Survey: Using Explainable AI (XAI) Techniques on a Data Privacy dataset

31 responses

Publish analytics

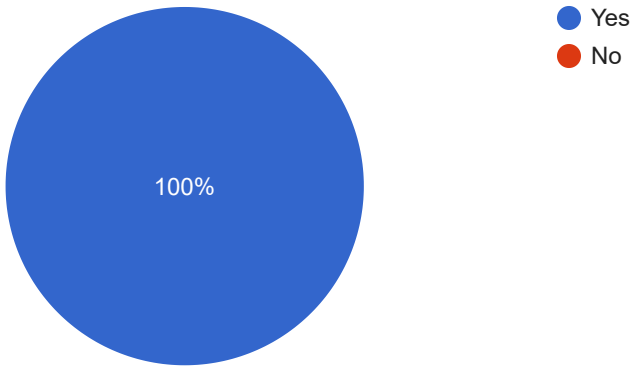
If you are a NUS / Yale-NUS student, are you 18 years old and above?

Copy

OR

If you are not a NUS student, are you above 21+ years old?

31 responses

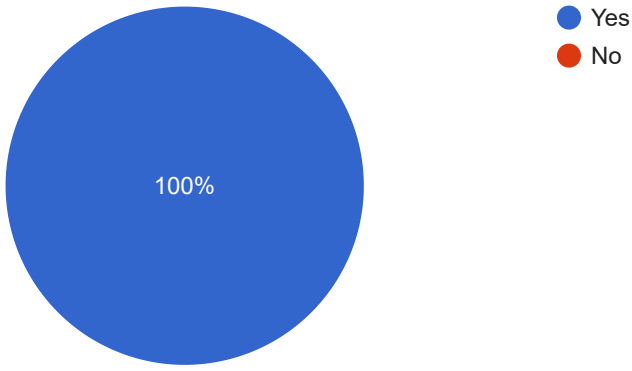


Participant Information Sheet

I have read about the purpose of this research study, agree to participate, and understand that I can withdraw at any time.

Copy

31 responses

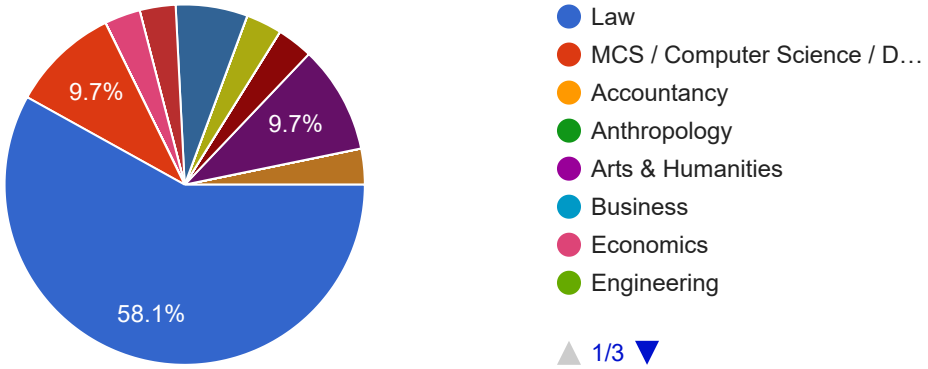


Part 1

What is your major / prospective major?

Copy

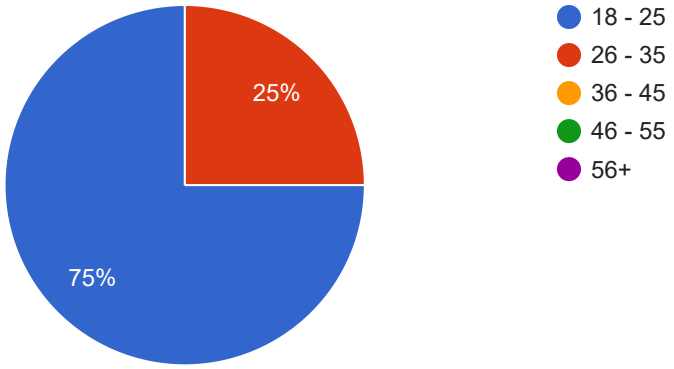
31 responses



What is your age?

Copy

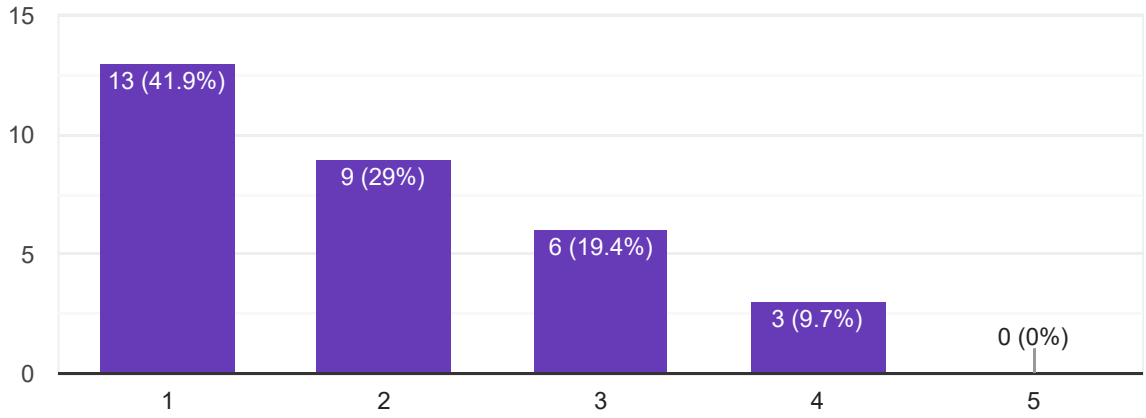
4 responses



Do you have any experience with AI / data science / programming?

Copy

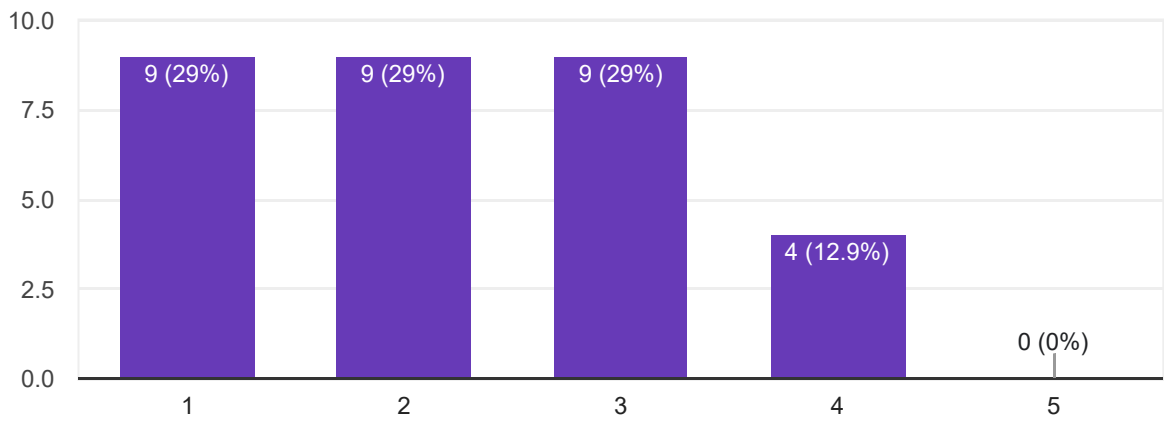
31 responses



Do you have any experience with regarding data privacy / law?

Copy

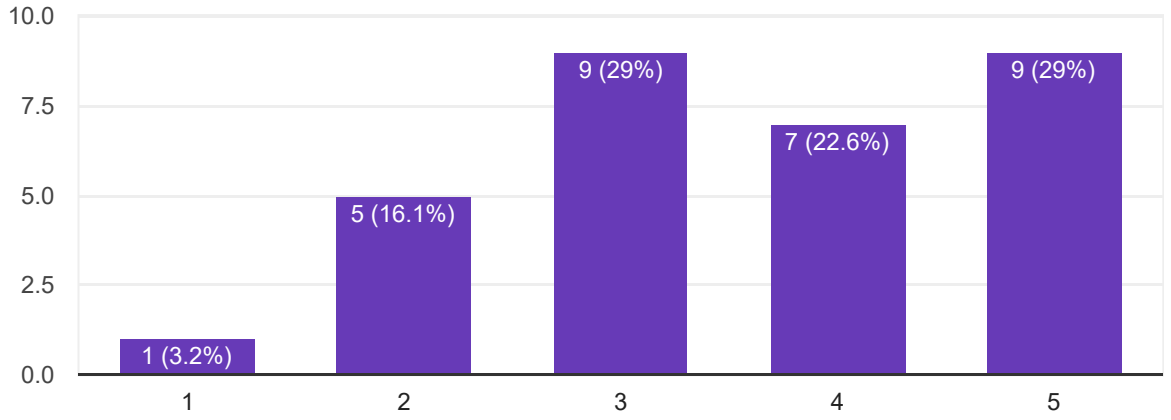
31 responses



Are you concerned about your data privacy?

Copy

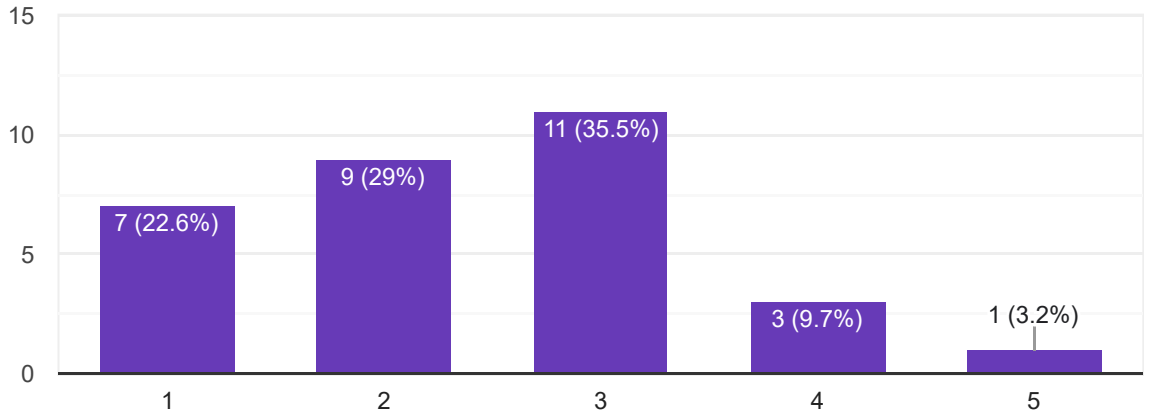
31 responses



How would you rate your capability in protecting your online data?

