

Hey Baby You Up? A Solution To Early Online Sexual Predator Detection

Haley Bloom, Mathematical & Computer Science Dept.

Advised by Dr. Thomas Montelione

Project Overview

Online sexual grooming is the process of coercing a minor into an emotional and/or sexual relationship through social media and other online platforms such as chat rooms.

The purpose of this research project is to develop an early online sexual predator detection framework to stop predators early in the grooming process. This was accomplished by training the deep learning model BigBird and utilizing the eSPD framework developed by Vogt et al. in 2021

Research Questions

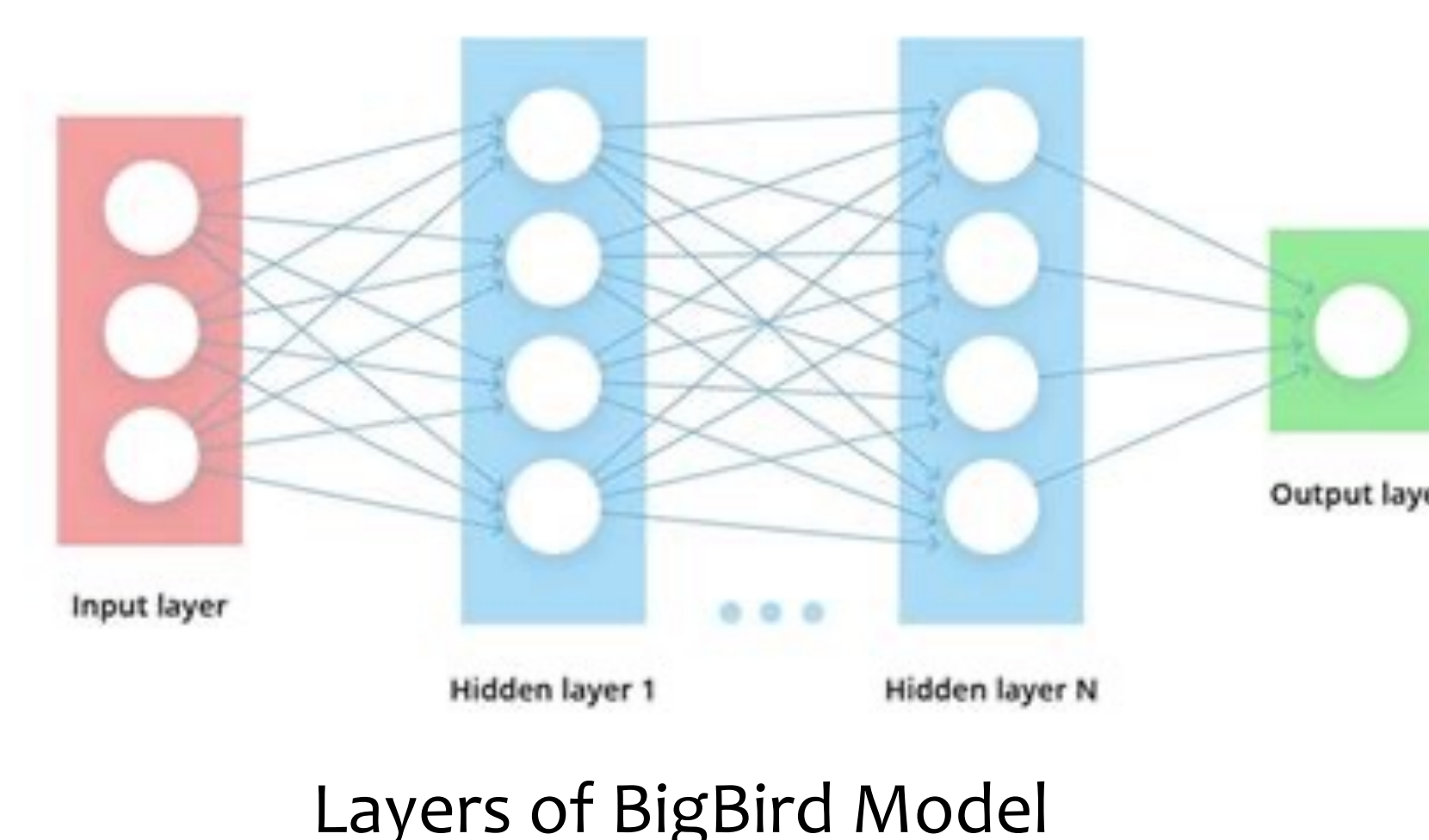
1. Investigate and utilize the publicly available eSPD framework and improve upon the initial results by Vogt et al.
2. investigate how accurately and early the BigBird model and fine-tuning detect predatory behavior.
3. Make the developed model as accurate as possible.

