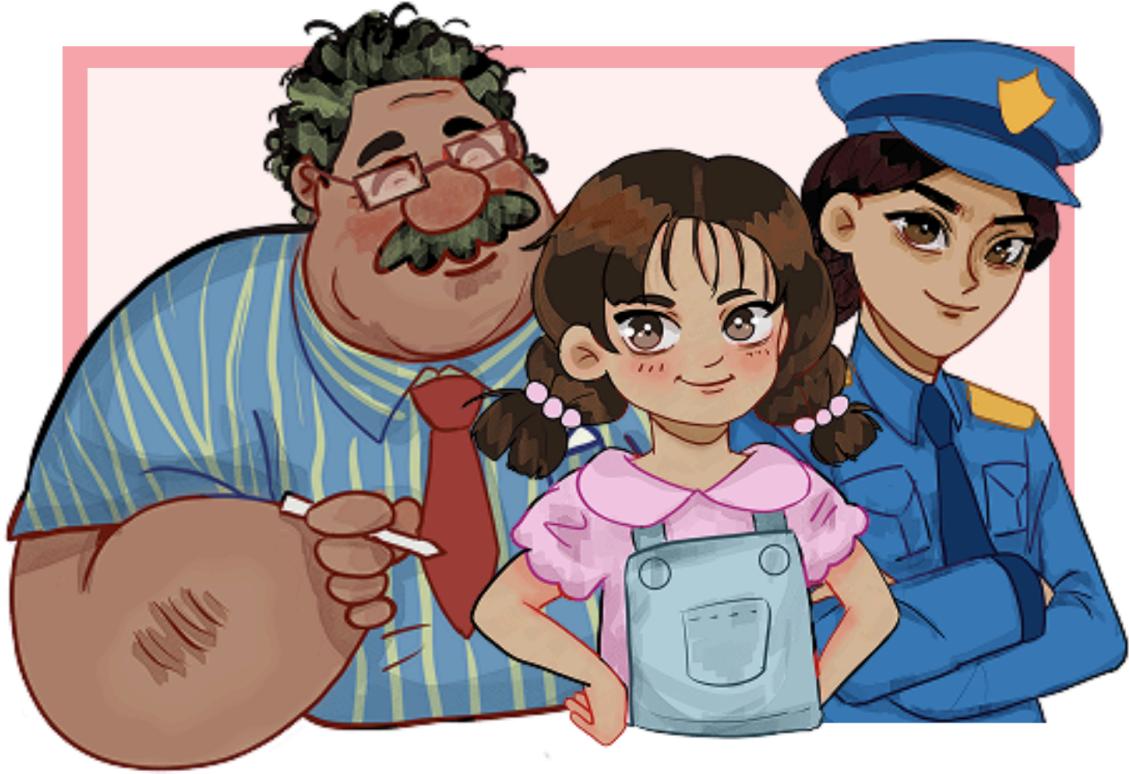


एकाइ ५

जिम्मेवार एआई : पूर्वाग्रहबाट मुक्त र विश्वसनीय प्रणाली निर्माण





Supported By:

ASML Foundation

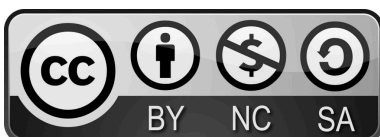


**Alternatives
4 children**

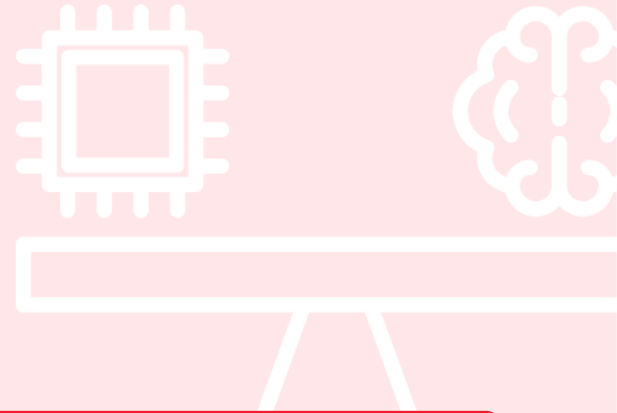
© Karkhana Samuha, 2025

Unless otherwise stated, this work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.

