Some Dogs Are More Equal Than Others:

A Text Mining Analysis of Dog Cruelty Cases in India

Kaushiki Chowdhury

Indian Institute of Technology (Indian School of Mines), Dhanbad

Course: Text Mining

Prof. Shanmugapriya T

Introduction

In 2025, as the month of March waned, and gave way to the early weeks of April, photos and videos emerging from Hyderabad Central University stared into the faces of all Indian citizens in a challenge against speciesism and as proof of humanity's *in*humanity. These images were tied to the ongoing student protests against the urban development of the Kancha Gachibowli land adjacent to the university. Although the land is not legally owned by the university, it has long been regarded as under its custodianship. More importantly, it has long been home to a rich ecosystem, including "237 bird species, including migratory birds, and hosts wild animals such as spotted deer, wild boars, Indian star tortoises, monitor lizards, and snakes like the Indian rock python." (*India Today* Environment Desk, 2025)

But the disregard of wildlife and cases of animal cruelty are not new phenomena. In recent times, though, they have begun to receive more attention and visibility. From the blatant destruction of habitats and encroachment into wildlife regions to the direct infliction of harm on animals, a country that so strongly claims to revere animals is failing to ensure their actual protection. The contradiction between symbolic reverence and everyday neglect has allowed this problem to persist largely unchallenged. Over time, as urbanization accelerated and traditional relationships with animals shifted, a new issue now frequents our news headlines—stray dogs and domestic canines are being regularly mistreated, and cases of mass poisoning, beating, burning, neglect, and even sexual violence have become alarmingly common.

Moreover, though India is certainly rich in biodiversity, the presence of stray dogs has become so commonplace on Indian streets that they have often slipped beneath public concern. Instead, they are ignored or treated as nuisance. It seems their ubiquity renders them almost

invisible in the public imagination, even as they face daily neglect, violence, and displacement, which highlights a deeper issue of societal apathy, where familiarity has truly bred contempt.

Given the urgency and significance of the issue, this study focuses on cases of cruelty against dogs in India. To examine how this topic is represented in the public discourse, media coverage was selected as the primary source of data, and text mining techniques were applied to articles from two major national newspapers.

Literature Review

It requires no formal review of scholarly literature to conclude that a humane agent is one who possesses the milk of human kindness, and to treat the creatures 'lower' than one with respect and dignity is one way to succeed in that avenue.

Further, man is an animal and he must coexist with other animals if he is meant to survive in society, and in the world. Domestic animals have made this survival a little more bearable—and with consideration of this particular study, dogs have provided man not only with service, but with companionship so great that if he forwent it, he would be robbed of much joy. Yet, he has consistently flouted these phantom rules of survival, he has caused, in many instances, irreparable damage to the ecosystem, and has injured helpless animals out of ignorance, or worse, in a subconscious machiavellian pursuit of power.

India is not only the 8th most mega biodiverse country in the world (Zoological Survey of India, 2020), but according to *EnviStats-India 2024: Environment Accounts*, it also has "[a]round 1,04,561 species of fauna…[having] been documented in the 10 biogeographic zones of the country." (Ministry of Statistics and Programme Implementation India, 2024).

India has also been a land where animals have been venerated, with many considered sacred or even divine. The country's religious and cultural history, recorded in its scriptures, myths, folktales, and festivals, is rife with deities and symbols embodying animals. However, the harsh reality of animal cruelty and exploitation in modern India presents a stark contrast to the imagined sanctity, and the dissonance underscores a profound paradox that can be solved by neither the existing policies, nor common practice.

Article 21 of the Indian Constitution, which ensures right to life for all individuals, has been interpreted by the Supreme Court to also include animals (Mathew, Chadha-Sridhar, 2014).

