A Century of Gender Bias in the Film Industry: Men have a Career whereas Women are Married

Anonymous ACL submission

Abstract

Movies play a crucial role in forming the perception and the mentality of masses about the society. Existence of bias and stereotypes in movies can harmfully reinforce and internalize negative biases in the audience. We present a study of gender bias prevalence in movie plots using a novel event chain-based feature extraction method. In our study, we model movies as sequences of events that correspond to the underlying characters and explore the effect of events on the way characters are described in movie plots. To increase our outreach, our study is performed over a century of data and includes movie plots from three most-watched film industries: Hollywood, Cinema of India, and Cinema of the United Kingdom. We demonstrate that our method quantitatively demonstrates existence of gender bias in movie plots as gendered characters evolve during movies. Our study is worrisome but at the same, our temporal analysis indicates that gender bias has a decreasing trend over the past century.

1 Introduction

The film industry plays an important role in forming societal norms and shaping the cultural attitudes and customs of people. The audience often internalize the qualities of popular movie characters as ideal qualities and are influenced by their portrayal. As a result, bias in movies can shape harmful attitude in the society and reinforce existing negative biases. For example, it is easy to find movies in which male protagonists are portrayed as highly qualified professionals such as a successful businessman, surgeon, lawyer etc. Whereas, female characters are portrayed as weak and fragile, and are associated with family-related tasks. Existence of biased movies along with dominance of men in the film industry (Lauzen, 2011) raises a natural question: Is gender stereotyping a common phenomenon in movie plots? For example, Are male characters generally portrayed with identifiable and remarkable achievements than female characters?

Modern natural language processing (NLP) provides powerful tools to study the existence of gender bias in large corpora and this problem has been explored in different contexts. For example, it has been demonstrated that gender bias exist in word embeddings (Zhao et al., 2019; Lewis and Lupyan, 2020). However, investigating the existence of gender bias in movie plots is an under-explored area in the NLP community. A few recent works have focused on evaluating gender bias in movies (Madaan et al., 2018), but these works are largely limited to discovering asymmetric association of certain appearance or occupation terms with each gender and do not consider the temporal nature of movies. A crucial aspect of movie plots that distinguishes them from many texts is portrayal of characters when major events occur. It is interesting to study how different characters evolve over time as certain events occur in a movie, e.g., portrayal of males vs females after marriage events. We study the presence of gender bias in movie plots as characters evolves during movies. Our contributions include:

- We contribute and tailor a dataset of movie plots from three major film industries for the years between 1930 and 2020.
- We propose an event chain method to detect gender bias. Movies are a sequence of events with each character having a different event chain. Analysis of these events help to properly reflect the underlying bias in temporal representation of different characters.
- We utilize Unigrams, Bigrams, and Weighted Log Odds Ratio to detect the different event-dependent words and phrases that are more prevalent in males versus females roles.
- We perform temporal analysis to study the trend of bias in movie plots in the past century.

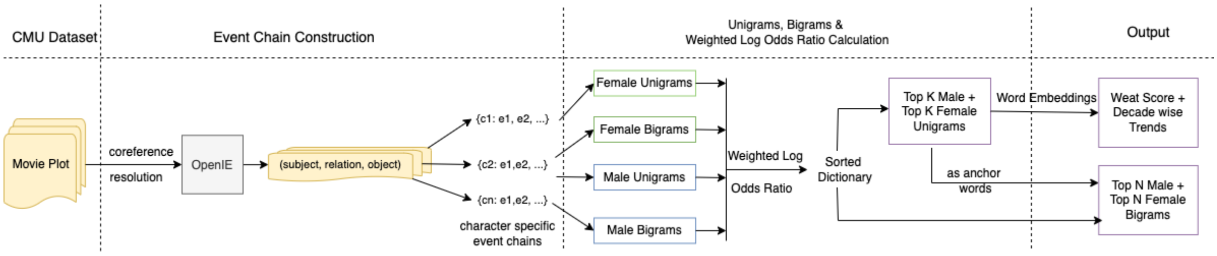


Figure 1: Architecture of the proposed processing scheme

2 Methodology

A block diagram-level visualization of our data processing pipeline is presented in Figure 1.

2.1 Datasets

We expand the CMU dataset (Bamman et al., 2013) for our study. It contains 42,306 movie plot summaries, extracted from Wikipedia on November 2^{nd}, 2012. In addition to the plots, the dataset contains the release date and genders of the actors who play different characters in the movie. Since, the CMU dataset contains movie plots until 2012, we separately extract the remaining plot summaries for the years 2013-2020 from Wikipedia. Our dataset includes 22901 Hollywood movies, 7644 movies from the cinema of India (CoIN), and 4837 movies from the cinema of the UK (CoUK) to include major film industries.

