Baseline S-RAD global model

NIGH GOLDENBERG RASO & VAUGHN

EXHIBIT
1822
07/23/25 CB

Ben Marchi
2023/04/03

ATTORNEY-CLIENT PRIVILEGED

Trial Exhibit No.
P-01681

Uber

Global S-RAD

- Aims
  - To establish a single S-RAD model capable of evaluating risks globally with at least comparable offline performance to current regional production models
  - To achieve performance using a subset of features likely to be approved for use by Regional Ops and PLC partners
  - To see evidence of sustained offline performance through time
    - *Still an open question*
  - To migrate model training pipelines to new Michelangelo platforms

Features - What is in and what is out

- Trip context features
  - Time of day
  - Day of week
  - City specific incident rates
- Rider features
  - Historic trip characteristics (cancel rate, weekend rate, etc.)
  - Star rating feedback
  - Incident rates
- Driver features
  - Historic trip characteristics (Trip count, weekend rate, etc.)
  - Incident rates

Included Features

- Driver/Rider gender
- Driver features
  - Cancellation history
    - Cancelled on rate
    - Cancellation rate
  - Star rating feedback
    - 1/2 star ratings given
    - 1/2 star ratings received

Excluded Features

Modeling challenges

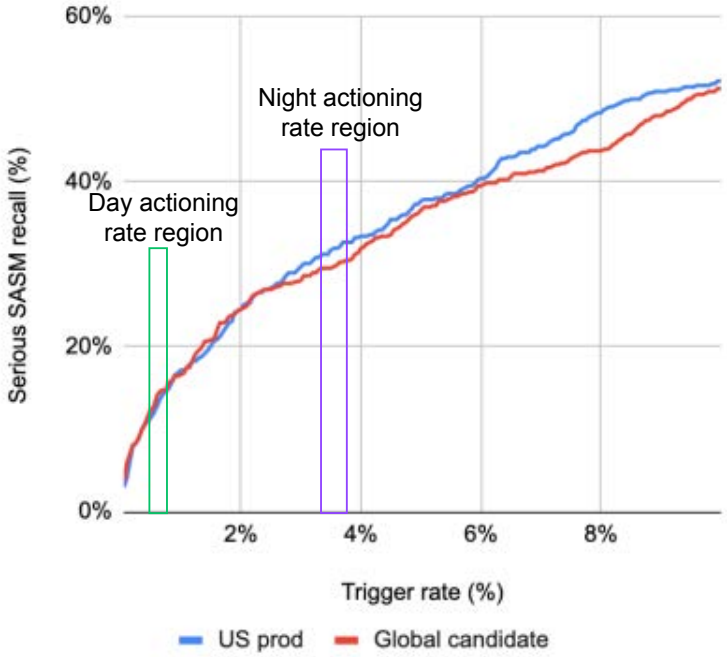
- Creating/verifying globally robust modeling features
- Global differences in incident data quality and reporting behaviors
- Global data volume
  - Training data volume increased from ~60M to ~136M observations
- Managing regional performance tradeoffs
- Maintaining offline performance with a reduced feature set

Regional performance comparison - Overall

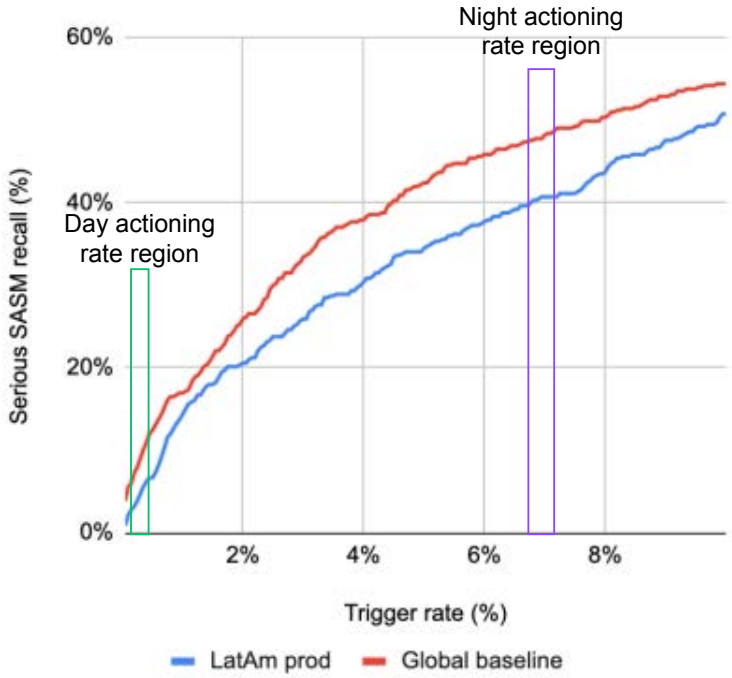
Region	Trigger rate	Prod. recall †	Global cand. recall	% difference
United States*	1.25%	18.3%	19.2%	5.1%
Brazil**	1.10%	15.6%	17.3%	11.0%
LatAm - Mexico**	1.10%	12.8%	16.6%	29.0%
LatAm - ACAC**	1.10%	11.5%	15.0%	30.2%
LatAm - South Cone**	1.10%	10.8%	14.6%	34.8%

