

# Improving LLMs with Presidio

Situation: LLMs have non-zero FNR rates, and we hope to make that zero

Task: Augment LLMs with Presidio outputs, since Presidio has zero FNR

## Summary

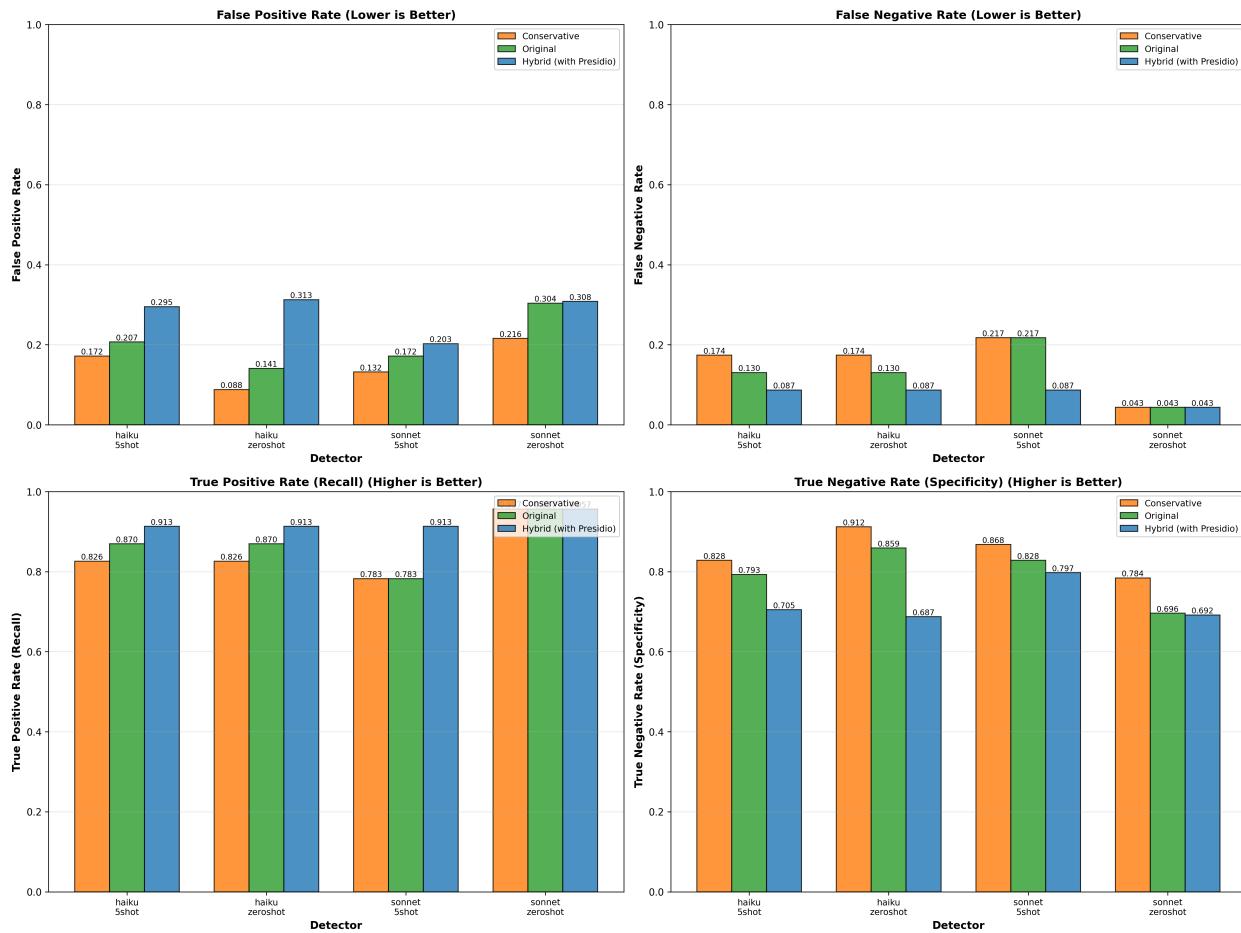
Including Presidio does lower FNR, but not to the zero-level. Sonnet 0-shot with or without Presidio still outperforms Haiku with Presidio in terms of FNR.

## Results

We find that adding Presidio outputs to the LLM does reduce FNR in Haiku models and the sonnet 5-shot model. However, none of the models reached 0 FNR. Note that there is a penalty hit to TNR.

Note that we will accept FNR=0.043 as the zero baseline, since we have identified in a previous report that one email has been wrongly classified as containing PII.

#### Classification Rates Comparison: Original vs Conservative vs Hybrid (with Presidio)



The failure cases for the FNR are email address for ID 7

Please cc the following distribution list with updates:

Phillip Allen (pallen@enron.com)  
 Mike Grigsby (mike.grigsby@enron.com)  
 Keith Holst (kholst@enron.com)  
 Monique Sanchez  
 Frank Ermis  
 John Lavorato

Thank you for your help

Phillip Allen

## Address for ID 47

Brenda,

I checked my records and I mailed check #1178 for the normal amount on August 28th. I mailed it to 4303 Pate Rd. #29, College Station, TX 77845. I will go ahead and mail you another check. If the first one shows up you can treat the 2nd as payment for October.

I know your concerns about the site plan. I will not proceed without getting the details and getting your approval.

I will find that amortization schedule and send it soon.

Phillip