eSPD Framework

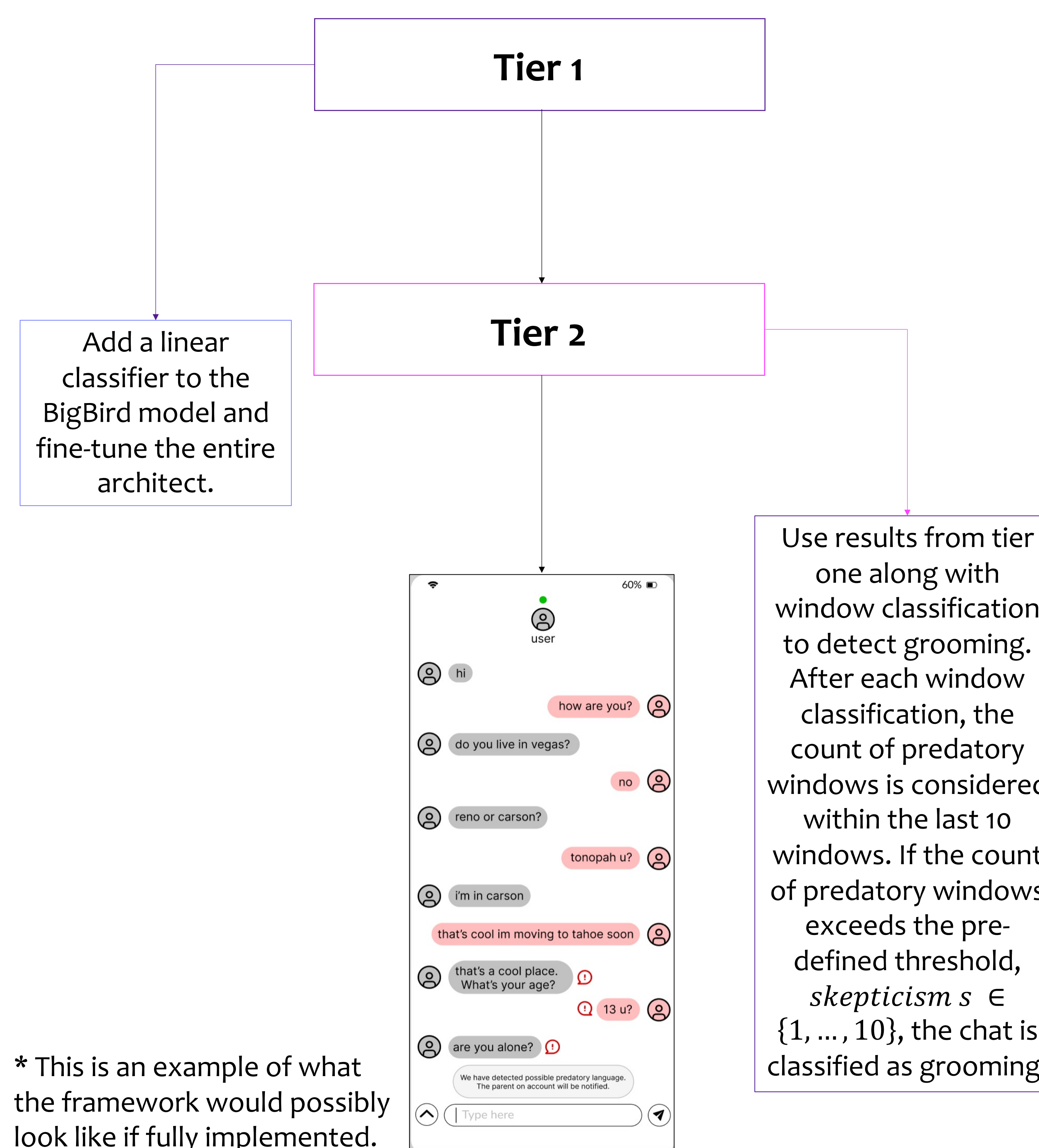
- Two-tier text classification framework for early sexual predator detection developed in 2021 by Vogt et al.
- Tier 1 moves a sliding window over the chats and classifies them using deep learning models.
- Tier 2 is a global tier that decides after each window prediction whether to raise a predator warning based in previous predictions.