India has also been lauded internationally for its animal protection legislation (Sekhar, 2014), and it is the perhaps the only country to explicit urge its citizens "to protect and improve the natural environment including forests, lakes, rivers and wildlife, and *to have compassion for living creatures* [emphasis added];" (*Constitution of India*, 1950, Article 51A(g)). In this vein, the government had also introduced the Prevention of Cruelty to Animals Act (1960) , and later the Wildlife Protection Act (1972) [henceforth referred to as PCA Act (1960), and WPA(1972) respectively] to safeguard the interests of animals—wild or domestic. Additionally, the Directive Principles of State Policy, and the Indian Penal Code (Sections 428-9) have provisions for handling animal cruelty cases. (Degal, 2024)

Yet, India actively participates in global markets that exploit animals—through trade of animal parts, captivity of wild animals for entertainment, and export of animal products. Elsewhere, unethical experiments, hard labour, poaching, ritual slaughter, forced animal fighting rings, dangerous transporting conditions, physical abuse, sexual abuse, and neglect persist. These problems are perhaps most visible through the treatment of stray dogs, and without proper way to address the issue, not only are the dogs at the shorter end of the stick but so are unwitting citizens who get into road accidents because of stray dogs, or are otherwise bitten, or even contract diseases borne by the unvaccinated dogs. (Rifai, 2023)

The issue of mistreatment of dogs has also emerged as a recurring subject of judicial scrutiny, prompting numerous Public Interest Litigations (PILs) before various courts and chief among them is *Animal Welfare Board of India v. A. Nagaraja* (2014): The case focused on Jallikattu and ruled that animals should be recognised animals as sentient beings with a right to live with dignity and free from unnecessary pain.

On the other end, the fact of the matter is that India lacks the actual resources to enforce justice. Even today, the punishment for criminal assault on animals has been largely fixed at a fine of only ₹50 (PCA Act, 1960, § 11). The act is not only outdated, it is simply weak and inefficient at enacting any proper justice. In a developing country like India, stray dogs abound and they are both loved and hated, but above all they are made to suffer—sometimes at the hands of natural circumstances, and sometimes due to manmade cruelty. While there are interventions such as the recent Animal Birth Control Rules (2023), large-scale application of the same remains limited.

Research Questions

For this paper, there were three main questions that the researcher hoped to answer that they felt could be useful to future research and could be used by policymakers to address requisite issues. The questions are as follows:

1. Which geographical regions (particular states and/or cities) are the most vulnerable to such cases of cruelty?

2. What recurring patterns emerge in reportage of cruelty against dogs?

3. How has public awareness, and reportage about this issue evolved over time?

The choice of these three research questions stems from the need to address both immediate and long-term concerns regarding dog cruelty in India, and each question is designed to explore a different aspect of the issue, thereby contributing to a comprehensive understanding of cruelty cases.

While the first question is critical for policymakers and welfare organisations to implement targeted interventions, the second question helps uncover common concerns, or key players in the issue, revealing societal attitudes which can potentially go on to influence advocacy efforts. The final question seeks to identify the shifting motifs in reportage, which may inform future interventions and strategies for raising public awareness.

Objectives

The primary objective of this research is to delve into the issue of dog cruelty in India through the lens of text mining techniques, aiming to extract meaningful patterns and insights from a diverse range of news articles. By leveraging innovative methods, the study seeks to deepen the understanding of the issue, and to provide actionable insights that can guide future interventions by policy makers, animal welfare organisations, and advocacy groups. The other objectives, apart from answering the research questions, are as listed below:

1. To provide actionable insights based on the findings that can inform future policy recommendations, welfare strategies, and advocacy efforts;

2. To assess the effectiveness of text mining techniques in analysing social issues;

3. To contribute to the field of digital humanities by combining animal welfare science with data analysis;

4. To integrate personal experiences and passion for welfare of dogs;

5. To assess the role of media discourse in shaping perceptions of dog cruelty.

The data-driven approach of this research can hopefully uncover the themes that dominate public discourse, and track the progression of public awareness. By bridging data analysis and social action, the research is ultimately aimed at fostering greater understanding of the problem, enabling all stakeholders to respond with greater precision and efficacy.

