

AMENDED Exhibit 1170

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

Case No.: 4:22-md-03047-YGR

MDL No. 3047

In Re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation

Using Age for Growth and Trust & Safety Scenarios

██████████
July 2022

Background

Age is an important signal for the Growth team for a variety of scenarios. However, we observe that users tend to lie about their ages (e.g., younger users claiming to be older, users choosing the default MM/DD when going through registration) similar to other internet services. This is problematic for multiple reasons: circumventing protections we've put into place for minors, Ads targeting for restricted categories such as alcohol, inaccurate internal analytics, etc.

The Data Science team has done deep dives into better [Understanding Age Trends and Aging Up](#) and [Age DAU reporting analysis / discrepancies](#). A few of the issues identified about using the [\[February 2019\] Inferred Age](#):

- The Growth team was using an inferred age variable that is not the recommended one
 - We used only the point estimate from the model, not the output of the entire pipeline which applies heuristics that the Ads team has come up with.
 - Search for “3. Review of age variable options” in [this doc](#).
- There is a [period of strange behavior](#) between 3/19 and 3/29 (which might be due to improper model training or scoring). After the strange period, we see a material jump in match rates for 21-24 (green) and 25-34 (red) age buckets that affects our DAU % change reporting.

Scenarios

In addition to the prior work above, there are a few scenarios that the Growth and Trust & Safety teams wish to or already using an inferred Age for.

| N2

Protecting Minors from surfacing in QuickAdd and Search

One protection that Snapchat has for minors is to prevent users who are under 18 from being shown up as recommendations in QuickAdd and in Search results unless the searcher and underage user ██████████ (see [Filter Underage Candidates in QA Launch Report](#)) in the ██████████. However, this protection relies on knowing the age of the user reliably and if an underage user lies about their age, then the protection can be bypassed.

Page 1 Comments

N1 Check with [REDACTED] about the company ASL for getting data from under 18
year users @ [REDACTED]
Assigned to [REDACTED]

8/10/2022 12:27 PM

N2 also chat with Claudia for Reg experiments

8/10/2022 12:29 PM

Unfortunately, [child sexual grooming](#) is a known problem on the platform. We have an [\[A/C Priv\] Anti-Grooming Tactical Plan](#) that aims to tackle this problem. However even with this protection in place, we are still not able to handle the cases like this one: [\[a/c priv\] notes on ZD #227034538](#)

- The victim's reported birthday is Redacted-PI 2000 (22 years old)
- The inferred age is $<>^{N3}$

Some ideas we'd like to explore:

- Increase recall of <18 or ≥ 18 years old prediction
 - Use existing model with a different threshold
 - Evaluate and/or train a model that formulates as a binary class problem
 - Ok with FPs since cost of FNs are much higher
- Use a confidence interval instead of the point estimate?
- Use the classification model output for 13-17 age bucket instead of the Regression model point estimate?
- Be less conservative for Growth scenarios when there is a lack of confidence in model prediction (e.g., similar likelihoods for multiple classes)
 - Instead of relying on reported age, Growth would still use model output

With the ideas above, we can directly use those to apply the existing protections. We can also make the protections more sophisticated and dependent on the model output (e.g., use different mutual friend count thresholds depending on the confidence of the model).

Analytics and Reporting

For Growth reporting, we care about age at the aggregate level instead of user-level predictions (like we do for T&S). Almost all of our reporting and analyses are done at the age bucket level, however as we've seen [here](#) that age-in and age-out matters a lot when examining engagement trends so point estimates (or mobility between them) are important. The Growth team has already started analyzing this and has some methods and recommendations to handle this during analysis.

One thing we would like is better predictions for very suspicious users who lie about their age (e.g., keep default age during Registration, those with birthdays in 2000) if the Ads team has some insights here.

However, we think that building, evaluating, and maintaining a model for point estimates of age for reporting purposes over time is quite expensive and with limited return.

Page 2 Comments

N3 [REDACTED] can you share the inferred age for this case that we talked about?
[REDACTED] 7/22/2022 03:46 PM

Using as a feature for QuickAdd and Search models

Today both QuickAdd and Search use the reported age and age bucket (and an inferred age bucket from UPSv2 in Registration Friending ranking) as features into different rankers. If there are any changes to the age model or output, we can experiment with using it and seeing if we can get gains.

Using it to hide that a user belongs to a High school community

We'd like to protect users from exposing that they belong to a High school community to 18+ y.o. Even if they're their friends.

Spam Detection

This is an area where we still need to get more clarity on the Growth side on how this signal is consumed by the Spam & Abuse team and the behaviors of spammers when it comes to reported ages.

This is another scenario where understanding and having better predictions for suspicious users (e.g., keep default age during Registration, those with birthdays in 2000) would be useful. If the Ads team has some insights here, would love to learn.

One observation we've noticed is that spammers often have the same birthday with little to no randomization. One hypothesis for this is that they use a 3rd party app for setup that fixes the date.

Age Up (25+ year olds)

Another focus area of the Growth team is Age Up efforts: to help the Snapchat user base grow for those aged 25 and older. In the future, if we have any marketing campaigns (or other targeting) to this user cohort, then the age model can be used here. However, similar to the protecting minors scenario, the exact age prediction is less important than the binary classification of whether the user is 25+ or not.

[Ignore] Working Notes

Background:

- [Understanding Age Trends and Aging Up](#)
- [Age DAU reporting analysis / discrepancies](#)
- [\[February 2019\] Inferred Age](#)

Scenarios

- Detecting minors and applying protections
- Analytics and Reporting (incl. AB and Husky breakdowns)
 - [David Boyle](#) conversation
 - Tag [REDACTED] on this one
 - For a given user it doesn't have that much stability. Longitudinal studies show the age value jump around a bunch, need something more stable
 - When model updates, don't retain a history of prior age inferences so they jump around a bunch
 - [REDACTED] conversation
 - For growth reporting, we care about aggregates not UU level predictions
 - [REDACTED] thinks it's lower priority for now, since building, evaluating, and maintaining a model over time is expensive
 - For growth reporting, it's ok to use reported age since we look at user aggregates over time
 - Reporting is done at age bucket level, but age-in/out is extremely important and a thorny problem
 - Chat with [David Levenson](#) about this
 - For very suspicious users who lie about their age (e.g. default age during Reg) do we have a way to better understand ages here
 - Bigger problem is understanding age-in and age-out rather than knowing exact age
 - We can handle this by doing cohort analysis where we fix the user's age
- Using as a feature for models (Friending, Search)
- Notifications/campaign targeting (?)
 - Not a priority
- Spam detection
 - [David Boyle](#) conversation
 - Chat with [REDACTED] about this, a lot of spammers set the exact same age (don't randomize the age very much)
 - [REDACTED] conversation
 - Similar scenario to: For very suspicious users who lie about their age (e.g. default age during Reg) do we have a way to better understand ages here
 - Don't know exactly how S&A team is dealing with this issue currently
 - Lot of users with birthday in 2000

- We also don't know why for certain users set the birthday as they do (one hypothesis is they use a 3rd party app for setup)
- Age up
 - If we build a model for Growth, then this could also benefit
 - Don't care about exact age here, but about whether or not about 25 or up
 - ■■■ heard that marketing team is doing some marketing campaigns for certain age users (check with ■■■■)