

Sentiment Analysis of User Interactions on Bollywood Celebrity Twitter Handles: Investigating Gender Stereotypes

BTech Project Report

Eslavath Meenakshi(B21CS028)

Bukya Krishnaveni(B21AI008)

Under the Supervision

Of

Dr. Sherin Sabu (SoLA)



Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur
November, 2024

1. Introduction

1.1 Gender and Intersectionality

Gender refers to the social, cultural, and behavioural characteristics and roles that society associates with being male, female, or other identities. It goes beyond biological differences and involves the roles, expectations, and behaviours that people adopt based on their identification as male, female, neither, or other gender identities.

In her 1972 essay *Is Female to Male as Nature is to Culture?*, Sherry B. Ortner talks about how women are seen as less important than men in societies around the world. Ortner argues that this hierarchy stems from symbolic associations that link women to “nature” and men to “culture.” According to her, women are frequently connected to natural elements, like the earth and life cycles. At the same time, men are associated with cultural domains, such as art, science, and politics—fields seen as more powerful and valued. Although beliefs vary across cultures, she observes that nearly all societies tend to rank women below men. For instance, some cultures revere female symbols, like goddesses, that reflect power but also emphasize women’s connection to nature, considered lesser than human-made culture. Because women are responsible for childbirth, childrearing, and nurturing, they are symbolically tied to nature, whereas men, who often take on roles of leadership and creativity, are linked to culture. This symbolic framework leads to perceptions of women’s roles, such as caregiving, as “lower” than men’s roles, like leading nations or managing businesses. Ortner further suggests that women are often seen as enigmatic, positioned between nature and culture—they are nurturing figures but can also be portrayed as mysterious, possessing “hidden powers” associated with figures like witches. Drawing on psychoanalytic theory, Ortner explains that people may view women as naturally suited to caregiving because of their bodies, an idea feminist psychoanalyst Nancy Chodorow (1978) extends by showing that family structures reinforce these roles, not just biology. Ortner concludes that since culture involves controlling nature, and women are symbolically closer to nature, societies often exclude women from influential cultural roles, assigning these roles to men. This reinforces a system where women are viewed as subordinate to men (Ortner, 1972).

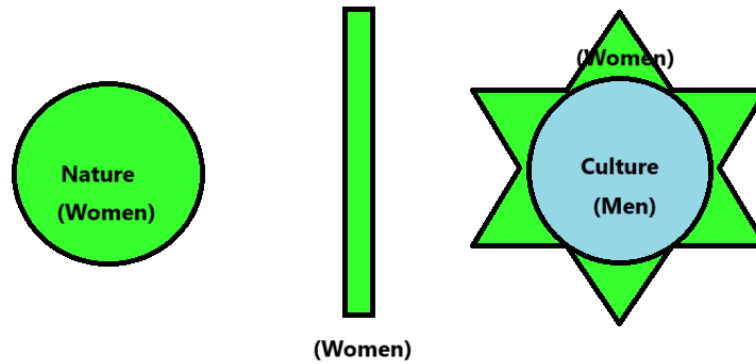


Fig. Representation of Women as Nature and cannot enter culture

Leela Dube explains that women play a crucial role in maintaining the caste system by adhering to family traditions and roles, such as childbirth, menstruation, and widowhood, which are closely tied to ideas of caste purity. She argues that women's work in traditional occupations reinforces caste-based roles. The caste system, Dube notes, is held together by strict rules that separate people into ranked groups and make different castes dependent on each other for certain tasks (Dube, 1996). Prem Chowdhry adds that caste-based marriage rules can lead to violence, especially if someone marries outside their caste. Such marriages can be seen as a threat to family honour, which is strongly connected to caste, particularly through control over women's sexuality and purity. This control often results in violence, particularly in rural areas where caste norms are stricter than in cities. According to Chowdhry, controlling women's marriage choices is a core aspect of the caste system, and any breach of these norms can have serious consequences (Chowdhry, 2007). Karin Kapadia observes that some non-Brahmin castes now value wealth and social class over caste alone when it comes to marriage. These groups prioritize financial and social status in potential matches, reflecting changes as people seek jobs and education. However, this shift has sometimes led to a decrease in women's status as traditional close-family marriages become less common. The movement from agricultural to paid work also changes family dynamics, affecting the value of women's work. Additionally, dowries are becoming more widespread, adding financial pressures on women (Kapadia, 1995). Together, these scholars highlight how caste and gender interact to shape marriage rules, family roles, and women's positions in both traditional and changing contexts.

Sylvia Walby argues that housewives and husbands should be understood as distinct social classes because of the different types of work they perform and their contrasting economic situations. According to Walby, gender roles involve more than class dynamics alone; they are deeply shaped by patriarchy, financial dependence, male authority, workplace hierarchies, and institutional policies (Walby, 1990). Walby views housework as real, valuable labour that is hard yet unpaid and should be recognized as legitimate work. Housewives not only support themselves but also contribute to their husband's and children's well-being. According to Walby, housework is not just a series of tasks but a type of work arrangement in which housewives experience an unequal exchange, often without the choice to change their "employer" (their husband), unlike regular wage workers. She also considers housewives a separate class because they are often economically dependent on and exploited by their husbands, although not all women are housewives. Walby further suggests that men and women can be viewed as distinct status groups, with housewives and husbands as distinct classes. Women can occupy a "dual class position" when they work as both housewives and wage earners, leading to mixed-class roles. Swasti Mitter (1986) adds that the global economy places women in challenging positions within the labour market, where they are often in low-paid jobs and simultaneously handling unpaid domestic labour. Economic shifts, such as Free Trade Zones, have divided the labour market into a small number of core workers and a larger group of peripheral workers, including many women in low-paid, part-time, or casual roles. This split happens partly because of societal expectations for women to also manage unpaid work at home, making it harder for them to secure stable jobs. Mitter describes women as part of a new "working class" that is largely low-paid and marginalized, facing unequal treatment in both paid and unpaid work settings. Together, these perspectives highlight how gender roles and class positions intertwine to shape women's labour and economic experiences in both domestic and global contexts.