2.2 Event Chain Construction

To study bias amongst different genders, we explore how different characters with different genders are portrayed. Rather than simply studying the association of terms with genders, we study the events that correspond to plot characters and their effects on characters' portrayals and evaluations during the entire movie plot. Our methodology reveal biases in a more meaningful way because association of some terms with genders might not be due to bias. Events can be any combination of verbs and adjectives describing the characters and other complex arguments about time, location, etc.

Prior to extracting character-specific events, we perform coreference resolution (Lee et al., 2017) on the plots so that the pronouns are associated with the corresponding character name. A simple English sentence has the (subject, verb, object) structure. In the sentences of movie plots, the subject can refer to a character and the verb and the object can refer to an event associated with that character. We use OpenIE (Mausam, 2016) tool to find insights about the events. We use it on every sentence to obtain tuples of the form (subject, relation, object) for all the events in a sentence. The relation and the object in the tuple are considered as events to construct subject-wise event chains.

The CMU dataset includes genders of actors as character metadata, but this gender mapping is not provided for all the movies' characters. To overcome this challenge, we use the Python builtin library, gender_predictor.GenderClassifier. The average accuracy of the predicted gender is 96%.

2.3 Unigrams, Bigrams and Weighted Log Odds Ratio

To detect bias, we are interested in words that occur distinctively either only for males or females. Hence, Odds Ratio can be effectively used to find the large frequency differences for females versus males. Note, however, a simple Odds Ratio calculations similar to prior works (Sun and Peng, 2021) is not applicable as it does not account for rare words that might be very frequent for a particular character or a movie plot, leading skewness in odds ratio counts. Moreover, simple odd ratio is unreliable for comparing small counts because a word may appear once for females and the same word may not even appear for males or vice versa. To mitigate these issues, we use Weighted Log Odds Ratio (WLOR) (Liberman, 2014). Let W_{w_g} be the frequency and α_{w_g} be the prior probability for each word, where w_g indicates gender. Also, let n_g be #words for each gender. Then, first calculate T_m and T_f and then the WLOR for each word:

$$
\begin{aligned}
T_g &= log[\frac{(W_{w_g} + \alpha_{w_g})}{(n_g + W_{w_g} + \sum_w \alpha_{w_g} + \alpha_{w_g})}] \\
odds_w &= \frac{T_m - T_f}{\frac{1}{W_{wm} + \alpha_{wm}} + \frac{1}{W_{wf} + \alpha_{wf}}}
\end{aligned} \quad (1)
$$

Prior to computing WLOR, we lemmatize using NLTK and remove stopwords. The result is

Industry	Bigrams for Males	Bigrams for Females
CoIN	rich businessman, software engineer, becomes successful, agrees marriage, brilliant student, film director, house wife, government officer, set fire, becomes stronger	becomes pregnant, commit suicide, avenge death, accepts proposal, wealthy lifestyle, reject marriage, love family, approves relationship, lawyer profession, friend travel
CoUK	security guard, college student, join force, commanding officer, proposes marriage, insane wife, love triangle, badly injured, war criminal, help escape	give birth, return home, estranged husband, begin affair, casual sex, lead group, home caregiver, leave house, marry love, go bed
Hollywood	lead group, love wife, commanding officer, win race, proposes marriage, join army, work life, crash party, basketball team	give birth, fall love, happily married, fashion designer, share kiss, becomes hysterical, pearl necklace, proposal marriage, asks husband, becomes jealous

Table 1: Events Bigrams from CoIN, CoUK and Hollywood

then fed into WLOR to extract the asymmetric Unigrams and Bigrams. The basic strategy is to first model word usage in the different female and male characters and then to investigate how does the word usage for male and female diverges from that in the whole combined event chain for both genders. WLOR enhances differences that are unexpected given the null hypothesis that both male and female event chains are random selections from the same vocabulary. Finally, we remove gendered words like mother, woman, lady for females and man, father for males. To strengthen the analysis, bigrams are also considered. In order to get meaningful bigrams, unigrams are used as anchors, i.e., bigrams that contain the unigrams are reported.