<https://creativecommons.org/licenses/by-nc-sa/4.0/>



जिम्मेवार एआई : पूर्वाग्रहबाट मुक्त र विश्वसनीय प्रणाली निर्माण



सिकाइका लक्ष्यहरू

यस पाठ एकाइको अन्त्यसम्ममा विद्यार्थीहरूमा निम्न सक्षमता विकास हुनेछ:

- पूर्वाग्रह (Bias) लाई परिभाषित गर्न र यसले मानव र एआई दुवैको निर्णय प्रक्रियालाई कसरी असर गर्छ भन्ने पहिचान गर्न ।
- एल्गोरिदम पूर्वाग्रह (Algorithmic Bias) : यो कसरी बन्छ, कस्तो प्रकारको डेटा त्रुटिले गर्दा हुन्छ, र यसले निष्पक्षता र समान अवसरलाई कसरी असर गर्छ भन्ने बुझ्न ।
- जिम्मेवार एआईका ६ वटा आधारभूत सिद्धान्तहरू सिक्न र तिनीहरूलाई वास्तविक जीवनका डिजिटल सन्दर्भहरूमा लागु गर्न ।
- प्रयोगकर्ताहरूले आफूले प्रयोग गर्ने डिजिटल प्रणालीहरूमा बुझेर दिइएको राय (informed feedback) मार्फत कसरी पूर्वाग्रहलाई चिन्न, प्रश्न गर्न र सक्रिय रूपमा कम गर्न सकिन्छ भन्ने बुझ्न ।

पूर्वाग्रह (Bias)

बुवा र छोरा एउटा गाडीमा जाँदै थिए । उनीहरू ठूलो दुर्घटना (Accident) मा परे ।

बुवाको त्यहीँ मृत्यु भयो, तर छोरालाई हतार-हतार अस्पताल लगियो ।

जब डाक्टरले त्यो बालकलाई देखे, डाक्टरले भने, “म यसको अपरेसन गर्न सकिदिनँ, **यो त मेरो छोरा हो !”**

अब भन्नुहोस्, “**त्यो डाक्टर को हो ?**”

- जवाफ : डाक्टर ती बच्चाकी आमा हुन् ।
- यो किन कठिन थियो : किनभने हाम्रो दिमागले छिटो निर्णयहरू गर्दा प्रायः हामी आफूसँग पहिल्यै भएको वा बनेको अधुरो वा अनुचित धारणाको प्रयोग गर्छौं । यसमा डाक्टर पुरुष नै हुनुपर्छ भन्ने हामीले ठान्यौं। यही नै मानव पूर्वाग्रह हो ।

पूर्वाग्रह (Bias) भनेको के हो ?

पूर्वाग्रह (Bias) भनेको कुनै ठोस प्रमाण वा सत्यता बिना नै कसैप्रति पहिले नै बनाइएको 'राम्रो' वा 'नराम्रो' धारणा हो। यसले गर्दा हामी सबैलाई समान र निष्पक्ष (Fair) व्यवहार गर्नुको सट्टा भेदभावपूर्ण व्यवहार गर्न पुग्छौं ।