Copy

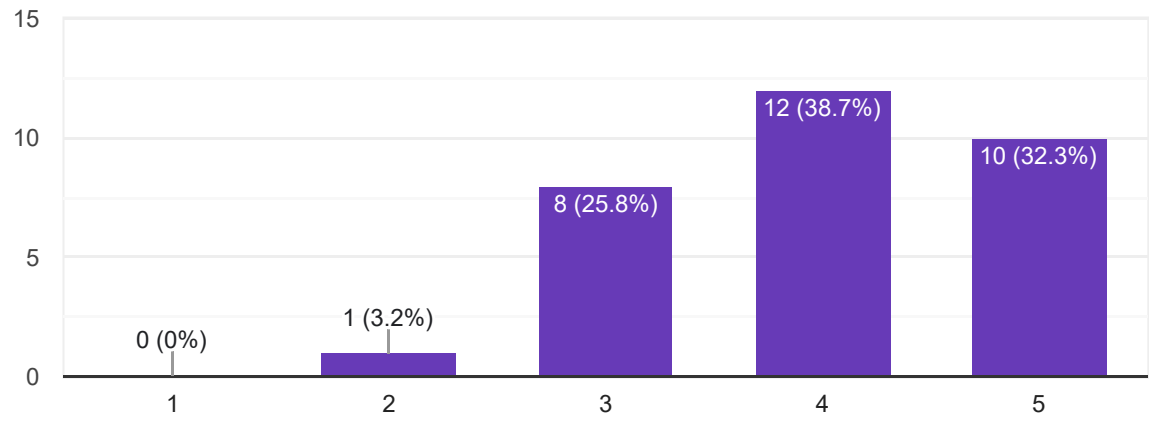
31 responses



Do you think decisions that are made by AI can be useful to society?

Copy

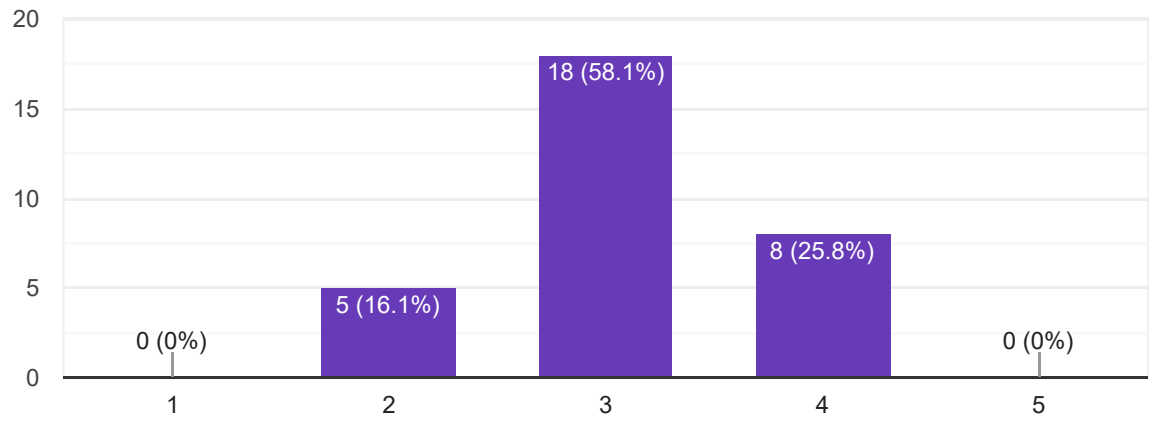
31 responses



Do you think decisions made by AI are fair?

Copy

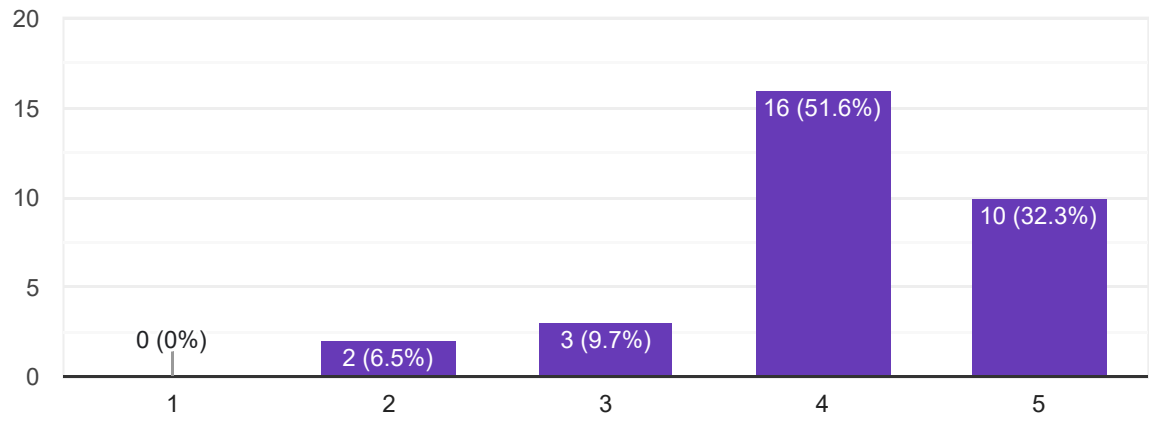
31 responses



Do you think decisions made by AI can be a risk to society?

Copy

31 responses

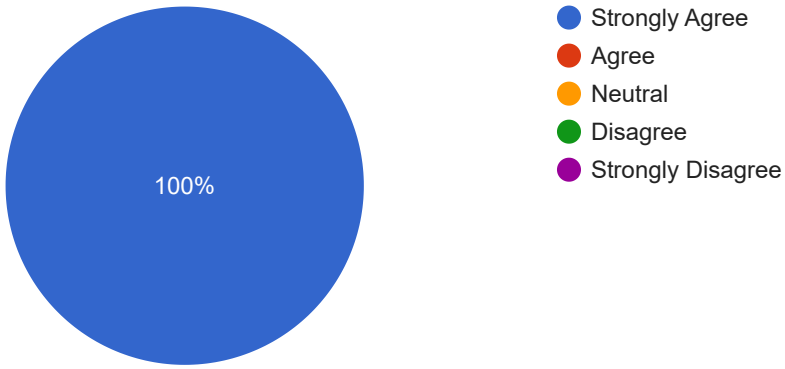


Part 2

Please select "strongly agree" to show that you are paying attention to this question.

Copy

31 responses



In this section, I will describe three different contexts with similar facts that relate to the use of the abovementioned model in analysing data privacy policies.

Each context corresponds with the perspective of an app developer, a member of the Personal Data Protection Commission (PDPC), and an user of the app.

I would then ask you questions to capture how your opinions on the use of AI in decision making would differ based on these three different perspectives.

Context 1: Imagine that you are an app developer. You are developing an app that uses cookies to track user activity online. To comply with the PDPA, you know that you need to include a sentence in your app's data privacy policy that notifies and asks for users' consent to use cookies.

Since you have no knowledge of the PDPA, you use the abovementioned model to analyse a pre-drafted data privacy policy that you found online. The model informs you that there is a sentence which states that cookies are being used.

You are deciding whether to rely entirely on the model's prediction, or pay costly legal fees to confirm with your friend who is a lawyer.

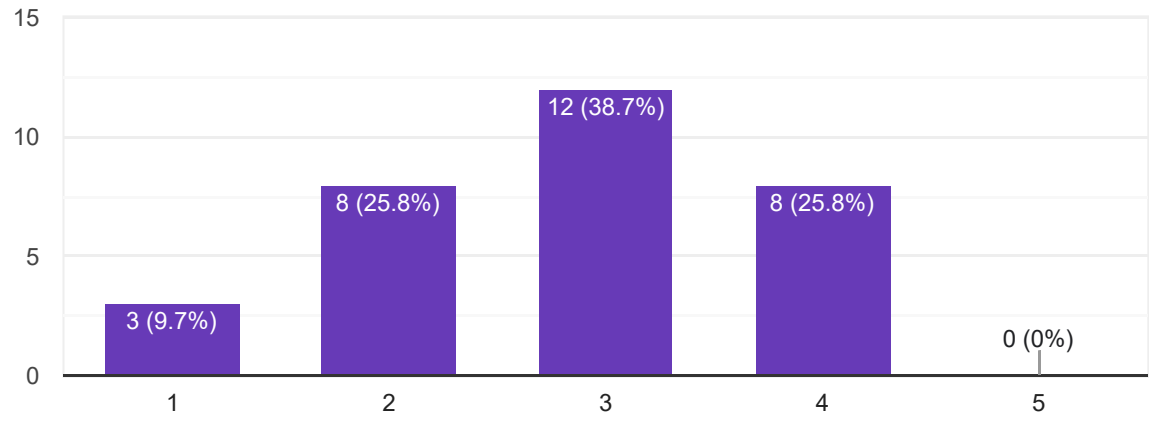
If the pre-drafted data privacy policy actually does not state that cookies are being used but your app uses cookies, you could face a fine of up to $10,000 in breach of the PDPA as you would have failed to notify your users.

How far do you, as the app developer:

Think that using the model is an effective method of identifying violations of the PDPA?

Copy

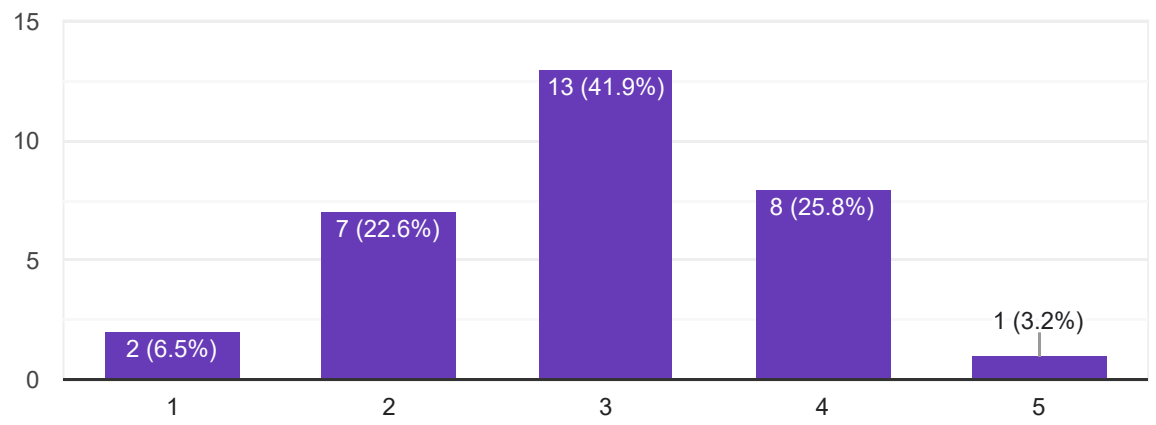
31 responses



Think that using the model is a fair method of identifying violations of the PDPA?

Copy

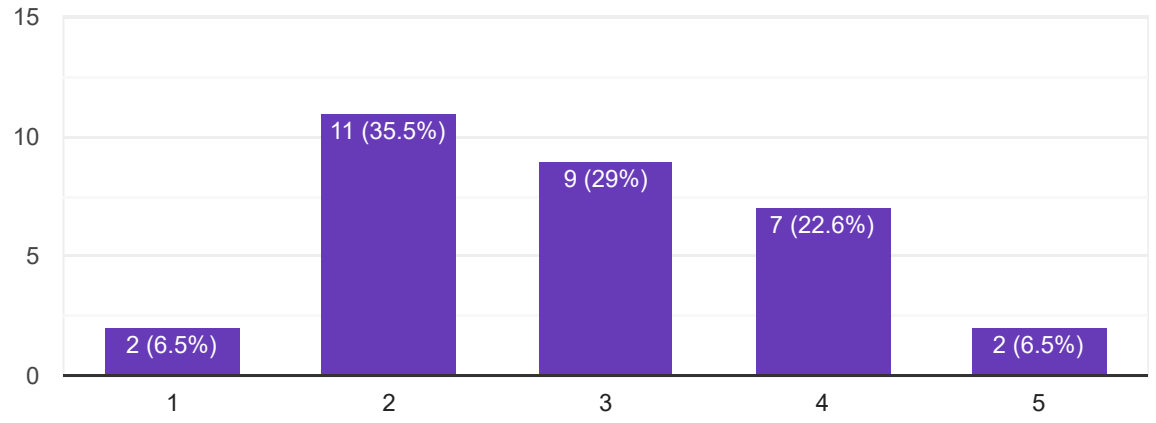
31 responses



Think that using the model is a method that could be a risk to society?