2.4 WEAT Score

The Word Embeddings Association Test (WEAT), as proposed by (Caliskan et al., 2017) is a statistical test to measure human bias in textual data. WEAT uses two lists of target words and two lists of attribute words. The first pair of lists correspond to terms we want to compare with each other and the second pair of lists represent the categories in which we believe bias can be present. [1]

3 Results and Discussions

We analyzed gender bias for the three film industries over time and male/female associated words.

3.1 Verifying the efficacy of our analysis methodology

The WEAT Scores are calculated for words associated with Male vs Female with gendered attributes taken from Caliskan et al. 2017 to follow the common trend in the literature. We train a word2Vec model on the movie plots separate for each film industry in order to get the word embeddings. Non-negative WEAT scores indicate existence of gender Bias in movie plots across all film industries.

3.2 Temporal Analysis of WEAT Score

Quite interestingly, we observe that bias increases in all three film industries from. The bias peaks around the 2000s and beyond that we observe that bias decreases. The exact reason behind this trend requires further analysis but we speculate this trend has occurred first due to less bias in the mentality of people as a result of improved education and less bias in the society. Second, this observation may stem from active efforts in the film industry to mitigate gender bias in movies. For this reason, we choose the year 2000 as the cutoff point for the coming discussion. An additional advantage of choosing the year 2000 is that it gives an approximately equal number of films before and after 2000 due to a boom in movie production in 2000s. We choose the top 250 words while calculating the WEAT score in our analysis. [2]

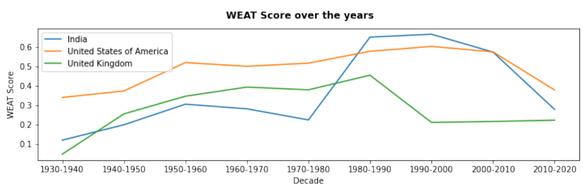


Figure 2: Temporal Analysis of WEAT Scores for CoIN, CoUK and Hollywood

Gender bias in the years before 2000: Men have a career whereas Women are married: We have presented results for each industry in Table 2. In general, Females are described by words “love”, “jealous”, “raped”, and “husband” which are at-

[1]For more details, please check App. A.1

[2]Check App. A.2 for unigram word clouds before and after 2000 for CoIN, CoUK and Hollywood

Country	Unigrams for Males	Unigrams for Females	WEAT score
CoIN	rich, police, businessman, challenge, criminal, fight, leader, honest, attack, successful	love, husband, birth, marry, pregnant, worship, suicide, widowed, necklace, beautiful, raped	0.768
CoUK	wife, guard, agent, chase, pirate, policeman, charge, order, mission, commander	husband, baby, pregnant, jealous, marry, sex, house, lover, prostitute, lonely	0.429
Hollywood	wife, army, officer, life, successful, battle, impress, journalist, command, professional	pregnant, husband, love, married, baby, passionate, raped, heartbroken, lonely, forbids	0.523

Table 2: Excerpt of Unigrams for the three industries in the years **before 2000** and their respective WEAT Scores

Country	Unigrams for Males	Unigrams for Females	WEAT score
CoIN	businessman, money, software, shot, kill, money, apology, leader, attack, mercy	love, husband, marriage, beautiful, birth, pregnant, respect, dance, actress, lawyer, wealthy	0.751
CoUK	army, attack, battle, catch, charge, proposes, care, escape, money, awarded	pregnant, witch, sex, birth, help, marriage, love, care, home, justice, bed	0.405
Hollywood	officer, girlfriend, fight, partner, government, successor, singer, humanity, nightclub, team	pregnant, husband, baby, lesbian, attractive, scholarship, ambassador, adventure, fashion, dance	0.504

Table 3: Excerpt of Unigrams for the three industries in the years **after 2000** and their respective WEAT Scores

tributes of romantic relation and submissive character. In contrast, males are described by words "leader", "police", "businessman", "attack" and "successful" which are attributes of career and dominance. This represents a significant amount of gender bias amongst all the three film industries.

Gender bias in the years after 2000: Bias still exists but women now have careers as well and Men are more caring: Table 3 presents our results for the 2000s years. We observe that females are still described by words "love", "beautiful", "husband", "marriage", "pregnant" which are attributes of appearance and romantic relations but we can also see career related words such as "actress", "lawyer", ans "ambassador" which is a positive sign and indicates less bias. On the other hand, words like "apology" have started to appear for describing men which indicates emergence of emotional attributes for men. Our analysis indicates that the struggle to tackle bias in the film industry is not over and more efforts are necessary. However, the decreasing bias in 2000s indicates the trend is heading towards the right direction.