Methodology

Declaration

This paper used artificial intelligence (AI) solely for code generation and debugging purposes. However, all code was reviewed by the researcher for their suitability to the research objectives. The use of AI was not intended to replace critical thinking or human interpretation, but to enhance the efficiency of the technical tasks involved.

Dataset

The dataset consisted of various news articles concerning animal cruelty—specifically, cases about dogs. Two major newspapers were used to find the articles. They were *The Times of India* and *NDTV*. They were chosen because of their repute and their efficient collation of such news as a separate section which enabled easy access and perusal of particular articles. Furthermore, many other newspapers, though inclusive of pertinent articles, could be hard to access on account of paywalls. The selection of these two newspapers was consistent with the information infrastructure available to the layman, as it is them who need to be made aware, and them who must activate their democratic agency to demand for change.

Collection

Web scraping was employed to directly access and gather relevant articles on dog cruelty cases. This method allowed for the automated extraction of urls to those articles, providing quick and efficient access to real-time data. Initially, to collect the links to articles from where the content is to be scraped, code was run to specifically detect those articles that contained any of the keywords "dog", or "dogs", "puppy", or "puppies". Naturally, the 'Requests' and the 'Beautiful Soup' libraries were used.

But owing to the author's deficiencies in coding skills, the issue of lazy-loading could not be addressed sufficiently at this instance. As such, in addition to the articles that were already loaded before clicking 'load more', more articles were chosen manually.

The criteria for collection focused on ensuring relevance by prioritizing recency, case-based text reports rather than videos, excluding articles on preventive measures or solutions, considering articles only relating to the Indian scenario, and avoiding those behind paywalls.

A total of 19 articles were collected via scraping from the relevant category pages. But upon reviewing, there remained only 10. But this is an obviously small corpus which would not point towards any statistically relevant finding. Thus, the author manually added a few more articles to this list to reach a target of 60 articles equally distributed across the two mentioned news websites. While doing this, in case of duplicates, only the most recently updated one has been taken. All the links are listed in the author's GitHub page which has been mentioned in the 'References'.

Methods

Text mining techniques have proven to be highly effective in analysing large volumes of structured text data, allowing for the extraction of valuable insights and patterns. In the context of researching dog cruelty cases, these methods facilitated the identification of key themes, trends, and relationships within vast amounts of news articles. By applying text mining methods it became possible to gain a deeper understanding of the discourse surrounding animal cruelty, and its societal implications.

For analysis and visualisation purposes the usual packages and modules were used—'Pandas', 'spaCy' [small English model], 'Counter' from 'collections', 'NLTK', 'Wordcloud', 'Matplotlib', 'Seaborn', 'GeoPandas', 'Geometry' from 'Shapely'.

As part of this paper, the author intentionally explored text mining methods that they had not previously used in practice and attempted to use this research, not only to answer specific questions about dog cruelty cases in India, but to serve also as an opportunity to expand their digital skills. The techniques thus selected were aligned directly with the research objectives while allowing trial of new tools and analytical frameworks. There were some editorial decisions that were taken for various ends and they are explained in detail in the ipynb files shared in the GitHub repository for this paper, and the same is mentioned in the 'References'.

Named Entity Recognition (NER) was used to identify geographical locations that appeared most frequently to pinpoint the regions where such cases were most prevalent, thereby grounding the analysis in spatial relevance. The main aim was to hopefully highlight areas that require more attention from the perspective of animal welfare concerns.

A Bigram Analysis was also conducted to find recurring word pairs within the collected text data, providing insights into common themes in public narratives. This method was chosen in part because the author had previously struggled with applying it effectively; revisiting it in this research provided an opportunity to improve that skill. Fortunately, this attempt proved fruitful, yielding meaningful patterns that enriched the overall analysis, and which have been discussed in the 'Analysis' section of this paper.

Additionally, a Timeline Analysis enabled the researcher to observe the growth in reportage over time, and to trace shifts in keyword frequency. These keywords were derived from the earlier bigram analysis, to ensure a cohesive approach. By exploring these temporal shifts, the study aimed to gauge public awareness and find any evolving patterns if there were any.

