Moreover, the most commonly occurring theme in non-inspiring posts is *bot*, encompassing textual content regarding Reddit rules and moderators.

In the UK data (Fig. 6, 8), the topic of *life* is amongst the most commonly occurring themes in inspiring posts, with discussions surrounding *career*, *luck*, and *pension*. On the other hand, the non-inspiring posts contain themes like *dark*, including conversations about *rain* and *winter*.

👥 vs. 🖥: The LLM-generated Indian data (Fig. 10, 11) often places a significant emphasis on the topic of *inspiring*, within which common words include *dedication* and *motivational*. There is also a significant number of posts surrounding *isro (Indian Space Research Organisation)*, featuring terms such as *space*, *mars*, and *mission*. In contrast, the real data from India contains mentions of *housing*, *summer*, and *living*, grouped under the category *live*. The topic of *movie* is also popular, containing words such as *hype*, *stardom*, and *socialize*, hinting at the culture surrounding the film industry.

In the LLM-generated UK data (Fig. 3, top left), the topic *nhs* is emphasized, where common words include *pandemic*, *heroes*, and *staff*. Additionally,
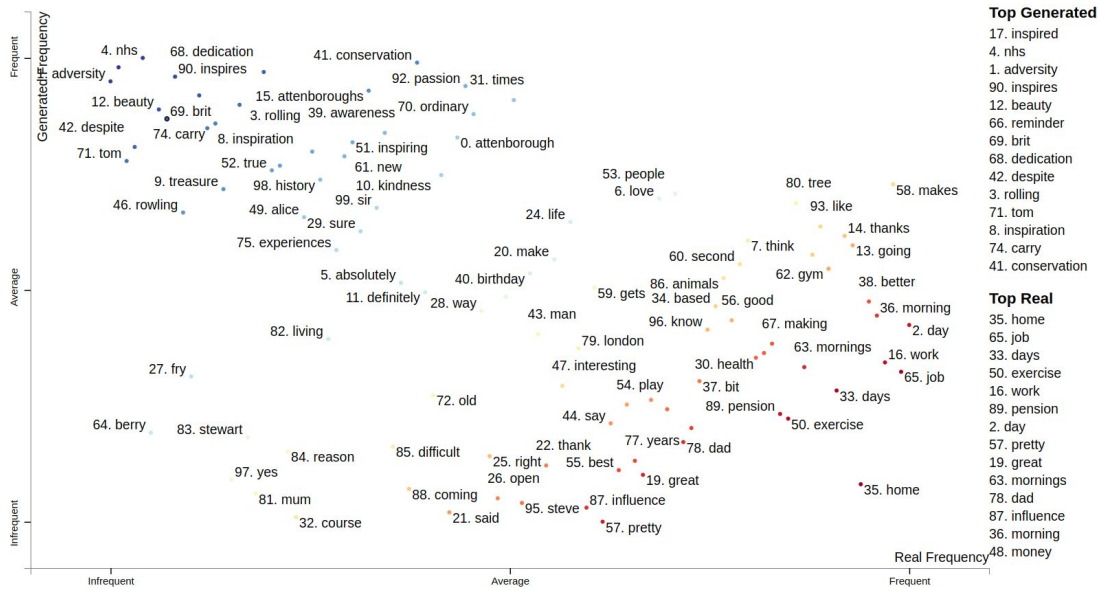
Figure 3: Visualization of **topics** used in the real and generated (👥 vs. 💻) inspiring posts from the **UK**. Points are colored red or blue based on the association of their corresponding terms with UK Real inspiring posts or UK LLM-Generated inspiring posts. The most associated topics are listed under **Top Generated** and **Top Real** headings. Interactive version: https://github.com/MichiganNLP/cross_inspiration.

discussions often revolve around *adversity*, with mentions of *resilience*, *determination*, and *spirit*. In contrast, in the real data from the UK (Fig. 3, bottom right), discussions related to *job* dominate, with mentions of *salary*, *savings*, and *time*. Moreover, there is also a significant focus on *exercise*, featuring words like *discipline*, and *motivate*.

LIWC Psycholinguistic Markers. We use Linguistic Inquiry and Word Count (LIWC), a gold-standard text analysis tool (Pennebaker et al., 2007, 2015), to obtain the words related to human cognitive processes from each post. Specifically, we use the LIWC2015 dictionary, which contains 6,400 words and word stems, each related to a cognitive category. As an example, the word "mother" is assigned the following cognitive categories: *female, family, social*.

Results. We display the top 24 categories and their LIWC scores, with the most significant differences across dimensions in Table 4. Across all data, we find that *male* words are more frequent than *female*, the most common pronoun is *I*, the most frequent tense is *present*, and the most common emotion is *positive*.

