

Research: Anti-Spam Algorithms Showed Political Bias During 2020 USA Elections

Originally published at <https://www.unite.ai/research-anti-spam-algorithms-showed-political-bias-during-2020-usa-elections/>



According to a new study, the spam filtering algorithms (SFAs) from three of the world's largest major email providers exhibited political bias during the United States' 2020 elections, with Google's Gmail skewing to the left, and Microsoft Outlook and Yahoo Mail favoring emails from right-wing candidates.

The paper states:

'Our [observations] revealed that all SFAs exhibited political biases in the months leading up to the 2020 US elections. Gmail leaned towards the left (Democrats) whereas Outlook and Yahoo leaned towards the right (Republicans). Gmail marked 59.3% more emails from the right candidates as spam compared to the left candidates, whereas Outlook and Yahoo marked 20.4% and 14.2% more emails from left candidates as spam compared to the right candidates, respectively.'

The authors analysis, they claim, demonstrates 'aggregate biases' in SFA activity.

The paper also acknowledges the possibility of 'farmed' spam marking, where actors seeking to silence opposition voices could solicit or obtain access to official communications from 'hostile' parties and affiliations with the intention of reporting the communication as spam, thereby influencing the algorithms that determine the likelihood of spam provenance from a particular sender.

However, the researchers observe, this does not explain the marked variations in the way that different email providers appear to have configured actions based on feedback from end users:

'Arguably, there is also this possibility that the SFAs of email services learnt from the choices of some voters marking certain campaign emails as spam and started marking those/similar campaign emails as spam for other voters. While we have no reason to believe that there were deliberate attempts from these email

services to create these biases to influence the voters, the fact remains there that their SFAs have learnt to mark more emails from one political affiliation as spam compared to the other.

'As these prominent email services are actively used by a sizable chunk of voting population and as many of the voters today rely on the information they see (or don't see) online, such biases may have an unignorable impact on the outcomes of an election.'

The [paper](#) is titled *A Peek into the Political Biases in Email Spam Filtering Algorithms During US Election 2020*, and comes from four researchers at the Department of Computer Science at North Carolina State University.

Round the Houses

The researchers' study covers a period of five months from July 2020 through to the end of November the same year, wherein they created 102 new email addresses across the three email platforms, and subscribed to two presidential, 78 senate and 156 house candidates email notification lists.

To discount demographic factors, the email accounts were created with varying demographic factors for each (fictitious) end user, and split across two strands: the first studied general bias trends in spam filtering algorithms across all the combined email services for presidential, House and Senate candidates; and the second examined the ways that various email interactions (such as marking or unmarking as spam by the end user) appeared to impact the behavior of algorithmic spam filters.

Several key observations came into focus during the study. The authors report that Gmail 'leaned towards the left', while Outlook and Yahoo leaned to the right. Yahoo retained 55.2% of all the political emails in the user's inbox, while Outlook filtered 71.8% of emails from political candidates of all stripes.

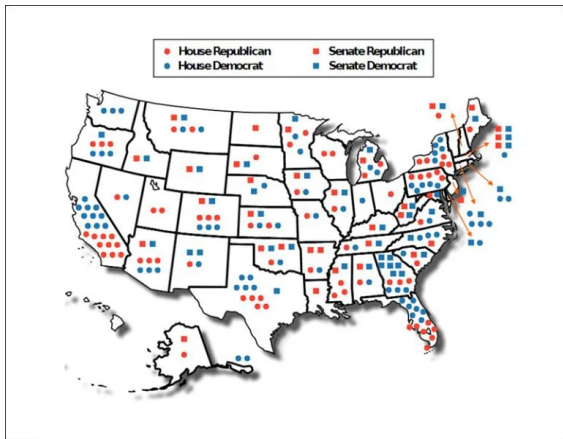
'Gmail, however, retained the majority of left-wing candidate emails in inbox (< 10.12% marked as spam) while [sending] the majority of right-wing candidate emails to the spam folder (up to 77.2% marked as spam).'

'We further observed that the percentage of emails marked by Gmail as spam from the right-wing candidates grew steadily as the election date approached while the percentage of emails marked as spam from the left-wing candidates remained about the same.'

Picking Candidates

While presidential candidates subscribed to for the study were limited to Joe Biden and Donald Trump, the researchers took care to make representative choices when considering subscribing to email communications from Senate and House candidates, for a number of reasons.

Firstly, states have varying numbers of seats in the House, based on the state's population count. Secondly, the number of Senate and House candidates across the two main political parties varies across states. Further, certain candidates were represented only by official.gov websites, which are statutorily prohibited from sending campaign emails; and finally, some of the candidates' subscription lists were protected by CAPTCHAs, which could not be automated by the researchers' custom data gathering framework.



Distribution of political affiliation of email subscriptions from Senate and House candidates.

Source: <https://arxiv.org/pdf/2203.16743.pdf>

To equalize the resulting imbalance between Democrat and Republican candidates, the researchers subscribed to campaign email information from the maximum number of candidates in any state where left and right candidates were equal in number, except in states such as Alaska, which only had one Republican Senate candidate.

