

KEYSTROKE DYNAMICS AUTHENTICATION SYSTEM

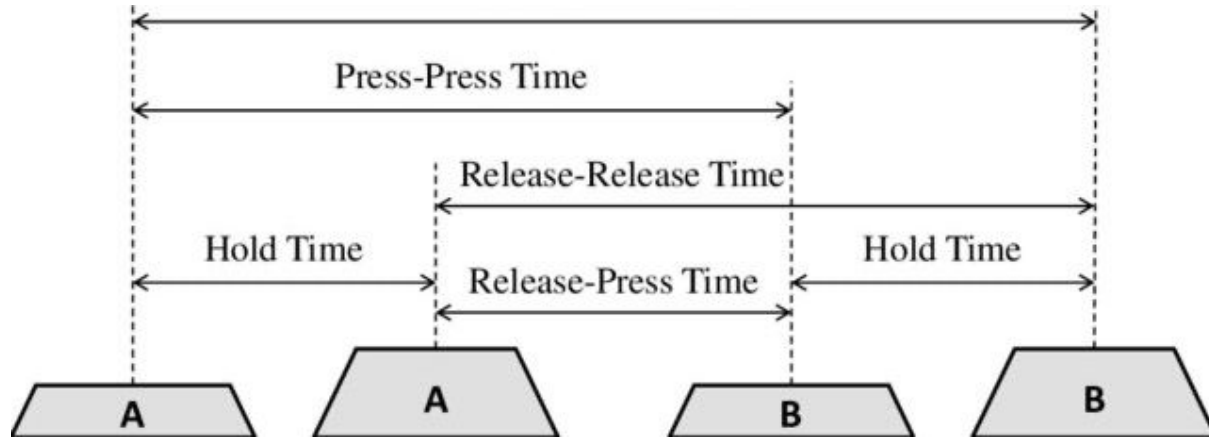
Biometric Systems
A. Y. 2022/2023

Marco Raffaele 1799912
Tommaso Battistini 1869913

INTRODUCTION



DATA ACQUISITION AND PREPROCESSING

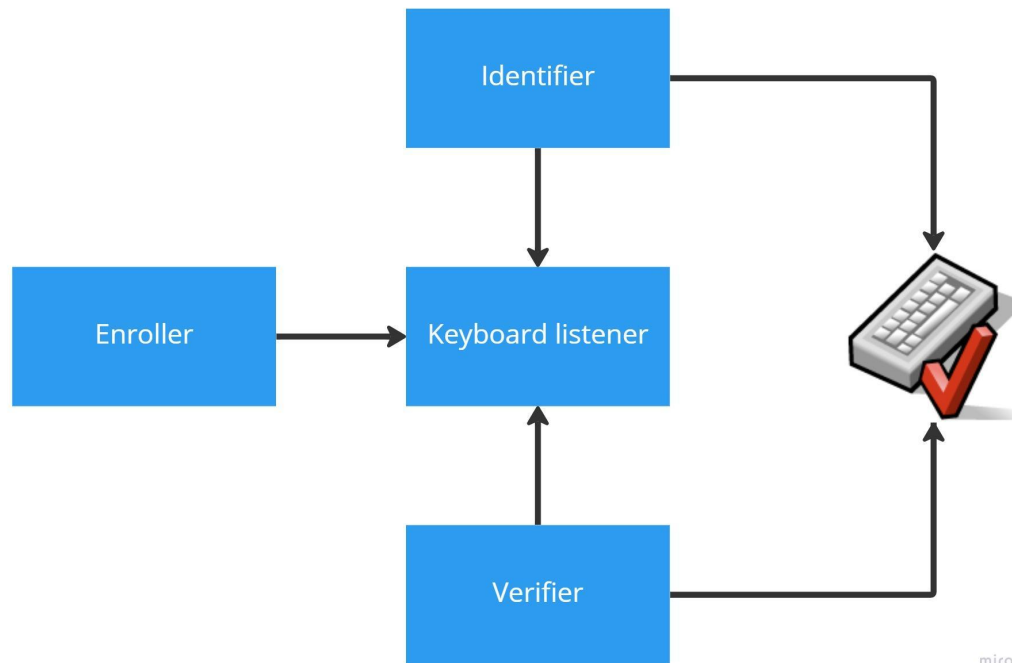


DATA ACQUISITION AND PREPROCESSING

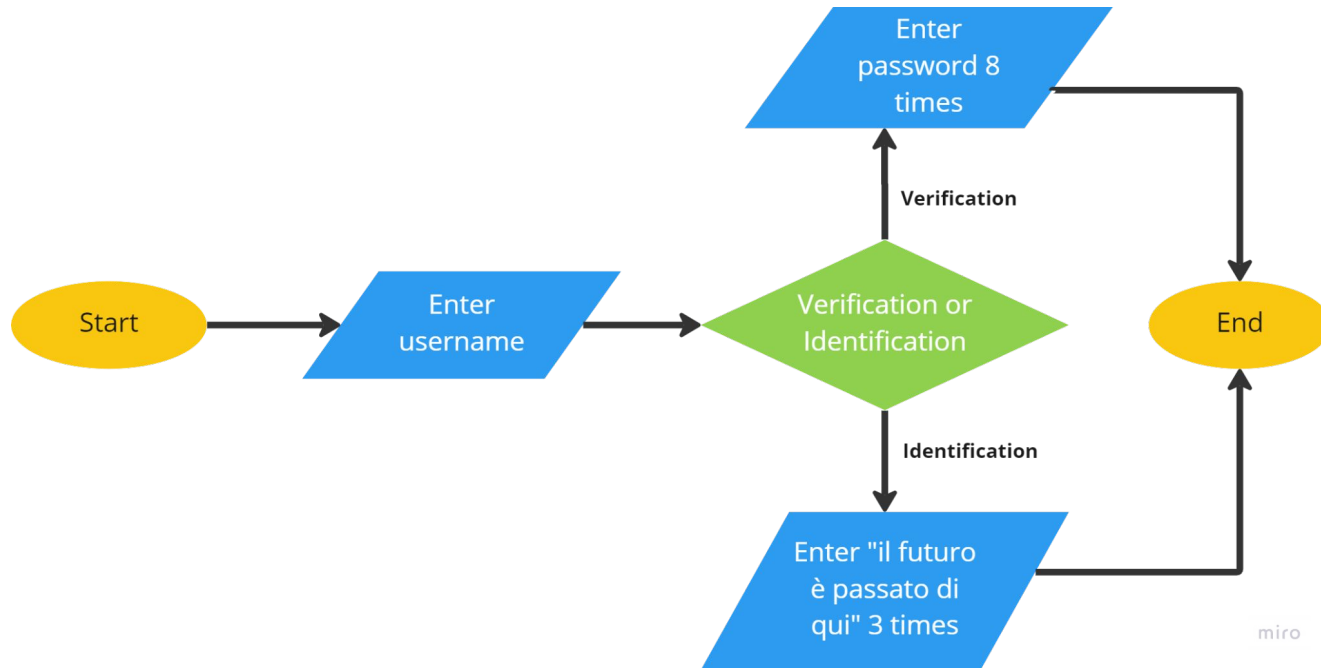
In addition to key **press** and **release** times, **slope** data are also saved.

```
[
  "0": {
    "hold_1": 0.08256649971008301,
    "press_press": 0.23795866966247559,
    "release_press": 0.15539216995239258,
    "release_release": 0.18987703323364258,
    "hold_2": 0.03448486328125,
    "total_time": 0.2724435329437256,
    "slope_h1": -0.04808163642883301,
    "slope_pp": -0.11260223388671875,
    "slope_rp": -0.06452059745788574,
    "slope_rr": -0.0780644416809082,
    "slope_h2": -0.013543844223022461,
    "slope_tt": -0.1261460781097412
  },
  "1": {
    "hold_1": 0.03448486328125,
    "press_press": 0.12535643577575684,
    "release_press": 0.09087157249450684,
    "release_release": 0.11181259155273438,
    "hold_2": 0.02094101905822754,
    "total_time": 0.14629745483398438,
    "slope_h1": -0.013543844223022461,
    "slope_pp": -0.1044154167175293,
    "slope_rp": -0.09087157249450684,
    "slope_rr": 0.015389442443847656,
    "slope_h2": 0.10626101493835449,
    "slope_tt": 0.0018455982208251953
  },
  "2": {
```

