

►

From Fragments to Facts: A Curriculum-Driven DPO Approach for Generating Hindi News Veracity Explanations

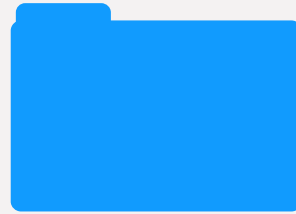
**Pulkit Bansal
2101MC48**

**Supervisor – Dr. Y.M. Tripathi
Co-Supervisor – Dr. Sriparna Saha**





When?



Why?



What?

When?



Task	Time
Project Started	03.01.25
Problem Statement Finalization	06.01.25
Literature Survey	07.01.25 – 11.01.25
Experiments	12.01.25 – 22.02.25

Why?



Why we done this work?






In our research we find that LLM can generate the explanation while deciding fake or real for the inputted news article but explanations are primarily focused on linguistic aspect.



We want to create an end to end model that can generate the explanation as we find on various fact checking website that we termed as Human aligned explanations.

Research Questions




-  RQ-1: How can automated systems reliably assess the veracity of Hindi news and generate human-like explanations that are coherent, contextually relevant, and factually accurate while explicitly justifying the model's veracity predictions?
-  RQ-2: Can Direct Preference Optimization, augmented with Fact-score and variance-based parameters, effectively align machine-generated explanations with human standards in constrained languages like Hindi, ensuring both factual accuracy and interpretability?
-  RQ-3: How can curriculum learning be integrated with DPO to refine veracity prediction in underrepresented languages, and what scalable methodologies can extend misinformation detection and explanation generation to other low-resource languages?

What?



What we have done?

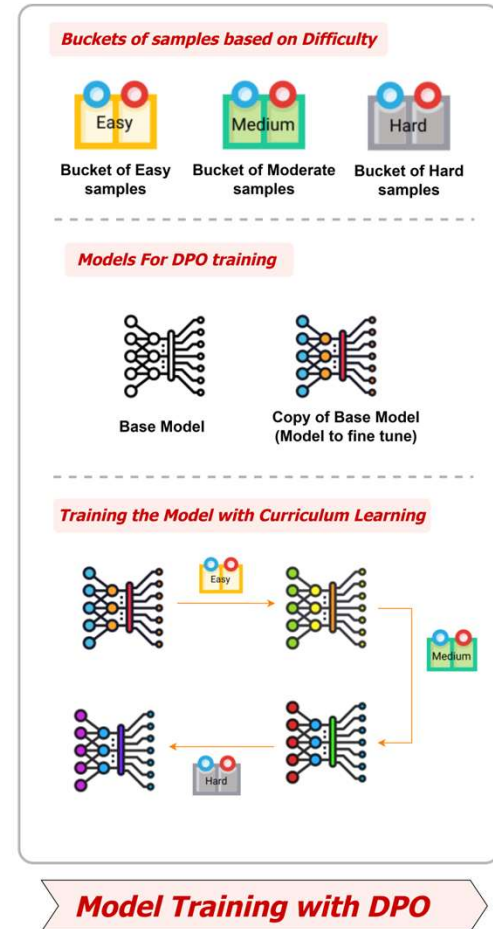
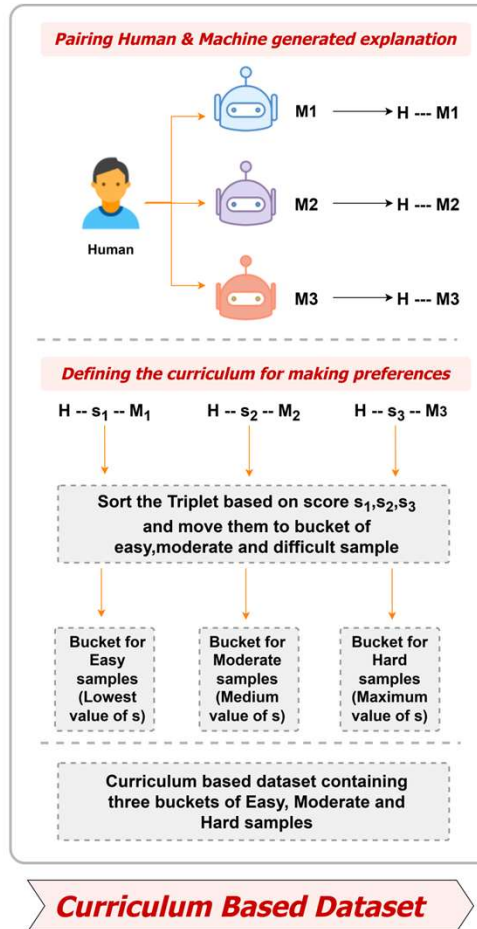
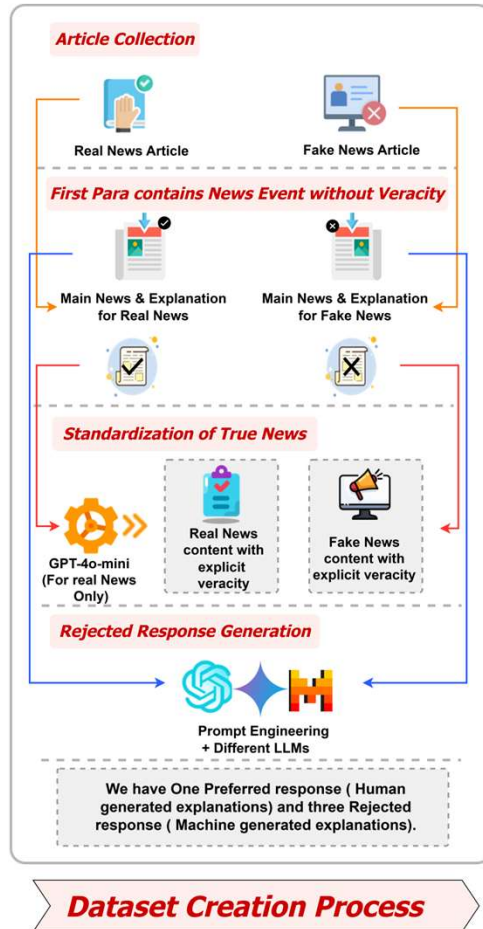
 For our task we have utilized Preference optimization technique named Direct Preference Optimization (DPO)