These approaches not only strengthened the analytical depth of study, but also contributed to the learning journey in this research. The integration of various text mining methods allowed for a comprehensive exploration of the data, offering nuanced insights that might have been difficult to achieve through traditional methods alone.

While AI tools were used to assist the analysis, they were not relied upon entirely. Instead, they served as a valuable aid, supporting the researcher's understanding and facilitating a more efficient and thorough examination of the data. This balance ultimately enhanced the overall rigor and impact of the study.

<div align="center">**Analysis**</div>

Named Entity Recognition (NER)

The findings reveal that Delhi is by far the most frequently mentioned location in the dataset, appearing as many as 60 times, indicating a concentration of cruelty cases in the capital. All the other locations, along with their frequencies, are listed in the table below:

Top Normalized Locations:

Places	Occurrences
Delhi	60
Mumbai	15
Punjab	13
Telangana	9
Rajasthan	7
Uttar Pradesh	6
Kolkata	5
Maharashtra	5
Karnataka	5
Bihar	4
Tripura	3
Gujarat	3
West Bengal	2
Madhya Pradesh	2
Jharkhand	2
Hyderabad	2
Andhra Pradesh	2
Jammu & Kashmir	1
Bhopal	1

Map of Locations Mentioned in Animal Abuse Articles
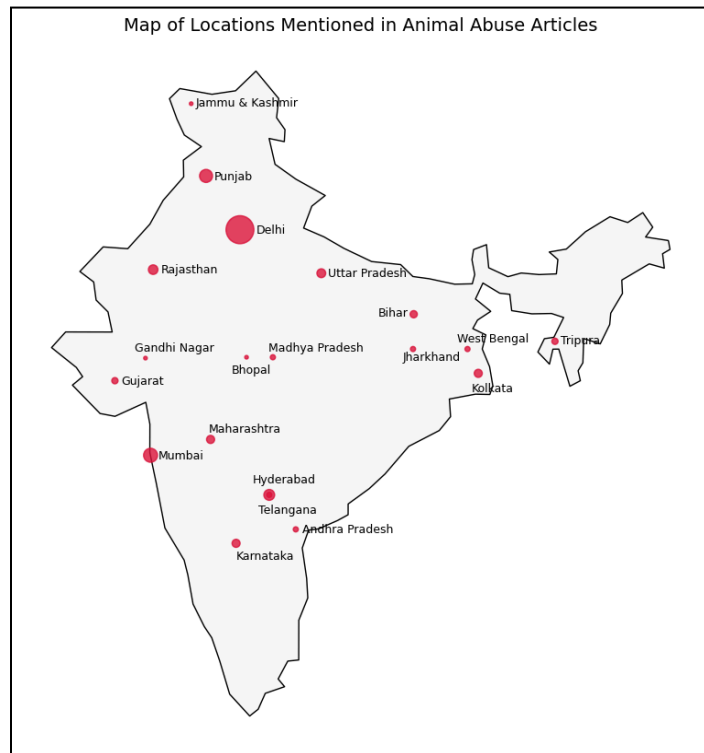
It is worth noticing that several Indian States were not mentioned at all within the dataset—notably, a large portion of the North-Eastern region, including Arunachal Pradesh, Assam, Manipur, Meghalaya, Nagaland, and Sikkim, was absent from the results.

Similarly, southern states such as Goa, Kerala, Tamil Nadu, Puducherry did not appear in the data. From the central and northern regions, Haryana, Himachal Pradesh, Uttarakhand, Chandigarh, Ladakh, and Chhattisgarh were not referenced, and in the east, Orissa went without any mention as well.

Although this finding may seem hopeful, it can be very likely that there is simply underreporting, especially considering national news' often inflated emphasis on northern and central India. Moreover, the bigger a state, larger would be its stray dog population, and conversely, larger the state, worse is the applicability of control measures.