3.3 Bigram Analysis

In order to make our analysis more event-based, we also analyze the differences in Bigrams for each of the genders. The main benefits of using bigrams is to understand the even-related context in which some of the words are used. For example, consider CoUK, where the unigram "wife" and "husband" is common for both male and female genders. But, one of the top bigrams is "insane wife" for men whereas "estranged husband" for females which indicates existence of extreme biased. Furthermore, "give birth" seems to remain a top bigram for females across all the three film industries whereas men are associated with more career related bigrams such as "rich businessman". [3]

4 Conclusions

We have established an event chain extraction based technique to detect gender bias in movie plots. Our results conclude that there is a significant amount of gender bias in movie plots in all primary film industries. Expectedly, the existing bias is more against females. However, our analysis of movies from the 21^{st} century indicates that descriptions of female characters have become less biased and with showcasing women characters as more "career" personalities than "household" personalities. However, there's still a significant focus on appearances and relations than merits and accomplishments for female characters. Since movies have a strong impact on society, it is essential we showcase an unbiased society in movie in hopes of this being a reality in future. Our analysis indicates we still are far away from having such a film industry.

[3] See App. A.3 for changes in Bigrams before and after 2000

Ethical Considerations

We have used the CMU Movie Dataset for the years 1930-2012 and for 2013-2020, we collected it from Wikipedia. However, we also acknowledge that potential biases present in movies plots may also vary with respect to movies dataset. The main aim of our work is create awareness regarding gender bias in movie plots. However, it can be misconstrued by some to create more biased texts.

Our work has a few limitations. Firstly, the movies that involve LGBTQ+ community are less which limits our work to two genders - Male and Female. Furthermore, it is possible that the movie does represent Male and Female characters with equal importance, but due to gender stereotypes present in the society might influence writing style of Wikipedia contributors. Finally, data collected from Wikipedia is limited to already available metadata, directors information, casting crew information, etc. To perform statistical and temporal analysis of gender bias in movies over decades we requires large number of movie plots. Our work can possibly contribute to the future work for detecting bias for non binary genders in movies, also can be used in detecting bias in various story line contents such as books.

References

David Bamman, Brendan O'Connor, and Noah A. Smith. 2013. Learning latent personas of film characters. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 352–361, Sofia, Bulgaria. Association for Computational Linguistics.

Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186.

Martha M Lauzen. 2011. *The Celluloid Ceiling: Behind-the-Scenes Employment of Women on the Top 250 Films of 2010*. Center for the study of women in television and film.

Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. End-to-end neural coreference resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 188–197, Copenhagen, Denmark. Association for Computational Linguistics.

Molly Lewis and Gary Lupyan. 2020. Gender stereotypes are reflected in the distributional structure of 25 languages. *Nature human behaviour*, 4(10):1021–1028.

Mark Liberman. 2014. Obama's favored (and disfavored) sotu words. language log.

Nishtha Madaan, Sameep Mehta, Taneea S. Agrawaal, Vrinda Malhotra, Aditi Aggarwal, Yatin Gupta, and Mayank Saxena. 2018. Analyze, detect and remove gender stereotyping from bollywood movies. In *FAT*.

Mausam Mausam. 2016. Open information extraction systems and downstream applications. In *Proceedings of the twenty-fifth international joint conference on artificial intelligence*, pages 4074–4077.

Jiao Sun and Nanyun Peng. 2021. Men are elected, women are married: Events gender bias on Wikipedia. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 350–360, Online. Association for Computational Linguistics.

Jieyu Zhao, Tianlu Wang, Mark Yatskar, Ryan Cotterell, Vicente Ordonez, and Kai-Wei Chang. 2019. Gender bias in contextualized word embeddings. *arXiv preprint arXiv:1904.03310*, page 629–634.

A Appendix

A.1 Details for Calculating WEAT Score

WEAT generally lies in the range of -2 to 2. For example, assume the target list to consist of different science and art fields, and the attribute list to consist of different pleasant and unpleasant words. The final WEAT test score is the degree to which science fields are more associated with the pleasant words, relative to art fields. A high positive score indicates that science careers are more associated with pleasant words, and a high negative score means that art careers are more associated with pleasant words.

A.2 Unigram Word Cloud Analysis

If we take a closer look at the word clouds for Hollywood shown in Figure (3). We see that before year 2000, females were more described by the words like "husband", "married", "heartbroken","baby", etc. On the other hand men are associated with more masculine and professional jobs like "officer", "army", "journalist"(Refer figures (3a) and (3c)). Over the years, as the society shed some light on gender inequality issue, movies were made to spread the word. As a result, even though gender inequality gap is not completely mitigated, we could see that in the Figure 3b female actress are portrayed as powerful characters on screen. Words like "adventure", "scholarship", "ambassador", "dance", etc were associated with females. Moreover, words like "girlfriend", "partner", "humanity" started appearing while describing male characters.