Gemma Tang Nain discusses how patriarchy—systems where men hold more power than women—and racism uniquely affect Black women. She explains that while both patriarchy and sexism are harmful, patriarchy encompasses broader power structures beyond simple gender discrimination. Nain emphasizes that Black women face compounded challenges due to their race and gender, often earning less than white men, Black men, and white women, even when equally or more qualified. She argues that in some workplaces, gender inequality poses greater barriers than race, limiting equal opportunities for Black women (Nain, 2001). Patricia Hill

Collins expands on these themes by examining how knowledge production often centres on men and white individuals, thereby marginalizing Black women's perspectives and contributions. Collins argues that Black women hold a unique position, often excluded from both feminist and Black community discourses. She highlights foundational themes in Black feminist thought, such as self-definition, intersectionality, and celebrating Black women's culture. Collins challenges sociologists to more critically consider whose experiences they include, advocating for a fuller, more inclusive understanding of oppression (Hill Collins, 2001). Together, Nain and Hill Collins illustrate how racism and patriarchy intersect to create specific barriers for Black women in both economic and cultural spheres.

1.2. Gender Stereotypes

Stereotypes, as a concept, were first introduced in the early 20th century by Walter Lippmann in his book *Public Opinion* (1922). He described stereotypes as “pictures in our heads”—mental images that help people make sense of the world by categorizing information but can also distort reality and create biases. Gender stereotypes are a specific type of stereotype involving ideas about how people should act based on whether they are male or female. These stereotypes are widespread in many societies and influence areas like work, school, and relationships.

From a young age, children are taught societal expectations for boys and girls through family, friends, and broader social influences. Media, including TV shows, movies, and advertisements, often portray traditional gender roles, reinforcing these stereotypes. Culture and religion also shape expectations for men and women, and schools may unknowingly support these ideas. Together, families, media, social norms, schools, and workplaces all contribute to reinforcing gender stereotypes, often pressuring people to act in ways deemed appropriate for their gender. The consequences of these stereotypes can be significant, affecting individuals' self-esteem, career choices, and mental health. They can create inequalities between men and women, limit people's freedom to be themselves, and reinforce power imbalances. By promoting discrimination and restricting roles, gender stereotypes uphold systems that limit personal freedom and reinforce societal hierarchies.

Masculine gender stereotypes refer to traits, behaviours, and roles traditionally associated with men, such as strength, assertiveness, leadership, and emotional restraint. Men are often expected to exhibit aggression, physical strength, and assertiveness. Masculinity is linked to leadership roles, dominance, and control in both social and professional settings. Men are also expected to suppress emotions, showing little vulnerability or emotional expression, and to be independent, rational, and logical decision-makers. Additionally, men are often seen as the primary financial providers, expected to fulfil the breadwinner role. Masculinity can also be associated with heterosexuality and a desire for sexual conquest. On the other hand, feminine gender stereotypes refer to traits, behaviours, and roles traditionally linked to women, such as nurturing, emotional expressiveness, passivity, and domesticity. Women are often expected to take on caregiving roles, nurturing children, and caring for others. Femininity is linked to emotional sensitivity, empathy, and expressiveness, and women are sometimes stereotyped as passive, submissive, and compliant to male authority. Women are also expected to take care of domestic work, cooking, and maintaining the home. There is often a strong focus on a woman's physical appearance, with an expectation for beauty, grooming, and fashion. Finally, women are sometimes seen as dependent on men for support and guidance rather than being self-sufficient.

The “good” woman stereotype is about women being caring, kind, and focused on their family or home. Women who fit this are seen as perfect mothers, wives, and caregivers and are expected to put family first over their own dreams or jobs. The “bad” woman stereotype is for women who go against these expectations, like those who are confident, focus on their careers, or are open about their sexuality. These women are sometimes called immoral, selfish, or seen as a threat to society. This stereotype punishes women who don’t follow the usual ideas of what a woman should be.

Various ***stereotypes*** depict women in limiting and often demeaning ways. A “gold digger” is a stereotype describing a woman believed to pursue relationships with wealthy men primarily for financial gain. A “bimbo” refers to a woman perceived as attractive but unintelligent, often valued more for her appearance than her abilities. “Arm candy” is a term used for a woman who accompanies a man as a visual accessory, admired for her looks but not seen for her individuality. Similarly, a “trophy wife” is portrayed as a younger or beautiful spouse whose presence is intended to elevate her husband’s social status, serving as a marker of his success.

Another stereotype is the “sugar baby,” typically a younger woman who forms a relationship with an older, wealthier individual, often in exchange for financial support or gifts.

Positive stereotypes are when people think certain groups have good qualities, like being good at math or being caring. This can make others expect people from these groups to act in certain ways. Even though these ideas sound nice, they can still be unfair because they expect everyone to be the same. For example, some people think Asian Americans are good at math or that women are always caring, but not everyone in these groups is the same. **Negative stereotypes** are when people think bad things about a group, like saying Black people are lazy or that women can't lead. These ideas are hurtful and unfair. They can lead to discrimination, making people feel excluded or treated badly. Negative stereotypes can also cause problems in places like schools, jobs, and politics by giving some people fewer chances because of unfair beliefs.

The **socialization** of stereotypes occurs in stages. Primary socialization is when individuals are first introduced to stereotypes, often through family or close relatives. At this stage, children begin to pick up on societal attitudes and beliefs. Secondary socialization occurs as people encounter and internalize stereotypes outside the family environment through friends, schools, and educational systems. Here, stereotypes are often reinforced, becoming part of the individual's social understanding. Tertiary socialization represents the broader societal acceptance of stereotypes, which influence systemic structures such as employment, income, and political representation. Mass media, including television, movies, and advertisements, plays a significant role in this stage by continually reinforcing these ideas.