*US production model used for comparison
**LatAm production model used for comparison
† Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Time of day actioning also impacts expected performance



United States



Brazil

Regional performance comparison - Overall

Region	Trigger rate	Prod. recall †	Global cand. recall	% difference
ANZ*	1.25%	10.8%	15.1%	40.1%
United Kingdom*	1.25%	16.5%	14.7%	-10.6%
EMEA - NEE*,**	1.25%	14.1%	22.1%	57.1%
EMEA - MEA*	1.25%	11.7%	17.0%	45.5%
EMEA - WSE*	1.25%	8.2%	11.8%	42.9%
India SA*×	1.25%	10.1%	13.8%	35.7%

*US production model used for comparison
**Excludes United Kingdom
×Includes India, Bangladesh, and Sri Lanka
†Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Regional performance comparison - TOD

Region	Time of day	Trigger rate	Prod. recall †	Global recall	% difference
United States*	Day	0.8%	14.8%	13.3%	-10.2%
	Night	3.5%	23.1%	23.6%	2.0%
Brazil**	Day	0.5%	8.0%	11.7%	46.2%
	Night	7.0%	21.7%	31.5%	45.2%
LatAm - Mexico**	Day	0.5%	3.4%	10.6%	208%
	Night	7.0%	29.8%	28.6%	-4.1%
LatAm - ACAC**	Day	0.5%	4.8%	8.3%	75%
	Night	7.0%	25.4%	32.0%	25.8%
LatAm - South Cone**	Day	0.5%	7.0%	8.7%	25.0%
	Night	7.0%	23.5%	29.6%	26.1%

*US production model used for comparison
**LatAm production model used for comparison
† Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Regional performance comparison - TOD

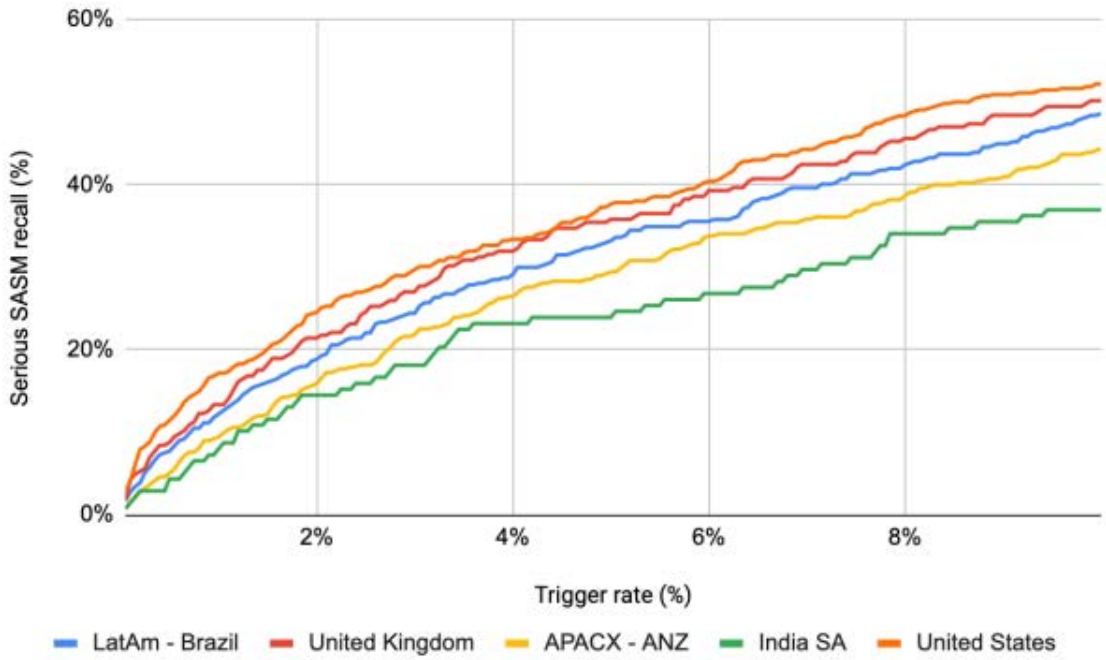
Region	Time of day	Trigger rate	Prod. recall †	Global recall	% difference
ANZ*	Day	0.8%	12.7%	13.2%	3.8%
	Night	3.5%	13.4%	18.2%	35.5%
United Kingdom*	Day	0.8%	18.3%	18.2%	-0.8%
	Night	3.5%	15.8%	14.5%	-7.7%
EMEA - NEE*,**	Day	0.8%	14.3%	23.2%	62.5%
	Night	3.5%	15.1%	24.7%	64.3%
EMEA - MEA*	Day	0.8%	6.7%	12.3%	83.3%
	Night	3.5%	19.4%	25.0%	28.6%
EMEA - WSE*	Day	0.8%	12.9%	14.3%	11.1%
	Night	3.5%	13.0%	12.0%	-7.7%
India SA*×	Day	0.8%	5.0%	6.9%	40.0%
	Night	3.5%	13.5%	27.0%	100%