For CoUK, the female unigrams are "husband", "marry", "baby", "lover", almost all the unigrams associated with relationships before the year 2000 (Refer figure (4a)). After the year 2000 (Refer figure (4b)), the female unigrams are still associated with relationship with different words like "justice", "help" shows a little bit decrease in gender bias. The male unigrams are associated with career oriented words with words like "agent", "policeman", "mission", "awarded", throughout the years, with some words like "proposes", "care" after 2000 which shows a little inclination towards life events. (Refer figures (4c) (4d))

As you can see in Figure (5a) and (5b), Female Unigrams for CoIN are "love", "birth", "marry", "raped" before the year 2000 but after the year 2000, words are "actress", "dance", "lawyer" which shows a decrease in household and submissive

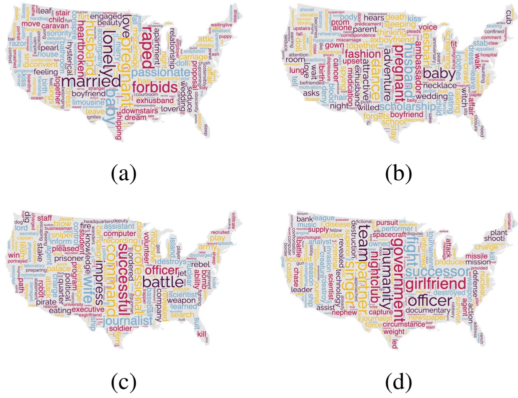


Figure 3: Word cloud mappings for Hollywood. Here, Figure(a) and Figure(b) depicts the based word cloud mappings for female for the movies before and after year 2000 respectively. Whereas, Figure(c) and Figure(d) depicts the based word cloud mappings for male for the movies before and after year 2000 respectively.

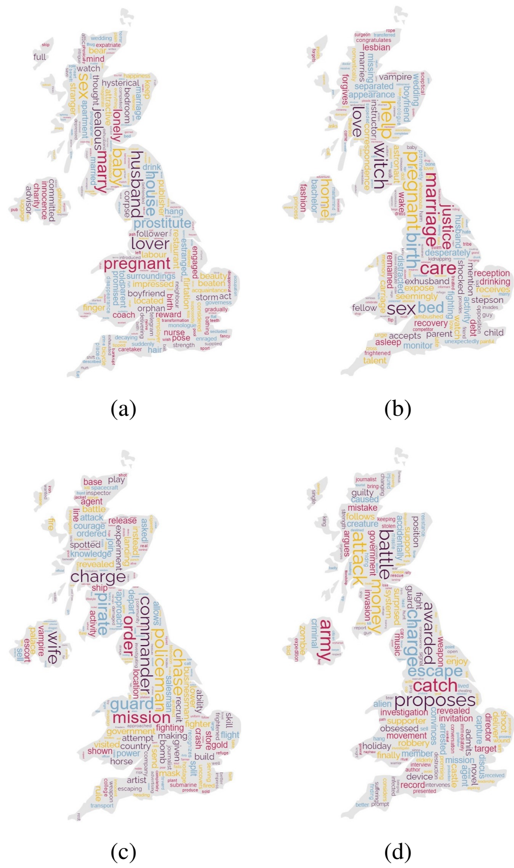
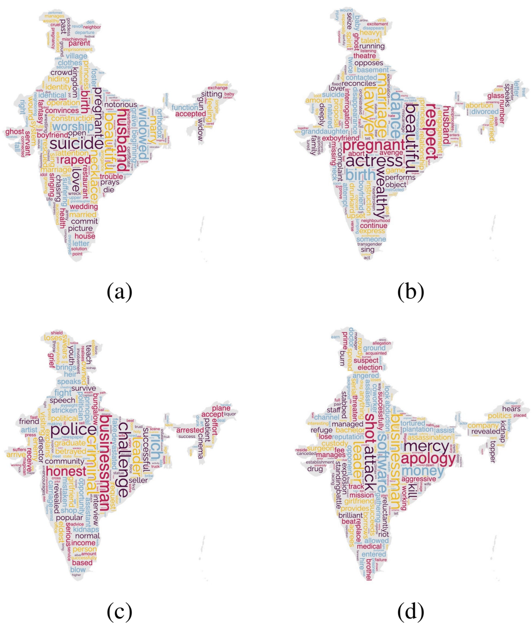