Stigmatization is when society disapproves of people because of their gender or other differences. This can stop women from advancing in their careers and limit the opportunities for both men and women. For example, women who challenge traditional gender roles, like being assertive or focused on their careers, may be negatively judged. The lasting power of stereotypes means these ideas, based on things like gender, age, race, or other traits, can continue for a long time and get passed down through generations. Gender stereotypes shape what people are expected to do, limit their behavior, and reduce opportunities over time. Role congruity theory (Eagly & Karau, 2002) says that people are seen as fitting certain roles based on gender. For women, roles that match traditional femininity, like caregiving, are seen as more suitable. In

contrast, roles connected to masculinity, like leadership, are seen as less appropriate for women, leading to bias or discrimination. This theory also suggests that society does not readily accept men who do not fit traditional masculine roles.

1.3.Bollywood Film Industry

Gender stereotypes have deeply shaped the Bollywood film industry, largely due to traditional societal beliefs and the underrepresentation of women in influential roles within the industry. Historically, Bollywood films often portrayed men as heroes and leaders, while women were cast in passive or secondary roles, often defined by their relationships with male characters. This reinforced traditional views of masculinity and femininity, presenting narrow ideas of what men and women should be like. The absence of diverse perspectives in the industry over time has perpetuated these stereotypes, limiting the complexity of gender portrayals on-screen and reinforcing traditional gender roles in society. As a result, these portrayals continue to influence public perceptions of gender, shaping societal expectations for both men and women.

1.3.1. Objectification of Women in Bollywood Films

Objectification refers to the treatment of women as objects rather than as complex individuals, reducing their value to physical appearance rather than their abilities, intelligence, or character. This form of representation is pervasive in Bollywood films, particularly in song sequences, where women are often positioned as spectacles for the pleasure of the viewer. Laura Mulvey's theory of the "male gaze" highlights how cinematic techniques are used to present women as objects of male desire, constructing them as passive subjects seen primarily through a lens of eroticism (Mulvey, 1975). Bollywood songs like "Tandoori Murgi," which compares a woman to a tandoori chicken, or "Jalebi," where a woman's body is likened to a sweet, overtly objectify women by reducing them to consumable objects of desire. In these portrayals, women are depicted not as individuals but as objects valued solely for their physical allure. Similarly, songs such as "Bidi Jalaye," "Zandu Balm," "Baby Doll," and "Param Sundari" cast women as objects of attraction, while "Cinema Hall," "Pingy Hu Mai," "Munni Badnam Hui," and "Sheila Ki Jawani" frame them as prizes to be admired or won, reinforcing the narrative that a woman's

value lies in her ability to captivate male attention. These representations perpetuate objectification by comparing women to inanimate items—foods, commodities, or trophies—conveying that they exist for visual pleasure and consumption. Such objectification implies that a woman’s worth is contingent upon her attractiveness, reducing her identity to physical appeal alone and disregarding her individuality. In treating women as mere objects, these songs reinforce cultural beliefs that deny women’s autonomy and human complexity, upholding limiting and dehumanizing stereotypes.

In many Bollywood films, women are depicted as symbols of beauty and sensuality, with camera work emphasizing body parts such as curves, cleavage, or body movements while sidelining their personalities or contributions to the narrative. This reinforces Mulvey’s concept of the male gaze, wherein women are portrayed not as fully developed characters but as objects to be looked at and enjoyed. This cinematic objectification perpetuates gender stereotypes by suggesting that women exist primarily for visual pleasure, reinforcing cultural beliefs that value women based on appearance and reinforcing barriers to recognizing their full humanity.

1.3.2. Fragmentation of Women in Bollywood

Fragmentation occurs when women are reduced to specific body parts by the camera lens, stripping them of their full identity as characters and focusing only on physical features like legs, breasts, or waist. In Bollywood and other film industries, directors frequently employ camera angles that emphasize particular parts of a woman’s body, especially in scenes where she is introduced or dancing. This technique aligns with Laura Mulvey’s concept of the “male gaze,” where the female body is segmented, or fragmented, into parts for the visual pleasure of the viewer, reinforcing the idea that women exist primarily to satisfy male desire (Mulvey, 1975). The camera often lingers on body parts, zooming in on a woman’s legs or chest, reducing her identity to an assemblage of attractive parts rather than a complete, complex character with a story. This objectifying camera technique diminishes the female character’s worth to her physical form, overshadowing her intelligence, personality, and strength. By fragmenting her body, the camera dehumanizes the woman and reinforces gender-based stereotypes, suggesting that her primary value lies in her appearance. This fragmentation not only shapes how the audience perceives female characters but also influences how other characters in the story view these

women. Female characters are often defined solely by their physical appearance, reducing them to objects of desire rather than individuals with agency, goals, and emotional depth.

1.3.3. Male Gaze and Bollywood

The male gaze refers to the way films are often made from a male point of view, with women being portrayed as objects of desire or sexual pleasure, aiming to please a presumed male audience. This idea was introduced by feminist film theorist Laura Mulvey in her 1975 essay *Visual Pleasure and Narrative Cinema*. The male gaze theory suggests that in many films, the audience is positioned to view female characters through the eyes of a heterosexual man, objectifying women in the process. In Bollywood, actresses like Katrina Kaif often perform roles that reflect this gaze, where their dance moves emphasize their bodies for male pleasure, even if their character is meant to show strength or independence. Visual techniques used to portray the male gaze often involve the camera lingering on a woman's body, focusing on her face, body movements, and clothing. This turns the woman into an object, and her worth in the story is often defined by how male characters or the audience react to her. The male gaze has a psychological impact on both men and women. For men, it reinforces the idea that women are objects to be looked at and desired. For women, it can make them believe their value is largely tied to their physical attractiveness, leading to pressure to fit certain beauty ideals.