 Taking inspiration from DPO framework and considering various aspects related to fake news detection we have introduced a new loss function for our training.

 We have modified the DPO loss function and introduce two new parameters named Finesse and Actuality to improve the performance of LLM on fake news related task.

 We have also introduced a new dataset that contains rank based preferences of explanations for the inputted news article.

Our Framework- De-factox



Dataset Creation

We have scraped True News and Fake news articles from various website. Then we have segregated the main news and Explanations from the website. These website are treated as Ground Truth (Positive samples) for our experiments.

For our training process we need negative samples as well for that we have utilized three state of the art LLMs and crated a synthetics ranking based preference dataset.

We have utilized. We have utilized gpt-4o-mini, Mistral-7B-v0.1, and gemini-1.5-flash. These models were selected due to their strong reasoning capabilities and proven performance in NLP tasks.



Explanation of Why the News is Fake

लेकिन सच ये है कि 'ये वायरल वीडियो दीया कुमारी का नहीं बल्कि गुजराती महिला निकिताबा राठोड़ का है। इस सच का खुलासा किया है आज तक चैनल ने, जिन्होंने गुजराती महिला निकिताबा राठोड़ से इस बारे में बात की और वायरल वीडियो का असली सच लोगों को बताया। निकिताबा राठोड़ ने 22 जनवरी यानी की अयोध्या में रामलला के प्राण-प्रतिष्ठा वाले दिन अहमदाबाद के नरोडा में आयोजित एक कार्यक्रम में हिस्सा लिया गया था, जहां लोगों ने प्रभु राम के अयोध्या वापसी पर 11 हजार दीए जलाकर दिवाली मनाई थी।

English Translation

But the truth is that this viral video is not of Diya Kumari but of a Gujarati woman Nikitaba Rathod. This truth has been revealed by Aaj Tak channel, who talked to Gujarati woman Nikitaba Rathod about this and told the real truth of the viral video to the people. Nikitaba Rathod had participated in a program organized in Naroda, Ahmedabad on 22 January, the day of Ram Lalla's Pran-Pratishtha in Ayodhya, where people celebrated Diwali by lighting 11 thousand diyas on the return of Lord Ram to Ayodhya.

Figure 2: Snippet of Fake news explanation with explicit reasoning for its veracity.

Rephrased Explanations

Explanation from original True News Article

लोकसभा चुनाव 2024 के लिए चुनाव आयोग द्वारा घोषित कार्यक्रम के बाद आज प्रधानमंत्री नरेंद्र मोदी ने आंध्र प्रदेश के पलनाडु जिले में एक रैली को संबोधित किया। इस दौरान पीएम मोदी ने टीडीपी चीफ चंद्रबाबू नायडू के साथ मंच शेयर किया और विपक्षी दलों पर निशाना साधा। इस रैली में जनसेना प्रमुख पवन कल्याण भी शामिल हुए। पीएम मोदी ने अपने संबोधन की शुरुआत तेलुगु में करते हुए लोगों को 'नमस्कारम्' कहा। इस दौरान उन्होंने कहा कि इस बार पूरा देश 400 पार का नारा लगा रहा है। प्रधानमंत्री मोदी ने रैली को संबोधित करते हुए कहा, 'कल ही देश में लोकसभा चुनाव का विगल बजा है और आज मैं आप सबके बीच आंध्र प्रदेश में हूँ। मुझे ब्रह्मा विष्णु और महेश का आशीर्वाद मिल रहा है। त्रिदेवों के आशीर्वाद से हमारी सरकार के तीसरे कार्यकाल में देश और भी बड़े निर्णय लेगा। इस बार चुनाव के परिणाम 4 जून को आने वाले हैं। पूरा देश कह रहा है 4 जून को 400 पार...' पीएम मोदी ने आगे कहा, 'हमारा NDA गठबंधन रीजनल एस्पिरेशंस और नेशनल प्रोग्रेस दोनों को साथ लेकर चलता है। इस चुनाव में बीजेपी के सहयोगी लगातार बढ़ रहे हैं, NDA की ताकत बढ़ रही है। चंद्रबाबू नायडू और पवन कल्याण दोनों लंबे समय से आप लोगों के हक और आंध्र के विकास के लिए दिन-रात आपके लिए काम करते रहे हैं। विपक्षी पार्टी कांग्रेस पर हमला बोलते हुए पीएम मोदी ने कहा कि NDA में हम सबको साथ लेकर चलते हैं, लेकिन दूसरी ओर कांग्रेस पार्टी है जिसका एक ही एजेंडा है- गठबंधन के लोगों को Use and Throw... आज कांग्रेस को भले ही मजबूरी में इंडी अलायंस बनाना पड़ा हो, लेकिन इनकी सोच वही है।