*US production model used for comparison
**Excludes United Kingdom
×Includes India, Bangladesh, and Sri Lanka
†Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Regional composite performance

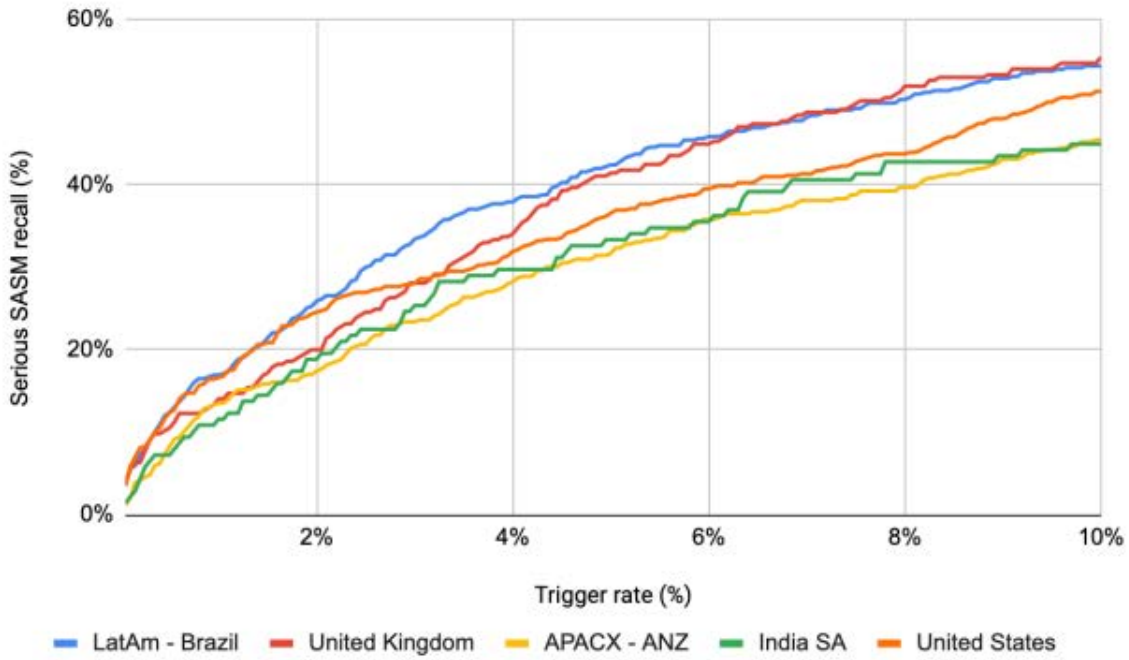
Region	Trigger rate	Prod. recall †	Global cand. recall	% difference
United States*	1.18%	18.3%	17.3%	-5.1%
Brazil**	0.97%	12.2%	17.8%	45.6%
LatAm - Mexico**	1.32%	13.8%	17.7%	27.9%
LatAm - ACAC**	1.23%	11.5%	16.0%	39.5%
LatAm - South Cone**	1.72%	14.6%	18.3%	25.8%
ANZ*	1.36%	15.6%	15.9%	1.5%
United Kingdom*	1.37%	16.1%	16.1%	0.0%
EMEA - NEE*,**	1.37%	20.8%	24.2%	16.1%
EMEA - MEA*	1.10%	12.0%	16.0%	33.3%
EMEA - WSE*	1.50%	12.4%	12.9%	4.8%
India SA*×	1.11%	10.9%	12.3%	13.3%