ARCHITECTURE



ENROLLEMENT



L2 NORM MODEL

The **L2 Norm** is a common similarity measure used in many fields, including computer vision, natural language processing, and information retrieval. Defined as follows

$$distance(X, Y) = ||X - Y||_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

We compute the L2 Norm of an artificial vector, computed by comparing the new samples with all the enrolled ones and picking the **most similar value** for every single feature

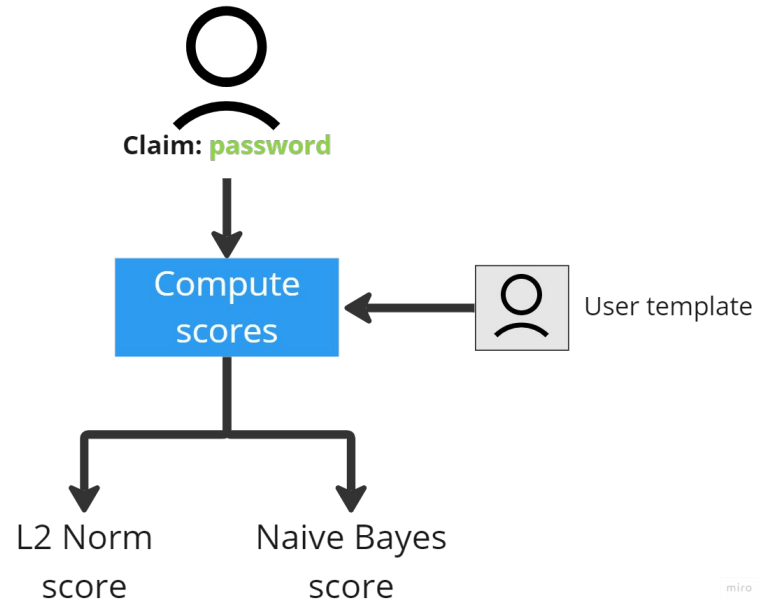
NAIVE BAYES MODEL

Naive Bayes is a probabilistic machine learning algorithm used for classification tasks by making predictions based on the probability of certain events happening, given certain conditions. We chose to use its Gaussian variation.. The used formula to estimate the underlying data distributions is defined as follows:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

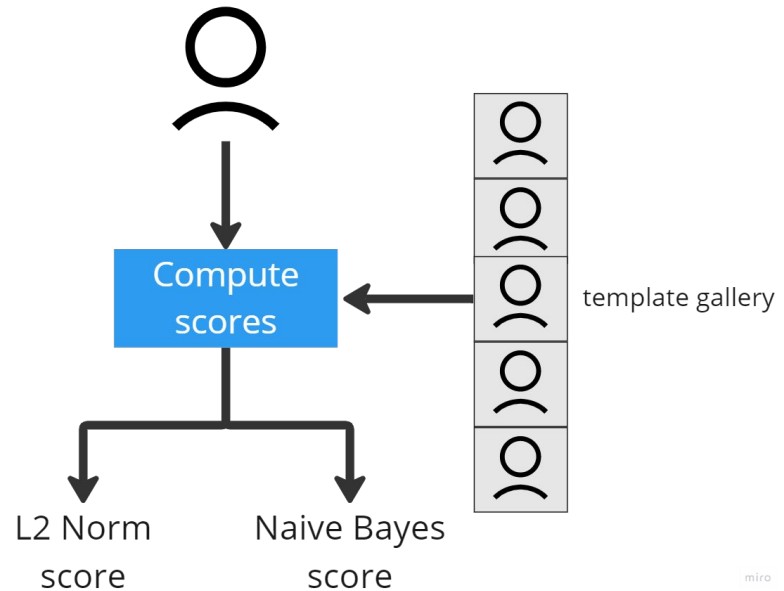
VERIFICATION

Verification is performed by comparing the typing patterns of a **password** with the data of the user associated with that password.