1.3.4. Gender Pay Gap in Bollywood

The gender pay gap in the film industry refers to the significant difference in earnings between male and female actors, even when they have similar or equally important roles. Female actresses often earn considerably less than male actors for the same or comparable roles despite the commercial success of films where female leads are central. For instance, a successful female lead in a Bollywood film might earn less than a male actor in a similar role, even if the female-led film performs well at the box office. This disparity is partly due to long-standing gender norms and male dominance in the industry, where male actors have historically been seen as the “main star,” leading to higher pay for them. Although women are increasingly taking on prominent lead roles, they still earn less than their male counterparts, reinforcing gendered hierarchies within the industry. The pay gap limits female actresses' financial freedom and causes the idea that men are more commercially capable than women. This also affects the types

of roles women are offered, often confining them to supporting roles while male actors continue to dominate leading roles with higher salaries. Ultimately, the wage disparity in the film industry reflects broader societal inequalities where women are undervalued in many professional fields, and the film industry is the best example of these social norms.

1.3.5. Professionalism and Women Actors

Professionalism in the film industry means treating people, especially actors, with respect and fairness in their careers. However, women often face gender biases that make it harder for them to be seen as true professionals. In the film industry, women are often judged more for their looks than their acting skills, while men are usually judged by their talent and experience. Women are also frequently typecast into specific roles, like the love interest or the supporting mother, limiting their career opportunities. Men, on the other hand, are given a wider range of roles. Female actors also have fewer opportunities to take on powerful or complex characters and are often limited to roles that focus on their physical appearance or sexuality. In addition, women in the film industry experience sexism and harassment, such as being asked to perform sexualized scenes or dealing with inappropriate comments from colleagues. This creates an unequal and uncomfortable working environment. And the film industry is still mostly controlled by male directors, producers, and writers, which means there are fewer female ones in storytelling. When women are in creative roles, their stories tend to be more diverse and reflective of real female experiences, but the lack of representation in key roles affects how women are seen on screen.

2. Methodology

2.1. Data Collection

The data for this project was extracted using the TwitterComment-Extract Chrome extension. This extension takes the URL of specific Bollywood celebrity Twitter accounts and extracts the comments from the respective handles. The task involved analyzing the interactions of 10 Bollywood actors (a convenient sample), consisting of 5 female actresses (Priyanka Chopra, Kiara Advani, Alia Bhatt, Anushka Sharma, and Deepika Padukone) and 5 male actors (Shah Rukh Khan, Salman Khan, Hrithik Roshan, Ranveer Singh and Akshay Kumar) total of around 25,000 comments were collected—12,500 comments for actresses and 12,500 for actors. These comments were analyzed to determine sentiment and to explore how these sentiments reflected potential gender stereotypes.

2.2. Data Preprocessing

The data preprocessing step involves cleaning and transforming raw textual data for sentiment analysis. The raw data for male and female Bollywood actors is loaded into a CSV file. Cleaned the data by removing unwanted elements from the text, such as URLs, Mentions (@username), Hashtags, Numbers, Special Characters, and Emojis, as these elements do not contribute meaningfully to sentiment analysis. Handling missing values and after preprocessing, the cleaned data is saved to new CSV files.

We used VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool to classify the sentiment of each tweet as positive, negative, or neutral based on the content of the tweet. For each tweet, it calculates a sentiment score using VADER and classifies the sentiment into three categories: Positive: Compound score ≥ 0.05 , Neutral: $-0.05 < \text{Compound score} < 0.05$, Negative: Compound score ≤ -0.05 . Sentiment Distribution counts the number of tweets in each sentiment category (positive, negative, neutral) for both male and female actors, as shown in (fig. 1)

2.3. Model Training and Evaluation Process

For our sentiment classification task, we trained three models—Logistic Regression, Naive Bayes, and Support Vector Machine (SVM)—on a combined dataset of tweets directed at Bollywood actors. Missing values in the tweet text were replaced by empty strings, stored in X, while sentiment labels from VADER analysis were stored in y. After an 80:20 train-test split, a TF-IDF vectorizer was fitted on the training set to ensure consistent feature representation across both sets.

Each model offered unique benefits: Logistic Regression handled binary and multi-class classifications with up to 1000 iterations for convergence; Naive Bayes (MultinomialNB), assuming feature independence, is well-suited for text classification; and SVM, using a linear kernel, enabled class probability estimation useful for imbalanced data. We trained each model on the TF-IDF-transformed data and evaluated them using accuracy, precision, recall, and F1 scores, as shown in (Table. 1). A heatmap confusion matrix visually highlighted the models' strengths and weaknesses, such as common misclassifications between sentiment classes as shown in (fig.2). This evaluation allowed us to compare models effectively, identifying which one best-balanced precision and recall across sentiment classes, offering the most accurate classification.

2.4. Stereotype Analysis on Gender

To explore gender stereotypes in tweets directed at Bollywood celebrities, a stereotype analysis identifies and counts the use of gendered words in user comments. The analysis defines two lists of stereotypical words: one reflecting masculine stereotypes (e.g., “strong,” “leader”) and the other reflecting feminine stereotypes (e.g., “caring,” “gentle”). This approach is applied separately to datasets of tweets for actors and actresses celebrities, tallying occurrences of each stereotype in user language. By summarizing these stereotype counts, as shown in (Table 2) and (fig.3), we gain insights into the prevalence and type of stereotypical language used for male versus female actors. This method helps highlight any patterns in user sentiment, revealing differences in how masculinity and femininity are discussed in public interactions with male and female celebrities.