English Translation

After the Election Commission announced the schedule for the 2024 Lok Sabha elections, Prime Minister Narendra Modi addressed a rally today in the Palnadu district of Andhra Pradesh. During this time, PM Modi shared the stage with TDP Chief Chandrababu Naidu and targeted the opposition parties. Jana Sena Chief Pawan Kalyan also participated in the rally. PM Modi began his speech in Telugu, greeting the people with "Namaskaram." During this, he said, "This time, the entire nation is chanting the slogan of crossing 400 seats." Addressing the rally, Prime Minister Modi said, "Just yesterday, the bugle for the Lok Sabha elections sounded, and today, I am here among you in Andhra Pradesh. I am receiving the blessings of Brahma, Vishnu, and Mahesh. With the blessings of the Trinity, our government, in its third term, will take even bigger decisions. This time, the election results are going to come on June 4. The whole country is saying, 'On June 4, 400 seats crossed!'" PM Modi further said, "Our NDA alliance moves forward, keeping both regional aspirations and national progress together. In this election, BJP's allies are continuously increasing, and the strength of the NDA is growing. Chandrababu Naidu and Pawan Kalyan have both been working day and night for your rights and the development of Andhra Pradesh for a long time. Attacking the opposition Congress party, PM Modi said, "In the NDA, we work by taking everyone along, but on the other side is the Congress party, which has only one agenda — to 'use and throw' its allies. Today, even though Congress has been forced to form the INDI Alliance out of compulsion, their mindset remains the same."

Rephrased Explanation with Veracity claims

यह खबर पूरी तरह से सत्य है। लोकसभा चुनाव 2024 के लिए चुनाव आयोग द्वारा घोषित कार्यक्रम के बाद, प्रधानमंत्री नरेंद्र मोदी ने आज आंध्र प्रदेश के पलनाडु जिले में एक रैली को संबोधित किया। इस दौरान, पीएम मोदी ने टीडीपी चीफ चंद्रबाबू नायडू के साथ मंच साझा किया और विपक्षी दलों पर निशाना साधा। इस रैली में जनसेना प्रमुख पवन कल्याण भी शामिल हुए। पीएम मोदी ने अपने संबोधन की शुरुआत तेलुगु में करते हुए लोगों को 'नमस्कारम्' कहा। उन्होंने कहा कि इस बार पूरा देश 400 पार का नारा लगा रहा है। प्रधानमंत्री मोदी ने रैली को संबोधित करते हुए कहा, "कल ही देश में लोकसभा चुनाव का विगल बजा है और आज मैं आप सबके बीच आंध्र प्रदेश में हूँ। मुझे ब्रह्मा विष्णु और महेश का आशीर्वाद मिल रहा है। त्रिदेवों के आशीर्वाद से हमारी सरकार के तीसरे कार्यकाल में देश और भी बड़े निर्णय लेगा। इस बार चुनाव के परिणाम 4 जून को आने वाले हैं। पूरा देश कह रहा है 4 जून को 400 पार" यह जानकारी भी सत्य है। पीएम मोदी ने आगे कहा, "हमारा NDA गठबंधन रीजनल एस्पिरेशंस और नेशनल प्रोग्रेस दोनों को साथ लेकर चलता है। इस चुनाव में बीजेपी के सहयोगी लगातार बढ़ रहे हैं, NDA की ताकत बढ़ रही है। चंद्रबाबू नायडू और पवन कल्याण दोनों लंबे समय से आप लोगों के हक और आंध्र के विकास के लिए दिन-रात आपके लिए काम करते रहे हैं। विपक्षी पार्टी कांग्रेस पर हमला बोलते हुए, पीएम मोदी ने कहा कि NDA में हम सबको साथ लेकर चलते हैं, लेकिन दूसरी ओर कांग्रेस पार्टी का एक ही एजेंडा है- गठबंधन के लोगों को Use and Throw। उन्होंने यह भी कहा, "आज कांग्रेस को भले ही मजबूरी में इंडी अलायंस बनाना पड़ा हो, लेकिन इनकी सोच वही है।" इसमें कोई भावनात्मक भाषा या भ्रामक तत्व नहीं है जो इसे फर्जी समाचार के रूप में चिह्नित करने के लिए पर्याप्त हो। इसलिए, उपलब्ध जानकारी के आधार पर, यह समाचार वास्तविक प्रतीत होता है।