IDENTIFICATION

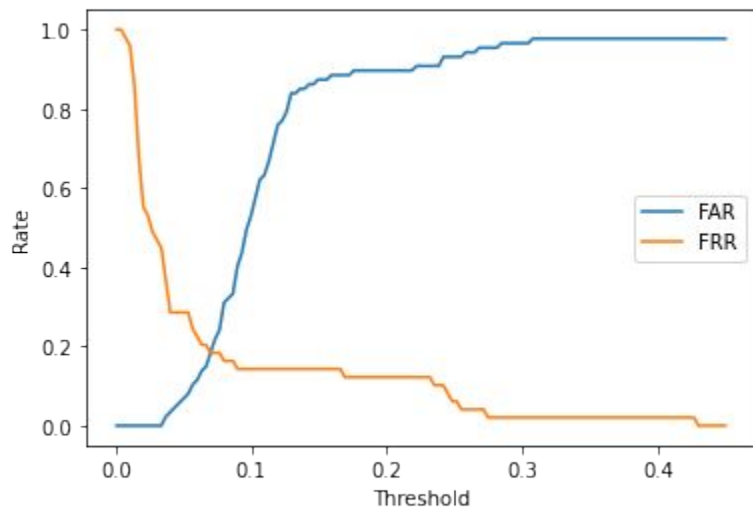
Identification is performed by comparing the typing patterns of a standard sentence ("il futuro è passato di qui") with the data from the same sentence in all user enrollments.