The analysis of the top 5 stereotypical keywords used in comments directed at Bollywood actors and actresses reveals distinct patterns in how masculinity and femininity are expressed, as shown in (Table. 3) and (fig. 4)

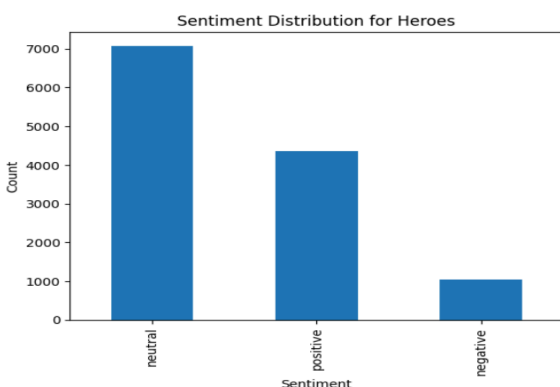
2.5. Stereotype Analysis on Individual actors and actresses

To deepen the analysis of gender stereotypes, we examine individual actors by counting words linked to both positive and negative stereotypes associated with masculine and feminine traits. These stereotypes, provided in both English and Hindi (in Roman script), capture a broader array of keywords relevant to sentiment analysis. For each comment directed at an actor or actress, four counts are calculated: positive masculine, negative masculine, positive feminine, and negative feminine (Table. 4). Visualization through bar charts displays these stereotype counts across categories, providing an accessible comparison of common stereotypes associated with actors versus actresses celebrities as shown in (fig. 5) . This approach helps reveal which types of stereotypical language are more prevalent for each gender in user interactions.

3. Results

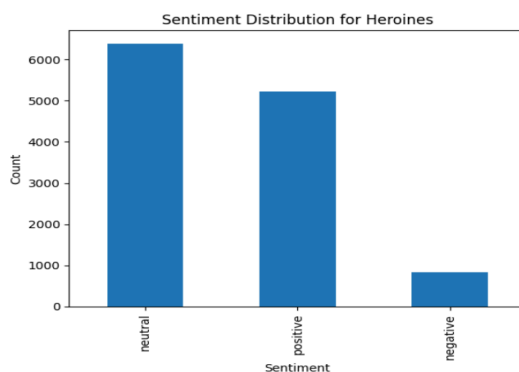
VADER_Sentiment for Heroes

neutral	7079
positive	4357
negative	1047



VADER_Sentiment for Heroines

neutral	6388
positive	5230
negative	828



2.2_Fig.1 shows that both Heroes and Heroines are predominantly discussed with neutral sentiment, followed by positive mentions, while negative sentiments are minimal. Heroes received 7,079 neutral, 4,357 positive, and 1,047 negative mentions, while Heroines received 6,388 neutral, 5,230 positive, and 828 negative mentions.

Results of Models:

Logistic Regression Accuracy: 0.9142

	Precision	Recall	f1-score	support
Negative	0.87	0.45	0.59	374
Neutral	0.90	0.99	0.94	2669
Positive	0.95	0.90	0.93	1943
Accuracy			0.91	4986
Macro Avg	0.90	0.78	0.82	4986
Weighted Avg	0.91	0.91	0.91	4986

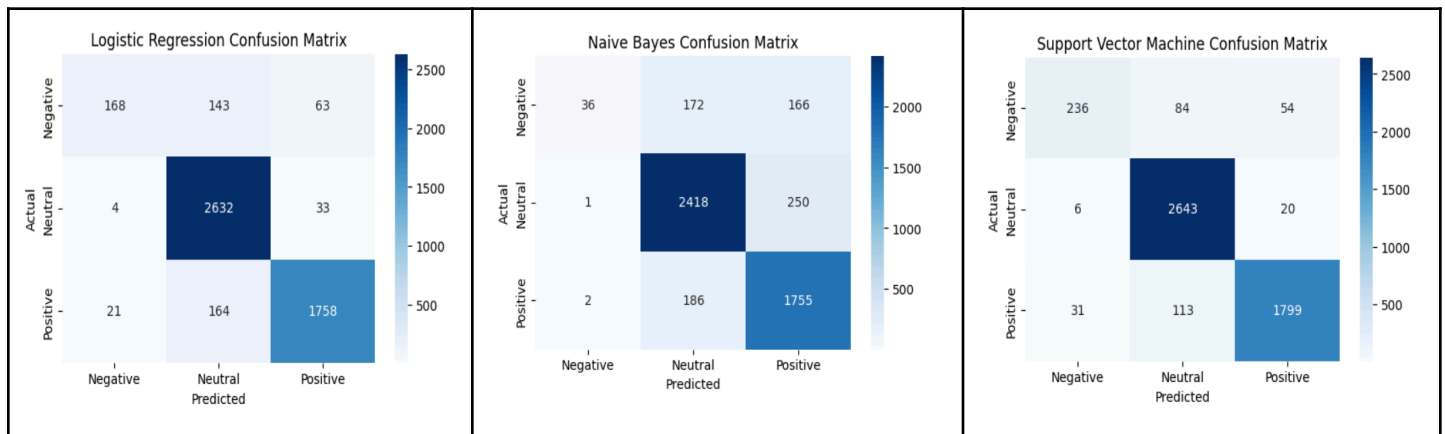
Naive Bayes Accuracy: 0.8442

	Precision	Recall	f1-score	support
Negative	0.92	0.10	0.17	374
Neutral	0.87	0.91	0.89	2669
Positive	0.81	0.90	0.85	1943
Accuracy			0.84	4986
Macro Avg	0.87	0.64	0.64	4986
Weighted Avg	0.85	0.84	0.82	4986

Support Vector Accuracy : 0.9382

	Precision	Recall	f1-score	support
Negative	0.86	0.63	0.73	374
Neutral	0.93	0.99	0.96	2669
Positive	0.96	0.93	0.94	1943
Accuracy				0.94
Macro Avg	0.92	0.85	0.88	4986
Weighted Avg	0.94	0.94	0.94	4986

2.3_ Table. 1 (Classification Report) shows that the *Support Vector Machine (SVM) model* **outperformed** both Logistic Regression and Naive Bayes, achieving the highest accuracy (93.82%) and most balanced performance across sentiment classes. While Logistic Regression also performed well, Naive Bayes struggled with detecting negative sentiments.