पूर्वाग्रहको नाम (Bias Name)	सरल अर्थ (Simple Meaning)	उदाहरण (Example)
ऐतिहासिक पूर्वाग्रह (Historical Bias)	पुरानो भेदभाव दोहोरिनु यदि हाम्रो समाज वा इतिहासमा पहिलेदेखि नै भेदभाव छ भने, डेटामा पनि त्यही देखिन्छ र कम्प्युटरले पनि त्यही सिक्छ ।	विगतमा पुरुषहरूले मात्र काम पाएका थिए भने कम्प्युटरले महिलाहरूको आवेदन सिधै अस्वीकार गर्न सक्छ ।
प्रतिनिधित्व पूर्वाग्रह (Representation Bias)	सबैलाई नसमेट्नु मेसिनलाई सिकाउने डेटामा सबै खालका मानिसहरू (जस्तै : विभिन्न जात, रङ, लिङ्ग) को जानकारी नहुनु	गोरो छाला भएका मान्छेको फोटोमा मात्र तालिम दिएको क्यामेराले कालो वा गहुँगोरो छाला भएको मान्छे चिन्न सक्दैन ।
मापन पूर्वाग्रह (Measurement Bias)	गलत कुरा मापन कुनै कुरा मापन गर्न गाह्रो हुँदा गलत 'तथ्यांक' (Data) प्रयोग गर्नु ।	"राम्रो स्वास्थ्य" मापन गर्नको लागि "अस्पतालको खर्च" हेर्नु । गरिब मान्छे अस्पताल जान नसक्ने हुनाले मेसिनले उनीहरूलाई "स्वस्थ" भनिदिन्छ ।

<p>पुष्टिकरण पूर्वाग्रह (Confirmation Bias)</p>	<p>आफ्नो सोचसँग मिल्ने जानकारी मात्र रोज्नु: एल्गोरिदम बनाउने मानिसहरूले आफ्ना विश्वास मिल्ने डेटा मात्र छान्नु</p>	<p>यदि इन्जिनियरलाई लाग्छ कि 'महँगो स्कुल पढेका मात्र जान्ने हुन्छन्', भने उसले त्यस्तै डाटालाई मात्र महत्व दिन्छ</p>
<p>अटोमेसन पूर्वाग्रह (Automation Bias)</p>	<p>मेसिनमा अन्धविश्वास मानिसले सोच्नुको सट्टा 'कम्प्युटरले भनेको हो, त्यसैले यो सधैं सही होला' भनेर विश्वास गर्नु</p>	<p>कम्प्युटरले गल्तीले 'यो मान्छे अपराधी हो' भन्यो भने, मान्छेले पनि विना अनुसन्धान त्यसैलाई स्वीकार गर्नु</p>

क्रियाकलाप - घटना अध्ययनहरू (Case Studies)

निर्देशन : तलका केस स्टडीहरू पढ्नुहोस् र यी कुन पूर्वाग्रहसँग सम्बन्धित छन्, त्यो पत्ता लगाउनुहोस् ।

घटना अध्ययन 'क'

एउटा कम्पनीले कर्मचारी छान्नका लागि १० वर्ष पुरानो डेटा प्रयोग गर्‍यो । पहिला त्यो कम्पनीमा प्रायः पुरुषहरूले मात्र काम गर्दथे । त्यसैले, कम्प्युटर सफ्टवेयरले 'महिलाहरू यो काम गर्न सक्दैनन्' भनेर सिक्‍यो। योग्य महिलाहरूको आवेदन फारम पनि छनोट भएन ।

घटना अध्ययन 'ख'

एउटा रोबोटलाई मानिसको आवाज बुझ्न सिकाइएको थियो, तर त्यसलाई काठमाडौँ र पोखराका मानिसको आवाजमा मात्र तालिम दिइएको थियो । जब दुर्गम गाउँका एक विद्यार्थीले आफ्नै लवजमा (Accent) बोल्न खोजे, रोबोटले त्यो आवाज चिन्न सकेन र काम गरेन ।

मानव पूर्वाग्रह VS एल्गोरिदम पूर्वाग्रह

पक्ष (Aspect)	मानव पूर्वाग्रह (Human Bias)	एल्गोरिदम पूर्वाग्रह (Algorithmic Bias)
कसमा हुन्छ?	मानिसहरूमा	मेसिनहरूमा
कसरी हुन्छ?	व्यक्तिगत राय, पालनपोषण, र सामाजिक मापदण्डहरू (जस्तै, कुनै निश्चित जात वा लिङ्गको लागि प्राथमिकता)	मानिसले सिर्जना गरेको पूर्वाग्रहपूर्ण डेटा वा अपूर्ण नमूनाहरूबाट सिक्ने
उदाहरण	एक शिक्षकले सही, तर फोहोर जवाफहरू भन्दा सफा हस्तलेखनलाई बढी अङ्क दिने	गाढा छालाका रंगहरू पहिचान गर्न असफल हुने वा विशिष्ट सांस्कृतिक पहिरनका वस्तुहरूसँग सङ्घर्ष गर्ने अनुहार पहिचान (facial recognition) एप