English Translation

This news is completely true. After the Election Commission announced the schedule for the 2024 Lok Sabha elections, Prime Minister Narendra Modi addressed a rally today in the Palnadu district of Andhra Pradesh. During this, PM Modi shared the stage with TDP Chief Chandrababu Naidu and targeted the opposition parties. Jana Sena Chief Pawan Kalyan also participated in the rally. PM Modi began his speech in Telugu, greeting the people with "Namaskaram." He said that this time, the entire nation is chanting the slogan "400+ seats." Addressing the rally, Prime Minister Modi said, "Just yesterday, the bugle for the Lok Sabha elections sounded, and today, I am here among you in Andhra Pradesh. I am receiving the blessings of Brahma, Vishnu, and Mahesh. With the blessings of the Trinity, our government, in its third term, will take even bigger decisions. This time, the election results are going to come on June 4. The whole country is saying, 'On June 4, 400+ seats!'" This information is also true. PM Modi further said, "Our NDA alliance moves forward by keeping both regional aspirations and national progress together. In this election, BJP's allies are continuously increasing, and the strength of the NDA is growing. Chandrababu Naidu and Pawan Kalyan have both been working day and night for your rights and the development of Andhra Pradesh for a long time." Attacking the opposition Congress party, PM Modi said, "In the NDA, we work by taking everyone along, but on the other side is the Congress party, whose sole agenda is to 'use and throw' its allies." He also said, "Today, even though Congress has been forced to form the INDI Alliance out of compulsion, their mindset remains the same." There is no emotional language or misleading element in this report to classify it as fake news. Therefore, based on the available information, this news appears to be genuine.

Negative samples

Main News

कांग्रेस नेता राहुल गांधी का एक वीडियो है, इसमें वो ये कहते दिखते हैं कि सत्याग्रह का मतलब है- सत्ता के रास्ते को कभी मत छोड़ो। उस वीडियो को पोस्ट करते हुए लोग राहुल पर तंज कस रहे हैं और कह रहे हैं कि उन्होंने सत्ता के मोह में महात्मा गांधी के सत्याग्रह आंदोलन का अर्थ ही बदल दिया। बीजेपी छात्तीसाढ़ के महासचिव ओ पी चौधरी ने ये वीडियो शेयरकरते हुए लिखा, "थिक्कर है शहजादे को, मैंने भी IAS की परीक्षा में फिलॉसफी सब्जेक्ट लिया था और थोड़ा-बहुत गांधीवाद पढ़ा था, लेकिन शहजादे का यह कौन सा गांधीवाद है? यह सत्याग्रह की यह कौन सी परिभाषा है जो कहती है कि सत्ता के रास्ते को कभी मत छोड़ो?" वहीं छात्तीसाढ़ के पूर्व गृह मंत्री और बीजेपी विधायक ब्रजमोहन अग्रवाल ने राहुल पर निशाना साधते हुए लिखा, "महात्मा गांधी के नाम से, सत्याग्रह की नई परिभाषा गढ़ @RahulGandhi ने उनके विचारों को तिलांजलि दे दी है."

English Translation

A video of Congress leader Rahul Gandhi has surfaced, where he is seen saying that the meaning of Satyagraha is "Never abandon the path to power." Sharing this video, people are mocking Rahul Gandhi and claiming that he has altered the very essence of Mahatma Gandhi's Satyagraha movement in his pursuit of power. BJP Chhattisgarh General Secretary O.P. Choudhary shared this video, writing, "Shame on the prince. I too had taken philosophy as a subject for the IAS exam and had studied Gandhian philosophy to some extent. But what kind of Gandhian philosophy is this of the prince? What is this definition of Satyagraha that says 'Never abandon the path to power'?" Similarly, Chhattisgarh's former Home Minister and BJP MLA Brijmohan Agrawal targeted Rahul Gandhi, writing, "By redefining Satyagraha in the name of Mahatma Gandhi, @RahulGandhi has completely discarded his ideas."

Explanation from GPT-4o-mini

इस समाचार को मैं वास्तविक मानता हूँ। इसकी पुष्टि विभिन्न स्रोतों और घटनाओं के माध्यम से की जा सकती है। राहुल गांधी का वीडियो बयान कांग्रेस के महाअधिवेशन के दौरान दिया गया है, और इस तरह के कार्यक्रमों में नेता अक्सर अपने विचार साझा करते हैं। इसके अलावा, समाचार में उल्लिखित बयान का संदर्भ भी स्पष्ट है, जिसमें उन्होंने 'सत्याग्रह' के अर्थ को अपनी शब्दावली में प्रस्तुत किया। साथ ही, उन्होंने बाद में अपने बयान को सुधारते हुए सही अर्थ प्रदान किया, जो कि राजनीतिक व्यवस्था के लिए सामान्य व्यवहार है। इसके अलावा, समाचार में अन्य राजनीतिक नेताओं की प्रतिक्रियाओं का भी जिक्र है, जो स्थिति को और प्रामाणिक बनाता है। जब कोई नेता किसी महत्वपूर्ण विषय पर गलतफहमी में होता है, तो उस पर प्रतिक्रिया होना सामान्य है। यह भी दर्शाता है कि राजनीतिक संवादात्मकता जारी है और नेता एक-दूसरे की बातों की व्याख्या करते हैं। इस प्रकार, समाचार की संपूर्णता और तथ्यात्मकता इसे वास्तविक बनाते हैं, इसलिए इसे फेक नहीं कहा जा सकता है।

English Translation

I consider this news to be authentic. It can be verified through various sources and events. Rahul Gandhi's video statement was made during the Congress plenary session, where leaders often share their views. Additionally, the context of the statement, where he initially misspoke about the meaning of 'Satyagraha' and later corrected himself, is clear and typical for political figures. The report also mentions reactions from other political leaders, adding credibility. Political discourse often involves interpreting and responding to such statements. Thus, the completeness and factual nature of this news make it genuine and not fake.