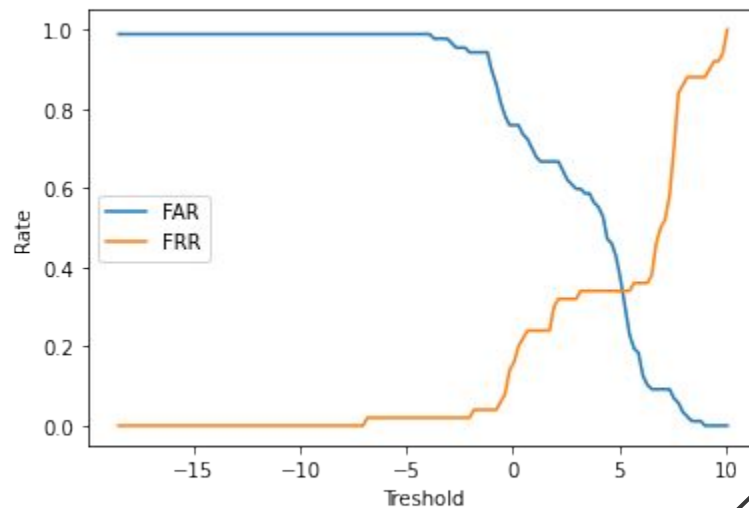
EVALUATION: VERIFICATION

False Acceptance and False Rejection Rates

L2 Norm



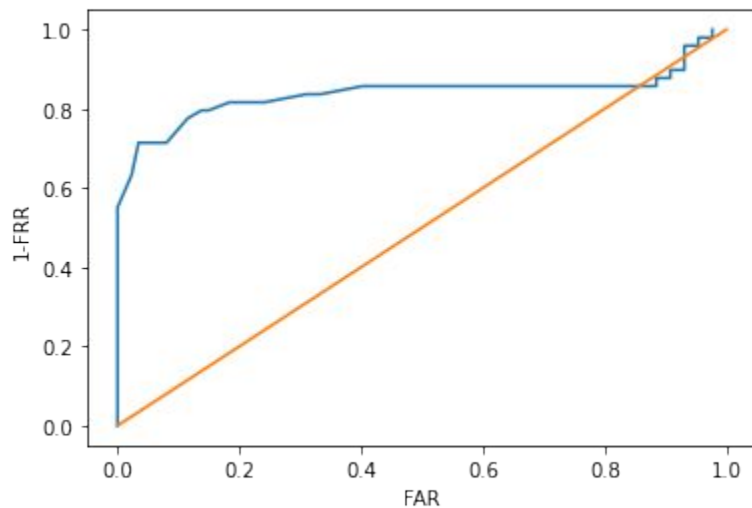
Bayes



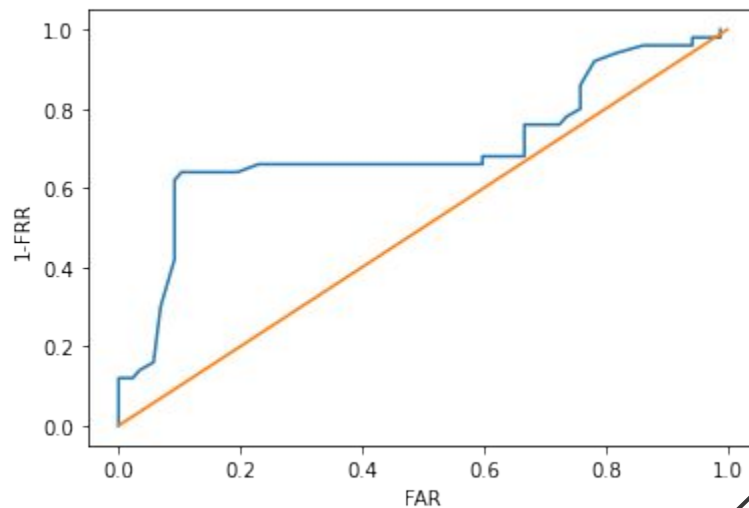
EVALUATION: VERIFICATION

Receiver Operating Characteristic (ROC)