Copy

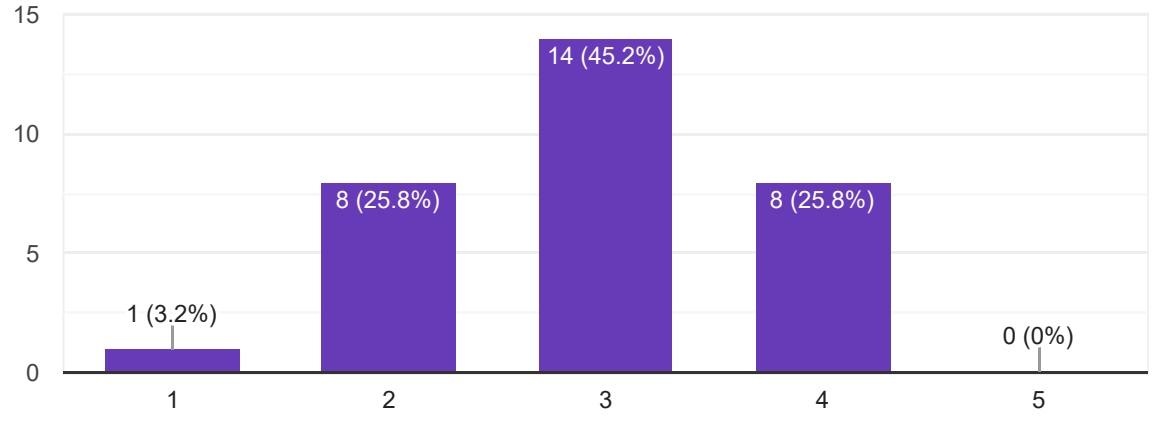
31 responses



Trust the prediction made by the model?

Copy

31 responses



Context 2: Imagine that you are a committee member part of the Personal Data Protection Commission (PDPC). A user of an app has informed you that an app is using cookies but has not notified its users.

Your team checks the code of the app and confirms that the app is indeed using cookies. Your team uses the abovementioned model and the model informs you that the data privacy policy does not contain any sentence that notifies its users that it uses cookies.

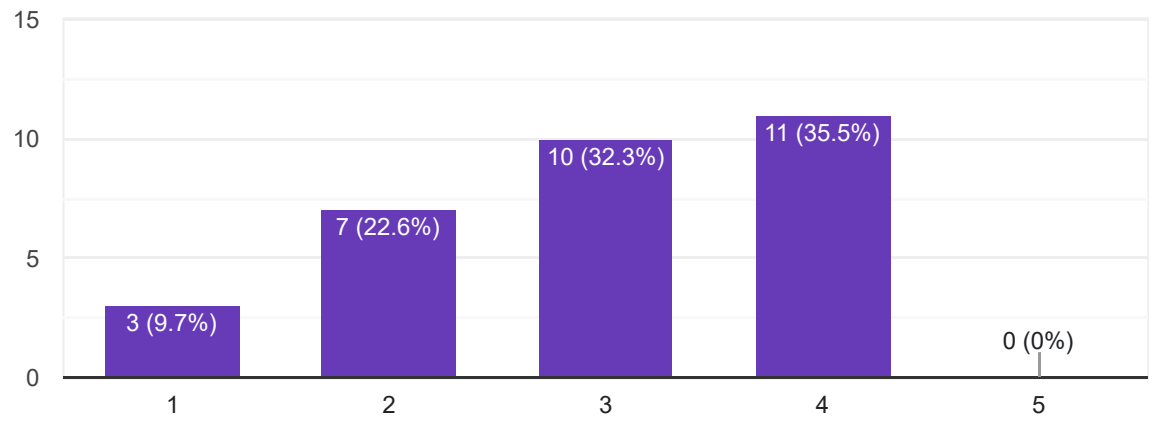
To increase the efficiency of the PDPC, your team is considering whether to adopt the abovementioned model to automate the analysis of data privacy policies. If this new method of analysis is adopted, the PDPC would rely entirely on the model's predictions to confirm whether app developers have breached the PDPA. The app developers would face a fine of up to $10,000 if they are found to have breached the PDPA.

How far would you, as a committee member of the PDPC:

Think that using the model is an effective method of identifying violations of the PDPA?

Copy

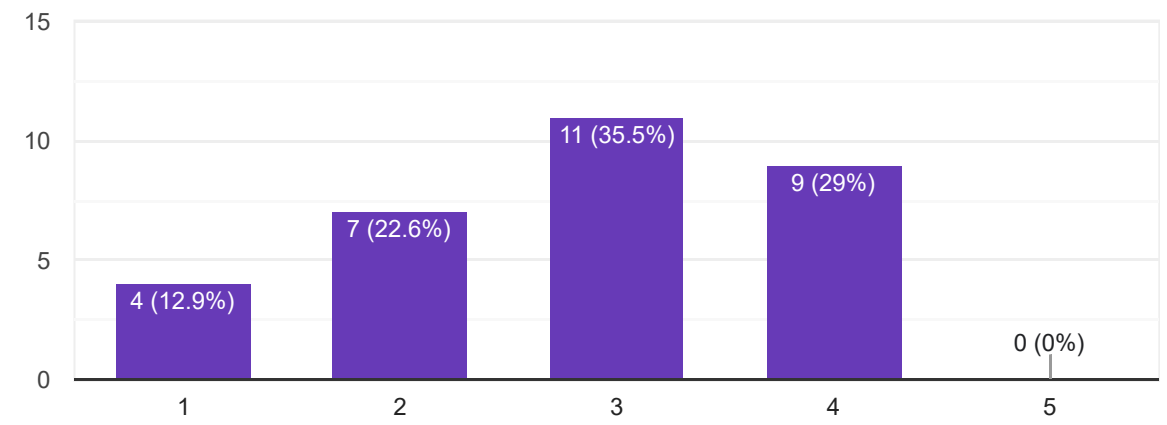
31 responses



Think that using the model is a fair method of identifying violations of the PDPA?

Copy

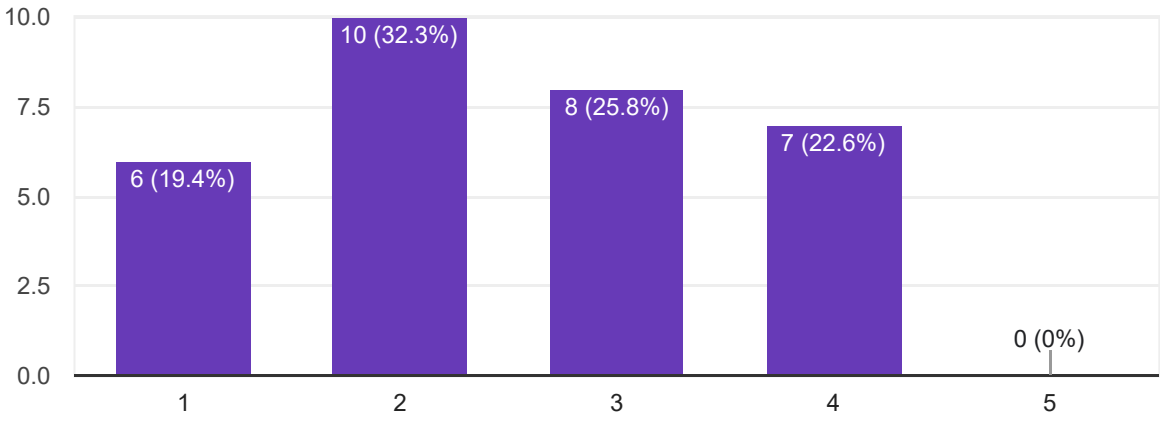
31 responses



Think that using the model is a method that could be a risk to society?

Copy

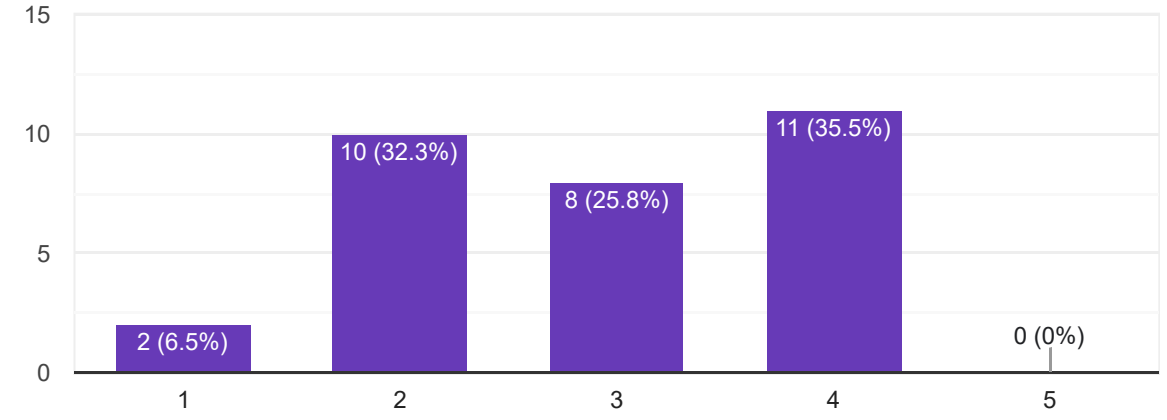
31 responses



Trust the prediction made by the model?

Copy

31 responses



Context 3: Imagine that you are a user of an app. You read in a forum where other users allege that the app uses cookies. You decide to analyse the data privacy policy of the app using the abovementioned model and the model informs you that the data privacy policy does not contain any sentence that notifies its users that it uses cookies.

You are deciding whether to submit this prediction as the only supporting piece of evidence to the PDPC to claim that the app has used cookies without notifying you.

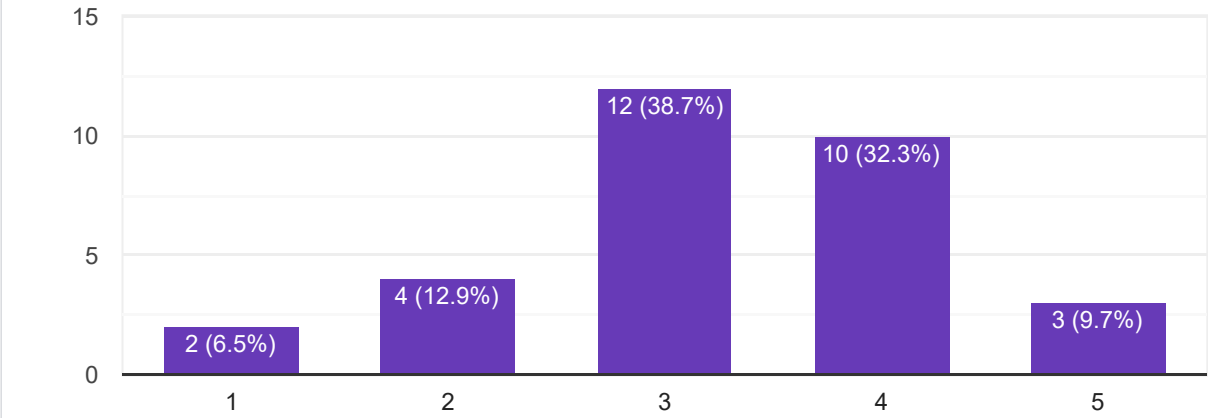
If the PDPC decides that the developer has indeed violated the PDPA, you could claim compensation from the app developer of up to $10,000.