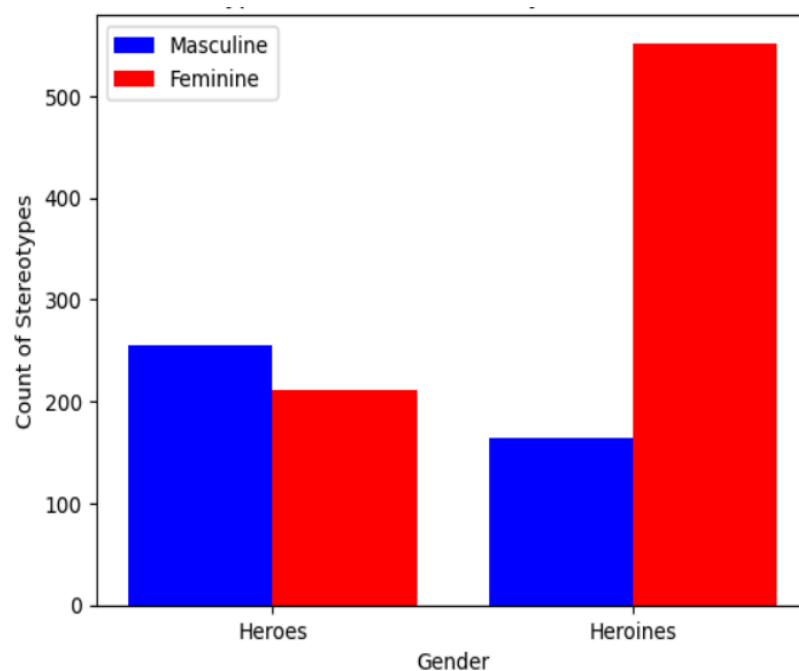


2.3_ Fig.2 Each Confusion matrix shows actual vs predicted sentiment classifications of each model, with darker blue indicating higher numbers of predictions. SVM appears to have slightly better performance than the other two models based on the higher numbers in the correct prediction.

Gender Stereotypes (actors and actresses):

	Masculine Count	Feminine Count	Total Tweets
Heroes Stereotype Summary	255	211	12483
Heroines Stereotype Summary	164	552	12446

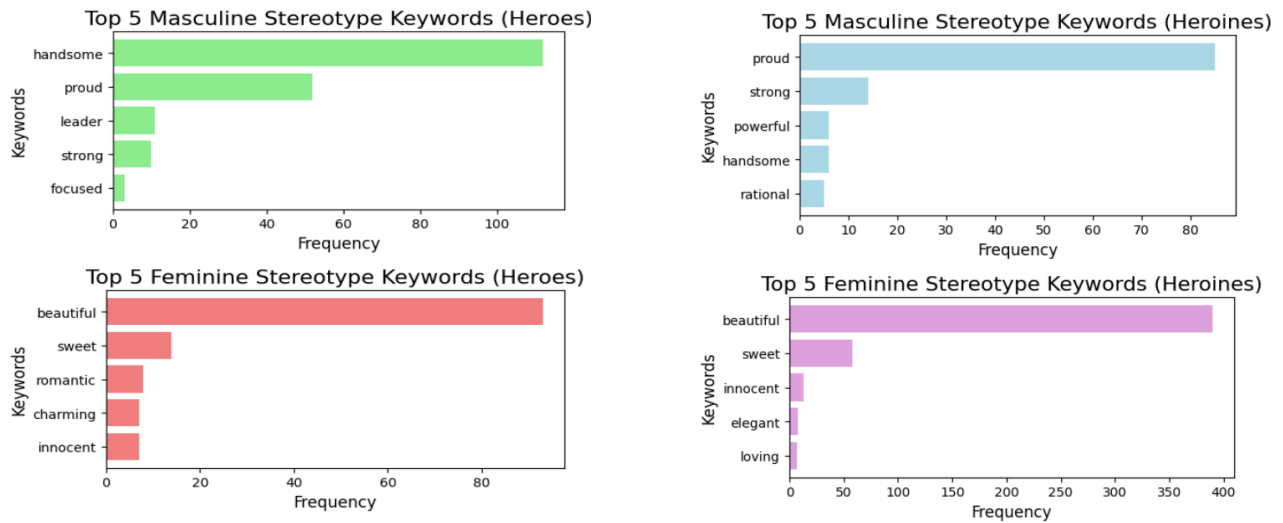
2.4_ Table. 2 represents the total Masculine and Feminine counts of Heroes and Heroines, and the Feminine count is higher for Heroines, where the Masculine count is higher for Heroes.



2.4_Fig.3 shows the visual representation of Gender Stereotypes in Tweets of Bollywood Actors and Actresses.

Masculine(Heroes)		Feminine(Heroes)		Masculine(Heroines)		Feminine(Heroines)	
Handsome	112	Beautiful	93	Proud	85	Beautiful	390
Proud	52	Sweet	14	Strong	14	Sweet	58
Leader	11	Romantic	8	Powerful	6	Innocent	13
Strong	10	Charming	7	Handsome	6	Elegant	8
Focused	3	Innocent	7	Rational	5	Loving	7

2.4_Table.3 shows the Top 5 Keywords of Gender Stereotypes, and This highlights a clear gender bias, with male celebrities primarily linked to masculinity and female celebrities to femininity.



