

Presidio gets better TNR

Situation: Presidio is fast and cheap for PII identification, and has strong FNR at zero. The problem is that TNR is only at 10%

Goal: we want to increase the TNR of Presidio while maintaining FNR at zero

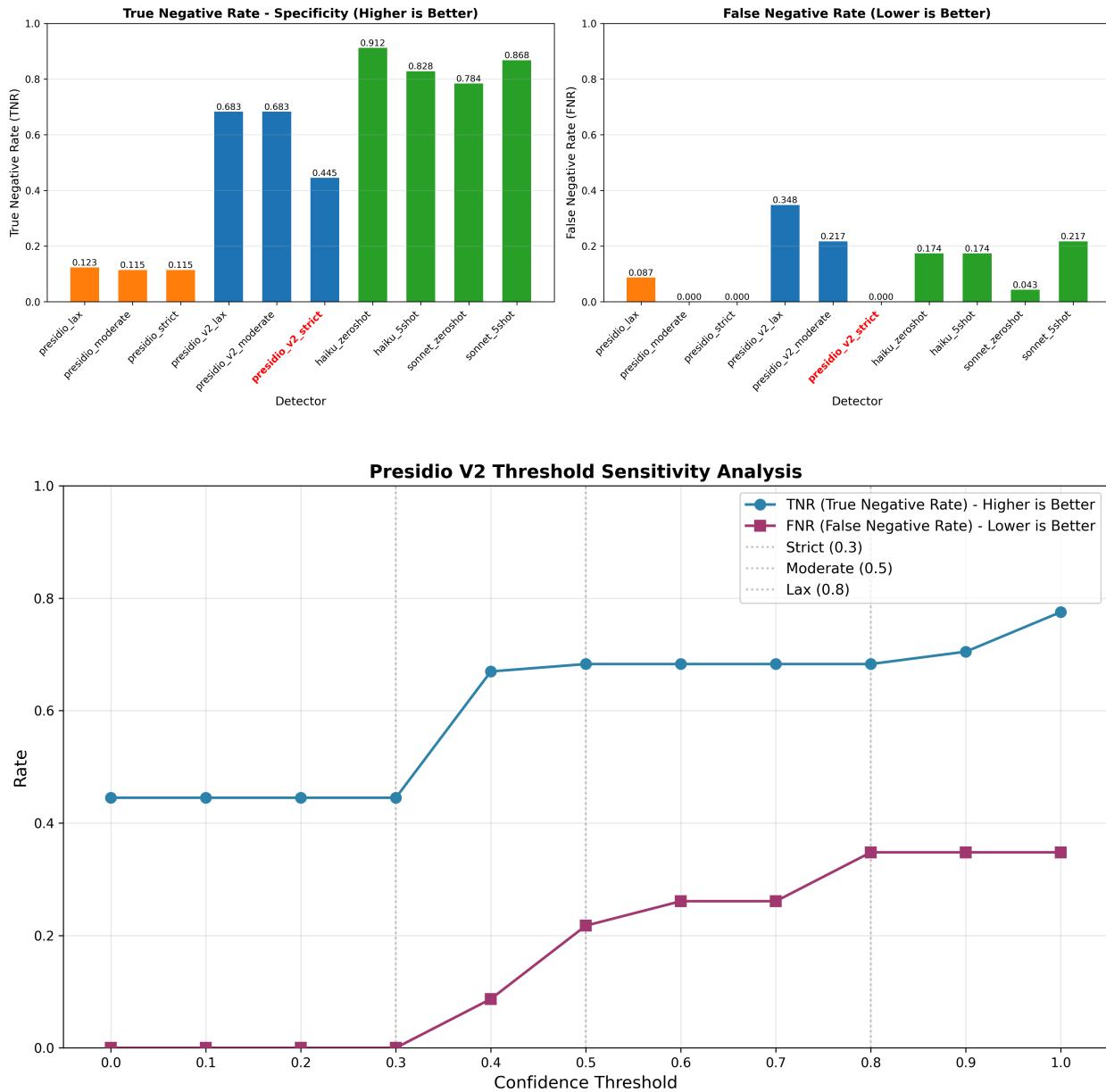
Methods

We first investigate the failure cases for Presidio and why the TNR is so low. We do so by inspecting the false positive cases for Presidio. Note that a single email can have multiple PII entities.

Reason for False Positive	Count	Percentage	Example Detection
Business/Professional Names	201	100.0%	Names like "Phillip K Allen", "John J Lavorato" in work email headers and signatures
Date/Time Information	144	71.6%	Scheduling info like "Tuesday at 11:45", "10/03/2000" in meeting coordination
URLs/Web Links	98	48.8%	Domain names like "hotmail.com", "austin.rr.com" in email addresses
Email Addresses	91	45.3%	Business emails like "stagecoachmama@hotmail.com", "pallen70@hotmail.com"
Location/Address	51	25.4%	Business locations like "Austin", city names, office locations
Phone Numbers/Extensions	5	2.5%	Business contact numbers like "512-748-7495", office extensions

We then worked on the Presidio program to make it reflect more closely to the specifications, especially only flagging name if there is a non-email address PII present, and removing screening for datetime and URLs. We also noticed that the current Presidio program didn't scan for CVV, passwords and usernames, and we wrote custom analyzers for them. We call this Presidio V2. See results below.

Results



Confidence threshold means how confident does Presidio think the entity is indeed a PII before it flags it. Note that FNR is zero from 0 to 0.3, and TNR is substantial at above 40%. We will choose a confidence threshold of 0 to stay as far away from a positive FNR.

Presidio V2 now has good TNR (~40%) with FNR maintained at 0%.