But in a more optimistic way, the findings might mean that these unmentioned states have found ways to successfully handle their dog population, and might offer the other states some guidelines to achieve the same future.
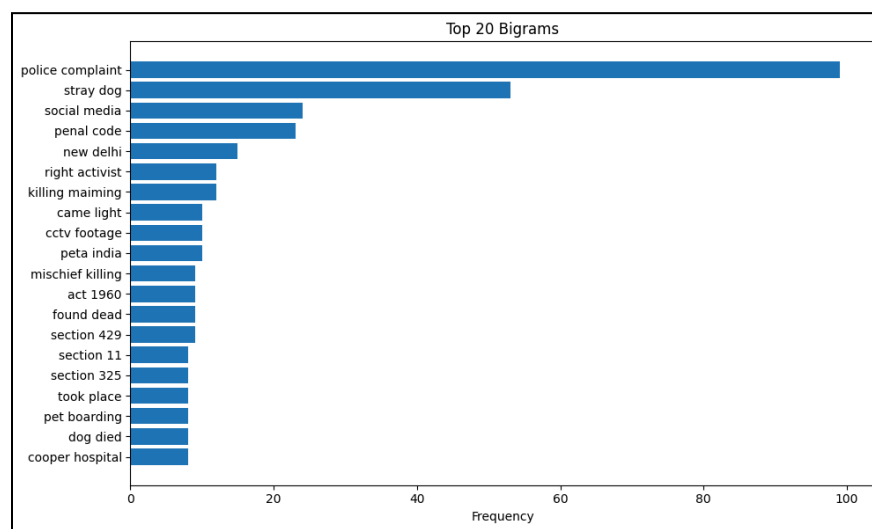
Bigram Analysis

The Bigram Analysis of the collected text data revealed several key recurring word pairs, providing insight into the prevalent themes in the discourse surrounding dog cruelty. Notably, the most frequent was "police-complaint" (99 occurrences), which indicates the significant role law enforcement plays in reporting and addressing these issues. Another prominent pair "stray-dog" (53 occurrences) is directly linked to the central subject of the research.

Delhi continues to be a major concern with 15 occurrences, while terms that reflect the severity of the cases—"killing-maiming" (12 occurrences), and "mischief-killing" (9 occurrences)—have also been discussed, further emphasising the tragic nature of the events. But the presence of "right-activist" (12 occurrences), and "peta-india" (10 occurrences) suggests the active involvement of animal rights organisations and activists in addressing and fighting issues

of cruelty. Social media (24 occurrences) also seems to play a crucial role in raising awareness and mobilizing efforts to combat this problem.

There are some bigrams including "pet-boarding" and "cooper-hospital" which refer to isolated incidences, and as such shouldn't be afforded the same weightage as the others. Still, they point towards interesting realities.

The bigrams are given in full in the appropriate ipynb file, and are visualised in the following barplot:


Timeline Analysis

The timeline analysis, based on the previously identified bigrams, revealed a notable peak in reportage and activity following the year 2024, and the surge aligns closely with the increased frequency of 'police' which emerged as the most recurring keyword. The rise in such reporting indicates a likely increase in cruelty cases, and indeed, the if public memory serves right, the late 2020s have had increased instances of barbaric treatment of all animals—in 2022, a Bengal monitor lizard was allegedly gang-raped in Maharashtra by four men! (Gayakwad, 2022)

Naturally, reportage increased in tandem with rising public concern, and it would seem law enforcement's involvement in these efforts has been cause for attention. At the same time,