In total, the authors had to account fairly for 11 such states, and finally ended up with all 50 states represented. 78 of the subscriptions across 36 states amounted to 44 Democrat and 34 Republican Senate candidate lists, while there were 156 subscriptions across 42 states for House candidates – 81 Democrats, and 75 Republicans.

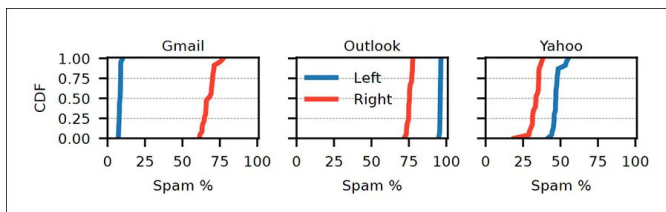
Analyzing the Data

The researchers collected 318,108 emails across the three email services in the study's active data gathering period which was truncated after November 20th due to the rapid fall-off in volume of mails after that date. Data content collected for each email included *MIME-Version*, *Content Type*, *Subject*, *From*, *To*, *Date*, *Message-ID*, *Delivered-To*, *Received-SPF*, and *Received-By*.

Due to the challenges involved in representing both political party communications fairly, Propensity Score Analysis (PSA) was chosen as the statistical method for the data. PSA generates *covariates* from unbalanced data that equalizes distributions in exceptional circumstances where control groups and traditional statistical splits are not easily applicable.

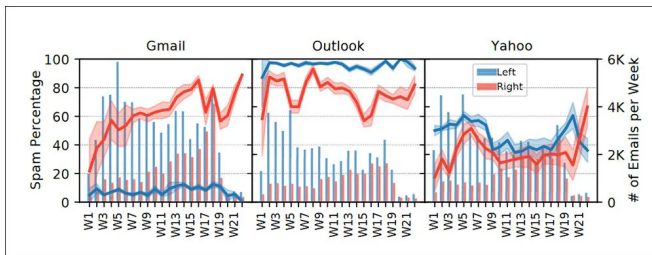
The authors conclude that SFAs for the email services studied exhibit political bias, and that early relative consistency across the services diverges into rather more specific behavior over time.

Gmail marks a higher percentage (67.6%) of right-wing political emails as spam, compared to a mere 8.2% of left-affiliated emails, but responds more dynamically to user interactions that unmark mails as spam than its cohorts. Outlook, instead, marks 95.8% of left-leaning political emails as spam, compared to 75.4% for right-wing emails, and Yahoo marks 14.2% more left-wing emails as spam than right-wing emails.



Cumulative distribution of the percentage of Democrat (blue) and Republican (red) emails that were marked as spam in email accounts of each service.

Furthermore, the results suggest that over the course of the period of the study, Gmail responds quite generically to an increased volume of emails across all political affiliations by increasingly marking them as spam, regardless of provenance. Yahoo consistently reported left-leaning emails as spam as the campaigns progressed, at the same time decreasing the number of right-wing mails marked as spam. Outlook seemed least affected by increasing volume of mails from either political party, maintaining a general right-wing bias.



Percentage of emails marked as spam across both political parties and all three email providers over the 153 days of the study period.

Response to User Interaction

When we mark a spam email as ‘Not spam’, the intention is to train the email system not to flag similar mails in future, though the underlying type of rule (email-based, content-based, etc.) is not always entirely clear.

The study’s results found that of the three email providers examined, only Gmail responded notably to a ‘not spam’ input from the user. By contrast, this user-driven spam-to-inbox ($S \rightarrow I$) interaction had very limited long-term effect in Outlook and Yahoo.

The researchers observe:

‘[Due] to the $S \rightarrow I$ interaction, the political bias in Gmail reduced significantly. However, unexpectedly, it increased in both Outlook and Yahoo because neither of the two services reacted noticeably to user’s desire to not mark the emails as spam that the two services were marking as spam.’

Conclusion

The authors conclude that Gmail responds ‘significantly’ to user interaction in comparison to Outlook and Yahoo, despite its own left-leaning predisposition.

The authors state:

‘While the political bias in Gmail stayed unchanged after the reading interaction, it decreased significantly due to the $I \rightarrow S$ and $S \rightarrow I$ interactions.’

And continue:

‘While the political biases changed in response to various interactions, Gmail maintained its left leaning while Outlook and Yahoo maintained their right leaning in all scenarios.’

The researchers acknowledge a general expectation by the end user that spam filters can and will adapt their behavior based on user intervention (such as moving an email from a spam folder to the inbox, or marking a mail as ‘not spam’), but that this mechanism is not reliable, and certainly is not consistent across the three email providers studied.

The paper notes:

‘[We] did not find any consistent actions that one could recommend to users to help them reduce the bias in the way the SFA treats political emails that are sent to them.’

First published 4th April 2022.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More