एआईले अनुचित काम गर्नलाई जानीजानी निर्णय गर्दैन । यसले हामीबाट मात्र सिक्छ । यदि मानवहरूद्वारा दिइएको डेटा पूर्वाग्रही छन् भने, एआई पनि पूर्वाग्रही हुन्छ । **हामी मेसिनलाई सिकाउने पाठहरूको लागि जिम्मेवार छौं।**

एल्गोरिदम पूर्वाग्रह (Algorithmic Bias) :

जब एआईले हाम्रो बानी (र गल्तीहरू) सिक्छ

एआईले मानवहरूले सिर्जना गरेको डेटाबाट सिक्छ । यदि डेटामा मानव पूर्वाग्रह समावेश छ भने, एआईले त्यो पूर्वाग्रह सिक्छ र यसलाई स्वचालित (automate) गर्छ । यसलाई एल्गोरिदम

पूर्वाग्रह भनिन्छ ।

वास्तविक संसारका एल्गोरिदम त्रुटिहरू :

- **"जागिर पोर्टल समस्या" (ऐतिहासिक पूर्वाग्रह) :** एक भर्ती प्लेटफर्मले पुरानो कम्पनी रेकर्डहरूमा तालिम पाएको एआई प्रयोग गर्छ । यदि व्यवस्थापकहरू ऐतिहासिक रूपमा ९०% पुरुष थिए भने, एआईले **"व्यवस्थापक" लाई "पुरुष" सँग जोड्न** सिक्छ र योग्य महिला उम्मेदवारहरूलाई अनुचित रूपमा कम सिफारिस गर्छ । प्रणालीले विगतको असमानतालाई स्वचालित बनाउँछ र बढाउँछ ।
- **"सौन्दर्य फिल्टर भ्रम" (अपूर्ण डेटा/सांस्कृतिक पूर्वाग्रह) :** एउटा सेल्फी एपले छालालाई उज्यालो बनाउँछ र आँखालाई ठूलो बनाउँछ । यो एआई प्रायः युरोकेन्द्रीय विशेषताहरू भएका अनुहारहरू (जस्तै गोरो छाला, सोझो कपाल, सानो नाक इत्यादि) लाई समर्थन गर्ने विश्वव्यापी डेटासेटहरूमा तालिम दिइएको थियो । यसले रङ्गभेदलाई मात्र बढावा दिँदैन, तर एआईलाई सौन्दर्यको साँघुरो एवम् पूर्वाग्रहपूर्ण बुझाइ सिकाएर युवा प्रयोगकर्ताहरूको आत्म-सम्मानलाई हानी पुर्याउँछ ।
- **"आवाज पहिचान अन्तर" (भाषिक पूर्वाग्रह) :** एउटा एआई आवाज सहायकले नेपाली वा भोजपुरी लबजहरूलाई, विशेष गरी ग्रामीण शब्द वा लबज, बुझ्न कठिन महसुस गर्छ । किन ? किनभने यसलाई विशेष रूपमा अङ्ग्रेजी र सहरी भाषा र उच्चारण डेटासेटहरूमा तालिम दिइएको छ । एआईले ती प्रयोगकर्ताहरूको प्रतिनिधि गर्न वा त्यो लबज ल्याउन असफल रहन्छ, जसको आवाज यसको सिकाइ सामग्रीबाट गायब छ (उदाहरणको रूपमा : "हमरा के भोजपुरी में फोटोसिन्थेसिस (Photosynthesis) के बारे में समझा दीं" भन्दा एआईले त्यो भाषामा उत्तर दिन नसक्नु) । यो पनि एक प्रकारको पूर्वाग्रह हो ।