Figure 4: Word cloud mappings for CoUK. Here, Figure(a) and Figure(b) depicts the based word cloud mappings for female for the movies before and after year 2000 respectively. Whereas, Figure(c) and Figure(d) depicts the based word cloud mappings for male for the movies before and after year 2000 respectively.

words and more career oriented words. While figure (5c) and (5d) showcases the Male Unigrams



(a) (b) (c) (d)

Figure 5: Word cloud mappings for CoIN. Here, Figure(a) and Figure(b) depicts the based word cloud mappings for female for the movies before and after year 2000 respectively. Whereas, Figure(c) and Figure(d) depicts the based word cloud mappings for male for the movies before and after year 2000 respectively.

for CoIN with words like "fight", "police", "businessman", "successful" before the year 2000 and "mercy", "apology", "leader" after the year 2000. Male are associated with career oriented words throughout the years but after 2000 words there are some unigrams towards emotions as well.

A.3 Bigram Analysis before and after 2000

Table 4 and Table 5 shows top bigrams for males and females from the movies in CoIN, CoUK and Hollywood before 2000 and after 2000 respectively.

In CoIN before 2000, bigrams associated with men are "rich businessmen", "becomes successful", "accepts challenge", etc. which potray men career oriented where as after 2000 along with career based bigarms such as "software engineer" and "make money" more family oriented terms such as "family life" and "house take" also appear. On the other hand, women before 2000 are associated with dependent and submissive traits such as "love marry", "wealthy man" and "get kidnapped", "skip school". However, after 2000 bigrams such as "friend travel", "enjoys life" and "secure job" which potray her to be independent and career oriented.

In CoUK before 2000, bigrams associated with males are "take control" and "join force" which describes their strong personality which are also seen in bigrams after 2000 such as "launch attack" and "point gun" along with some other bigrams such as "save life" and "start relationship". For female characters, there is no significant difference in the bigrams before and after 2000.

In Hollywood before 2000, bigrams describing men are "police inspector", "factory worker", "baseball player", etc which indicates that men are physically fit and take on-site jobs, and after 2000, there are bigrams like "ask wife", "meet girlfriend", "get attention" which show their interest in household matters along with the career and self-dominant oriented words. For females, while the bigrams, for both before and after 2000, focuses on home, marriage and dependant on other characters, there are bigrams like "lead charge", "cabaret singer", "cheer leader", "lesbian relationship", after year 2000, which tells about the leadership, independent and forward thinking life of women.

Country	Bigrams for Males	Bigrams for Females
India	illegal activity, still alive, rich businessman, killed accident, self esteem, becomes successful, meet women, take blame, seek revenge, help friend, realize mistake, accepts challenge	love marry, life village, give birth, commit suicide, wealthy man, diamond necklace, borrows money, family meet, get kidnapped, receives letter
CoUK	take control, film star, secret agent, join force, give chase ,take advantage, career successful, forced job, train driver, go work	give birth, estranged husband, search work, house mistress, economy dependent, skip school, go church, engagement ring, becomes jealous, run across
Hollywood	police inspector, factory worker, kill wife, commanding officer, becomes manager, go help, proposes marriage, join army, baseball player, give life	give birth, strip artist, move house, engage marry, becomes pregnant, husband love, prepare wedding, becomes jealous, married wife, work nurse

Table 4: Bigrams of Events from CoIN, CoUK and Hollywood for the years before 2000

Country	Bigrams for Males	Bigrams for Females
India	childhood friend, rich businessman, need money, make money, arrives scene, house take, software engineer, take hostage, find evidence, family life	avenge death, leave village, reject marriage, secure job, accepts proposal, friend travel, get selected, marriage father, win case, enjoys life
CoUK	heart attack, security guard, attempt escape, point gun, start relationship, former girlfriend, save life, personal assistant, launch attack, use money	give birth, come home, committed suicide, breast cancer , fall ill, go missing, need help, arranged marriage, falling love, help friend
Hollywood	get attention, arm dealer, lead army, basketball team, lead detective, ask wife, meet girlfriend, prom night, music producer, must fight	fall love, get married, happily married, head cheerleader,lead charge, lesbian relationship, stay parent, birth baby, cabaret singer, coaching position

Table 5: Bigrams of Events from CoIN, CoUK and Hollywood for the years after 2000