Existing S-RAD US



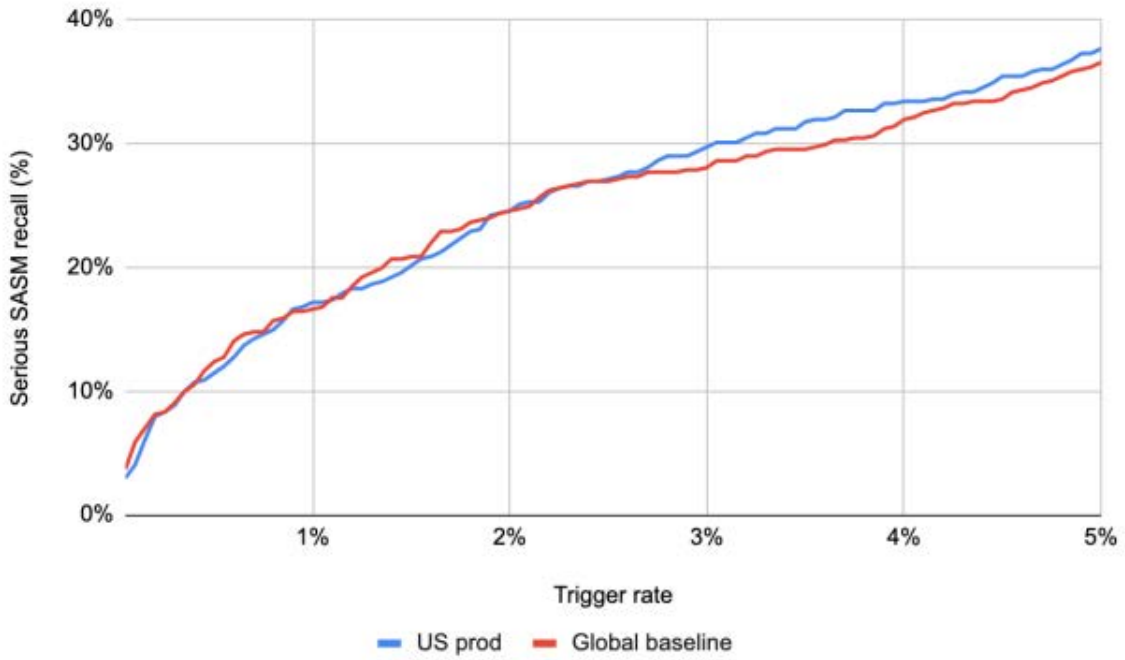
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Baseline candidate global model



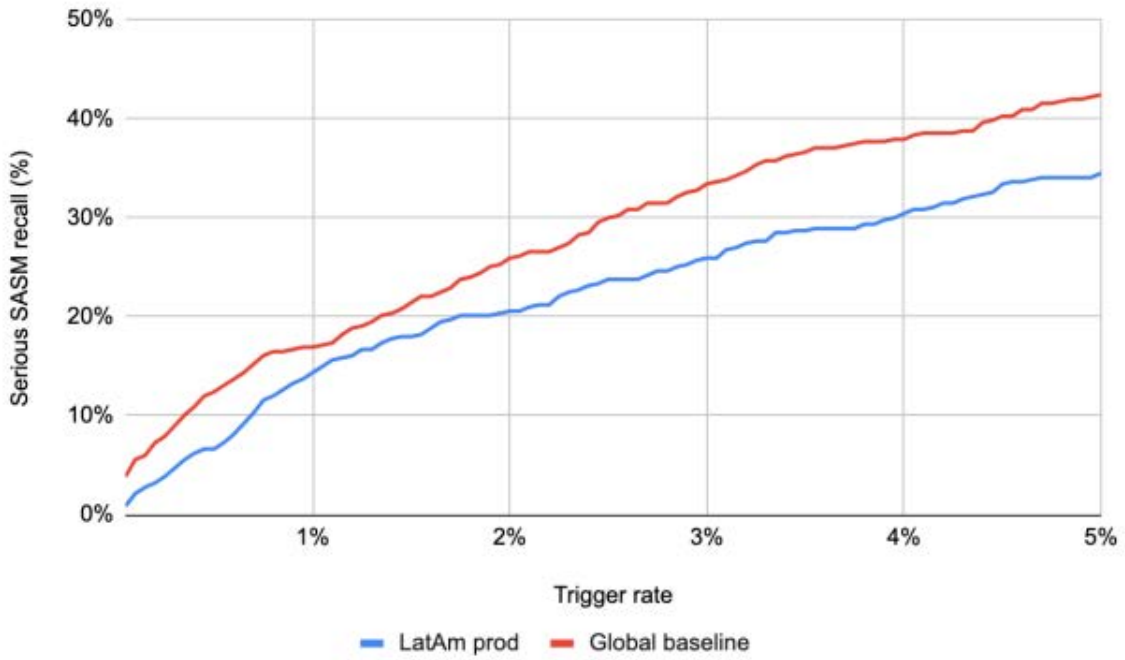
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