the recurrence of 'act' may highlight the public's renewed interest in the now outdated PCA Act. Other keywords include 'stray', 'killing', and 'media'.
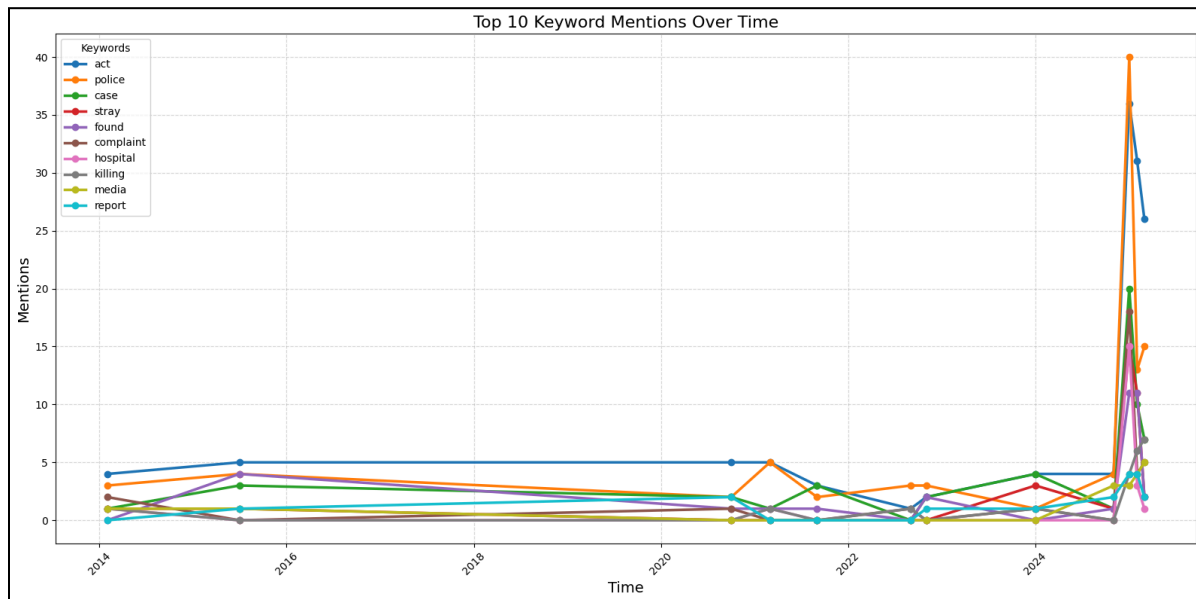
But the line graphs are noticeably overlapping, and are hard to read to clearly, thus individual mentions have been plotted as below:
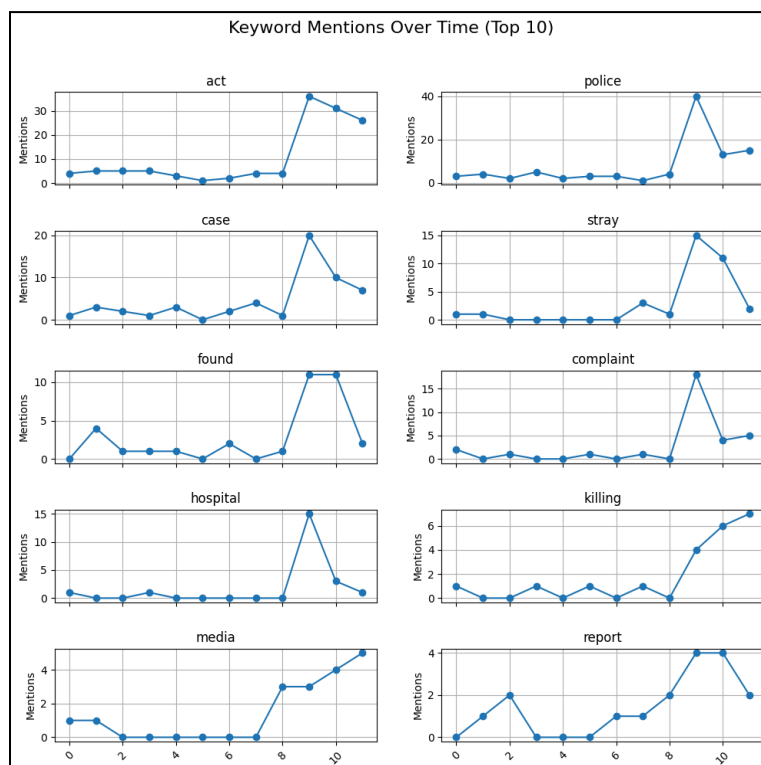
It is interesting to note that the graphs for 'police', 'media', and 'killing' are upward progressing, suggesting these terms may remain prominent going forward. Particularly pertinent is that the occurrence of 'media' highlights the critical role that traditional and social media platforms continue to play in amplifying awareness and drawing public attention to these issues.