Explanation from Gemini-1.5-flash

यह समाचार वास्तविक लगता है। लेख में राहुल गांधी के एक भाषण का उल्लेख है जिसमें उन्होंने सत्याग्रह की अपनी व्याख्या प्रस्तुत की, जिस पर बाद में उन्होंने सुधार किया। यह घटना एक विशिष्ट राजनीतिक घटना है जिसकी खबर विभिन्न मीडिया माध्यमों में आ सकती है। लेख में कांग्रेस के महाअधिवेशन, तरीकों और स्थान का उल्लेख है, जो इस और अधिक विश्वसनीय बनाता है। भाषण के एक अंश के गलत होने और बाद में सुधार किए जाने की जानकारी भी समाचार की प्रामाणिकता को इंगित करती है, क्योंकि ऐसी गलतियाँ और उन पर होने वाली प्रतिक्रियाएँ वास्तविक राजनीतिक घटनाओं में आम हैं। यह सब मिलाकर, लेख में दी गई जानकारी और घटनाओं का विवरण एक वास्तविक समाचार लेख के सभी लक्षण दर्शाता है। हालाँकि, किसी भी समाचार की सत्यता की पूरी पुष्टि के लिए अन्य विश्वसनीय स्रोतों से इसकी जाँच करना आवश्यक है।

English Translation

This news appears to be authentic. The article mentions a speech by Rahul Gandhi in which he presented his interpretation of 'Satyagraha,' later correcting himself. This is a specific political event that could be reported across various media outlets. The mention of the Congress plenary session, along with its dates and location, adds to its credibility. The report of a misstatement in the speech followed by its correction also indicates the authenticity of the news, as such errors and subsequent reactions are common in real political events. Overall, the details provided in the article and the description of events exhibit all the traits of a genuine news piece. However, to fully confirm the authenticity of any news, it is essential to verify it through other reliable sources.

Explanation from Mistral-7b

यह समाचार लेख संभवतः फर्जी है क्योंकि यह सरल और स्पष्ट शैली में लिखा गया है, जिसमें विशिष्ट विवरण और जानकारी की कमी है। सामग्री किसी भी वर्तमान घटनाओं या समाचारों से संबंधित नहीं है, और लेख की भाषा समाचार लेखों के लिए सामान्य नहीं है। यह एक औपचारिक और पेशेवर तरीके से लिखा गया है, जो सामान्य समाचार लेखों के अनुरूप नहीं है। इसलिए, यह लेख संभवतः फर्जी हो सकता है।

English Translation

This news article is possibly fake because it is written in a simple and clear style, lacking specific details and information. The content does not seem to relate to any current events or news, and the language of the article is not typical of standard news reports. It is written in a formal and professional manner, which does not align with the usual tone of news articles. Therefore, this article is likely to be fake.

Dataset Algorithm

Algorithm 2 Dataset Creation from Fact-Checked News Articles

Require: Scraped news article A containing main news N and explanation E , Large Language Models $\{LLM_1, LLM_2, LLM_3\}$, Scoring function S

Ensure: Dataset D with explanations categorized by quality

- 1: **Step 1: Segregate Explanation and Main News**
- 2: Extract main news N and ground-truth explanation E_{GT} from article A
- 3: **Step 2: Generate LLM Explanations**
- 4: **for** each model LLM_i in $\{LLM_1, LLM_2, LLM_3\}$ **do**
- 5: Provide N as input to LLM_i and obtain predicted explanation E_i
- 6: **end for**
- 7: **Step 3: Compute Scores for Explanations**
- 8: **for** each predicted explanation E_i **do**
- 9: Compute score $S(E_i)$ using scoring function S
- 10: **end for**
- 11: **Step 4: Rank Explanations**
- 12: Sort explanations $\{E_1, E_2, E_3\}$ in ascending order based on $S(E_i)$
- 13: **Step 5: Bucketize Explanations**
- 14: Define three score-based categories:
 - **Low-quality bucket** $B_L \leftarrow$ Explanations with lowest scores
 - **Medium-quality bucket** $B_M \leftarrow$ Explanations with mid-range scores
 - **High-quality bucket** $B_H \leftarrow$ Explanations with highest scores
- 15: **Step 6: Construct Final Dataset**
- 16: Form dataset D by concatenating explanations in the order:

$$D = B_L \cup B_M \cup B_H$$

17: **return** D

Finesse Score

Why we done this work?