United States comparison



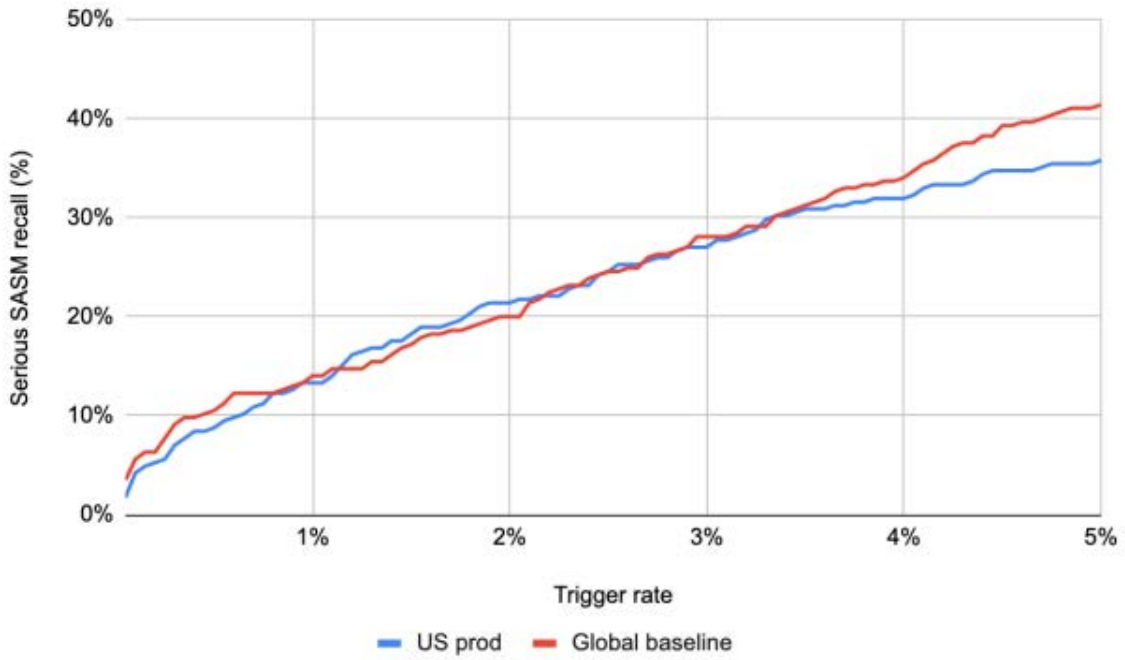
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Brazil comparison



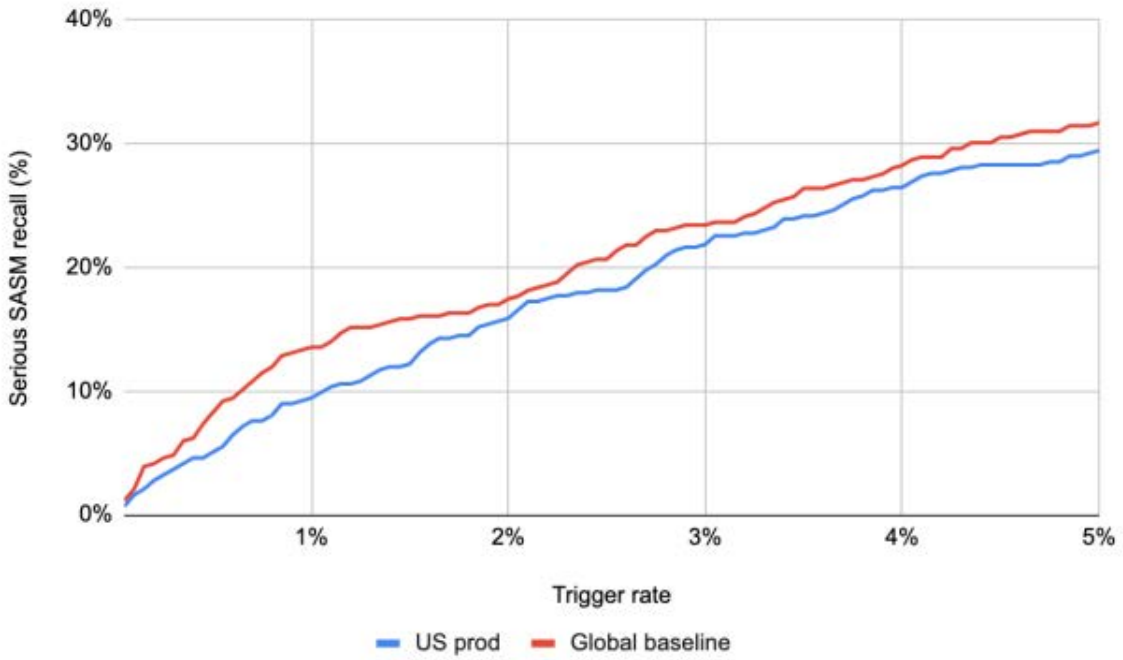
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