L2 Norm



Bayes

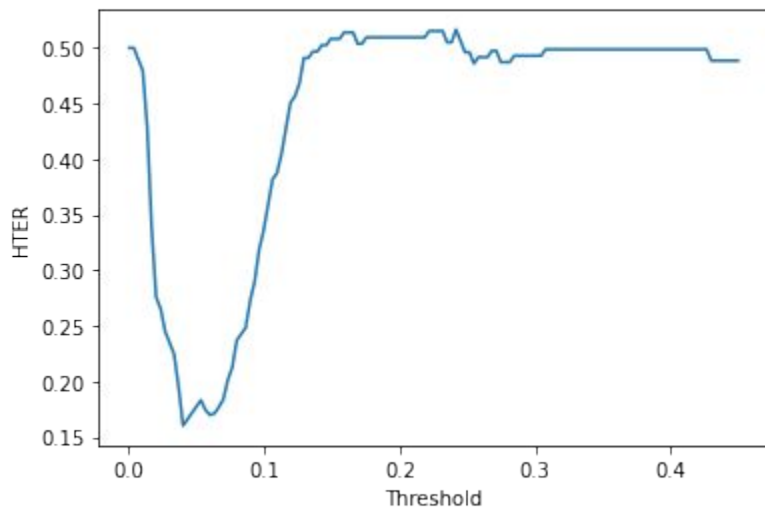


EVALUATION: VERIFICATION

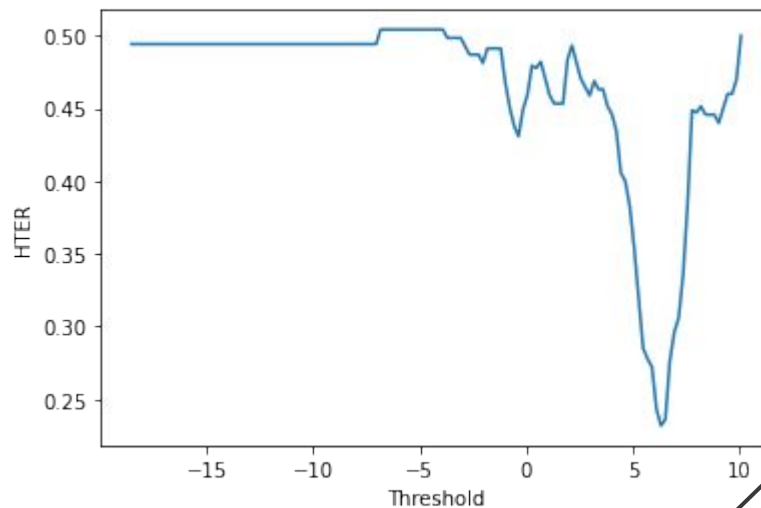
Half Total Error Rate (HTER)

$$HTER = (FAR + FRR)/2$$

L2 Norm



Bayes



EVALUATION: IDENTIFICATION

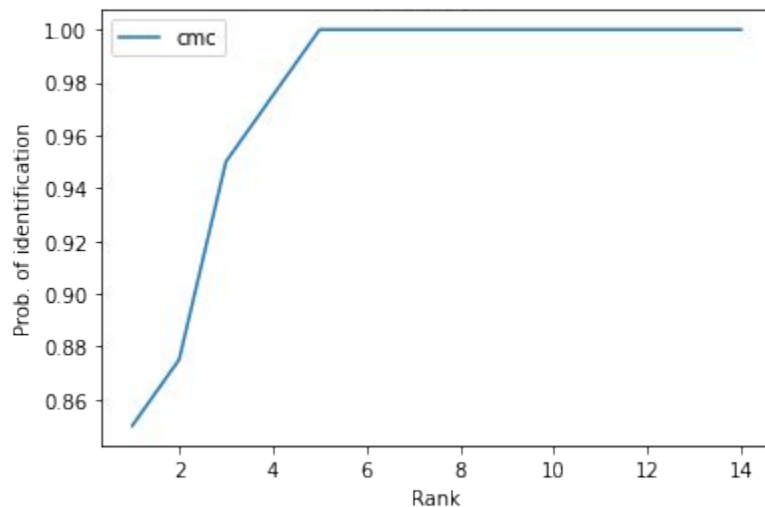
Accuracy

Bayes	0.875
L2 Norm	0.926

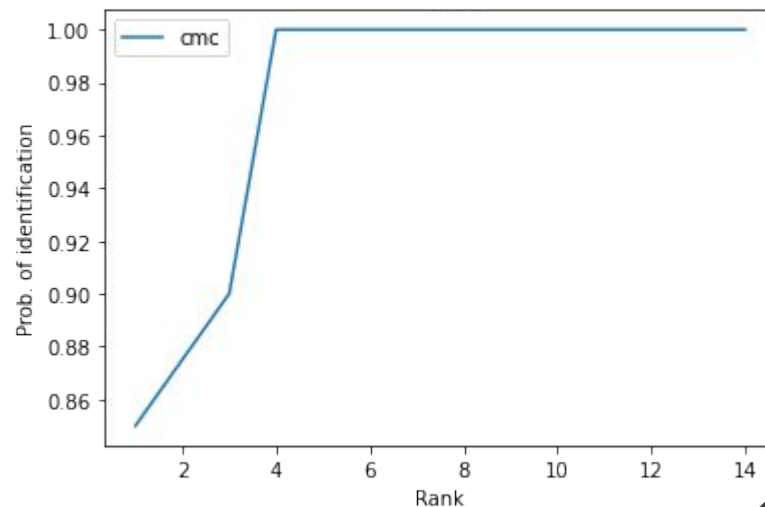
EVALUATION: IDENTIFICATION

Cumulative Match Characteristics (CMC)

L2 Norm



Bayes



CONCLUSIONS



- Recap



- Applicability



- Issues and Limitations