In veracity prediction and explanation, model hallucination leads to inconsistent probability distributions for identical inputs, especially in ambiguous cases. To address this, we introduce the Finesse score, which quantifies variance in probability distributions, directly measuring model uncertainty.



The Finesse parameter was computed by generating five responses for each preferred explanation using random temperature values between 0.6 and 0.9. The mean-variance across these probability distributions served as the Finesse score, capturing the model's output uncertainty



This score was integrated into the DPO loss function to scale log probabilities, prioritizing explanations with high consistency and lower uncertainty for improved factual reliability

Actuality score

Our approach penalizes factual errors, ensuring that only rigorously accurate explanations receive higher ranks, unlike traditional methods that may favor well-written but speculative responses.

The Actuality score provides a transparent and systematic measure of reliability, crucial for high-stakes domains like journalism and policy-making.

Its reference-free nature enables adaptability across diverse topics, maintaining factual consistency in fields such as politics, economics, science, and health.

Prompt for Actuality Score

Task: You will be given a news article. Follow these steps:

1. Extract up to 15 of the most important and factually relevant sentences from the article.
2. For each extracted sentence, assess its factual correctness:
 - Label each sentence as **1** if it is factually accurate.
 - Label it as **0** if it contains factual errors.
3. Compute the **average** of all the labels (1s and 0s).

Output: Return only the factual consistency score as a single numerical value (e.g., 0.75). Do not include any additional explanations, calculations, or extracted sentences. **Here is the news article:** {article}
Answer:

Hin-DPO Loss function

$$L_{\text{Hin-DPO}}(\pi_{\theta}; \pi_{\text{ref}}) = -\mathbb{E}_{(x, y_w, y_l) \sim D} \left[\log \sigma \left(\beta \left(\frac{(1 + s_w) \log \left(\frac{\pi_{\theta}(y_w|x)}{\pi_{\text{ref}}(y_w|x)} \right) - \max(0.01, s_l) \log \left(\frac{\pi_{\theta}(y_l|x)}{\pi_{\text{ref}}(y_l|x)} \right)}{v + \epsilon} \right) \right) \right] \quad (1)$$

y_w : Winning sample
 y_l : Losing Sample
 v : Variance calculated using Monte Carlo method and
 s : Factcheck score.
 π_{θ} : Model under Update
 π_{ref} : Reference Model
 x : Inputted Prompt

Gradient Analysis of Hin-DPO

B.1 Objective Function

The Hin-DPO objective is given by:

$$L_{\text{Hin-DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(x, y_w, y_l) \sim \mathcal{D}} \left[\log \sigma \left(\frac{\beta ((1 + s_w)r_w - \max(0.01, s_l)r_l)}{v + \epsilon} \right) \right],$$

where

$$r_w = \log \frac{\pi_\theta(y_w|x)}{\pi_{\text{ref}}(y_w|x)}, \quad r_l = \log \frac{\pi_\theta(y_l|x)}{\pi_{\text{ref}}(y_l|x)}.$$

B.2 Loss function formulation

$$r(x, y) = \beta \log \frac{\pi_\theta(y|x)}{\pi_{\text{ref}}(y|x)} + \beta \log Z(x)$$

In our method we use two different type reward modelling for Preferred and rejected response.
For Preferred response

$$r_w(x, y) = \beta(1 + s_w) \log \frac{\pi_\theta(y|x)}{\pi_{\text{ref}}(y|x)} + \beta \log Z(x)$$

For Rejected response

$$r_l(x, y) = \beta \max(0.01, s_l) \log \frac{\pi_\theta(y|x)}{\pi_{\text{ref}}(y|x)} + \beta \log Z(x)$$

$$P(y_w \succ y_l \mid x) = \frac{\exp(r_w(x, y_w))}{\exp(r_w(x, y_w)) + \exp(r_l(x, y_l))} = \frac{1}{1 + \exp(r_l(x, y_l) - r_w(x, y_w))}$$

$$P(y_w \succ y_l \mid x) = \frac{1}{1 + \exp \left(\beta \max(0.01, s_l) \log \frac{\pi_\theta(y_l \mid x)}{\pi_{\text{ref}}(y_l \mid x)} - \beta(1 + s_w) \log \frac{\pi_\theta(y_w \mid x)}{\pi_{\text{ref}}(y_w \mid x)} \right)}$$

$$P(y_w \succ y_l \mid x) = \sigma \left(\beta(1 + s_w) \log \frac{\pi_\theta(y_w \mid x)}{\pi_{\text{ref}}(y_w \mid x)} - \beta \max(0.01, s_l) \log \frac{\pi_\theta(y_l \mid x)}{\pi_{\text{ref}}(y_l \mid x)} \right)$$

B.3 Gradient Derivation

Let u denote the argument of the sigmoid function:

$$u = \frac{\beta((1 + s_w)r_w - \max(0.01, s_l)r_l)}{v + \epsilon}.$$

The gradient of $L_{\text{Hin-DPO}}$ is:

$$\nabla_\theta L_{\text{Hin-DPO}} = -\mathbb{E}_{(x, y_w, y_l) \sim \mathcal{D}} [(1 - \sigma(u)) \nabla_\theta u].$$