United Kingdom comparison



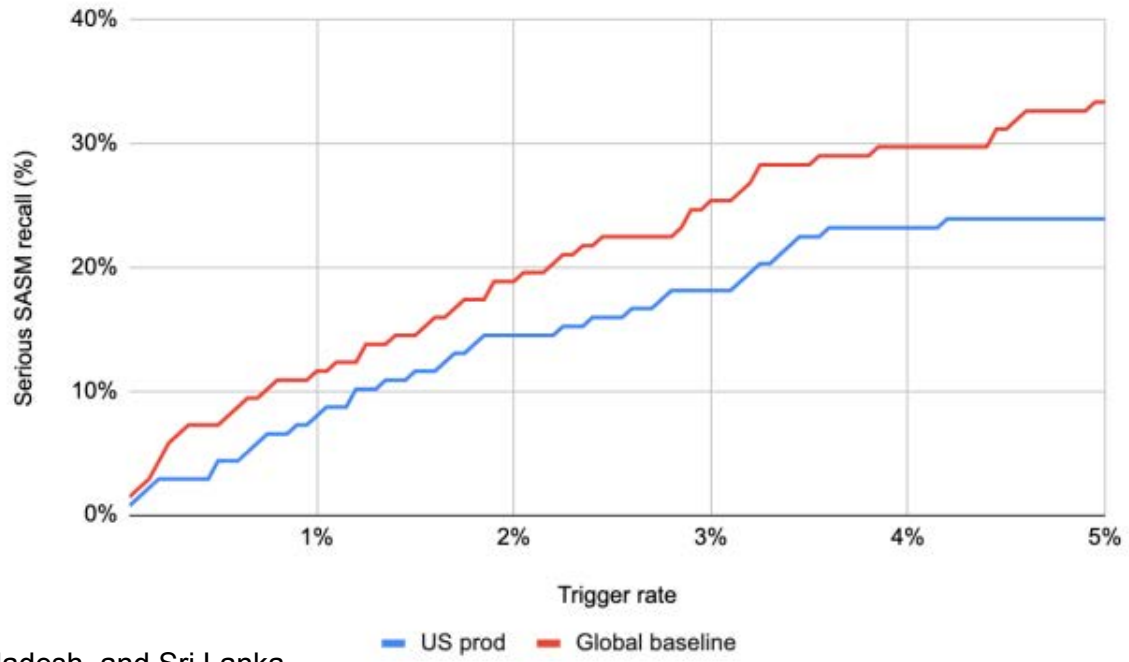
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

ANZ comparison



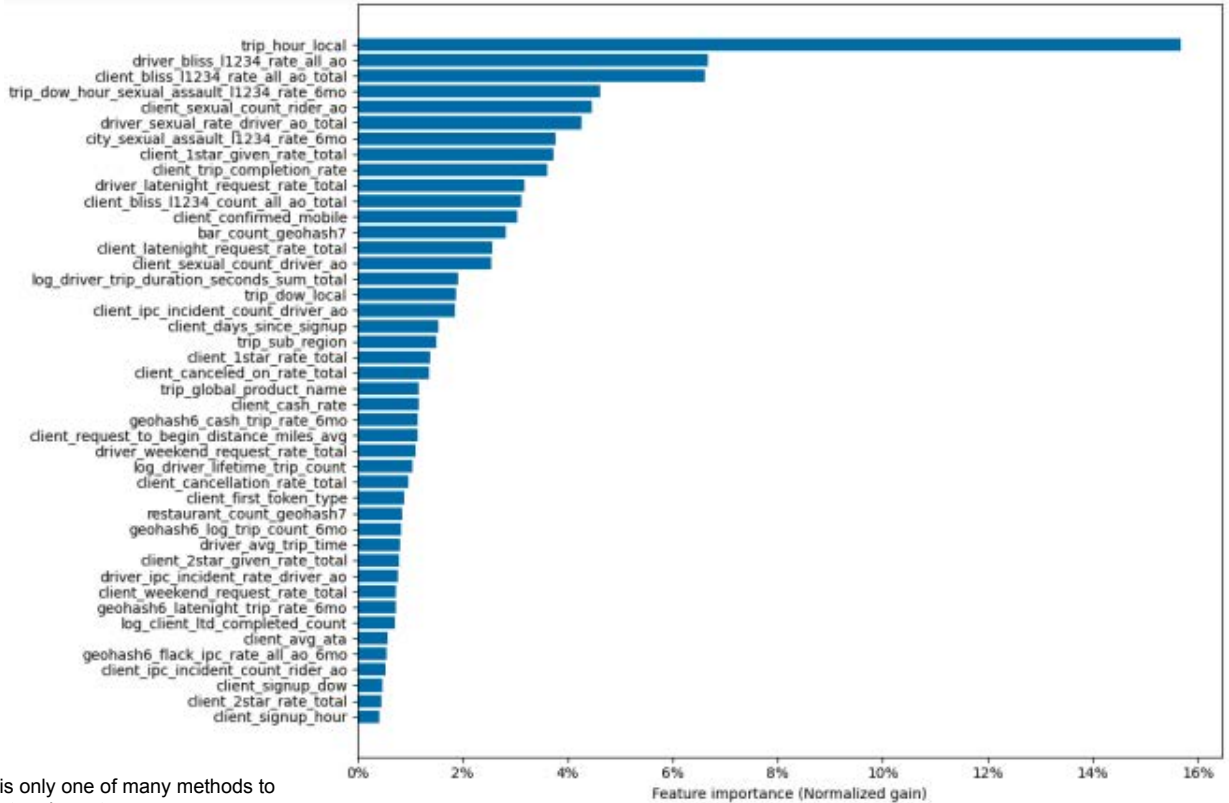
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

