

How Nib tells a **real** signature from a forgery.

The principles behind Unforged Nib, how we test nib v2.2 on public datasets, and where its limits are.

nib v2.2 the model tested	3 public datasets	none customer or benchmark signatures used for training
-------------------------------------	-----------------------------	---

Summary

Unforged Nib checks whether a handwritten signature on a document was made by the person it claims to be from. It compares the signature against that person's own enrolled reference signatures and returns one of six decisions, a calibrated probability with a band, and plain-language reasons.

We test nib v2.2 on public datasets, straight out of the box. Results will be published once the dataset owners have agreed. It is designed to be strict: when it is unsure, it sends a signature to a person rather than approving it. This paper explains how it works, how we test it and where its limits are.

1. The problem

A signature is a behaviour, not a picture. Nobody signs the same way twice: size, slant, spacing and pressure drift from one signature to the next. A good verifier has to accept that natural variation while still catching a forger who has studied the real thing.

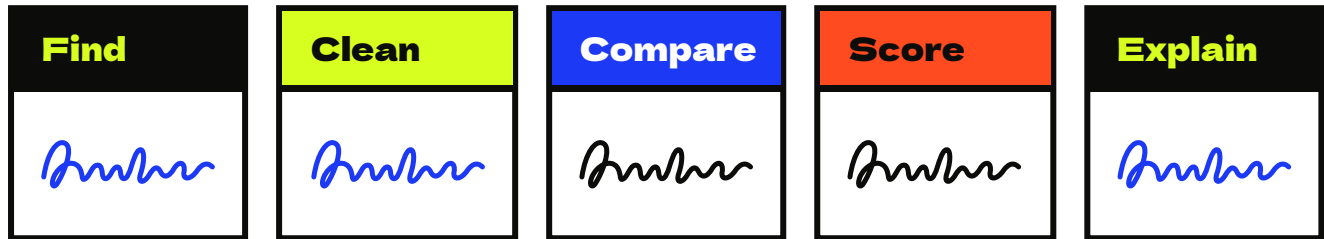
Forgeries come in two kinds. A **random forgery** is someone else's signature, or a guess at a name. A **skilled forgery** is made by someone who has practised the genuine signature. Skilled forgeries are the hard case, and the one that matters in fraud, so they are what we report first.

2. Comparing a person with themselves

Nib is writer-dependent: it never asks "does this look like a signature?", only "does this look like *this person's* signature?". At enrolment you give it up to 10 genuine reference signatures. From them Nib learns both what the

person's signature looks like and how much it normally varies. A questioned signature is judged against that range, so a naturally messy signer is not penalised for being messy, and a very consistent signer is protected by their own consistency.

3. The pipeline



Find. Locate the signature on the page, whether it is on a form, a PDF, a scan or a phone photo taken at an angle. No signature, or an image too poor to read, is caught here, and is free.

Clean. Remove ruled lines, printed boxes, stamps and paper tint so that only the ink is measured. A form's layout never counts for or against a signature.

Compare. Measure letter shapes, stroke flow and proportions against each enrolled reference, several independent ways, and normalise each measurement by how much this person usually varies.

Score. Combine the evidence into one calibrated probability that the signature is genuine, with a band showing how sure the estimate is.

Explain. Report the decision and the reasons behind it in plain language, so a person reviewing it knows where to look.

4. Probability, band and policy

The probability is calibrated, so it reads as a likelihood rather than a raw similarity score, and it means the same thing whoever the signer is. The band shows the uncertainty; more and better references narrow it.

A policy turns the probability into a decision. The default policy fails a signature below 0.08 and approves it from 0.95; everything between is flagged for a person. Customers can set a stricter or looser policy, but every figure in this paper uses the default.

5. Six decisions, people in the loop



Every check ends in one of six decisions: AUTO APPROVE, APPROVE, FLAG, FAILED, UNREADABLE and NO SIGNATURE. FLAG sends the document to a person. FAILED

always goes to a person too: Nib never rejects anyone automatically. UNREADABLE and NO SIGNATURE are free, because there was nothing to verify.

6. Your data

Customer signatures are never used for training. If you switch on reference learning, accepted signatures become extra references for that one person, and nothing else. Benchmark signatures are never used for training either: Nib is tested only on data it was never trained on.

7. How we test

Datasets. CEDAR (55 writers, Latin script), BHSig260-Hindi (160 writers, Devanagari) and BHSig260-Bengali (100 writers). Each writer has 24 genuine signatures and a set of skilled forgeries.

Protocol. For each writer we draw k reference signatures at random ($k = 3, 5$ and 10 ; 10 is the product maximum), check the remaining genuine signatures and the skilled forgeries, and repeat the draw three times. There is no training on the benchmark and one global policy for every dataset.

Sealed and development sets. Bengali is our development set: we look at individual results there to find and fix problems. CEDAR and Hindi are sealed: we only see the headline summary, only at milestone runs, and every sealed run is logged. CEDAR's flag breakdown was viewed once, before this protocol existed; we disclose it here.

Measures. Equal error rate (EER) and area under the ROC curve (AUC) summarise how well genuine and forged separate. Under the default policy we also count the skilled forgeries approved and the genuine signatures approved automatically.

Skilled forgeries approved. Of the forgeries made by someone who studied the real signature, how many the default policy would approve.

Equal error rate. The point where false accepts and false rejects are equally likely. Lower is better.

Genuine signatures approved automatically. How many real signatures go straight through, and how many are sent to a person.

8. Results

Results will be published once the dataset owners have agreed. Until then we publish the method, not the numbers, so you can judge how the results will be made before you see them.

9. Limits

Strict by design. The default policy approves only from 0.95 . Many genuine signatures go to a person rather than straight through.

Scripts and capture. We test Latin, Devanagari and Bengali signatures on scanned paper. The service reads Latin-script signatures today; phone photos and other capture types are not yet benchmarked here.

References matter. Accuracy depends on the number and quality of references. More is better, up to ten.

Not an expert opinion. Nib is decision support. It does not replace a forensic document examiner in court.

References

Kalera, M. K., Srihari, S. and Xu, A. (2004). Offline signature verification and identification using distance statistics. *International Journal of Pattern Recognition and Artificial Intelligence*, 18(7). (CEDAR)

Pal, S., Alaei, A., Pal, U. and Blumenstein, M. (2016). Performance of an off-line signature verification method based on texture features on a large Indic-language dataset. *12th IAPR Workshop on Document Analysis Systems*. (BHSig260)

Unforged · unforged.ai · nib v2.2, default policy default@1. Results will be published once the dataset owners have agreed.