How far would you, as a user of the app:

Think that using the model is an effective method of identifying violations of the PDPA?

Copy

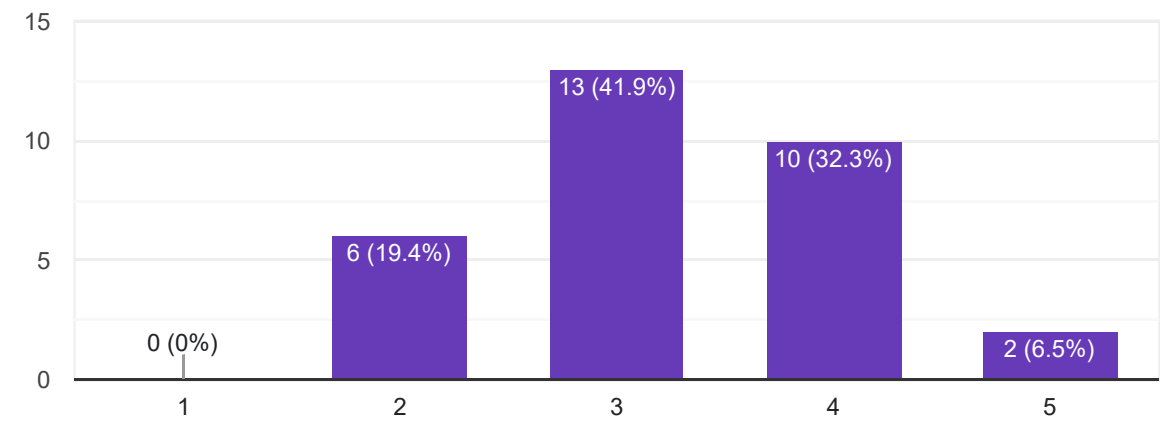
31 responses



Think that using the model is a fair method of identifying violations of the PDPA?

Copy

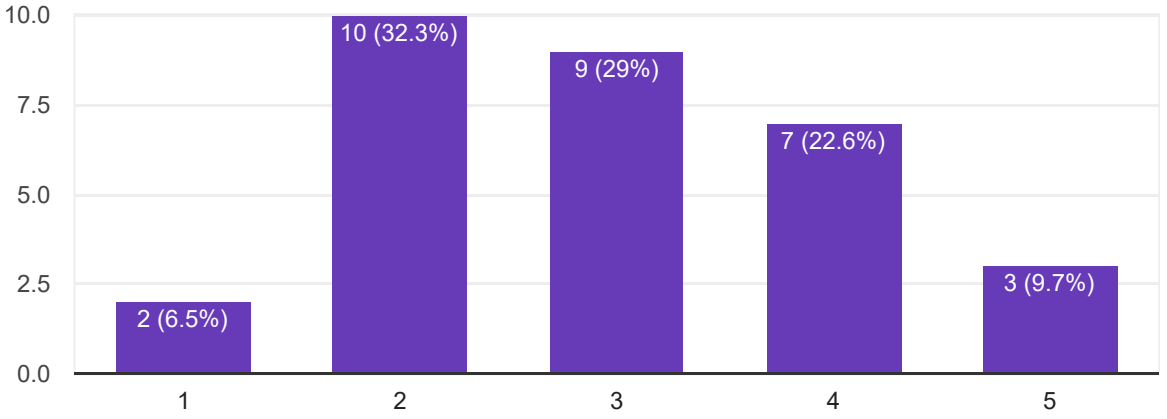
31 responses



Think that using the model is a method that could be a risk to society?

Copy

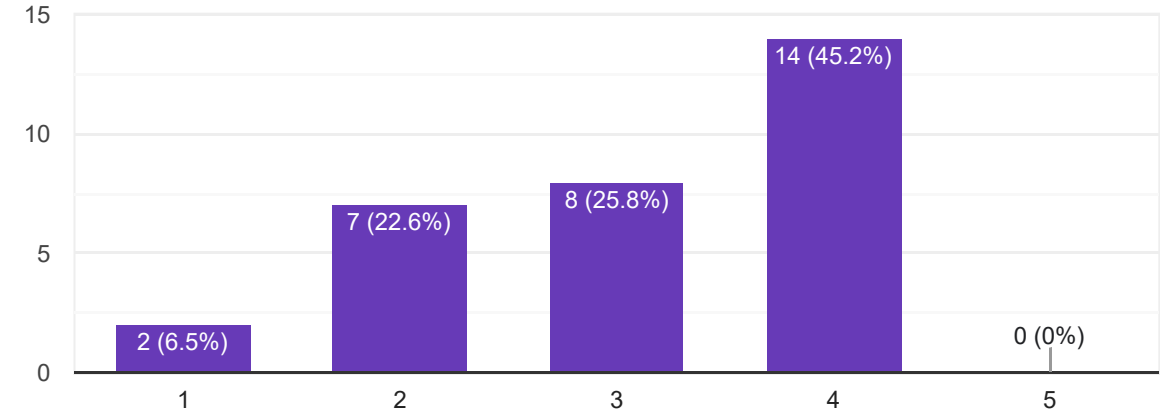
31 responses



Trust the prediction made by the model?

Copy

31 responses

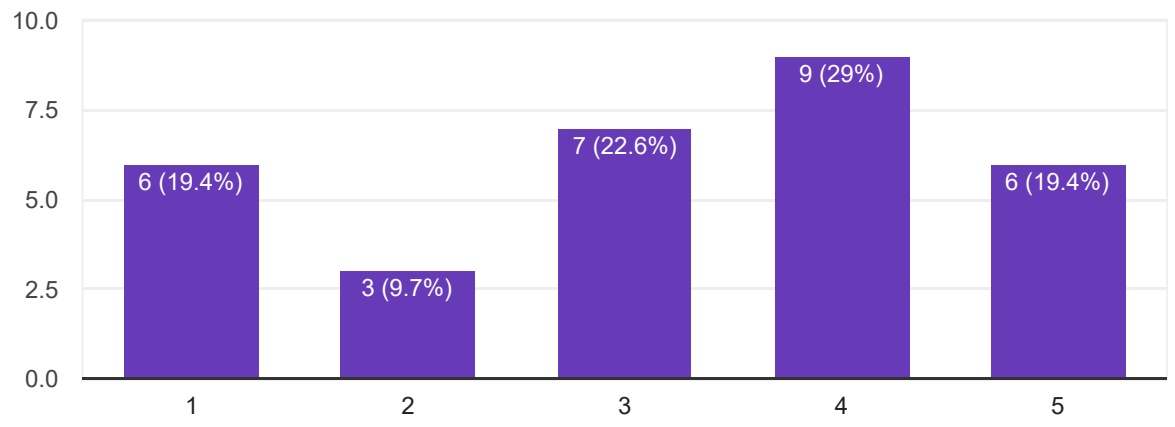


Part 3

Do you understand why the model made the prediction?

Copy

31 responses



Why do you think the model made this prediction?

31 responses

Cookies appears 4 times and technologies as well

Key words came up

There is a lot of mention of the word cookies

All the words have positive influence on the model, especially numerous mentions of the word cookies

Cookies and and technologies were used

HELP

Based on the data given to the model during the machine learning phase of the model, the model makes decisions regarding the sample data based on similar or identical data in the training data.

A lot of words that were identified with cookies were flagged

frequency of apperance of "cookies"

The model has been trained to match words like "collection" or "data" and if those appear in the text, it will categorise the input accordingly. "

cookies are mentioned 4 times

cos a cookie was identiifed

I have no clue man

The data input.

The prediction model identified many orange highlighted words which accurately corresponded to the actual practice.

by identifying words in silo

The key words were picked up by the model in one specific domain

The clause mentions cookies with no mention of sending infromation to other 3rd parties

No idea

The model picked up the relavant keywords in the terms and conditions

not sure

It's honing in on the word cookie

Key words easy to identify and set correctly.

Probably because it’s code recognised the issue

cookies

many words present which AI knew to be associated w that cookie

Has words "cookies" - which contributed significantly to the prediction of "Identifier_Cookie_or_Simiar_Tech_1st_Party" and does not have any words that identify with NOT "Identifier_Cookie_or_similar_Tech_1stParty". This resulted in "Identifier_Cookie_or_similar_Tech_1st Party" having the highest probability.

The word "cookie" is a good predicator when identifying whether an app uses cookies

Word "cookies" is used several times

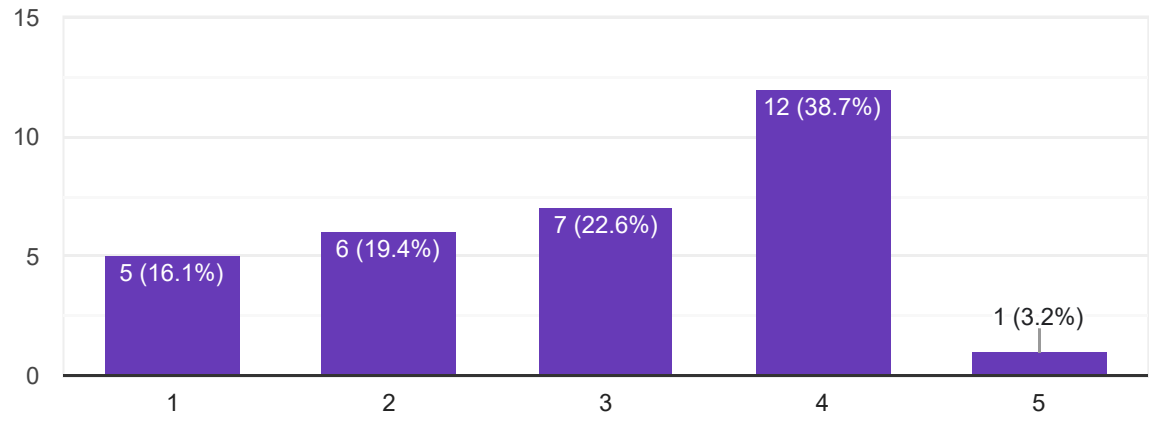
More words associated with the target practice.

Not sure, could be recognising the word COOKIES

Did you find the visualisation easy to interpret?