सचेत प्रयोगकर्ता गीता
(Mindful User Gita)

“जब मैले एआई प्रयोग गरेर अनुवाद गर्न खोजेँ, यसले हाम्रा स्थानीय उखान र लवजहरू बुझ्नै सकेन । यो एआई खराब भएर होइन, यसले हाम्रो संस्कृतिबारे सिक्ने मौका नै पाएको थिएन । यसले के देखाउँछ भने, एआईले त्यही मात्र जान्दछ, जुन उसलाई सिकाइएको छ । यसले हाम्रो आफ्नै भाषा र रहनसहनको असल डेटाहरू पाएको छैन र हाम्रो संस्कृति सिक्न पाएको छैन । त्यो नै एआईको सबैभन्दा ठूलो कमजोरी वा पूर्वाग्रह हो । त्यसैले एआई समावेशी हुनुपर्छ ताकि यसले हरेकको संसारलाई चिन्न सकोस् ।”

एल्गोरिदम पूर्वाग्रहका कारणहरू

कारण (Cause)	विवरण (Description)
अपूर्ण डेटा	यदि डेटासेटमा विविधताको कमी छ भने (जस्तै, केवल सहरी फोटोहरू, केवल पुरुष डाक्टरहरू) ।
ऐतिहासिक पूर्वाग्रह	एआईले पुरानो सामाजिक ढाँचाहरूबाट सिक्छ, जसमा जातीय, भाषिक, लैंगिक र अन्य पूर्वाग्रह भएको डेटा हावी हुन्छ ।
प्रतिक्रिया चक्रहरू	एआईले मानिसहरूले पहिले नै धेरै क्लिक गरेका चर्चित कुराहरूलाई मात्र बारम्बार देखाइरहन्छ । यसले गर्दा चर्चित कुराहरू झन् चर्चित हुन्छन्, तर नयाँ वा फरक खालका राम्रा सामग्रीहरू कसैले नदेख्ने गरी ओझेलमा पर्छन् र पूर्वाग्रह झन् बढ्दै जान्छ ।

एल्गोरिदम पूर्वाग्रह बारे थप सामग्री : 📌 [३ प्रकारका एल्गोरिदम पूर्वाग्रह](#) 📌 [Khan academy - Algorithmic Bias](#)

जिम्मेवार एआईका ६ सिद्धान्तहरू (The 6 Principles)

यदि पूर्वाग्रह असमानताको ऐना हो भने, जिम्मेवार एआई हामीलाई पूर्वाग्रह हेर्न, चिन्न र सच्याउन मद्दत गर्ने माध्यम हो । राम्रो एआई बनाउनुको अर्थ सबैलाई प्रतिनिधित्व गर्ने प्रणालीहरू (System) डिजाइन गर्नु हो । यो ६ वटा मुख्य सिद्धान्तहरूद्वारा निर्देशित छ :

१. निष्पक्षता (Fairness)

- अर्थ : एआईले सबै प्रयोगकर्ताहरूलाई समान रूपमा व्यवहार गर्नुपर्छ । जाति, लिङ्ग वा वर्गमा आधारित भएर भेदभाव गर्नु हुँदैन ।
- वास्तविक जीवनमा : जागिरका लागि आवेदन दिने व्यक्तिहरूलाई मूल्याङ्कन गर्दा एआईले उनीहरूको लिङ्ग (महिला वा पुरुष) हेर्नु हुँदैन । यसको सट्टा, एआईले केवल उनीहरूको सीप, योग्यता र क्षमतालाई मात्र आधार मान्नुपर्छ ।

२. पारदर्शिता र स्पष्टता (Transparency & Explain ability)