2.4_Fig.4 shows the Visual representation of the Table.3

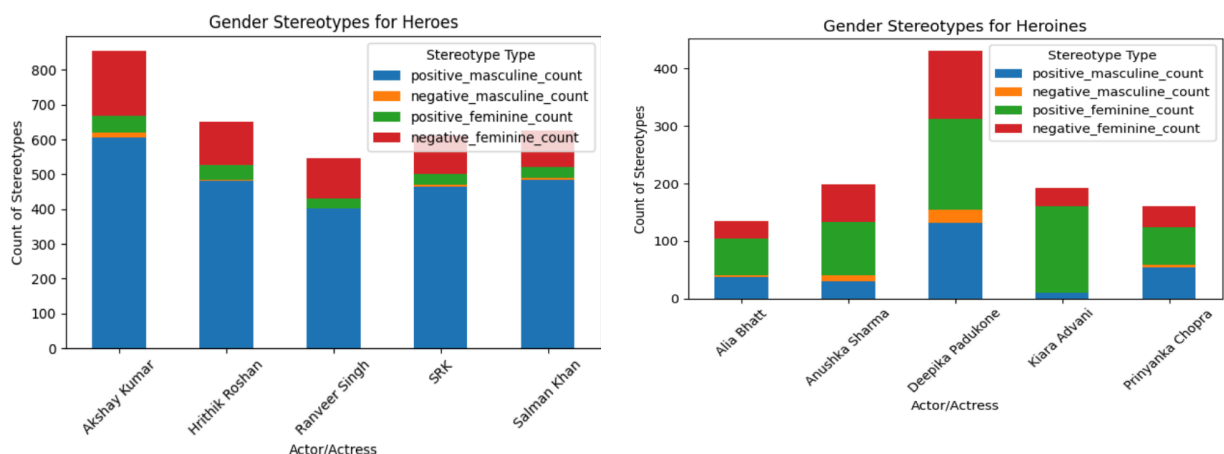
Heroes Stereotype Summary :

Name	Positive Masculine Count	Negative Masculine Count	Positive Feminine Count	Negative Feminine Count
Akshay Kumar	607	12	49	186
Hrithik Roshan	481	4	41	126
Ranveer Singh	401	0	29	118
SRK	464	6	31	110
Salman Khan	485	5	30	106

Heroines Stereotype Summary :

Name	Positive Masculine Count	Negative Masculine Count	Positive Feminine Count	Negative Feminine Count
Alia Bhatt	38	1	63	30
Anushka Sharma	30	10	93	64
Deepika Padukone	131	18	157	116
Kiara Advani	10	0	149	31
Priyanka Chopra	54	3	65	36

2.5_ Table. 4 shows the Positive and negative Masculine and Feminine counts of each hero and heroine.



2.5_ Fig. 5 shows the visual representation of the Table.4

4. Discussion

The sentiment analysis of Bollywood celebrities on Twitter reveals distinct patterns in how actors and actresses are perceived, highlighting differences in sentiment distribution, model performance, and gender stereotypes. This analysis uncovers an interesting mix of expected and surprising portrayals of Bollywood figures on social media. The sentiment analysis shows that both heroes and heroines are primarily discussed with neutral sentiment, followed by positive mentions and a lower frequency of negative sentiments. Specifically, heroes received 7,079 neutral, 4,357 positive, and 1,047 negative mentions, while heroines received 6,388 neutral, 5,230 positive, and 828 negative mentions. This trend indicates that public discourse surrounding both groups tends to be neutral, with a strong presence of positive sentiment, reflecting the public admiration and celebrity status of these figures. Negative mentions are relatively low, possibly due to fans' tendency to engage positively with well-known celebrities. Notably, heroes attract more negative sentiment than heroines, potentially due to controversies, such as extramarital affairs or career setbacks, which may impact their public image more than that of heroines.

In terms of model performance, the Support Vector Machine (SVM) model outperformed both the Logistic Regression and Naive Bayes models, achieving the highest accuracy (93.82%) and a balanced performance across sentiment classes. This suggests that SVM effectively captures the nuances of sentiment classification for Bollywood-related tweets. In comparison, Naive Bayes struggled with detecting negative sentiment, as shown by its lower precision and recall for the negative category. Logistic Regression performed well but still trailed behind SVM, underscoring SVM's robustness for this type of classification task.

Analysis of gender stereotypes in tweets reveals notable differences in how heroes and heroines are perceived. Male celebrities are predominantly associated with traditional masculine stereotypes, with 255 mentions of masculine traits versus 211 mentions of feminine traits. Conversely, heroines are strongly associated with feminine stereotypes, with 552 mentions of feminine traits and only 164 mentions of masculine traits. This disparity highlights the persistence of traditional gender roles in Bollywood, where male celebrities are celebrated for qualities like strength, leadership, and pride. In contrast, female celebrities are valued for their beauty, kindness, and innocence. This finding aligns with **Sherry B. Ortner's *Nature and Culture theory***, which associates men with cultural accomplishments and leadership roles and women with beauty and nurturing qualities, reflecting a traditional view of gender roles.

The gender stereotype analysis in Table 3 shows clear patterns in how Bollywood heroes and heroines are characterized based on stereotypical traits. Male celebrities, or “heroes,” are frequently linked to positive masculine attributes, such as being “handsome” (112), “proud” (85), “strong” (14), and a “leader” (11). These descriptions suggest that male heroes are celebrated for qualities like pride, strength, and leadership, fitting conventional masculine stereotypes. Although male celebrities occasionally attract feminine descriptors like “sweet” (14) and “innocent” (7), these are less common, supporting the expectation that men should embody toughness and authority.