Copy

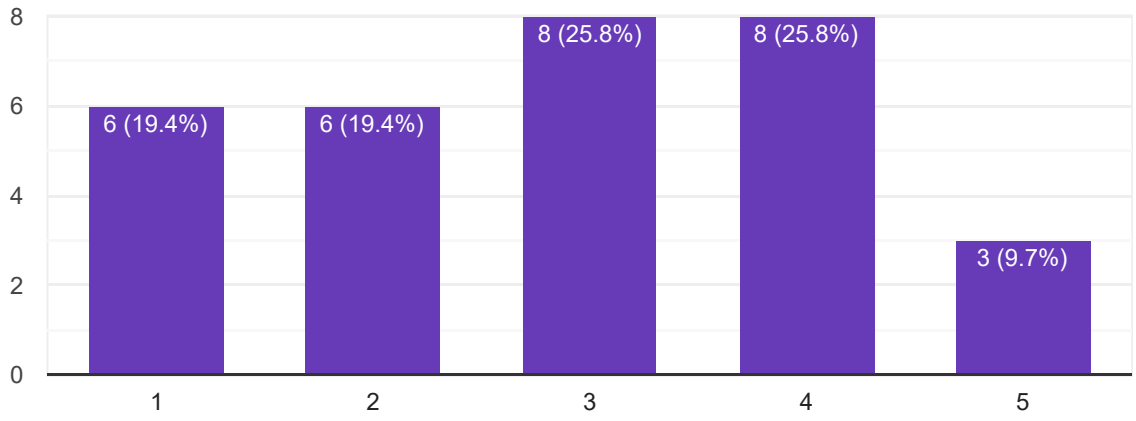
31 responses



Do you understand why the model made the prediction?

Copy

31 responses



Why do you think the model made this prediction?

31 responses

cookies appears

Not sure about why 'and' is identified

The usage of words that would classify it as the actual practice was more than the not actual practice

Lack of strong words that would suggest a negative influence (mostly just connectors like ‘and’), in contrast the presence of the word ‘cookies’.

Cookies were used

SEND HELP

Same as previous answer

Had the word cookies

verbs and conjunctions related to "cookies"

Same answer as above.

cookies is mentioned but I don't get why certain words are tagged as blue

a cookie was used

No clue again sorry

Highlighted words?

The prediction model accorded greater weight to certain orange highlighted words which accurately corresponded to the actual practice.

idk

The model picked up the key words of 2 or more specific sreas

The clause mentions cookies with no mention of sending infromation to other 3rd parties

No idea

Even though there were more words relavant to 3rd party, 1st party words like cookies have a larger waitage and made the model predict 1st party cat instead

not sure

Again, it looks for cookie

Found the right key words

Not sure

cookies

comparatively more words present (taking into account strength of link and not just frequency esp for "cookies") which AI knew to be associated w that cookie

Despite having higher number of words that suggest a "NOT Identifier_Cookie..._1stParty", the word "Cookie" contributed more significantly to the model as compared to the other words, resulting in the eventual (right) prediction

Not sure

Word "cookies" is used several times

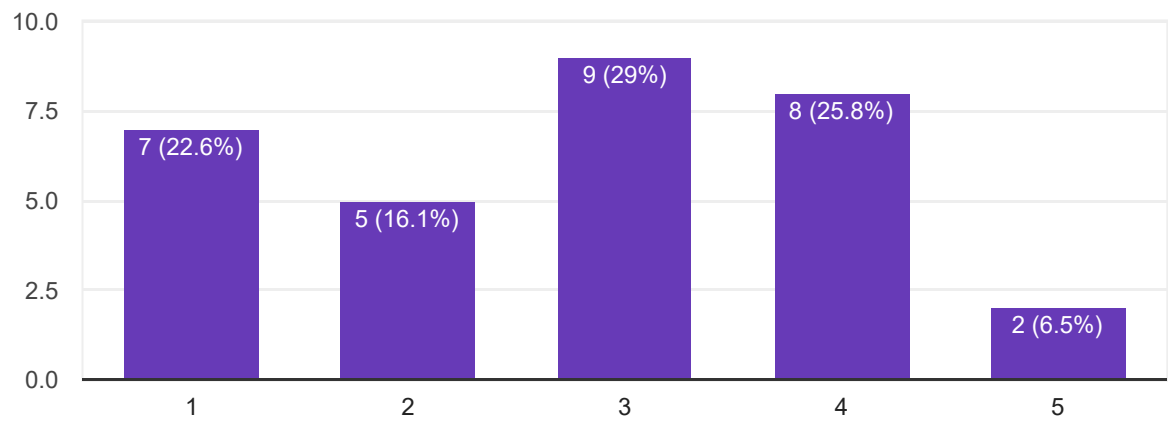
Overall highest score for the target practice, with more words sugessting the target then not.

Not sure, could be recognising certain keywords

Did you find the visualisation easy to interpret?

Copy

31 responses

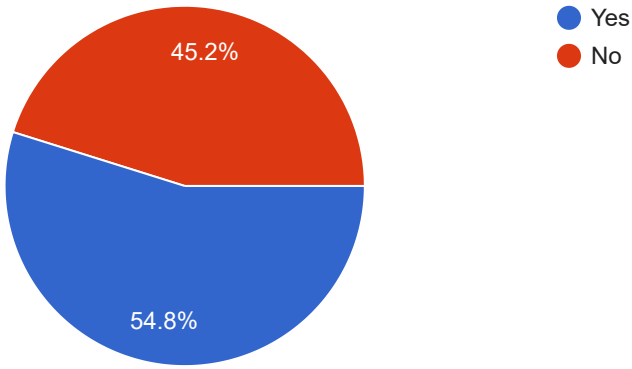


Based on your current understanding, do you think that the sentence below would be predicted to be "Identifier_Cookie_or_Similar_Tech_1stParty"?

"We also use **tracking technologies** to keep records, store your preferences, improve our advertising, and collect Non-Identifying Information, including Device Data and information about your interaction with the Site and our Business Partners' web sites."

Copy

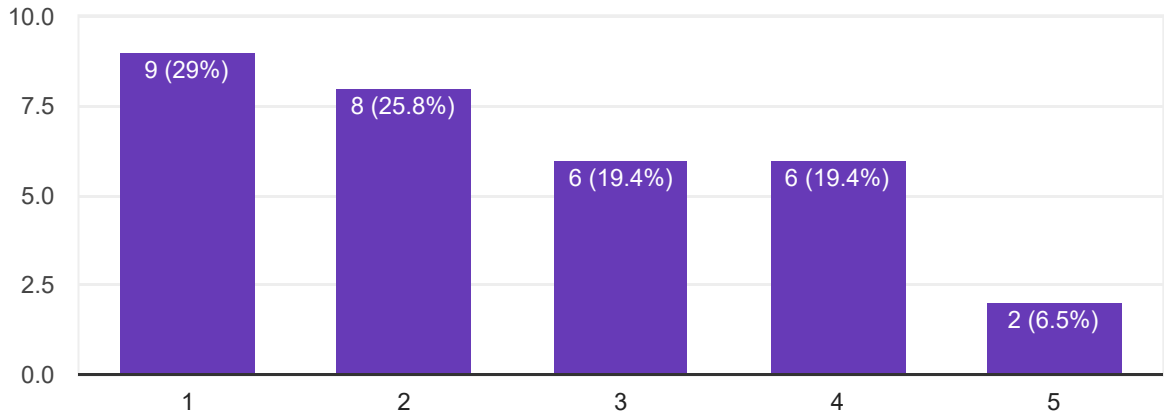
31 responses



Do you understand why the model made the prediction?

Copy

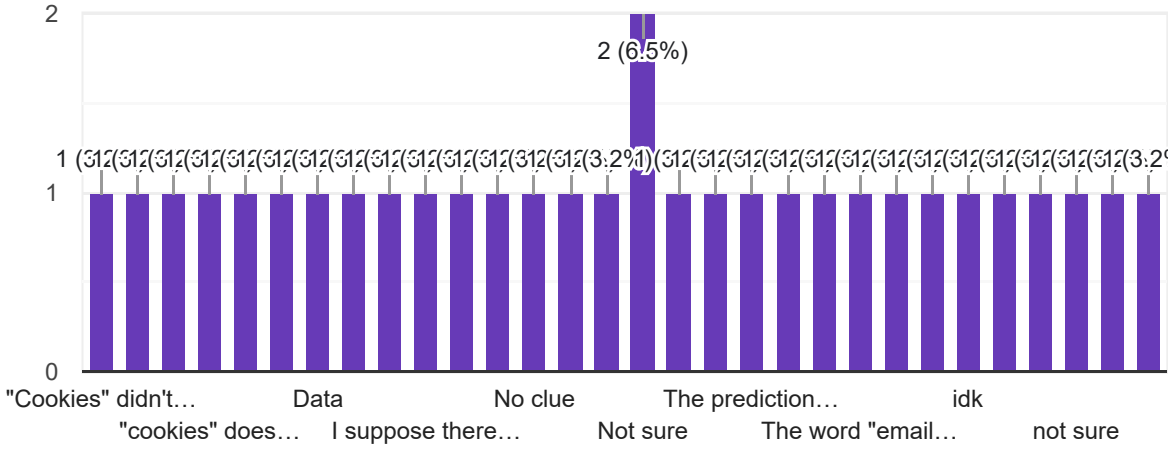
31 responses



Why do you think the model made this prediction?

Copy

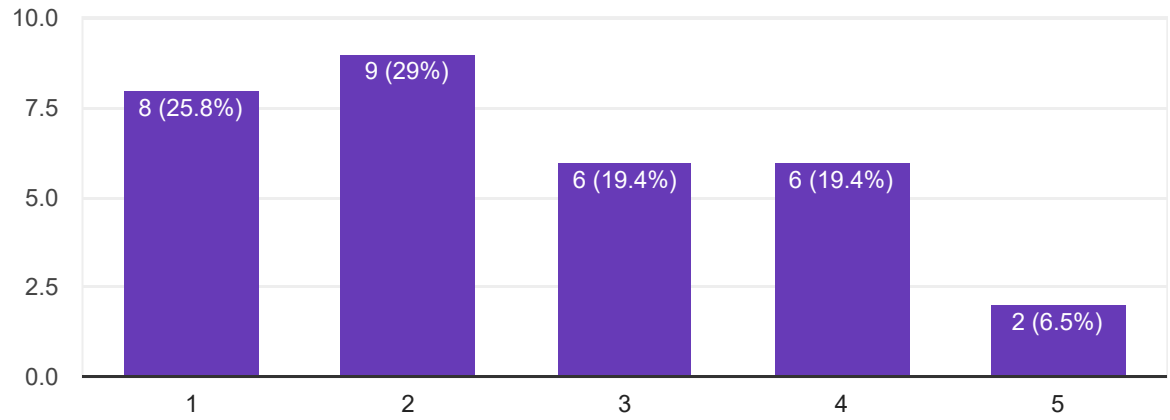
31 responses



Did you find the visualisation easy to interpret?

Copy

31 responses

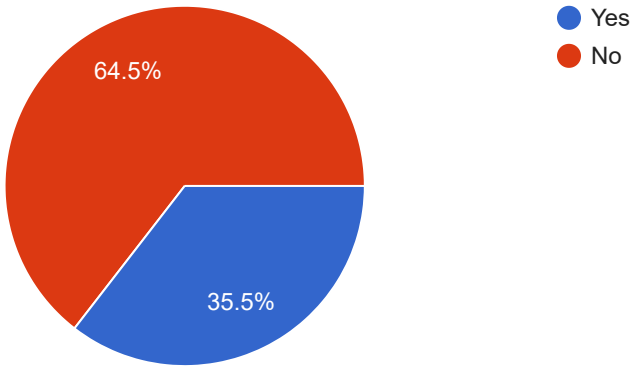


Based on your current understanding, do you think the sentence below would be predicted to be in "Identifier_Cookie_or_Similar_Tech_1stParty?"

Copy

"As explained above, you may either volunteer to us certain information (such as your **phone number**), or we may automatically collect certain information, such as through the use of your mobile device system's permissions, or through the use of cookies or similar tracking technologies."