BigBird Model

The BigBird model is a pre-trained transformer model that takes in the chat data and trains the model to recognize patterns in the chats that predators display. The model outputs the accuracy percentage and the speed in which it took the model to output the results.

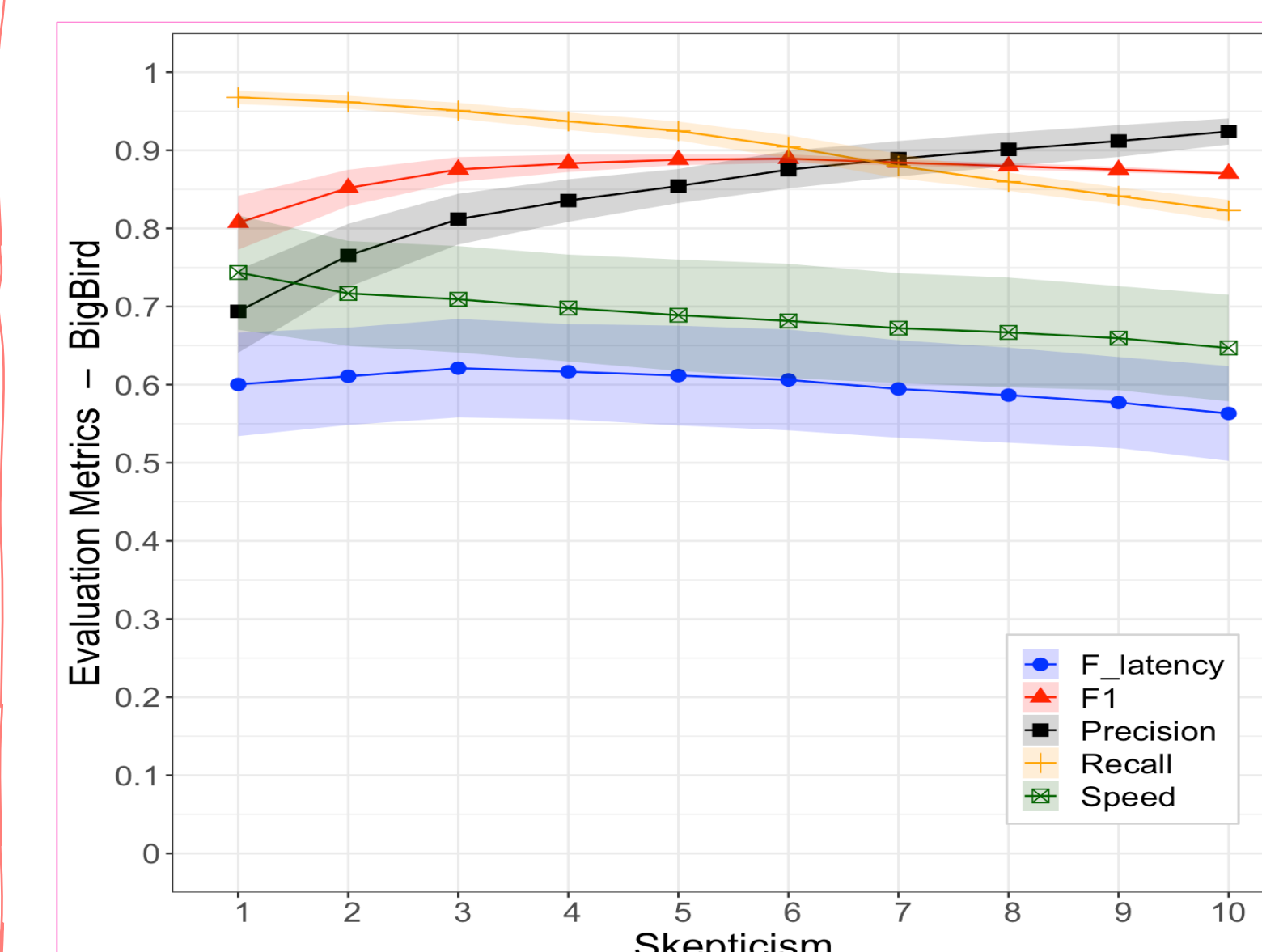


Methodology



* This is an example of what the framework would possibly look like if fully implemented. This is not an actual application.