- अर्थ : एआईले कसरी निर्णय गर्‍यो भन्ने कुरा बुझ्न सकिने हुनुपर्छ । यसको अर्थ हामीले एआईले प्रयोग गरेको डेटा र एल्गोरिदम हेर्न सक्नुपर्छ ।
- वास्तविक जीवनमा : एउटा स्कूल एआई च्याटबोटले यदि विद्यार्थीको कापी जाँच गर्छ भने, विद्यार्थीहरूलाई गलत जवाफका कारणहरू र अझ कसरी दिइयो भनेर स्पष्ट रूपमा व्याख्या गर्नुपर्छ ।

३. सुरक्षा र दृढता (Safety & Robustness)

- अर्थ : एआई प्रणालीहरू सधैं सुरक्षित रूपमा सञ्चालन हुनुपर्छ र ह्याकिङ (hacking) वा गल्तीहरूबाट सुरक्षित हुनुपर्छ ।

वास्तविक जीवनमा : स्वचालित ड्राइभिङ कारहरूले सधैं मानिसको सुरक्षालाई

प्राथमिकता दिनुपर्छ ।

४. समावेशीता (Inclusiveness)

- **अर्थ** : एआई प्रणालीहरू सबै मानिसका लागि उपयोगी हुनुपर्छ । उमेर, भाषा, क्षमता, वा पृष्ठभूमिका आधारमा कसैलाई पनि छुट्याउनु हुँदैन ।
- **वास्तविक जीवनमा** : आवाजबाट काम गर्ने एआई उपकरणहरूले बोल्न गाह्रो हुने वा फरक तरिकाले बोल्ने मानिसहरूलाई पनि सहज रूपमा प्रयोग गर्न सकिने हुनुपर्छ ।

५. गोपनीयता र सुरक्षा (Privacy & Security)

- **अर्थ** : एआईले प्रयोग गर्ने वा यसलाई दिइने प्रयोगकर्ताको तथा अन्य गोप्य डेटालाई सुरक्षित राख्नुपर्छ ।
- **वास्तविक जीवनमा** : प्रयोगकर्ता डेटाको सुरक्षा गर्ने र दुरुपयोग हुनबाट रोक्ने ।

६. उत्तरदायित्व (Accountability)

- **अर्थ** : जब एआईले गल्ती गर्छ, त्यसको जिम्मेवारी लिने कोही हुनुपर्छ ।
- **वास्तविक जीवनमा** : यदि कुनै एआईले एउटा अस्पतालमा गल्ती गर्छ भने, एआई विकास गर्ने कम्पनी वा अस्पतालमध्ये कोही एक जिम्मेवार हुनुपर्छ ।

थप सामग्री : 📌 [Responsible AI Principles](#)

तपाईंको भूमिका : पूर्वाग्रह चिन्ने, प्रश्न गर्ने, र न्यूनीकरण गर्ने

एआईलाई निष्पक्ष बनाउनका लागि तपाईं कम्प्युटर वैज्ञानिक हुनु आवश्यक छैन, तपाईं एक विद्यार्थी वा सामान्य प्रयोगकर्ता भएर पनि मद्दत गर्न सक्नुहुन्छ । तपाईंका सचेत कार्यहरू (Conscious actions) एआईलाई नैतिक रूपमा सुधार गर्नका लागि महत्त्वपूर्ण छन् ।


चरण १ : एआईमा अनुचितता पत्ता लगाउने (डिजिटल जासूस)

डिजिटल जासूसको रूपमा, तपाईंले एआईले कहिले असन्तुलित सामग्री देखाउँछ भनेर पत्ता लगाउन सिक्नुपर्छ ।

- सोध्नुहोस् : “के कमी छ ?”, मानौं तपाईंले एआईलाई "नेपालको एउटा मुख्य चाडको खानाको बारेमा बताऊ" भनेर सोध्नुभयो। यदि एआईले सधैं सेलरोटी र गुन्द्रुकको मात्र कुरा गर्छ, तर मधेशमा खाइने ठकुआ, भुसवा वा बगिया जस्ता परिकारहरूको नामै लिँदैन भने, त्यहाँ 'डेटाको पूर्वाग्रह' (**Data Bias**) छ। यदि एआईले नेपालका सबै क्षेत्रको खानपान र संस्कृतिलाई बराबर महत्व दिँदैन भने, यसको मतलब उसले हाम्रो विविधतालाई अझै राम्रोसँग सिक्न बाँकी छ ।
- अवलोकन गर्नुहोस् : “कुन विचारहरू लुकेका छन् ?” के तपाईंको भिडियो एपले सधैं अधिकारमा एकै प्रकारका व्यक्तिलाई मात्र देखाउँछ (जस्तै, प्रविधि भिडियोहरूमा केटाहरू मात्र वा विज्ञापनहरूमा गोरो छालाका मोडलहरू मात्र) ? यसको अर्थ एआईले केवल एक प्रकारको **सामान्यीकृत धारणा (Stereotype)** लाई बढावा दिइरहेको छ ।
- तुलना गर्नुहोस् : “के यो सम्पूर्ण सत्य हो ?” यदि एउटा एपले पोखरा नयाँ राजधानी हो भन्छ भने, विश्वास गर्नु अघि प्रमुख समाचार साइटमा जाँच गर्नुहोस्! **उत्तेजकपूर्ण (Sensational)** कुरामाथि प्रश्न गर्नुहोस् । एउटा स्रोतमा मात्र विश्वास नगर्नुहोस् ।

एआईको गल्ती र पूर्वाग्रह सुधारणै तरिका

एआईले कहिलेकाहीं गलत वा भेदभावपूर्ण कुरा भन्न सक्छ। यस्तो अवस्थामा यी ३ कदम चाल्नुहोस्:

- **सत्यता जाँचुहोस् (Verify):** एआईले दिएको जानकारी शंकास्पद लागेमा तुरुन्तै किताब वा शिक्षकसँग सोधेर पुनर्पुष्टी गर्नुहोस्। एआईले सधैं सही कुरा मात्र भन्छ भन्ने हुँदैन ।
- **प्रतिक्रिया दिनुहोस् (Feedback):** एआईको गलत उत्तरमा 'थम्ब्स डाउन' () वा 'Report' बटन थिच्नुहोस्। यसले एआई बनाउने कम्पनीलाई आफ्नो गल्ती सुधारणै मद्दत पुग्छ ।
- **सुझाव दिनुहोस् (Re-prompt):** यदि एआईले पूर्वाग्रहपूर्ण कुरा गर्यो भने, उसलाई "तपाईंको उत्तर निष्पक्ष छैन, कृपया सबैलाई समेट्ने गरी जवाफ दिनुहोस्" भनी स्पष्ट निर्देशन दिनुहोस् ।

चरण २ : एआईलाई राम्रो बनाउने (तपाईंको प्रतिक्रियाको शक्ति)

तपाईंको डिजिटल कार्यहरूले एआईलाई अझ **निष्पक्ष** र **समावेशी (Inclusive)** हुन सिकाउँछन्।

- **प्रतिक्रिया दिनुहोस्** : यदि कुनै अनुवाद एप वा आवाज उपकरणले तपाईंको स्थानीय शब्दहरू गलत बुझ्छ भने, "**सुझाव सम्पादन (suggest edit)**" वा "**गल्ती रिपोर्ट गर्नुहोस् (report mistake)**" मा क्लिक गर्नुहोस् ।
 - उदाहरण : तपाईंले नेपालीमा 'ओहो! कति राम्रो!' भन्नुभयो, तर अनुवादकले त्यसलाई 'Oh no! How good!' भनेर गलत अनुवाद गर्‍यो । तपाईंले 'सुझाव सम्पादन' मार्फत सही अनुवाद सच्याउन सक्नुहुन्छ । त्यो सरल कार्यले एआईलाई तपाईंको भाषा सिक्न मद्दत गर्छ र भाषिक पूर्वाग्रह कम गर्छ ।
- **स्थानीय ज्ञान पोस्ट वा शेयर गर्नुहोस्** : जब तपाईं आफ्नो स्थानीय संस्कृति वा परम्पराहरूको बारेमा फोटो, कथाहरू, वा जानकारी पोस्ट गर्नुहुन्छ, तपाईंले इन्टरनेटको डेटा पूलमा **वास्तविक, महत्त्वपूर्ण विविधता** थपिरहनुभएको हुन्छ ।
 - उदाहरण : गुगल म्याप्स (Google Maps) मा आफ्नो गाउँको नयाँ बाटोको फोटो अपलोड गर्नु वा स्थानीय चाडपर्व (Festival) को बारेमा विकिपिडियामा जानकारी थप्नु । यसले एआईलाई नेपालको विविधता राम्ररी चिन्न सिकाउँछ ।
 - **थप जानकारी** : यहाँ एउटा भिडियो लिङ्क छ, जसले Google Maps मा जानकारी अपलोड गर्ने प्रक्रिया देखाउँछ : [YouTube video about how to add missing places or suggest edits on Google Maps] (यसको लागि उपयुक्त YouTube लिङ्क खोजेर यहाँ राख्न सकिन्छ ।)