💡 vs. ✗: Comparing inspiring to non-inspiring real posts, we find that, in the Indian data, inspiring posts are more likely to be related to *socializing, insight, feelings, perception, affection*, and contain *more positive emotions*. Additionally, Indian inspiring posts tend to have *more comparisons* and use

more words related to *achievement, health, reward*, and *work*.

UK real inspiring posts contain more words related to *affection, comparisons, feelings, achievement, health, home, leisure, money, reward*, and *work* than non-inspiring posts. Furthermore, compared to Indian inspiring posts, UK inspiring posts have fewer words related to *family, socializing, affection, perception, religion, less positive emotions* and more words related to *achievement, health, home, leisure, money, rewards*, and *work*.

👥 vs. 💻: Comparing real to LLM-generated inspiring posts, we find that real Indian posts tend to include more words related to *family, social interactions, comparisons, feelings, perceptions* as well as *home, leisure* and *rewards*. Conversely, real Indian posts contain fewer words related to *affection, insight, achievement, health*, and *religion*.

UK real inspiring posts contain more words related to *comparisons, feelings, health, home, leisure, money, rewards* and *work* than LLM-generated UK posts. Conversely, real UK posts contain fewer words related to *socializing, affection, insight, perception, achievement* and *religion* than LLM-generated UK posts. Furthermore, compared to LLM-generated Indian posts, LLM-generated UK posts have fewer words related to *socializing, achievement, leisure, money, religion, work* and more words related to *affection, perception* and *more positive emotions*.

LIWC	India 💡👥	✗👥	💡🖥	UK 💡👥	✗👥	💡🖥
Social Processes						
FAMILY	0.4	0.4	0.2	0.2	0.1	0.2
FRIEND	0.3	0.4	0.2	0.2	0.4	0.2
FEMALE	0.4	0.3	0.2	0.4	0.4	0.2
MALE	1.1	1.0	0.9	0.9	1.3	1.7
SOCIAL	9.3	7.7	8.6	7.4	8.1	8.1
Affective Processes						
AFFECT	6.1	4.3	8.4	5.4	4.8	9.4
NEGEMO	1.4	1.3	1.0	1.3	1.6	0.9
POSEMO	4.5	3.0	7.3	4.1	3.1	8.3
Cognitive Processes						
COMPARE	3.0	2.4	1.8	3.0	2.7	1.8
INSIGHT	2.3	2.0	4.8	2.2	2.1	4.9
Perceptual Processes						
FEEL	0.6	0.4	0.4	0.7	0.3	0.4
PERCEPT	2.4	2.0	2.1	2.1	2.1	2.4
Personal Concerns						
ACHIEV	1.9	1.0	3.5	2.4	1.6	2.6
HEALTH	0.7	0.4	0.9	1.0	0.8	0.8
HOME	0.3	0.2	0.1	0.7	0.4	0.2
LEISURE	1.5	1.4	1.2	1.7	1.4	0.9
MONEY	0.7	0.8	0.6	1.2	0.8	0.4
RELIG	0.4	0.5	1.1	0.1	0.1	0.6
REWARD	2.0	1.1	1.0	2.2	1.6	1.1
WORK	3.0	1.9	3.0	3.9	2.7	1.9

Table 4: Comparing LIWC scores across cultures (India and the UK) in inspiring vs. non-inspiring (💡 vs. ✗) and real vs. LLM-generated posts (👥 vs. 🖥). We present the top 24 categories with the most significant differences across these dimensions.

Figure 4: Multi-label classification test results with the *few-shot* and *default* setups with the XLM-RoBERTa base (Conneau et al., 2019) model.

91.1% across all labels, while the *default* setup achieves an accuracy of 89.1% and F1 score of 96.3%. Figure 4 displays the disaggregated results for each label. We find that even with very few training data (600 posts), in the *few-shot* setup, the XLM-RoBERTa model learns to accurately distinguish inspiring content across cultures (India, UK) and data sources (real and generated). Furthermore, we do not find significant differences in performance across cultures and data sources. Finally, the model can accurately distinguish between real and generated inspiring posts in both *few-shot* and *default* training setups.

5 Cross-Cultural Inspiration Detection across Real and LLM-generated Posts

To answer our last research question – *RQ3: Can detection models effectively differentiate inspiring posts across diverse cultures and data sources?* – we fine-tune a multi-label classification model to identify if a post is inspiring, represents India or the UK, and whether it is real or LLM-generated.

Implementation Details. We use the XLM-RoBERTa base model (Conneau et al., 2019), pre-trained on 2.5TB of filtered CommonCrawl (Wenzek et al., 2020). We fine-tune the model on our dataset for five epochs with a learning rate of $2e-5$ and a batch size of 8, while monitoring the validation performance and selecting the best model checkpoint based on the F1 metric.