Female celebrities, or “heroines,” are predominantly associated with traditional feminine qualities, with descriptors such as “beautiful” (390), “sweet” (58), and “innocent” (13) appearing frequently. This emphasis on appearance and gentleness reflects conventional feminine ideals, indicating that women are often valued more for their physical beauty and perceived kindness

than for attributes linked to strength or autonomy. While a few female actors, such as Deepika Padukone, are occasionally associated with traditionally masculine traits like “strong” (10) and “focused” (3)—likely due to her roles in action-oriented films such as *Pathan* and *Jawan*—these instances remain relatively infrequent, highlighting Bollywood's continued adherence to gendered representations.

Table 4 shows individual male gender stereotypes that show how Bollywood heroes like Akshay Kumar, Hrithik Roshan, Ranveer Singh, SRK (Shah Rukh Khan), and Salman Khan are seen in terms of gender stereotypes. These actors are mostly linked to positive masculine traits, like strength, pride, and leadership, which is reflected in their high “positive masculine count.” For example, Akshay Kumar has the highest count (607), indicating that he is often portrayed as a strong, proud, and reliable hero because of the movie concept that he takes. Other actors, like Hrithik Roshan (481) and Salman Khan (485), also have highly positive masculine traits, showing that they are seen as powerful, confident, and dominant figures. The “negative masculine count” is relatively low for these actors, suggesting that they are not often shown with negative masculine traits, such as being arrogant or aggressive. If that is the case, then also society takes it as masculine behavior. For instance, Hrithik Roshan has just 4 negative masculine traits, showing that he is primarily seen in a positive light due to his handsome face and dance skills. When it comes to feminine traits, the counts are much lower for these male celebrities. For example, Akshay Kumar has only 49 positive feminine traits, showing that he is rarely associated with traditionally feminine qualities like gentleness or nurturing. The negative feminine traits, such as being weak or vulnerable, are more noticeable. Akshay Kumar has a significantly higher count of 186 negative feminine traits, which reflects the traditional stereotype that men should not be soft or emotional, potentially influenced by the comedic or empathetic roles that he usually plays. For instance, in the movie *Padman*, his portrayal aligns with traits often categorized as “feminine,” such as empathy, caregiving, and sensitivity to community needs.

Looking at individual heroines, Alia Bhatt’s positive feminine traits (63) are higher than her positive masculine traits (38), showing that her portrayal leans heavily on feminine stereotypes,

with a few masculine traits. Her negative feminine traits (30) suggest that she may be seen as innocent or vulnerable in some roles or due to nepotism. Anushka Sharma has a relatively balanced portrayal, with high counts of positive feminine traits (93) and a notable amount of positive masculine traits (30). This could mean that her roles sometimes show both feminine and masculine strengths, and her negative feminine traits (64) are high because she is the wife of Virat Kohli, and his fans don't want her to do romantic roles in movies and wear revealing clothes. Deepika Padukone has the highest positive trait counts, with 131 masculine and 157 feminine, suggesting that she has a mix of both types of traits, reflecting strong, complex characters. These positive traits may be because of her recent pregnancy, hit movies, and also because she is a self-made actress. Kiara Advani has mostly feminine (149 positive feminine traits) with almost no masculine traits because of her characters or roles in the movies like traditional feminine qualities like beauty, charm, and sensitivity. Priyanka Chopra balances both sides, with 54 positive masculine and 65 positive feminine traits, representing strength and independence alongside traditional beauty. However, her lower count overall may reflect her recent absence from Bollywood due to her international projects, such as the upcoming film *The Bluff* alongside Karl Urban.

Although actresses like Deepika Padukone and Priyanka Chopra show both masculine and feminine traits, Bollywood has been slow to create more complex characters that break gender stereotypes. As society changes, we may see more balanced roles, but traditional gender roles still strongly influence Bollywood films.

5. Conclusion

In conclusion, this analysis of Twitter comments on Bollywood celebrities highlights the persistent influence of gender stereotypes in public sentiment and digital discourse. Among the machine learning models employed—Logistic Regression, Naive Bayes, and SVM—SVM emerged as the most effective in capturing sentiment nuances within Bollywood-related content, reinforcing its utility for sentiment classification in social media contexts. The gender stereotype analysis reveals a clear bias: male actors are predominantly associated with traits like strength

and leadership, while female actors are often linked to beauty and kindness. This pattern mirrors entrenched societal narratives about gender roles, indicating that social media continues to reflect and reinforce traditional stereotypes. As digital platforms increasingly shape public perceptions, this study emphasizes the importance of addressing gender bias in online discourse. By promoting more balanced representations, there is potential to foster a more equitable digital environment that challenges rather than perpetuates outdated gender norms.

References

- Chodorow, N. (1978). *The reproduction of mothering: Psychoanalysis and the sociology of gender*. University of California Press.
- Chowdhry, P. (2007). *Contentious marriages, eloping couples: Gender, caste, and patriarchy in Northern India*. Oxford University Press.
- Eagly, A. H., & Karau, S.J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>
- Dube, L. (1996). *Caste and women*. In M. N. Srinivas (Ed.), *Caste: Its twentieth-century avatar* (pp. 1-27). Viking.
- Hill-Collins, P. (2001). *Learning from the outsider within: the sociological significance of Black feminist thought*, From *Gender: A Sociological Reader*.
- Lippmann, W. (1922). *Public opinion*. Harcourt, Brace and Company.
- Nain, G. (2001) *Black women, sexism and racism*, From *Gender: A Sociological Reader*.
- Kapadia, K. (1995). *Siva and her sisters: Gender, caste, and class in rural South India*. Westview Press.
- Mitter, S. (1986). *Common fate, common bond: Women in the global economy*. Pluto Press.
- Mulvey, L. (1975). Visual pleasure and narrative cinema. *Screen*, 16(3), 6–18. <https://doi.org/10.1093/screen/16.3.6>
- Ortner, S. B. (1972). *Is female to male as nature is to culture?* *Feminist Studies*, 1(2), 5-31. <https://doi.org/10.2307/3177638>
- Walby, S. (1990). *Theorizing patriarchy*. Basil Blackwell.