विद्यार्थीले आवाज उठाउने विभिन्न तरिकाहरू

(Different Ways Students Can Speak Up)

तपाईं केवल रिपोर्ट गर्नेभन्दा बढी गर्न सक्नुहुन्छ । तपाईंको 'आवाज' नै सबैभन्दा शक्तिशाली उपकरण हो :

- 1. कक्षामा बहस (Classroom Debates) :** पूर्वाग्रहबाटपूर्ण एआईका केस स्टडीहरू (Case Studies) प्रस्तुत गर्नुहोस् र यसको नैतिक (Ethical) प्रभावबारे बहस गर्नुहोस् ।
- 2. सकारात्मक टिप्पणी (Positive Commenting) :** यदि तपाईंले कसैलाई दुर्व्यवहार (Harassment) भएको देख्नुभयो भने, उत्पीडकलाई जवाफ दिनुको सट्टा, पीडित व्यक्तिलाई निजी सन्देश पठाएर वा पोस्टमा सकारात्मक, समर्थनमूलक टिप्पणी गरेर समर्थन गर्नुहोस् ।
- 3. सामग्री छनोट र व्यवस्थापन (Content Curation) :** आफ्नो सामाजिक सञ्जाल (Social Media) मा तपाईंले मन पराउने विविध (Diverse), निष्पक्ष (Fair), र शैक्षिक (Educational) सामग्रीहरू मात्र पोस्ट वा शेयर गर्ने बानी बसाल्नुहोस् ।
- 4. प्लेटफर्महरूलाई अनुरोध (Request to Platforms) :** अनुवाद वा आवाज उपकरणमा 'गल्ती रिपोर्ट गर्नुहोस्' मार्फत निरन्तर प्रतिक्रिया (Consistent Feedback) पठाउनुहोस्, यो सामूहिक दबाबले कम्पनीहरूलाई आफ्ना उत्पादनहरूमा सुधार गर्न बाध्य पार्छ ।

थप सामग्री: [👉 पूर्वाग्रह चिन्ने, प्रश्न गर्ने, र न्यूनीकरण गर्ने](#)



सचेत प्रयोगकर्ता गीता
(Mindful User Gita)

“मैले एक पटक याद गरें कि एउटा प्रसिद्ध फोटो एपको फिल्टरले मेरो साथीको अनुहार चिन्न सकेन किनभने उनको छालाको टोन गाढा थियो। एप छोड्नुको सट्टा, मैले एउटा छोटो सुझाव सन्देश लेखें, “कृपया आफ्नो एप सुधार गर्नुहोस्, सबै छालाका टोनहरू महत्त्वपूर्ण हुन्छन्।” यो सानो लाग्न सक्छ, तर हामीजस्ता प्रयोगकर्ताहरूको इमानदार प्रतिक्रिया नै एल्गोरिदमहरूलाई निष्पक्ष बनाउने एक मात्र तरिका हो।”

केस स्टडी उत्तर :

- ऐतिहासिक पूर्वाग्रह (Historical Bias)
- प्रतिनिधित्व पूर्वाग्रह (Representation Bias)