Limitations

Despite the valuable insights generated through the application of text mining techniques, this study is not without its limitations.

The corpus for this research was compiled using only two major news websites, and while these sources are reputable, the restricted scope may have excluded diverse perspectives or region-specific reportage from smaller or vernacular publications. Additionally, the dataset comprised only 60 articles, which may not be fully representative of the broader media discourse relating to this issue. Moreover, all articles, being in English, inherently limits the research to urban and English-speaking narratives. The lengths of these articles were also different, which would inevitably affect implementation of text mining tools.

While NER is a robust technique, its accuracy is dependent on the quality of the model used. Some smaller towns or rural areas might not have been identified correctly at all. In fact, the normalisation process of the locations incorporated editorial discretion, and thus could be subject to human error. Alternatively, the chosen tools were only used to answer specific questions; they could have been used in a more generalised way to identify larger patterns and trends.

Further, the research focuses only on the 2014-2025 period, and the articles are not evenly distributed across multiple years, and peaks could just indicate isolated events that attracted major attention, rather than being actual trends. This imbalance will have likely skewed the analysis.

Most obviously, the study is focused on dogs and will not reflect the broader welfare trends. While the narrow lens allows for concentrated analysis of a pressing issue, other areas of concern—chief among them is the cattle industry—are not accounted for.

Last, the case for researcher bias, and editorial biases, sensationalism, selective coverage of the media still persist as a grave limitation.

However, actionable steps to address these limitations can be devised, and the research itself can be expanded and modified to approach from numerous angles and answer both specific or general research questions.

Future Work

Despite the limitations of the research, it still can be useful for future purposes. From the lens of research, future work in this matter may involve focusing on specific states or cities to allow more expansive analysis. Moreover, given the limited corpus, providing access to more newspapers—both national, and state focused—will allow for more thorough research. In fact, the research may be expanded to include more animals, or even just specific animals—given India's cattle industry's unfortunate history, that may provide a necessary area for study. Elsewhere, research may even be specified to highlight the more endangered species to inquire into *their* status.

From the lens of inclusive democracy, as outlined by thinkers like Takis Fotopolous and John Sargis, animal exploitation must be eliminated to become a truly democratic nation. Moreover, a systematic overhaul of the legal institution and legislation is absolutely crucial to bring about lasting change. Rukmini Sekhar suggests a "three-pronged administrative change"—establishment of a separate Ministry for Animals, and an enforcement agency akin to UK's Royal Society for the Prevention of Cruelty to Animals (RSPCA), and enactment "speedy justice" and "legal or punitive deterrents." She, along with several other legal professionals also call for a review of the archaic PCA (1960), in favour of more effective animal welfare acts. There have even been attempts to repeal the PCA Act in favour of more effective bills like the Draft Animal Welfare Act (2011), and later the PCA (Amendment) Bill (2016)—both by the Animal Welfare Board of India—but they are yet to be adopted in full.

On the other hand, public figures, animal rights activists, and common citizens have also called for the increment of fines for animal cruelty from a measly ₹50, or ₹100, or at most ₹200 to more hefty amounts. To this extent, the hashtag #NoMore50 has taken off across social media

platforms in reference to this, and even bills have been introduced in the parliament relevant to this issue. (Agarwal, 2024; Degal, 2024).

Conclusion

The findings of this paper have been mostly depressing, and one might be tempted to believe that there is no positive end to this story—the capital, with the most resources at its disposal to address the issues of animal cruelty, is ironically the hub of this cruelty. The increasing reports of incidents highlight a positive trend for a negative issue, and the plight of animals become more gruesome by the day. But one should take this very study as proof that India is waking up to the issue and with social media campaigns backed by influential figures, there might come about a day where all animals in India can be exalted to the revered status they once enjoyed.