Compute $\nabla_{\theta} u$ as:

$$\nabla_{\theta} u = \frac{\beta}{v + \epsilon} [(1 + s_w) \nabla_{\theta} r_w - \max(0.01, s_l) \nabla_{\theta} r_l].$$

Substituting $\nabla_{\theta} u$ back, the gradient becomes:

$$\nabla_{\theta} L_{\text{Hin-DPO}} = -\mathbb{E}_{(x, y_w, y_l) \sim \mathcal{D}} \left[\frac{\beta^2 \sigma(u)}{v + \epsilon} [(1 + s_w) \nabla_{\theta} r_w - \max(0.01, s_l) \nabla_{\theta} r_l] \right].$$

Hin-DPO Algorithm

Algorithm 1 Direct Preference Optimization (DPO) Training Algorithm

Require: Training dataset and Dataloader D with win and lose samples (paired or unpaired), initial model parameters θ_0 , reference model π_{ref} , number of iterations T , scaling factor β , temperature parameter τ

```
1: Initialize model  $\pi_\theta$  with parameters  $\theta_0$ 
2: Set  $\pi_\theta$  to training mode and  $\pi_{\text{ref}}$  to evaluation mode
3: for iteration = 1 to  $T$  do
4:   for each batch in  $D$  do
5:     Initialize running mean  $\mu \leftarrow 0$  and running variance  $\sigma^2 \leftarrow 0$  {Running statistics for probability distribution}
6:     Set num_iter = 5 {Number of iterations for variance computation}
7:     for iter = 1 to num_iter do
8:       Compute logits for the preferred response:
9:       logits  $\leftarrow \pi_\theta(\text{pref\_ids}, \text{pref\_mask}).\text{logits}$ 
10:      Compute probabilities:
11:      probs  $\leftarrow \exp(\log\_probs(\text{logits}, \text{pref\_ids}))$ 
12:      Update Mean and Variance:
13:       $\mu \leftarrow \mu + \frac{\text{probs} - \mu}{\text{iter}}$ 
14:       $\sigma^2 \leftarrow \sigma^2 + (\text{probs} - \mu) \times (\text{probs} - \mu)$ 
15:    end for
16:    Compute final variance:
17:     $\sigma^2 \leftarrow \frac{\sigma^2}{\text{num\_iter} - 1}$ 
18:    Get log probabilities for preferred and dispreferred responses using  $\pi_\theta$ :
19:    model_pref_log  $\leftarrow \log\_prob(\pi_\theta(\text{pref\_ids}, \text{pref\_mask}), \text{pref\_ids})$ 
20:    model_dispref_log  $\leftarrow \log\_prob(\pi_\theta(\text{dispref\_ids}, \text{dispref\_mask}), \text{dispref\_ids})$ 
21:    Get log probabilities for preferred and dispreferred responses using reference model  $\pi_{\text{ref}}$ :
22:    ref_pref_log  $\leftarrow \log\_prob(\pi_{\text{ref}}(\text{pref\_ids}, \text{pref\_mask}), \text{pref\_ids})$ 
23:    ref_dispref_log  $\leftarrow \log\_prob(\pi_{\text{ref}}(\text{dispref\_ids}, \text{dispref\_mask}), \text{dispref\_ids})$ 
24:    Compute DPO loss:
25:    loss  $\leftarrow \text{DPO\_loss}(\text{model\_pref\_log}, \text{model\_dispref\_log}, \text{ref\_pref\_log}, \text{ref\_dispref\_log}, \sigma^2, \beta)$ 
26:    Backpropagate loss:
27:    loss.backward()
28:    Update model parameters:
29:     $\theta \leftarrow \text{optimizer.step}()$ 
30:   end for
31: end for
```

EXPERIMENTS: SETUP & BASELINES

We fine-tuned five models, including three Large Language Models (LLMs) and two Pre-trained Language Models (PLMs).

The LLMs used were **Gemma-2-2B-It**, **Llama-3.2-3B-Instruct**, and **Mistral-7B-Instruct-v0.3**. Additionally, we fine-tuned two PLMs: **mBART-large-50** and **mT5-base**.

The quality of the generated explanations was assessed using three key metrics: **BERTSCORE**, **ROUGE-1,2**, **L** score and **METEOR** score.