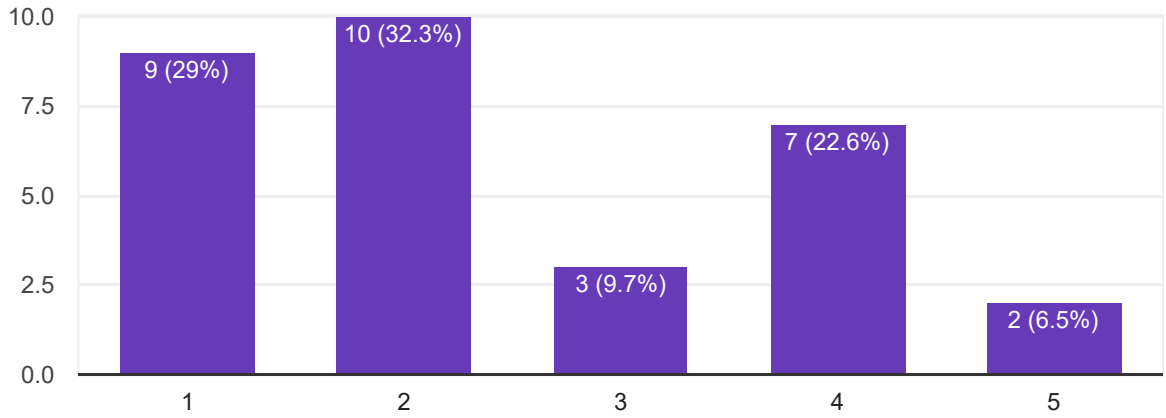
31 responses



Do you understand why the model made the prediction?

Copy

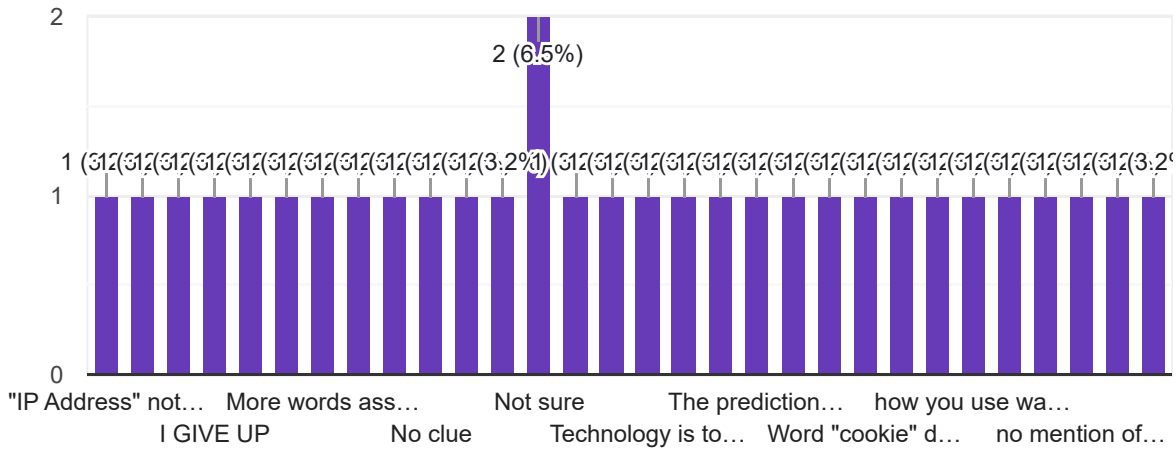
31 responses



Why do you think the model made this prediction?

Copy

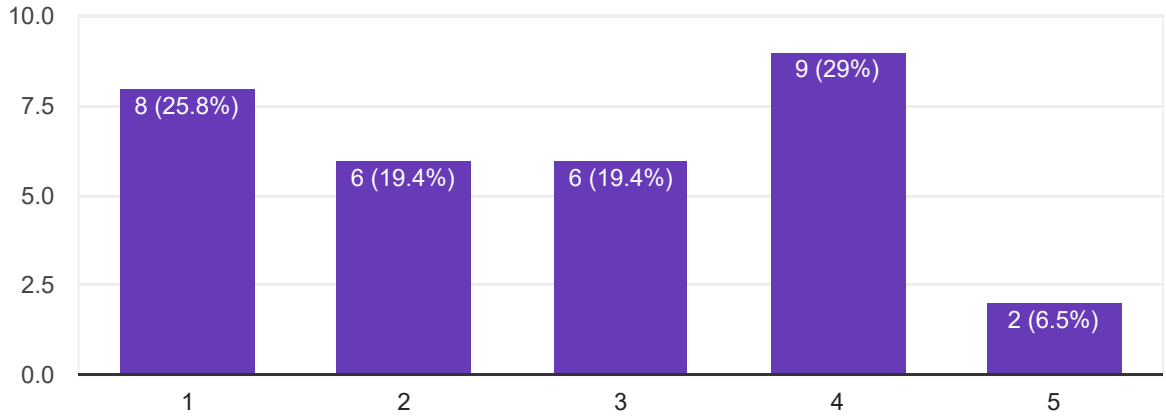
31 responses



Did you find the visualisation easy to interpret?

Copy

31 responses

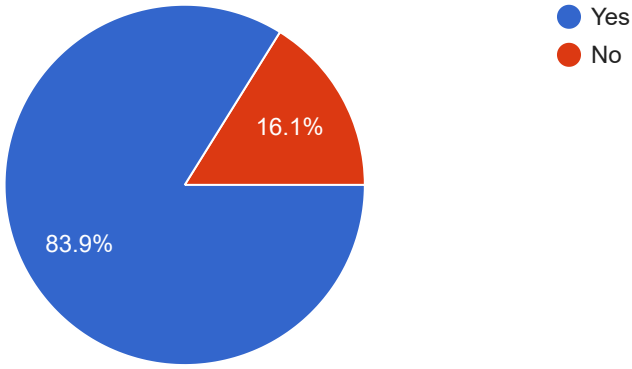


Based on your current understanding, do you think the sentence below would be predicted to be in "Identifier_Cookie_or_Similar_Tech_1stParty?"

"These **cookies and other such tracking technologies** allow the collection of data, such as your device's model, operating system and screen size, the other applications installed on your device, and information about how you use our services."

Copy

31 responses

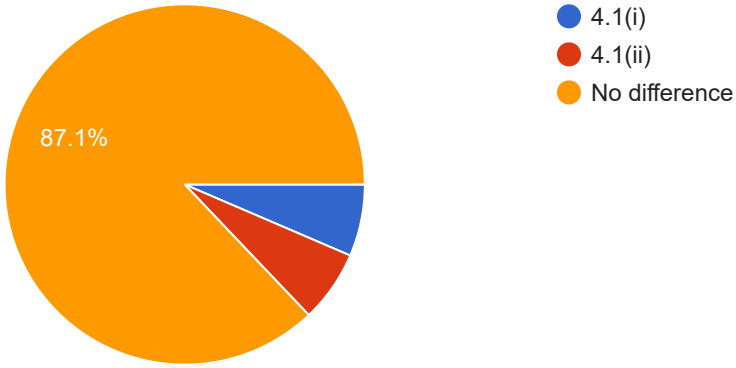


Part 4

Which explanation did you find easier to interpret?

Copy

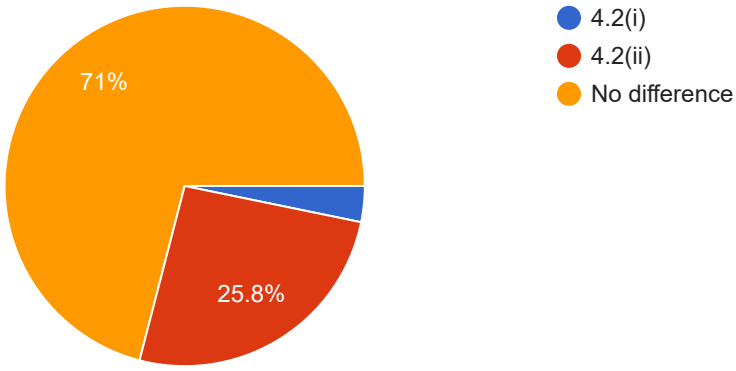
31 responses



Which explanation did you find easier to interpret?

Copy

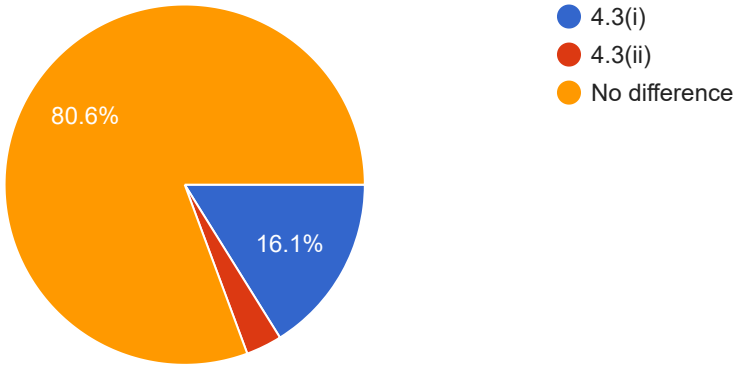
31 responses



Which explanation did you find easier to interpret?

Copy

31 responses

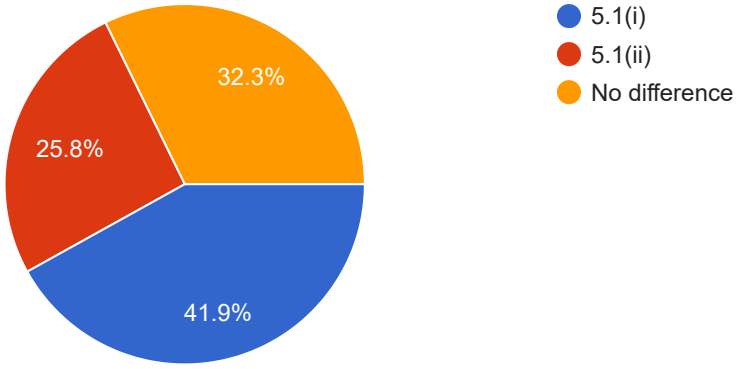


Part 5

Which explanation did you find easier to interpret?

Copy

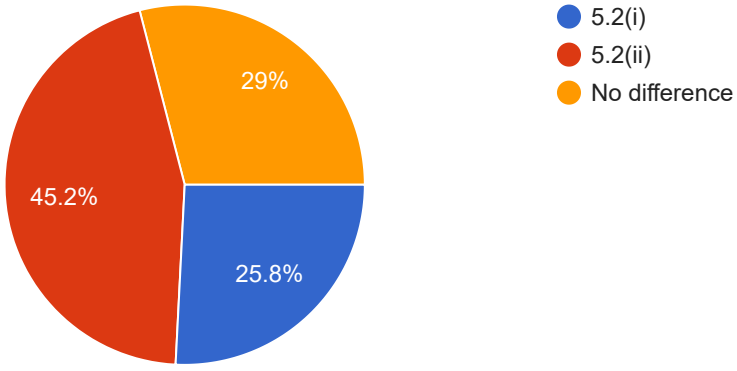
31 responses



Which explanation did you find easier to interpret?