Sri Aurobindo Ghosh famously noted that the assumed difference between animals and humans is an ideological construct created only for the latter's advantage. This is in line with speciesist perspectives that the animal philosopher Richard Ryder claims to be "unintelligent and out of date" (Ryder quoted in Sekhar, 2014).

Finally, Indians must recall that "[t]he greatness of a nation and its moral progress can be judged by the way its animals are treated." (Gandhi) Divided though we may be—on grounds ever irreconcilable, and matters perpetually polarising—we must not forget that man too is an animal, and should we ever truly accept that "some animals are more equal than others" (Orwell, 1945), then our backs too shall bend under the weight of the whip, our names too shall be erased clumsily, and we too shall perish in our burning hubris.

References

Github Repository: https://github.com/kaushikichowdhury/Assignments

    Files:

1. https://github.com/kaushikichowdhury/Assignments/blob/main/Webscraping.ipynb

2. https://github.com/kaushikichowdhury/Assignments/blob/main/Analysis.ipynb

Agarwal, S. (2024b, February 7). Cruelty to animals New Bill proposes increase in fine from Rs

    50 to Rs 1 lakh. *Lawstreet.co*.

    https://lawstreet.co/executive/cruelty-to-animals-new-bill-proposes-increase-in-fine-from

    -rs-50-to-rs-1-lakh

Degal, D. (2024). *Animal abuse in India*. *African Journal of Biological Sciences, 6*(Special

    Issue-2), 1170–1177. https://doi.org/10.33472/AFJBS.6.Si2.2024.1170-1177

Gayakwad, R. (2022b, April 14). Four held for "raping" Bengal monitor lizard in Maharashtra

    forest. *The Times of India*.

    https://timesofindia.indiatimes.com/city/mumbai/four-held-for-raping-bengal-monitor-liz

    ard-in-maharashtra-forest/articleshow/90846235.cms

India Today Environment Desk. (2025, April 3). These animals will be affected by destruction of

    Kancha Gachibowli forest. *India Today*.

    https://www.indiatoday.in/environment/story/kancha-gachibowli-forest-tree-felling-anima

    ls-affected-university-of-hyderabad-protest-2703511-2025-04-03

Mahawar, S. (2022, June 27). *An overview of animal protection laws in India - iPleaders*.

    iPleaders.

https://blog.ipleaders.in/overview-animal-protection-laws-india/#Draft_Animal_Welfare_
Act_2011

Mathew, J. T., & Chadha-Sridhar, I. (2014). Granting animal rights under the Indian
Constitution: A misplaced approach? An analysis in light of *Union of India v. A.
Nagaraja*. *NUJS Law Review, 7*(3), 349.

Ministry of Statistics and Program Implementation. (2022). *EnviStats India 2022, Vol II:
Environment Accounts*. In *EnviStats India* (Chapter 7: Biodiversity). Government of
India.

https://mospi.gov.in/sites/default/files/reports_and_publication/statistical_publication/Env
iStats/Chap7-Biodiversity_envst22.pdf

Ministry of Statistics and Program Implementation. (2024). *EnviStats India 2024: Environment
Accounts* (Chapter 5: Biodiversity, p. 86). Government of India.

https://www.mospi.gov.in/sites/default/files/publication_reports/EnviStats_India_2024.pd
f

Rifai, A. K. M. (2023). *Examining the stray dog menace in India through a multi-perspective
approach*. *International Journal of Novel Research and Development, 8*(12), a661–a668.
https://ijnrd.org/papers/IJNRD2312083.pdf

Sekhar, R. (2014). The cruelty of speciesism: From the point of view of inclusive democracy and
ecology, the issue of animal protection cannot be simplified into a human versus animal
dichotomy. *Economic and Political Weekly, 49*(40), 71–72.

http://www.jstor.org/stable/24480829

Zoological Survey of India, Ministry of Environment, Forest, and Climate Change, Government of India. (2020). *Animal Discoveries: New Species, New Records of 2020*. Government of India.