Results

Metrics:	ROUGE-1	ROUGE-2	ROUGE-L	METEOR	BERTSCORE
Model	mBART				
Base-model	13.03	6.12	9.72	18.23	69.94
DPO	13.96	7.02	10.25	18.98	72.12
DPO+ <i>Actuality</i>	14.10	6.92	10.45	19.25	73.07
DPO+Finesse	13.71	6.15	10.29	19.09	72.50
Hin-DPO	14.66	7.12	10.77	19.73	74.01
Model	mT5				
Base-model	15.13	6.84	9.93	19.16	70.14
DPO	16.24	7.45	10.69	19.63	72.29
DPO+ <i>Actuality</i>	16.31	7.42	10.81	19.95	73.06
DPO+Finesse	16.11	7.31	10.74	19.66	72.30
Hin-DPO	16.40	7.58	11.12	20.03	74.19
Model	mistral-7b				
Base-model	23.01	13.24	19.44	23.12	71.11
DPO	24.76	14.02	20.78	25.04	73.14
DPO+ <i>Actuality</i>	24.95	14.45	20.94	26.11	75.17
DPO+Finesse	24.79	14.11	20.79	25.12	73.87
Hin-DPO	26.21	15.04	21.93	26.55	76.24
Model	Llama3.2-7b				
Base-model	29.32	17.27	22.56	27.45	74.24
DPO	31.15	18.68	24.17	28.71	75.23
DPO+ <i>Actuality</i>	32.41	19.25	24.79	29.45	76.18
DPO+Finesse	31.78	18.76	24.43	28.98	76.02
Hin-DPO	33.89	20.04	25.72	30.47	76.96
Model	gemma2-2b				
Base-model	28.17	17.36	21.86	27.68	72.12
DPO	30.11	18.72	23.41	29.17	76.11
DPO+ <i>Actuality</i>	30.83	19.01	24.15	29.84	78.73
DPO+Finesse	30.25	18.87	23.58	29.20	76.56
Hin'-DPO	31.12	19.78	24.68	31.74	80.02

Table 1: Performance comparison across models. Bold values denote the best performance.

Conclusion

- + In conclusion, this work presents a novel framework, DeFactoX, that effectively addresses the challenges of misinformation detection and explanation generation in Hindi.
- + By creating a synthetic Hindi preference dataset and leveraging advanced techniques like DPO, factscore integration, and curriculum learning, we have demonstrated a scalable method for generating coherent, contextually relevant, and factually accurate explanations.

References

Gargi Joshi, Ananya Srivastava, Bhargav Yagnik, Mohammed Hasan, Zainuddin Saiyed, Lubna A Gabralla, Ajith Abraham, Rahee Walambe, and Ketan Kotecha (2023) – Explainable misinformation detection across multiple social media platforms. IEEE Access, 11:23634–23646.

Neema Kotonya and Francesca Toni (2020) – Explainable automated fact-checking: A survey. Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Sudhanshu Kumar and Thoudam Doren Singh (2022) – Fake news detection on Hindi news dataset. Global Transitions Proceedings, 3(1):289–297.

Xin Lai, Zhuotao Tian, Yukang Chen, Senqiao Yang, Xiangru Peng, and Jiaya Jia (2024) – Step-DPO: Step-wise preference optimization for long-chain reasoning of LLMs. arXiv preprint arXiv:2406.18629.

Chin-Yew Lin (2004) – ROUGE: A package for automatic evaluation of summaries. Text Summarization Branches Out, pages 74–81.

Sewon Min, Kalpesh Krishna, Xinxi Lyu, Mike Lewis, Wen-tau Yih, Pang Koh, Mohit Iyyer, Luke Zettlemoyer, and Hannaneh Hajishirzi (2023) – FActScore: Fine-grained atomic evaluation of factual precision in long-form text generation. Proceedings of EMNLP 2023, pages 12076–12100, Singapore. Association for Computational Linguistics.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D. Manning, Stefano Ermon, and Chelsea Finn (2024) – Direct preference optimization: Your language model is secretly a reward model. Advances in Neural Information Processing Systems, 36.

Jon Roozenbeek, Claudia R. Schneider, Sarah Dryhurst, John Kerr, Alexandra L.J. Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden (2020) – Susceptibility to misinformation about COVID-19 around the world. Royal Society Open Science, 7(10):201199.

References

Corby Rosset, Guoqing Zheng, Victor Dibia, Ahmed Awadallah, and Paul Bennett (2023) – Axiomatic preference modeling for long-form question answering. Proceedings of EMNLP 2023, pages 11445–11475. Daniel Russo, Serra Sinem Tekiroglu, and Marco Guerini (2023) – Benchmarking the generation of fact-checking explanations. Transactions of the Association for Computational Linguistics, 11:1250–1264.

Pranab Sahoo, Ayush Kumar Singh, Sriparna Saha, Vinija Jain, Samrat Mondal, and Aman Chadha (2024) – A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv preprint arXiv:2402.07927.

Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, Jiatao Gu, and Angela Fan (2020) – Multilingual translation with extensible multilingual pretraining and finetuning.

Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien Vincent, Zhufeng Pan, Shibo Wang, et al. (2024) – Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. arXiv preprint arXiv:2403.05530.

Sean Trott and Pamela Rivi re (2024) – Measuring and modifying the readability of English texts with GPT-4. Proceedings of TSAR 2024, Miami, Florida, USA. Association for Computational Linguistics.

Bram Wallace, Meihua Dang, Rafael Rafailov, Linqi Zhou, Aaron Lou, Senthil Purushwalkam, Stefano Ermon, Caiming Xiong, Shafiq Joty, and Nikhil Naik (2024) – Diffusion model alignment using direct preference optimization. Proceedings of CVPR 2024, pages 8228–8238.

Sizhe Wang, Yongqi Tong, Hengyuan Zhang, Dawei Li, Xin Zhang, and Tianlong Chen (2024) – BPO: Towards balanced preference optimization between knowledge breadth and depth in alignment. arXiv preprint arXiv:2411.10914.



THANK YOU