India SA comparison



Includes India, Bangladesh, and Sri Lanka
Recall = # serious SASM incidents flagged / # total serious SASM incidents
Recall computed using all non S-RAD actioned, UberX trips with valid driver_uuid from 2022-11-01 to 2023-01-31

Feature importance



*Total gain is only one of many methods to measure feature importance

Effect of training data size and recency

Region		Baseline US: 2 yr, 11 m BR: 2 yr, 10 m	1 year			2 year			3 year		
	Training data end date	US: 2020-10-31 BR: 2020-09-30	2020-12-31	2021-12-31	2022-08-30	2020-12-31	2021-12-31	2022-08-30	2020-12-31	2021-12-31	2022-08-30
	Eval trigger rate										
United States	1.25%	18.3%	15.5%	15.7%	18.5%	18.6%	18.1%	19.2%	17.5%	17.7%	18.8%
Brazil	1.10%	15.6%	14.1%	16.1%	17.6%	16.9%	15.6%	18.0%	16.1%	17.3%	17.8%
ANZ	1.25%	10.8%	8.9%	9.9%	14.0%	12.2%	12.6%	14.0%	12.6%	13.8%	14.0%
United Kingdom	1.25%	16.5%	9.8%	15.8%	14.7%	13.3%	13.7%	14.7%	13.7%	15.1%	15.8%
India	1.25%	10.1%	8.0%	5.8%	13.0%	10.1%	10.1%	13.0%	10.1%	10.1%	11.6%

Color legend
Max regional recall
± 5% prod recall

For a fixed training window, models tend to perform better with more recent training data

Models with 2 years of training data seems to perform best, though peak COVID observations may be biasing certain training runs

Effect of regularization

Adding additional constraints to limit overfitting appears to result in models with more **sustained** offline performance

Region		No additional regularization			With terminal leaf regularization			Percent difference		
	Training data start date	2018-01-01	2019-01-01	2020-01-01	2018-01-01	2019-01-01	2020-01-01	2018-01-01	2019-01-01	2020-01-01
	Eval trigger rate									
United States	1.25%	17.5%	18.6%	15.5%	18.1%	18.6%	17.2%	3.2%	0.0%	10.7%
Brazil	1.10%	16.1%	16.9%	14.1%	17.3%	18.4%	17.3%	8.0%	8.9%	22.7%
ANZ	1.25%	12.6%	12.2%	8.9%	15.4%	14.5%	13.1%	21.8%	18.9%	46.2%
United Kingdom	1.25%	13.7%	13.3%	9.8%	15.1%	14.7%	15.4%	10.3%	10.5%	57.1%
India	1.25%	10.1%	10.1%	8.0%	13.0%	10.9%	11.6%	28.6%	7.1%	45.5%

Training end date: 2020-12-31
Evaluation period: 2022-11-01 to 2023-01-31

Effect of regularization

Adding additional constraints **does not meaningfully impact** immediate post training recall