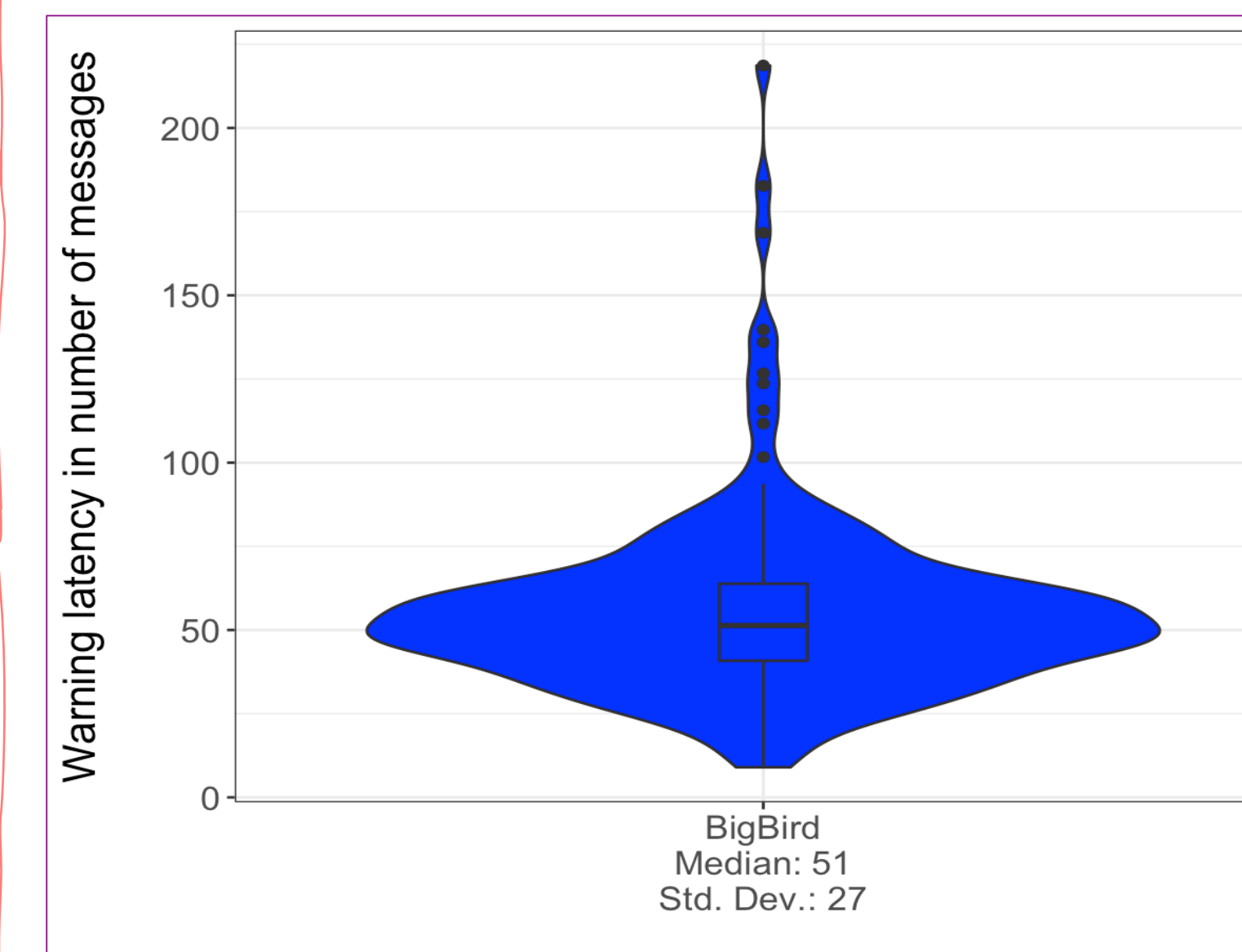
Results



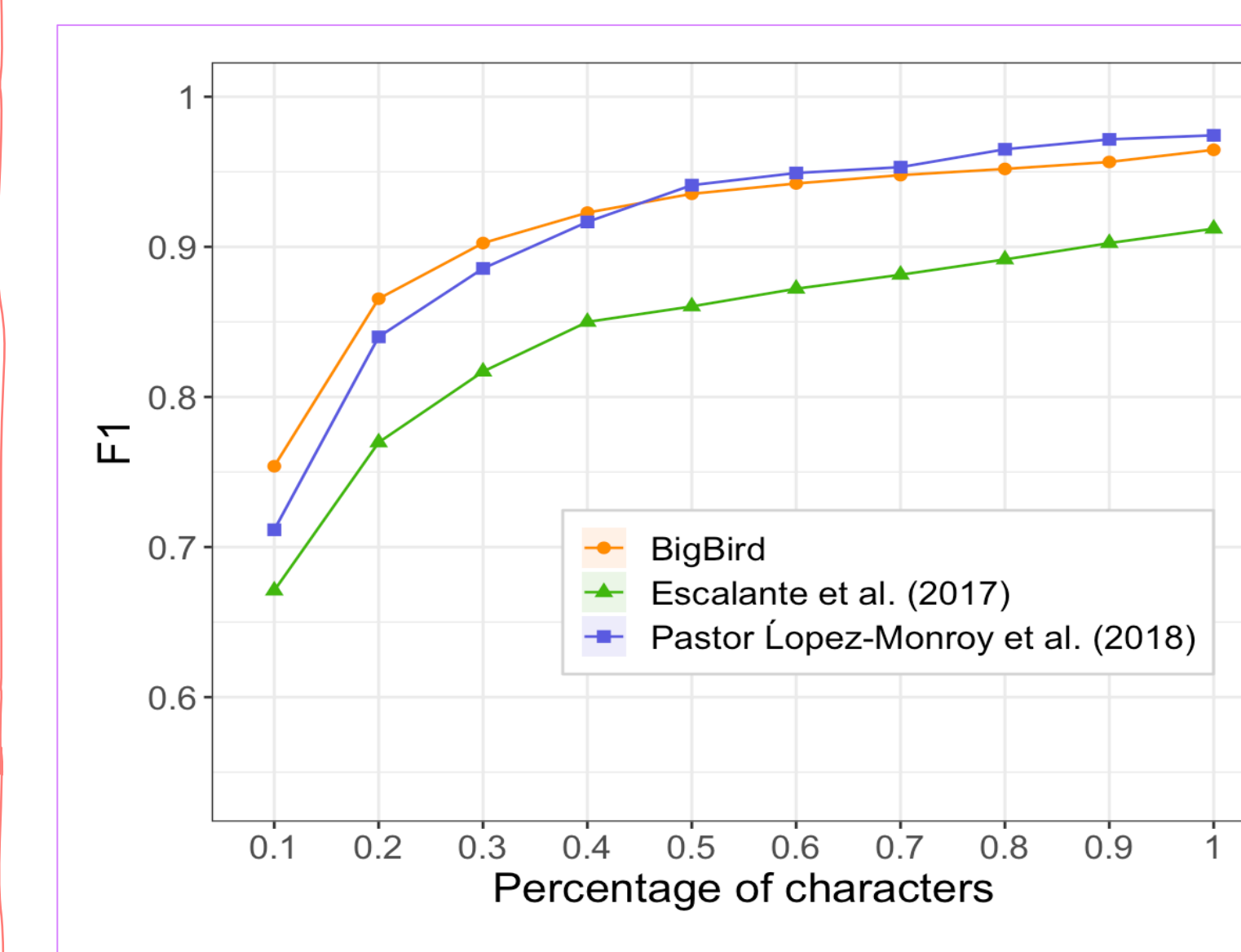
Plot of BigBird Evaluation Metrics By Skepticism

	F1-score	Precision	Recall	Speed	$F_{latency}$
Final model	0.88 (± 0.019)	0.87 (± 0.066)	0.89 (± 0.053)	0.63 (± 0.053)	0.56 (± 0.022)

Warning Accuracy score on PANC dataset as mean with standard deviation



Plot of BigBird Warning Latency in Number of Messages



Plot of BigBird and other SOTA solutions By Percentage of Characters

- The mean, F_1 score, and $F_{latency}$ were calculated for each level of skepticism (1-10)
- A skepticism level of 5 showed the best result for the weighted F_1 score.
- The final F_1 score at level 5 skepticism was .89 or 89%
 - This means that the model was 89% accurate with identifying sexual predators when the skepticism level is 5.

- Mean warning latency is the average count of how many messages were exchanged between predator and victim before a warning is raised.
- The mean warning latency was 51 with a standard deviation of 27.

- Comparison of the BigBird model results to previous solutions
- The BigBird solution beat the solution by Escalante et al. by eight percent.
- The solution by Pastor Lopez-Monroy et al. beat the BigBird model by no more than two percent accuracy.