

## MISOGYNISM IN SOCIAL MEDIA

A prediction report directed at preventing misogynism and discriminatory behavior against women online

#### **ABSTRACT**

A lot of opiniated comments are being shared on social media platforms of which some demographics may be targets of involuntary segregation; Misogynism against the female gender. The purpose of this reports findings is to protect women from said degrading language and to make sure that social media platforms are safe to traverse.

Troels Ryberg Andersen, Simon Søndergaard & Kristian Thorndahl -SDSAAUB A/S

SDS, Applied Data Science and Machine Learning



### Misogynism; a reoccurring subject

- Misogynism has been expertly defined as "directs abuse at women or a closely related gendered group (e.g. feminists)"
- Misogynism has been flourishing on the anonymity of the internet, and certain a few of these observations were gathered and discussed by the organization <u>Amnesty</u><sup>2</sup>, in which they base the incriminating behavior online as disregarding UN conventions for "<u>Elimination of</u> <u>Discrimination against Women</u>" created back in 1979.
- This unfortunate reoccurring subject needs to be regulated by social media platforms as not to be preliminary towards a broader acceptance of gender segregation.



Figure 1: Reddit data, commonly occurring words in misogynistic comments. Word size = occurrences

A worryingly amount of misogynistic comments on the online community based platform 'Reddit' has been explored for your convenience, and the results of common words used in these comments is as shown above.

Though these may not seem to be incriminating, they have been converted to a

common denominating word for a more lenient experience.

But that beas the question; how would one differentiate a hateful,

misogynistic comment from a non-misogynistic comment?

<sup>&</sup>lt;sup>2</sup> TOXIC TWITTER - WOMEN'S EXPERIENCES OF VIOLENCE AND ABUSE ON TWITTER Chapter 3, Amnesty international (2018)



An Expert Annotated Dataset for the Detection of Online Misogyny (Ella Guest et. al., 19-Apr-2021)

# What solutions does SDSAAUB A/S offer?

Though, not all misogynism may be obvious to the individual and neither be identical in its form, SDSAAUB A/S has been hard at work to create a computer model which can learn and identify misogynistic behavior in online 'So-Me'<sup>3</sup> platforms by using state-of-the-art **machine learning algorithms**.

Thereby, <u>differentiating</u> common language use and, with time, satirical comments<sup>4</sup> from real <u>discriminatory</u> language as not to *incriminate* on the *freedom of speech* of 'So-Me' outlet users.

<sup>&</sup>lt;sup>4</sup> Training machine learning models takes a lot of data, a larger sample size may be needed to ensure satirical or other innocent humorous content not to be wrongfully flagged.



<sup>&</sup>lt;sup>3</sup> So-Me is a commonly used abbreviation for **social media** 

#### The solution...

With the <u>reddit</u> and <u>twitter</u> dataset, by using Machine Learning
 techniques, we have been able to create a model that could <u>predict</u> misogynistic language on a multitude of 'So-Me' platforms.<sup>5</sup>

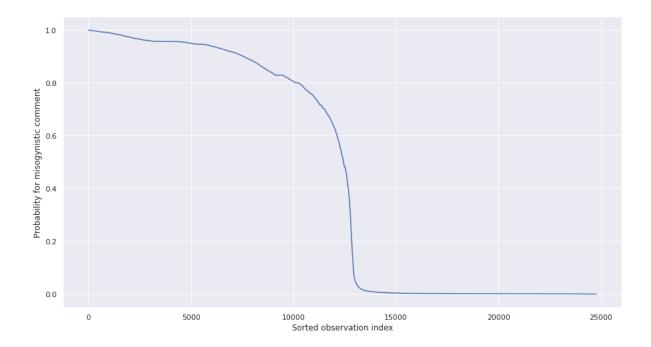


Figure 2 Misogynistic probable tweets versus improbable

- As seen in the graph above our machine learning model has, at the
  discretion of our excellent software developers, predicted the probability of
  misogynistic language in a selection of tweets. The outcome was an
  astonishing result of around 10.000 tweets being within 80% percent probable.
- These tweets can be flagged as not deemed suitable for the platform and can be removed or reviewed at the 'So-Me' moderators' discretion. Though the sensitivity range of detecting misogyny can be modified as required.
- To fight anonymous aggressors, 'So-Me' platforms must always one step ahead, and we at SDSAAUB A/S are here to help you make that step:

With a goal of creating a safer online experience!6

<sup>&</sup>lt;sup>6</sup> To learn more on our pricing model or to book an exam or meeting, visit us on: <a href="https://www.sdsaaub.org/fightmisogyny/contact">www.sdsaaub.org/fightmisogyny/contact</a>



<sup>&</sup>lt;sup>5</sup> Trialed and tested on twitter and reddit within 95% certainty on normalized language data.