Copy

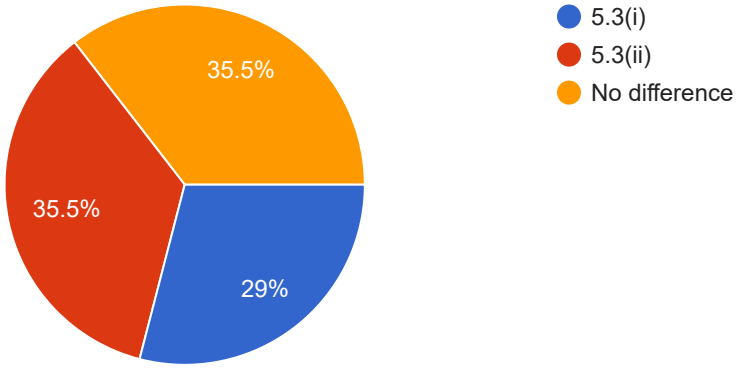
31 responses



Which explanation did you find easier to interpret?

Copy

31 responses



Part 6

Context 1: Imagine that you are an app developer. You are developing an app that uses cookies to track user activity online. To comply with the PDPA, you know that you need to include a sentence in your app's data privacy policy that notifies and asks for users' consent to use cookies.

Since you have no knowledge of the PDPA, you use the abovementioned model to analyse a pre-drafted data privacy policy that you found online. The model informs you that there is a sentence which states that cookies are being used.

You are deciding whether to rely entirely on the model's prediction, or pay costly legal fees to confirm with your friend who is a lawyer.

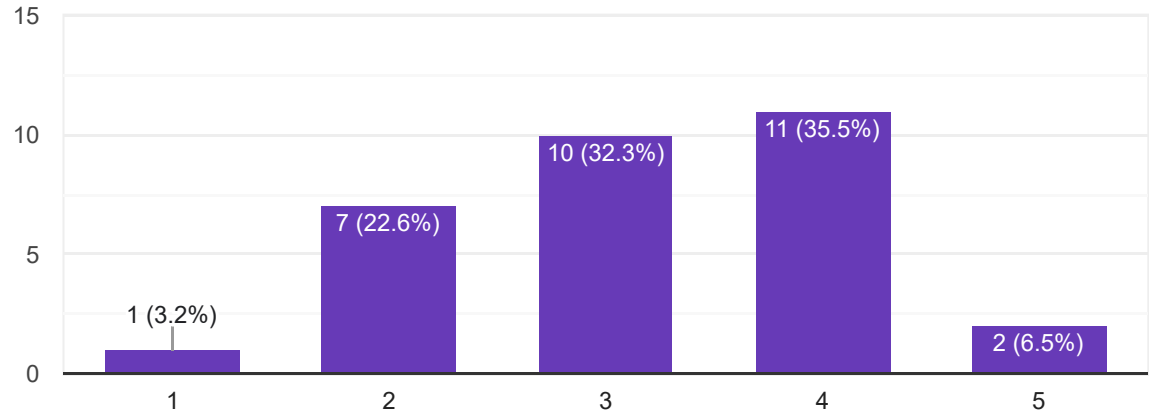
If the pre-drafted data privacy policy actually does not state that cookies are being used but your app uses cookies, you could face a fine of up to $10,000 in breach of the PDPA as you would have failed to notify your users.

How far do you, as the app developer:

Copy

Think that using the model is an effective method of identifying violations of the PDPA?

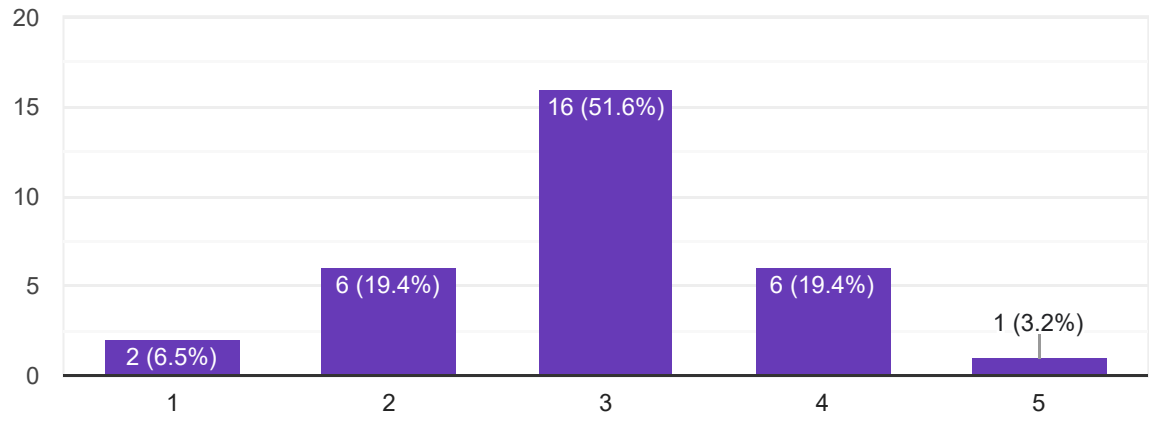
31 responses



Think that using the model is a fair method of identifying violations of the PDPA?

Copy

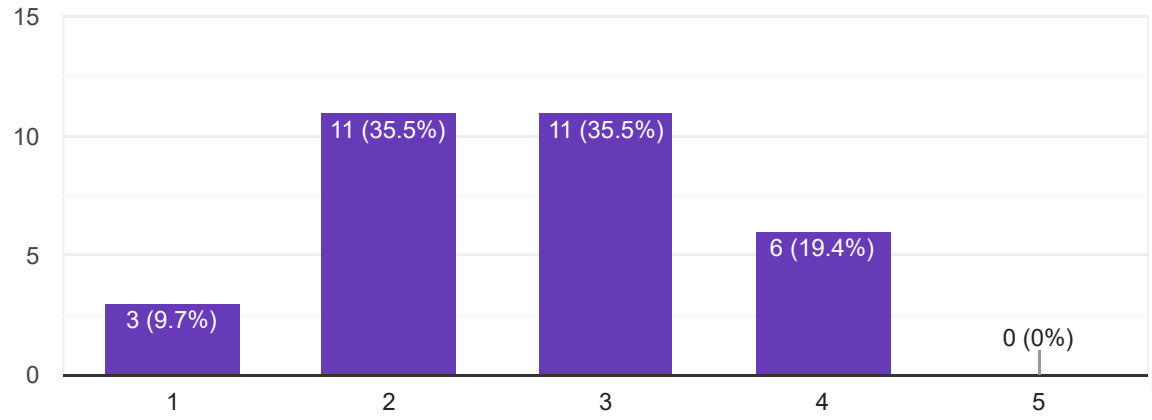
31 responses



Think that using the model is a method that could be a risk to society?

Copy

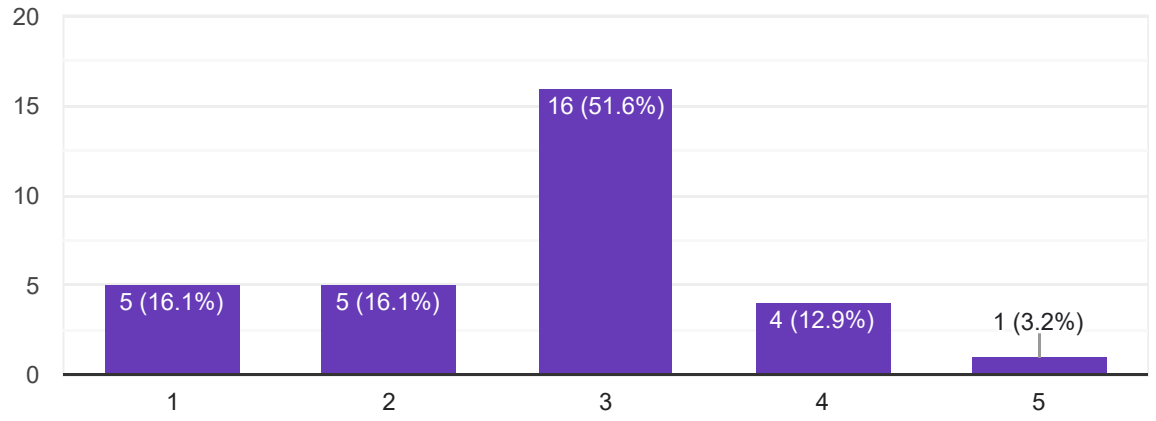
31 responses



Trust the prediction made by the model?

Copy

31 responses



Context 2: Imagine that you are a committee member part of the Personal Data Protection Commission (PDPC). A user of an app has informed you that an app is using cookies but has not notified its users.

Your team checks the code of the app and confirms that the app is indeed using cookies. Your team uses the abovementioned model and the model informs you that the data privacy policy does not contain any sentence that notifies its users that it uses cookies.

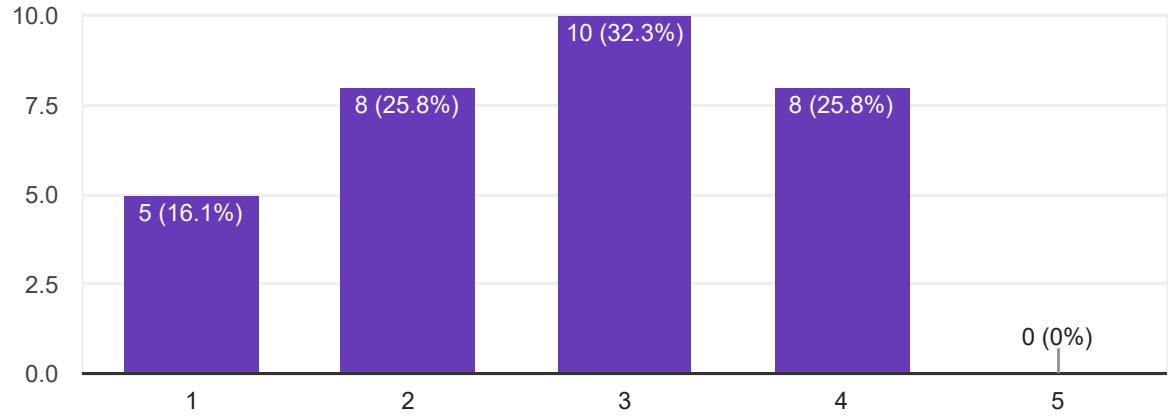
To increase the efficiency of the PDPC, your team is considering whether to adopt the abovementioned model to automate the analysis of data privacy policies. If this new method of analysis is adopted, the PDPC would rely entirely on the model's predictions to confirm whether app developers have breached the PDPA. The app developers would face a fine of up to $10,000 if they are found to have breached the PDPA.

How far would you, as a committee member of the PDPC:

Copy

Think that using the model is an effective method of identifying violations of the PDPA?