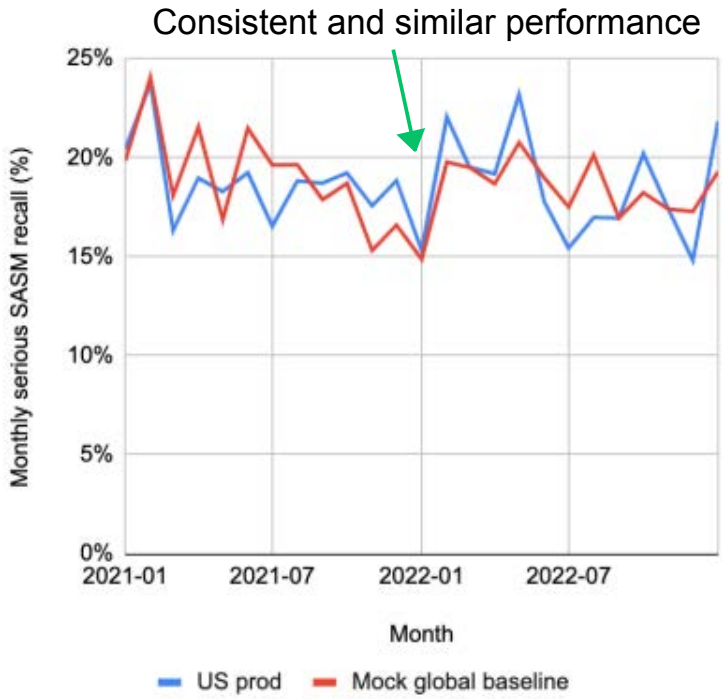
Region		No additional regularization			With terminal leaf regularization			Percent difference		
	Training data start date	2019-09-01	2020-09-01	2021-09-01	2019-09-01	2020-09-01	2021-09-01	2019-09-01	2020-09-01	2021-09-01
	Eval trigger rate									
United States	1.25%	18.8%	19.2%	18.5%	18.6%	19.2%	18.6%	-1.0%	0.0%	1.0%
Brazil	1.10%	17.8%	18.0%	17.6%	18.6%	17.3%	17.3%	4.8%	-3.6%	-1.2%
ANZ	1.25%	14.0%	14.0%	14.0%	14.9%	15.1%	15.4%	6.6%	8.2%	9.8%
United Kingdom	1.25%	15.8%	14.7%	14.7%	15.1%	14.7%	16.5%	-4.4%	0.0%	11.9%
India	1.25%	11.6%	13.0%	13.0%	12.3%	13.8%	13.0%	6.2%	5.6%	0.0%

Training end date: 2022-08-30
Evaluation period: 2022-11-01 to 2023-01-31

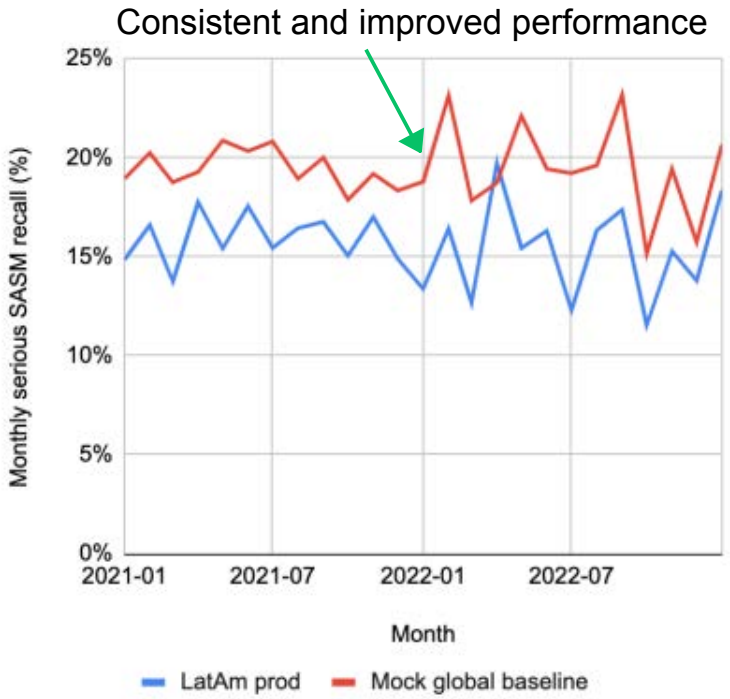
Impact of excluded variables

Region	Eval trigger rate	Prod recall	Global model recall			
			Baseline	Driver cancel vars	Driver star rating vars	Driver cancel + star rating vars
United States	1.25%	18.3%	18.8%	18.8%	18.5%	18.5%
Brazil	1.10%	15.6%	17.3%	17.3%	18.8%	19.5%
ANZ	1.25%	10.8%	9.9%	14.7%	16.6%	15.4%
United Kingdom	1.25%	16.5%	15.8%	16.8%	18.6%	18.9%
India	1.25%	10.1%	12.3%	11.6%	14.5%	13.0%

Expected model stability



United States



Brazil

Additional checks

- Overlap analysis between flagged users
- Online/offline score comparison
- Marketplace impact
  - If we need to change flagging rates, what does that mean for performance?
- Driver fairness
- Plan swapability