**Dataset:**

Each column represents specific metrics or features extracted from keystroke dynamics when the keyword "roberto" is typed by different users. Let me explain each column briefly:

1. **Usuario (User):**

This column represents the user or individual who typed the keyword "roberto". In this dataset, the user is consistently labeled as "Lilia".

1. **r\_H, o\_H, b\_H, e\_H, r\_H, t\_H, o\_H:**

These columns represent the duration or time taken for pressing and releasing each of the letters in the keyword "roberto". Each column corresponds to one of the letters in the word "roberto" (r, o, b, e, r, t, o).

1. **or\_DD, bo\_DD, eb\_DD, re\_DD, tr\_DD, ot\_DD:**

These columns represent the time intervals or durations between pressing one key and pressing the next adjacent key in the sequence of typing "roberto". For example:

* + or\_DD: Duration between pressing 'o' and releasing 'r'.
  + bo\_DD: Duration between pressing 'b' and releasing 'o'.
  + And so on for the rest of the columns.

**Model:**

After the data collection, this data was used to train different ML models. We choose the following models:

* KNeighborsClassifier
* GradientBoostingClassifier
* LogisticRegression
* Linear SVC (Support Vector Classifier)
* RandomForestClassifier

The Usuario column from the dataset was taken as the target or Y column when training the model and the other columns were as the X features. As the Usuario column is in string form, we cannot directly pass that to our models when training them, we have to do the encoding. Converting some string data to a numerical form is called encoding.

After encoding the Usuario columns, which convert the users into numerical representations of the text data like 0, 1, 2, and so on, that data is then passed to the models. Furthermore, each model is trained and evaluated on different metrics. Finally, we chose the best model among all the models we tested.

At the end of the whole process, we export the best model along with the encoder we used. We can then use this model and the encoder in our web app.