Model Training Setup. We experiment with two setups to split the data into training, validation, and test sets: a *default* train-val-test split of 64-16-20% and a *few-shot* split of 8-2-90%.

Results. Since the dataset is evenly distributed across classes, a random baseline results in a 50% accuracy score.

The XLM-RoBERTa model with *few-shot* setup achieves an accuracy of 75.0% and F1 score of

6 Conclusion

In this paper, we introduced the task of cross-cultural inspiration detection and generation in social media data. To facilitate research in this domain, we released INSPAIRED, a dataset of 2,000 real inspiring posts, 2,000 real non-inspiring posts, and 2,000 LLM-generated inspiring posts, evenly distributed across India and the UK. We performed extensive linguistic data analyses to gain insight into what topics inspire each culture and compare AI-generated inspiring posts to real inspiring posts across various linguistic dimensions. Despite the difficulty humans have in distinguishing between real posts and those generated by LLMs, we found that these posts have noticeable differences in style, structure, and semantics and that, even with little data, fine-tuned models accurately distinguish inspiring content across cultures and data sources.

We hope our work will enable the exploration of various applications to improve creativity and motivation, including storytelling, advertising, and social media, as well as therapy and coaching. Our dataset can be used to test, fine-tune, and analyze other models, and it is publicly available at `https://github.com/MichiganNLP/cross_inspiration`.

Limitations

1. A more fine-grained data split

Our posts are divided into two main categories - inspiring posts from India and inspiring posts from the UK. Initially, we tried to create a more detailed classification at the region or city level, but we faced difficulties in finding annotators from those specific regions. However, we encourage future research to explore this direction and investigate how inspiration varies within a country or region, as well as explore other demographic information, such as language, age, gender, or income.

Study case for India. The extensive discussions on Indian subreddits, like r/AskIndia and r/India, offer straightforward access to text-based motivational narratives. Additionally, region-based sub-Reddits, like r/Chennai and r/Mumbai, provide diverse insights into localized experiences and discussions. The nature of posts on this topic is mostly anecdotal. Reddit users contribute by sharing personal experiences that have inspired them.

2. Limited number of cross-cultural inspiring posts

It is crucial to recognize that inspiration is not always explicitly articulated through words like "inspiring" or "motivation", especially within cultures that aren't as open about talking about such topics. In many instances, it manifests more subtly, embedded within narratives, imagery, or cultural expressions. This implicit nature of inspiration adds another layer of complexity to the data collection process.

The language barrier also introduces an additional barrier, as it requires linguistic fluency and cultural understanding to interpret and analyze inspiring content effectively. Posts in languages other than English may contain nuances and cultural references that are not easily translatable. Moreover, it becomes more challenging to collect inspiring content when data collection is restricted to certain countries. That is why we collected fewer posts compared to Ignat et al.'s "general" inspiring posts.

3. Relevance of LLM-based data to current times

In our reliance on LLM-based data, it's imperative to recognize the temporal limitations inherent in its training corpus. While these models offer remarkable capabilities in understanding and generating text, they might not fully capture the current cultural and societal context in their outputs. This limitation can be primarily attributed to the fact that these models are trained on data only up to a specific cutoff date, i.e., GPT-4 only learned from data dated up to September 2021.

References

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Cindy K Chung and James W Pennebaker. 2008. Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of research in personality*, 42(1):96–132.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Unsupervised cross-lingual representation learning at scale. *CoRR*, abs/1911.02116.

Samuel Rhys Cox, Ashraf Abdul, and Wei Tsang Ooi. 2023. Prompting a large language model to generate diverse motivational messages: A comparison with human-written messages. In *Proceedings of the 11th International Conference on Human-Agent Interaction*, pages 378–380.

David Dale, Anton Voronov, Daryna Dementieva, Varvara Logacheva, Olga Kozlova, Nikita Semenov, and Alexander Panchenko. 2021. Text detoxification using large pre-trained neural models. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 7979–7996, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Katherine R Dale, Arthur A Raney, Sophie H Janicke, Meghan S Sanders, and Mary Beth Oliver. 2017. Youtube for good: A content analysis and examination of elicitors of self-transcendent media. *Journal of Communication*, 67(6):897–919.

Katherine R Dale, Arthur A Raney, Qihao Ji, Sophie H Janicke-Bowles, Joshua Baldwin, Jerrica T Rowlett, Cen Wang, and Mary Beth Oliver. 2020. Self-transcendent emotions and social media: Exploring the content and consumers of inspirational facebook posts. *New Media & Society*, 22(3):507–527.

Andrew J Elliot and Todd M Thrash. 2002. Approach-avoidance motivation in personality: approach and avoidance temperaments and goals. *Journal of personality and social psychology*, 82(5):804.

Joseph L Fleiss and Jacob Cohen. 1973. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and psychological measurement*, 33(3):613–619.

Rudolph Flesch. 1948. A new readability yardstick. *Journal of applied psychology*, 32(3):221.

Yi R Fung, Tuhin Chakraborty, Hao Guo, Owen Rambow, Smaranda Muresan, and Heng Ji. 2022. Normsage: Multi-lingual multi-cultural norm discovery from conversations on-the-fly. *arXiv preprint arXiv:2210.08604*.

Jing Huang and Diyi Yang. 2023. Culturally aware natural language inference. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 7591–7609.

Oana Ignat, Y-Lan Boureau, Jane A. Yu, and Alon Y. Halevy. 2021. Detecting inspiring content on social media. *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8.

Ali Shariq Imran, Sher Muhammad Daudpota, Zenun Kastrati, and Rakhi Batra. 2020. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets. *Ieee Access*, 8:181074–181090.

Maurice Jakesch, Jeffrey T Hancock, and Mor Naaman. 2023. Human heuristics for ai-generated language are flawed. *Proceedings of the National Academy of Sciences*, 120(11):e2208839120.

Kayla N Jordan, Joanna Sterling, James W Pennebaker, and Ryan L Boyd. 2019. Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proceedings of the National Academy of Sciences*, 116(9):3476–3481.

Elise Karinshak, Sunny Xun Liu, Joon Sung Park, and Jeffrey T Hancock. 2023. Working with ai to persuade: Examining a large language model's ability to generate pro-vaccination messages. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1):1–29.

Jason Kessler. 2017. Scattertext: a browser-based tool for visualizing how corpora differ. In *Proceedings of ACL 2017, System Demonstrations*, pages 85–90, Vancouver, Canada. Association for Computational Linguistics.

J.H. Leavitt and P.J. Leavitt. 1997. *Poetry and Prophecy: The Anthropology of Inspiration*. Studies in Literature and Science. University of Michigan Press.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *ArXiv*, abs/1907.11692.

Marina Milyavskaya, Iana Ianakieva, Emily Foxen-Craft, Agnes Colantuoni, and Richard Koestner. 2012. Inspired to get there: The effects of trait and goal inspiration on goal progress.

Soichiro Murakami, Sho Hoshino, and Peinan Zhang. 2023. Natural language generation for advertising: A survey. *Preprint*, arXiv:2306.12719.

Victoria C. Oleynick, Todd M. Thrash, Michael C. LeFew, Emil G. Moldovan, and Paul D. Kieffaber. 2014. The scientific study of inspiration in the creative process: challenges and opportunities. *Frontiers in Human Neuroscience*, 8.

Bo Pang, Lillian Lee, et al. 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, 2(1–2):1–135.

Jiaxin Pei, Aparna Ananthasubramaniam, Xingyao Wang, Naitian Zhou, Apostolos Dedeloudis, Jackson Sargent, and David Jurgens. 2022. Potato: The portable text annotation tool. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*.

James W. Pennebaker, Roger John Booth, and Martha E. Francis. 2007. Linguistic inquiry and word count (LIWC2007).

James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate G. Blackburn. 2015. The development and psychometric properties of LIWC2015.

James W Pennebaker, Cindy K Chung, Joey Frazee, Gary M Lavergne, and David I Beaver. 2014. When small words foretell academic success: The case of college admissions essays. *PloS one*, 9(12):e115844.

Aida Ramezani and Yang Xu. 2023. Knowledge of cultural moral norms in large language models. *arXiv preprint arXiv:2306.01857*.

Juan Ramos et al. 2003. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*, volume 242, pages 29–48. Citeseer.

Arthur A Raney, Sophie H Janicke, Mary Beth Oliver, Katherine R Dale, Robert P Jones, and Daniel Cox. 2018. Profiling the audience for self-transcendent media: A national survey. *Mass Communication and Society*, 21(3):296–319.

Diana Rieger and Christoph Klimmt. 2019. The daily dose of digital inspiration: A multi-method exploration of meaningful communication in social media. *New Media & Society*, 21(1):97–118.

Student. 1908. The probable error of a mean. *Biometrika*, pages 1–25.

Ruixiang Tang, Yu-Neng Chuang, and Xia Hu. 2023. The science of detecting llm-generated texts. *arXiv preprint arXiv:2303.07205*.

Todd M. Thrash and Andrew J Elliot. 2003. Inspiration as a psychological construct. *Journal of personality and social psychology*, 84 4:871–89.

Todd M. Thrash and Andrew J Elliot. 2004. Inspiration: core characteristics, component processes, antecedents, and function. *Journal of personality and social psychology*, 87 6:957–73.

Todd M. Thrash, Laura A. Maruskin, Scott E. Cassidy, James W. Fryer, and Richard M. Ryan. 2010. Mediating between the muse and the masses: inspiration and the actualization of creative ideas. *Journal of personality and social psychology*, 98 3:469–87.

Bertie Vidgen, Tristan Thrush, Zeerak Waseem, and Douwe Kiela. 2021. Learning from the worst: Dynamically generated datasets to improve online hate detection. In *ACL*.

Guillaume Wenzek, Marie-Anne Lachaux, Alexis Conneau, Vishrav Chaudhary, Francisco Guzmán, Armand Joulin, and Edouard Grave. 2020. CCNet: Extracting high quality monolingual datasets from web crawl data. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4003–4012, Marseille, France. European Language Resources Association.

Junchao Wu, Shu Yang, Runzhe Zhan, Yulin Yuan, Derek F. Wong, and Lidia S. Chao. 2023. A survey on llm-generated text detection: Necessity, methods, and future directions. *ArXiv*, abs/2310.14724.

Yuta Yanagi, Ryohei Orihara, Yuichi Sei, Yasuyuki Tahara, and Akihiko Ohsuga. 2020. Fake news detection with generated comments for news articles. In *2020 IEEE 24th International Conference on Intelligent Engineering Systems (INES)*, pages 85–90. IEEE.

Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: story writing with large language models. In *27th International Conference on Intelligent User Interfaces*, pages 841–852.

A Appendix

A.1 Fine-tuning Process

XLM-RoBERTa Model. Implementation Details. We experiment with three different fine-tuned classification models for the purpose of weakly labeling the presence of inspiration in the dataset. The first attempt involved fine-tuning the model on the general inspiration dataset introduced in (Ignat et al., 2021). In each of the experiments, the data is split into three subsets: training, validation, and test sets, using an 80:10:10 ratio. The fine-tuning process involved training the model for 5 epochs

with a learning rate of $2e-5$ and a batch size of 8. We monitored the model performance on a separate validation set and selected the best model checkpoint based on accuracy metric.

In the second experiment, we focused on a subset of the dataset used in annotation, consisting of 200 posts. These posts were labeled as inspiring if at least one user considered them to be so. We fine-tuned the base XLM-RoBERTa model using this subset, maintaining the same training configurations as in the initial experiment. In the last experiment, we leveraged the user-annotated posts to further refine the model trained on the general inspiration dataset from the first experiment. This approach aimed to leverage the insights gained from the initial fine-tuning process and apply them to the user-annotated subset, thereby refining the model to better capture the nuances present in this specific dataset.

A.2 Topic Analysis

Figure 5: Scattertext visualization of **unigrams** used in the real inspiring and non-inspiring (💡 vs. ✗) Reddit posts from **India**. Points are colored in red or blue based on the association of their corresponding terms with Indian Non-inspiring posts or Indian inspiring posts. The most associated terms are listed under "Top inspiring" and "Top Non-inspiring" headings.
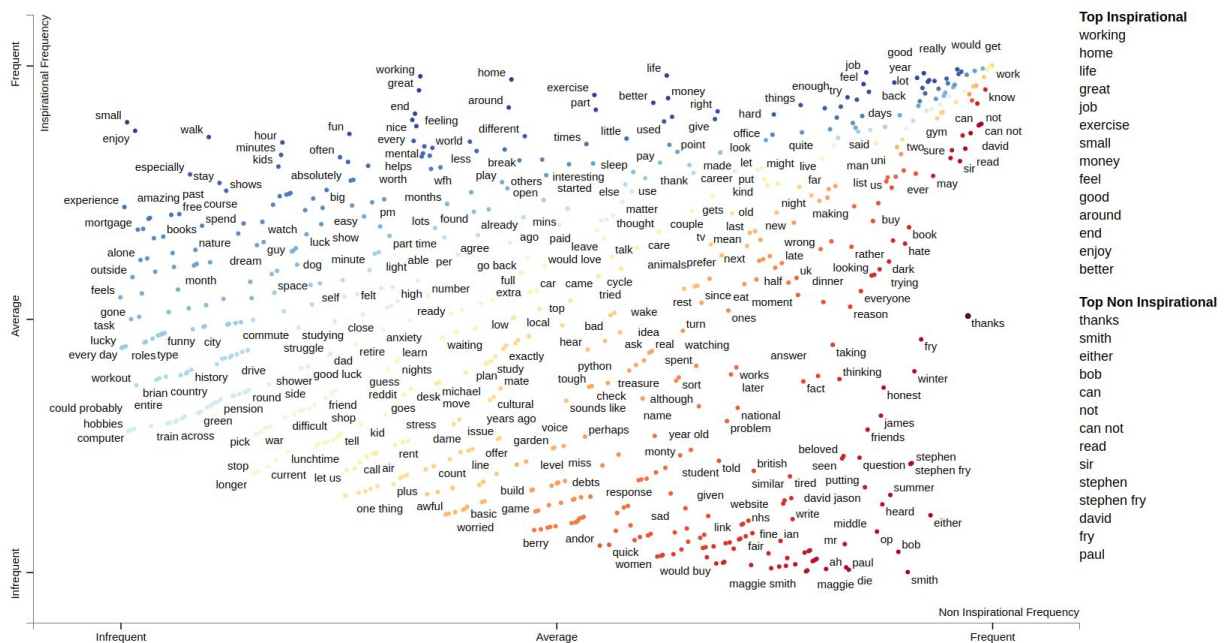
Figure 6: Scattertext visualization of **unigrams** used in the real inspiring and non-inspiring (💡 vs. ✗) Reddit posts from the **UK**. Points are colored in red or blue based on the association of their corresponding terms with UK Non-inspiring posts or UK inspiring posts. The most associated terms are listed under "Top inspiring" and "Top Non-inspiring" headings.

Figure 7:

Inspirational Frequency — Frequent / Average / Infrequent

Top Inspirational
84. people
20. idukki
25. roads
86. far
70. way
12. abroad
5. hope
52. therapy
49. little
1. sharing
42. changing
50. pradesh
67. ksrtc
77. visit

Top Non Inspirational
37. bot
54. kappa
48. use
44. depots
31. grown
75. happened
79. curry
43. dude
33. know
19. big
47. played
80. college
85. day
4. cute

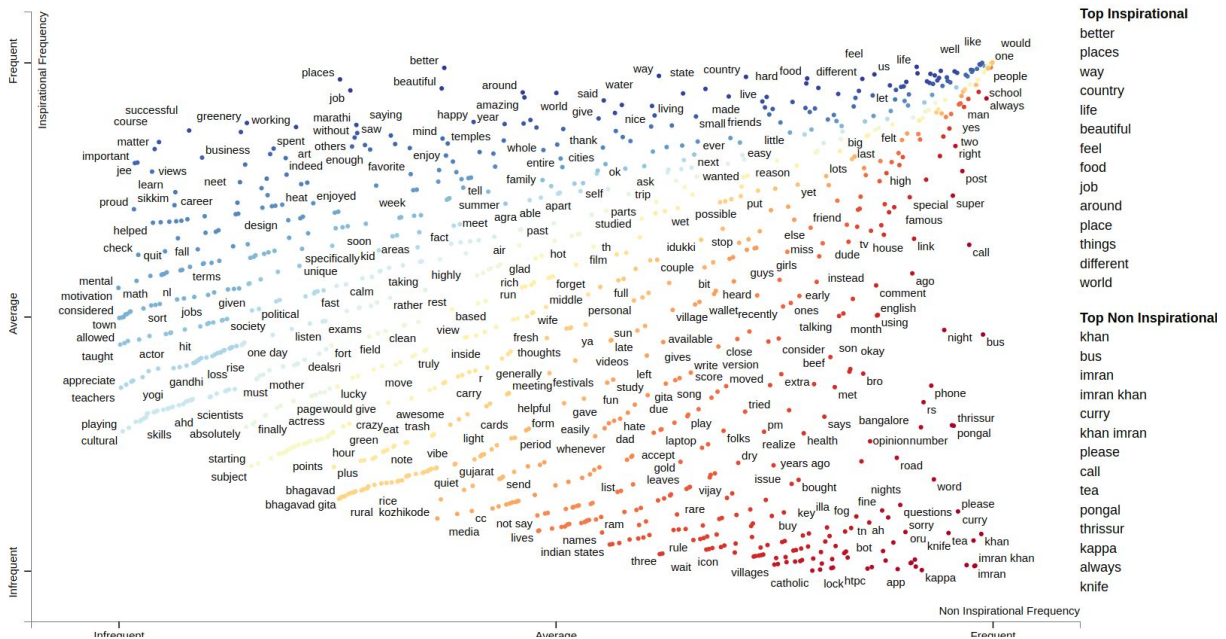Non Inspirational Frequency — Infrequent / Average / Frequent

Figure 7: Scattertext visualization of **topics** used in the real inspiring and non-inspiring (vs.) Reddit posts from **India**. Points are colored in red or blue based on the association of their corresponding terms with India Non-inspiring posts or India inspiring posts. The most associated topics are listed under "Top inspiring" and "Top Non-inspiring" headings.

Figure 8:

Inspirational Frequency — Frequent / Average / Infrequent

Top Inspirational
70. great
80. life
13. exercise
75. rich
64. nice
50. buy
92. drugs
46. right
57. young
17. house
41. interesting
36. job
62. wish
65. mum

Top Non Inspirational
51. smith
86. dark
40. wrong
76. wolf
26. use
94. beckham
71. guns
5. mortimer
52. pretty
49. know
16. heard
42. try
79. kind
89. list

Non Inspirational Frequency — Infrequent / Average / Frequent

Figure 8: Scattertext visualization of **topics** used in the real inspiring and non-inspiring (vs.) Reddit posts from the **UK**. Points are colored red or blue based on the association of their corresponding terms with UK Non-inspiring posts or UK inspiring posts. The most associated topics are listed under "Top inspiring" and "Top Non-inspiring" headings.
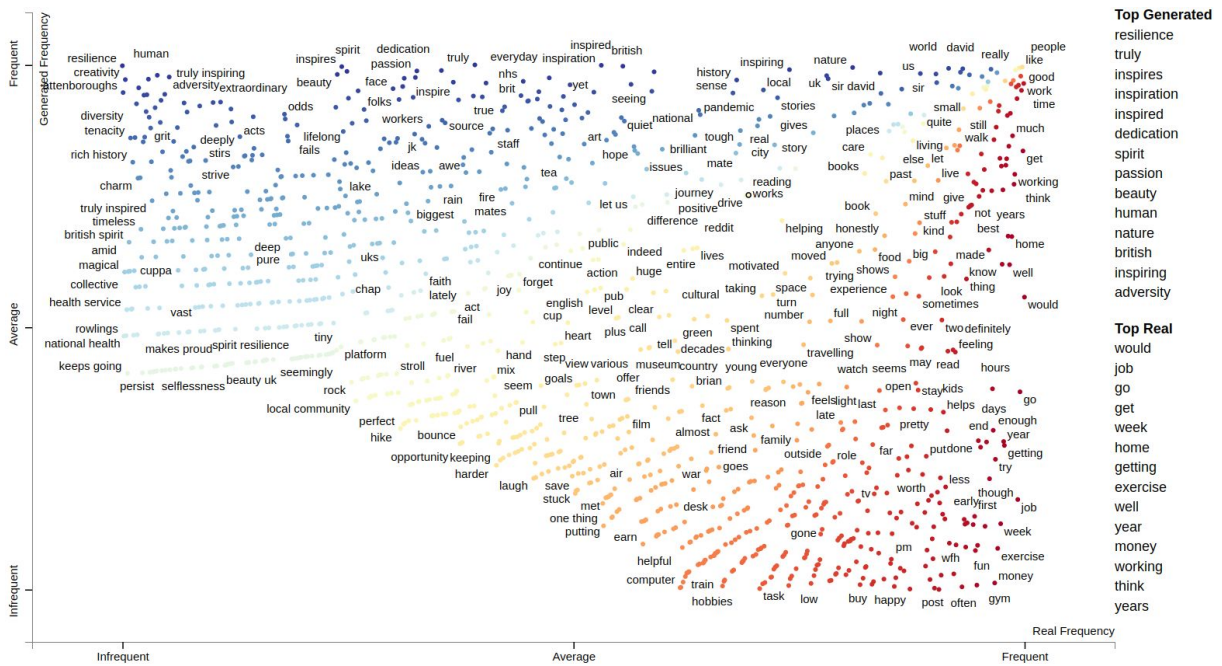
Figure 9: Scattertext visualization of **unigrams** used in the real and generated (👥 vs. 💻) inspiring posts from the **UK**. Points are colored red or blue based on the association of their corresponding terms with the UK Real inspiring posts or the UK LLM-Generated inspiring posts. The most associated topics are listed under **Top Generated** and **Top Real** headings.
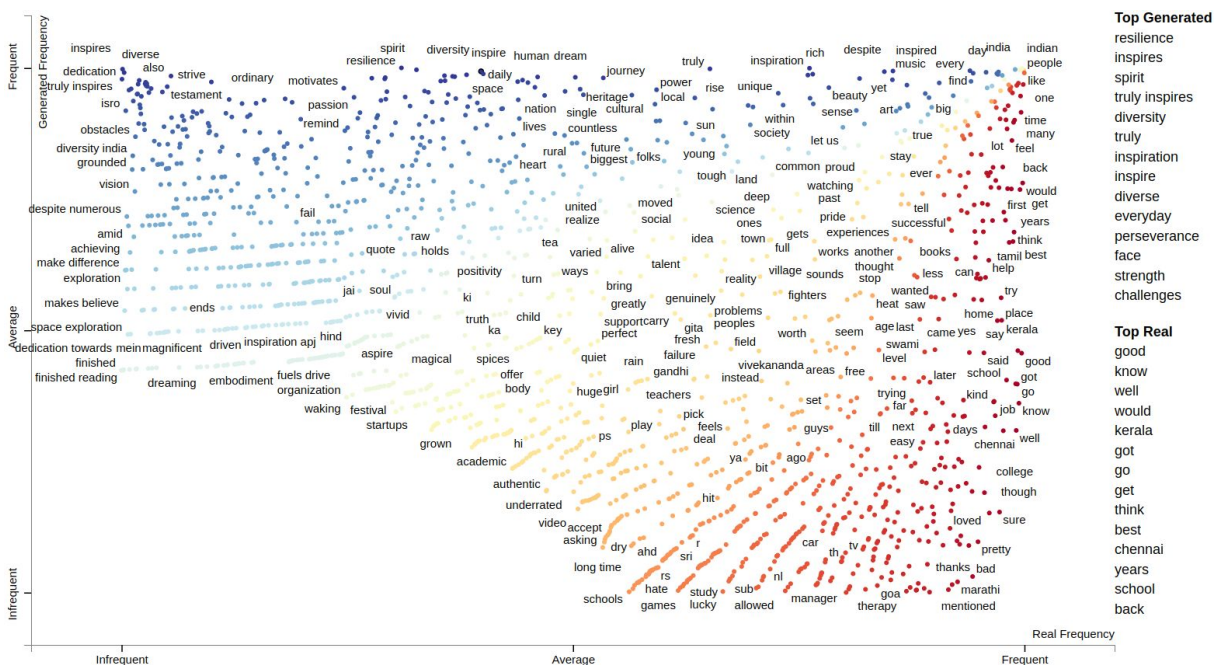
Figure 10: Scattertext visualization of **unigrams** used in the real and generated (👥 vs. 💻) inspiring posts from **India**. Points are colored red or blue based on the association of their corresponding terms with India Real inspiring posts or India LLM-Generated inspiring posts. The most associated topics are listed under **Top Generated** and **Top Real** headings.
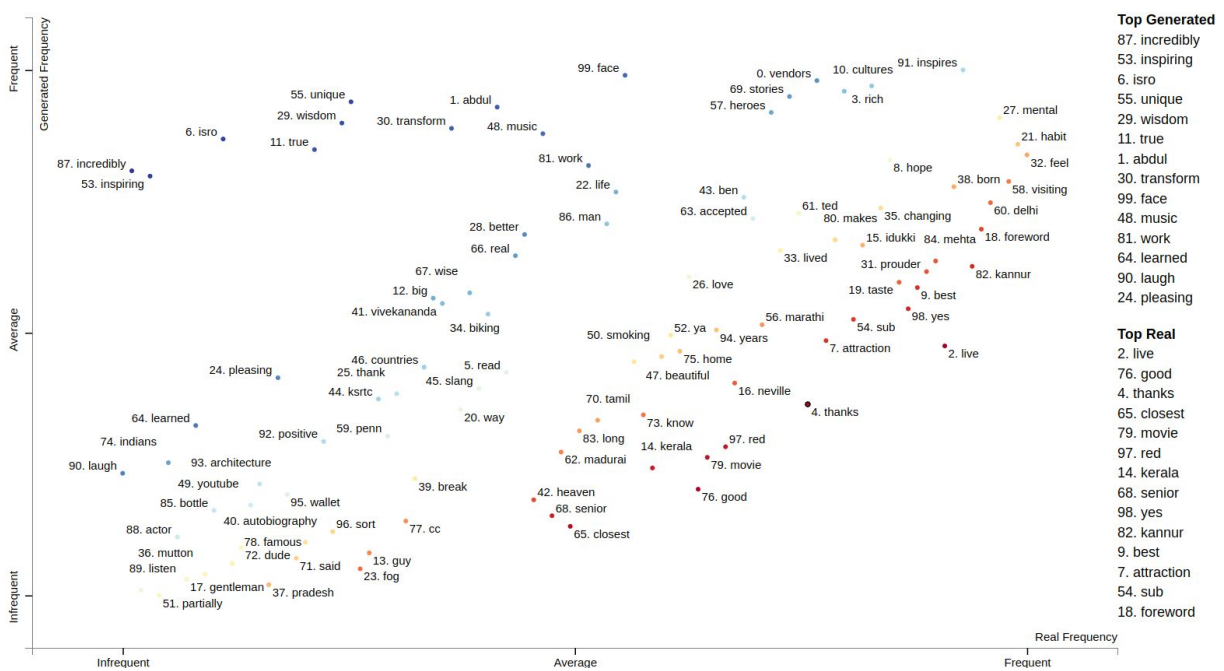
Figure 11: Scattertext visualization of **topics** used in the real and generated (👥 vs. 🖥) inspiring posts from **India**. Points are colored red or blue based on the association of their corresponding terms with India Real inspiring posts or India LLM-Generated inspiring posts. The most associated topics are listed under **Top Generated** and **Top Real** headings.