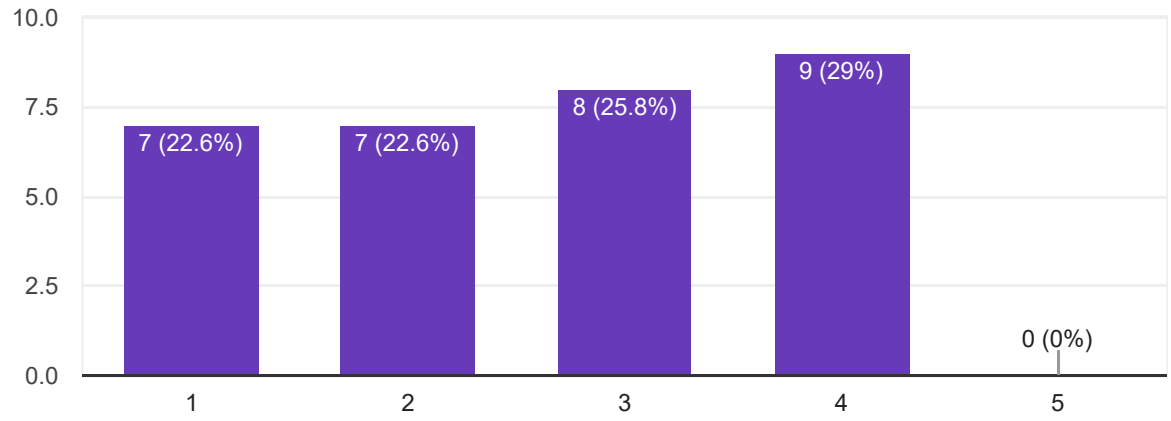
31 responses



Copy

Think that using the model is a fair method of identifying violations of the PDPA?

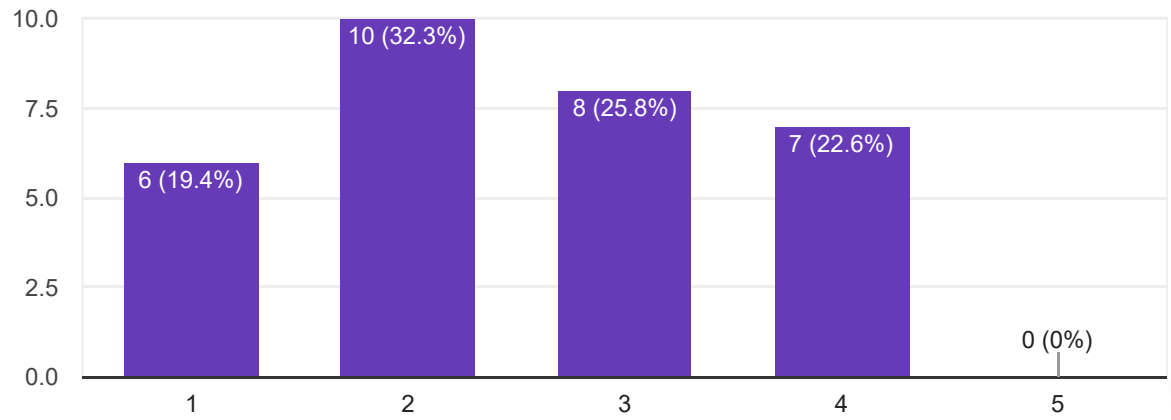
31 responses



Copy

Think that using the model is a method that could be a risk to society?

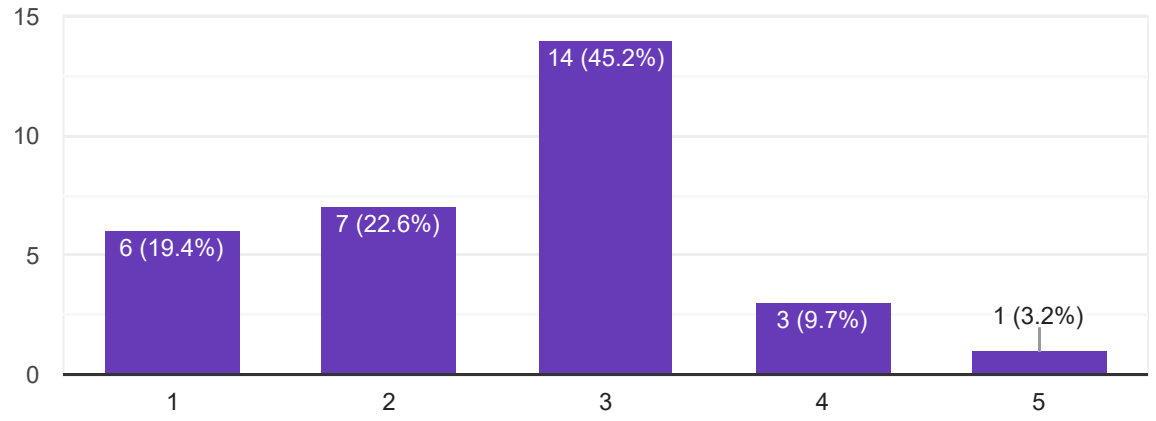
31 responses



Trust the prediction made by the model?

Copy

31 responses



Context 3: Imagine that you are a user of an app. You read in a forum where other users allege that the app uses cookies. You decide to analyse the data privacy policy of the app using the abovementioned model and the model informs you that the data privacy policy does not contain any sentence that notifies its users that it uses cookies.

You are deciding whether to submit this prediction as the only supporting piece of evidence to the PDPC to claim that the app has used cookies without notifying you.

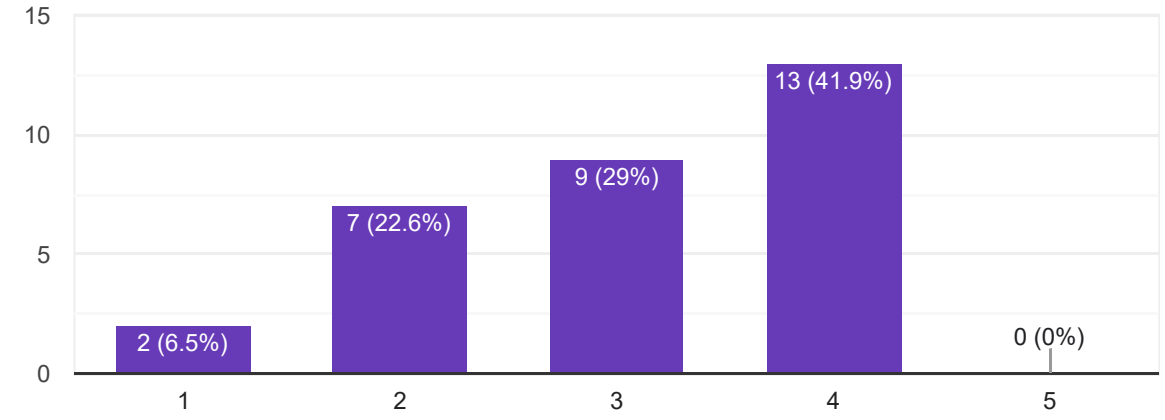
If the PDPC decides that the developer has indeed violated the PDPA, you could claim compensation from the app developer of up to $10,000.

How far would you, as a user of the app:

Copy

Think that using the model is an effective method of identifying violations of the PDPA?

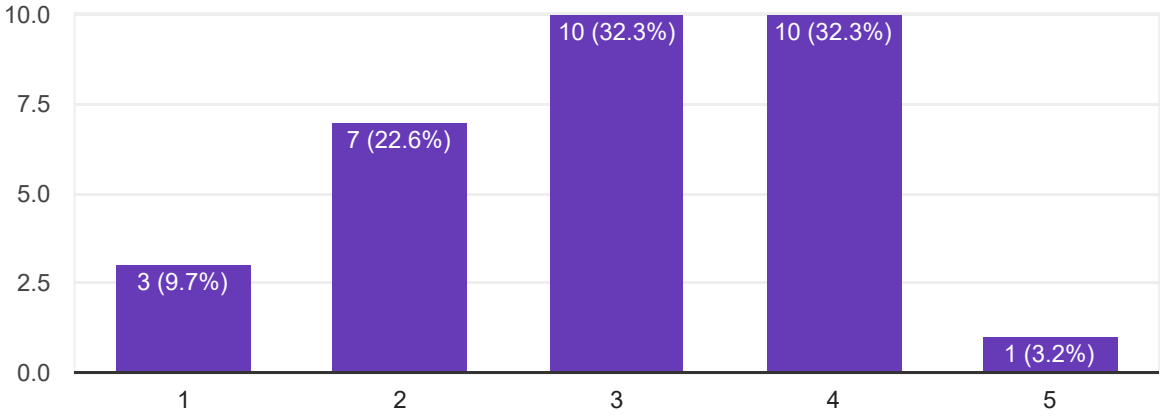
31 responses



Think that using the model is a fair method of identifying violations of the PDPA?

Copy

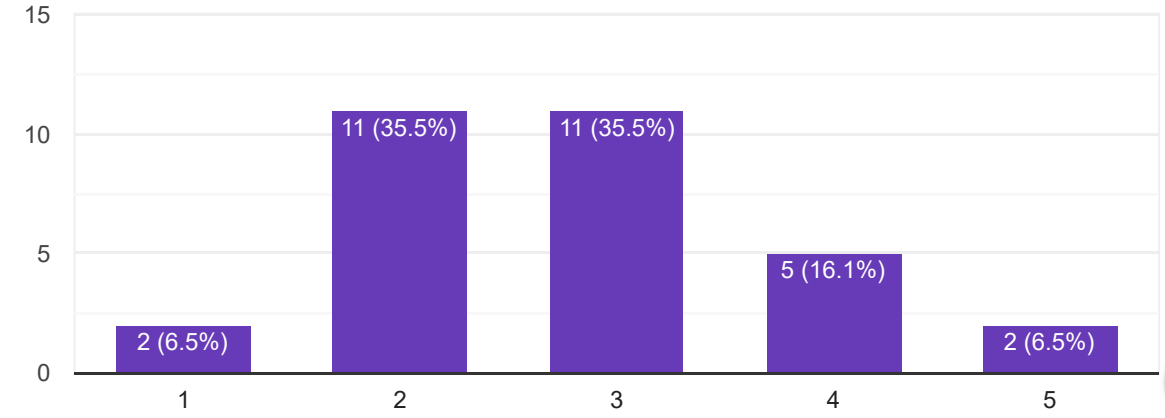
31 responses



Think that using the model is a method that could be a risk to society?

Copy

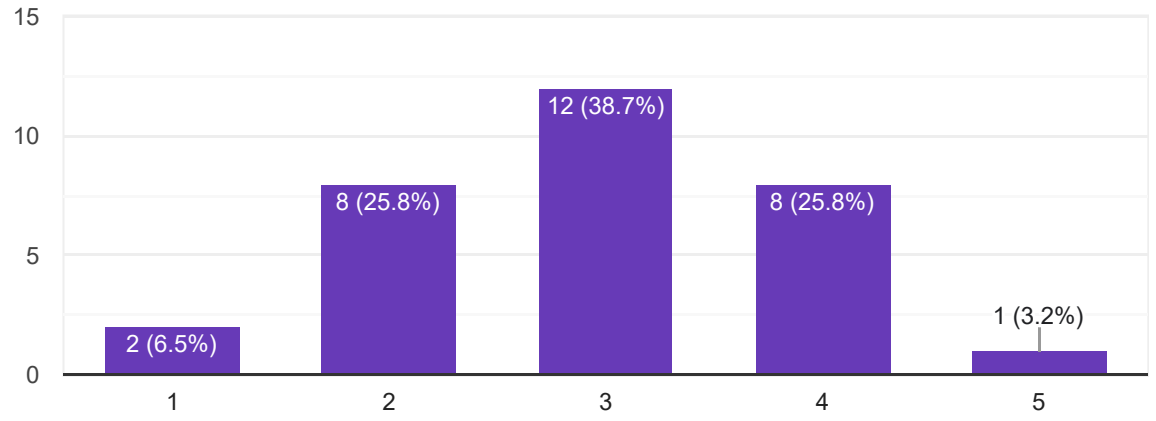
31 responses



Trust the prediction made by the model?

Copy

31 responses



This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms