बैंक ऑफ़ बड़ौदा
Bank of Baroda

Technology Partner
Microsoft Azure

Bank of Baroda Hackathon - 2022

Team Name :1001001

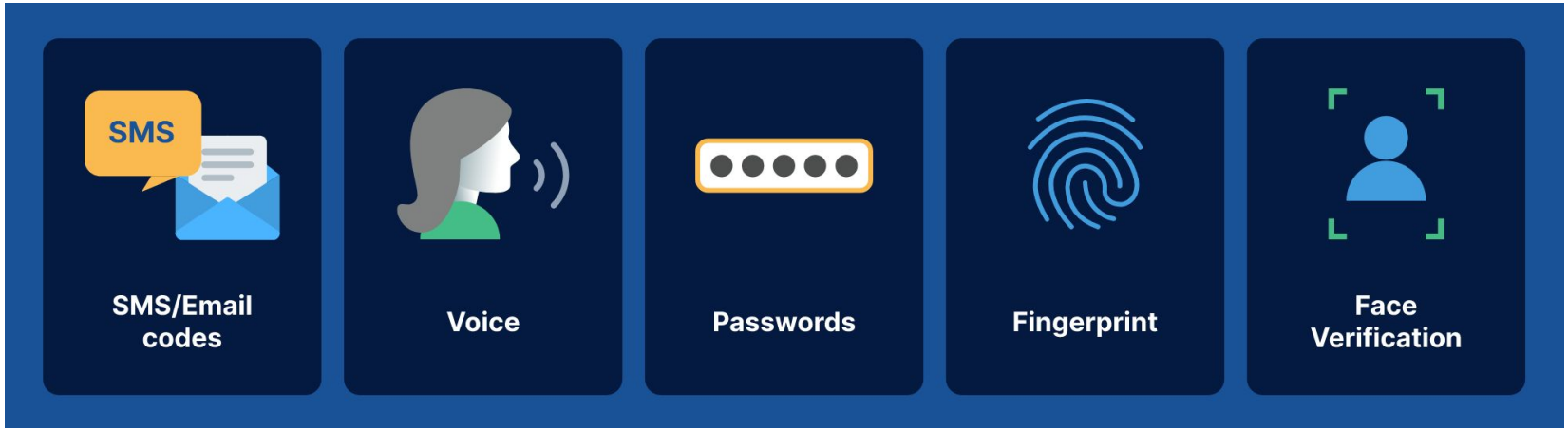
Title : **Keystroke.io**

Team bio : Final year CSE students from M S Ramaiah Institute of Technology, Bangalore

Date : 20/09/2022

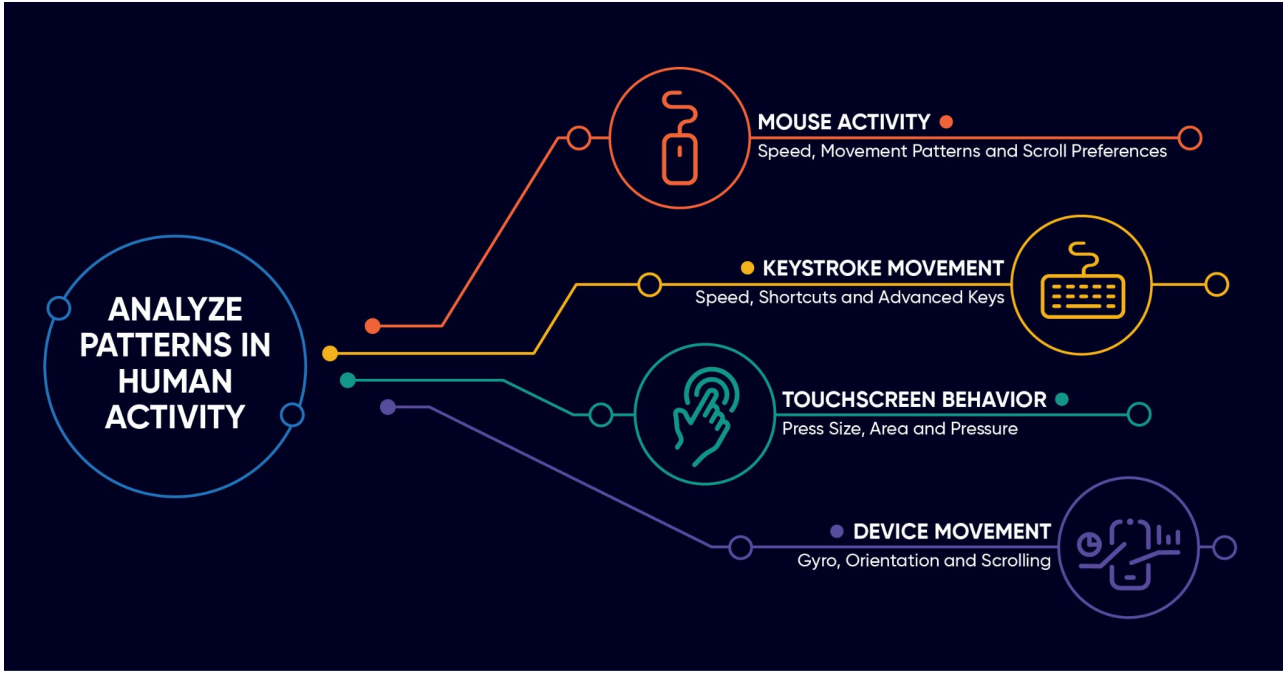
Alternate Authentication (Use case - 6)

Traditional Authentication Mechanisms-



Username and password pairs as authentication factors are as weak as they are ubiquitous. Usernames and passwords can be "phished," stolen, discovered, and cracked in a number of ways. Therefore, a neuro-physiological approach would enhance the current authentication methods compared to only facial recognition or fingerprints which can be cracked in number of ways.

Behavioral Biometrics



Behavioral biometrics analyzes a user's digital physical and cognitive behavior to distinguish between cybercriminal activity and legitimate customers.

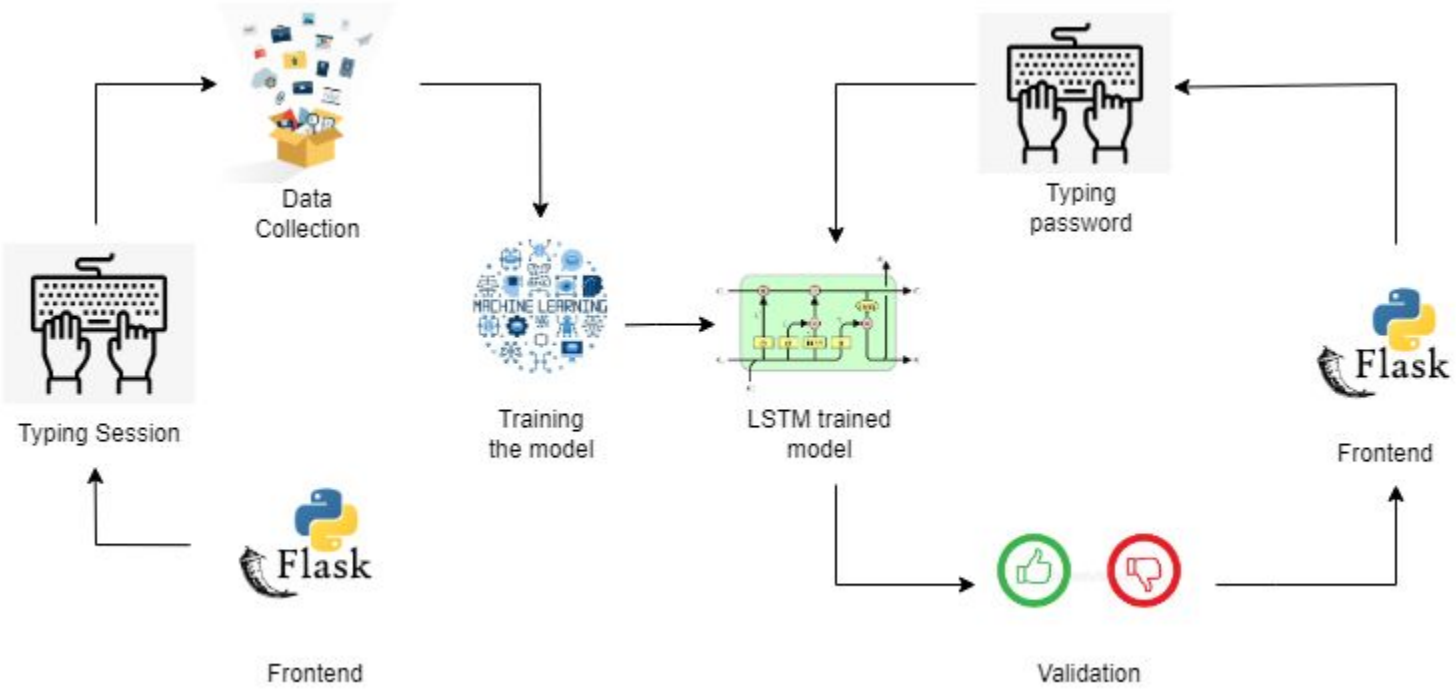
Problem Statement

More than ever before the Internet is changing computing as we know it. Unfortunately, with these advances in technology comes increased chances of malicious attack and intrusion and have unveiled new threats to computer system security.

As we press into the twenty-first century, Digital banking has become the single most effective channel for financial institutions to drive growth, increase revenue and attract new customers. However, advanced safeguards against fraud and impersonation, as well as more foolproof measures against unauthorized access to computer resources and data are now being sought.

We present Keystroke.io, a safeguard mechanism which authenticates access by recognizing certain unique and habitual patterns in a user's typing rhythm.

High-level Implementation



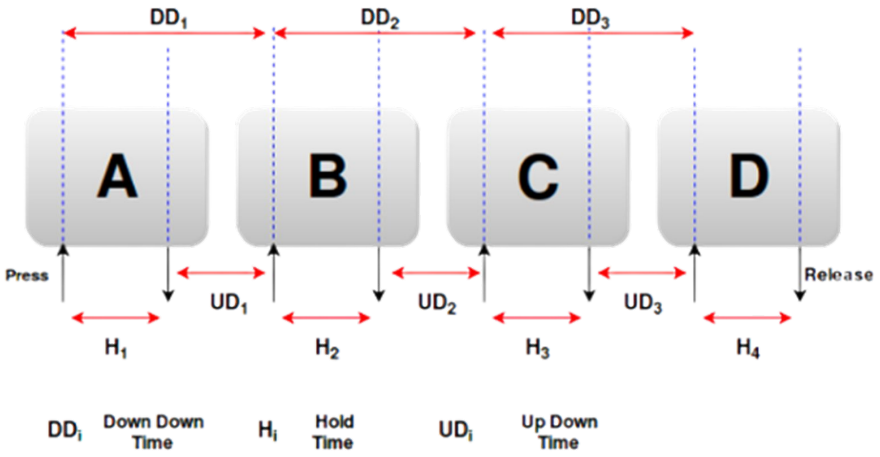
Features

The following are the features which would be extracted during the data collection process:

Hold time (dwell time):
The duration for which the key is held down,i.e. the time between pressing key1 and releasing key1.

Down-Down time:
The time between the pressing key1 and pressing key2.

Up-Down Time:
The time interval between pressing key1 and releasing key2.



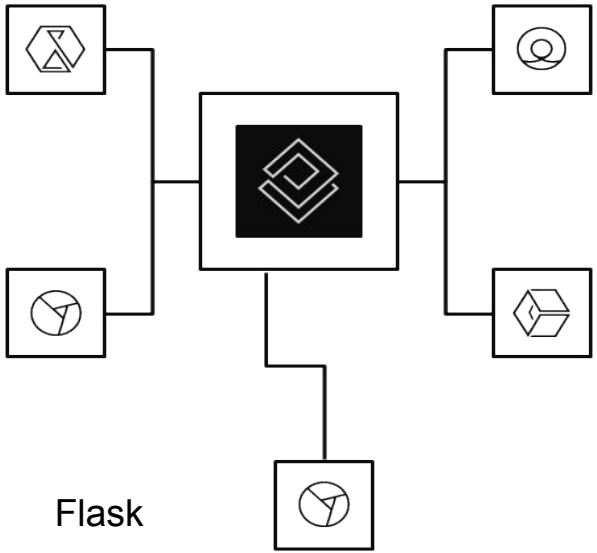
Technology Stack

PyQt5

PyQt is a Python binding for Qt, which is a set of C++ libraries and development tools providing platform-independent abstractions for graphical user interfaces (GUIs)

Microsoft Azure & MySQL

For storing dataset and training the model



Tensorflow

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks

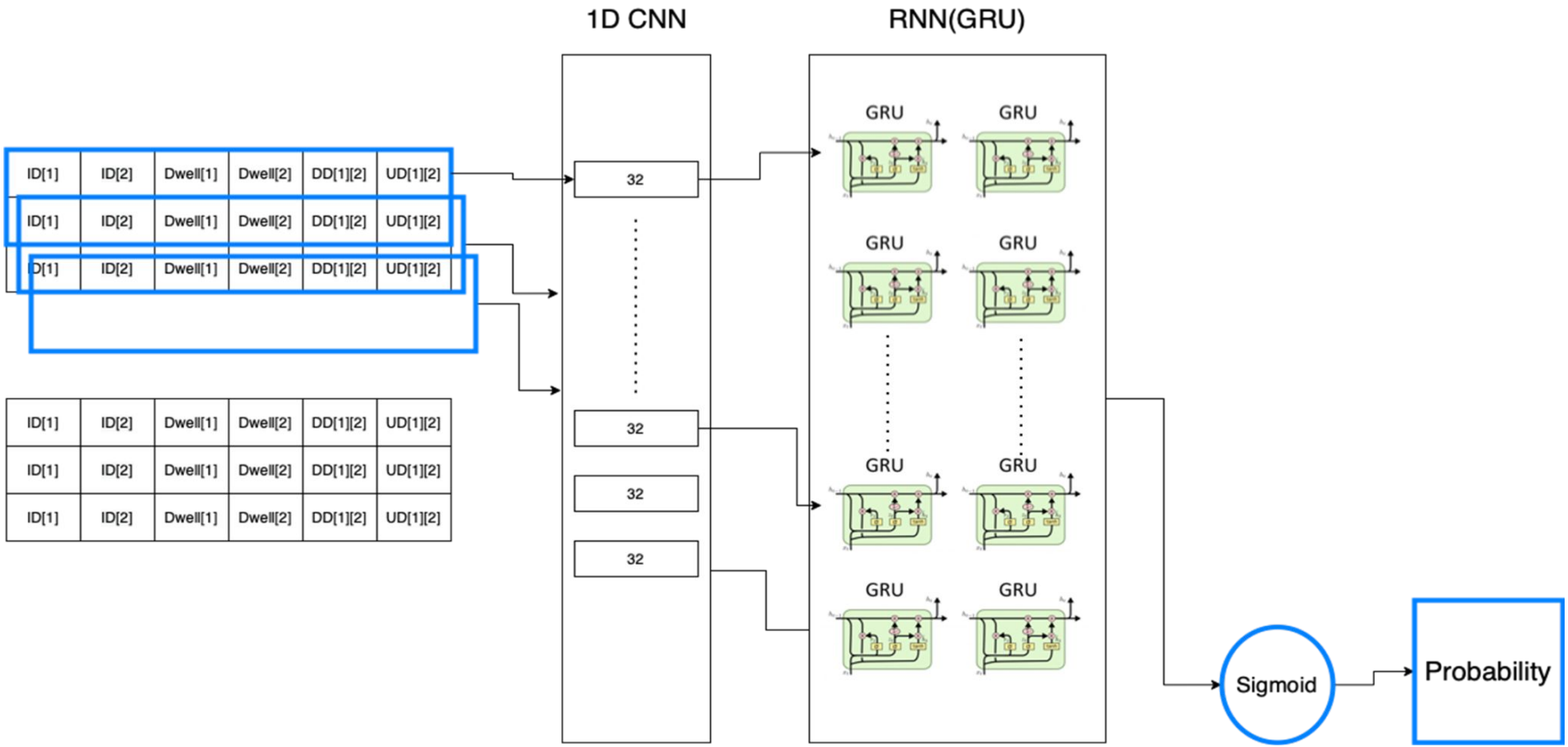
Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics

Flask

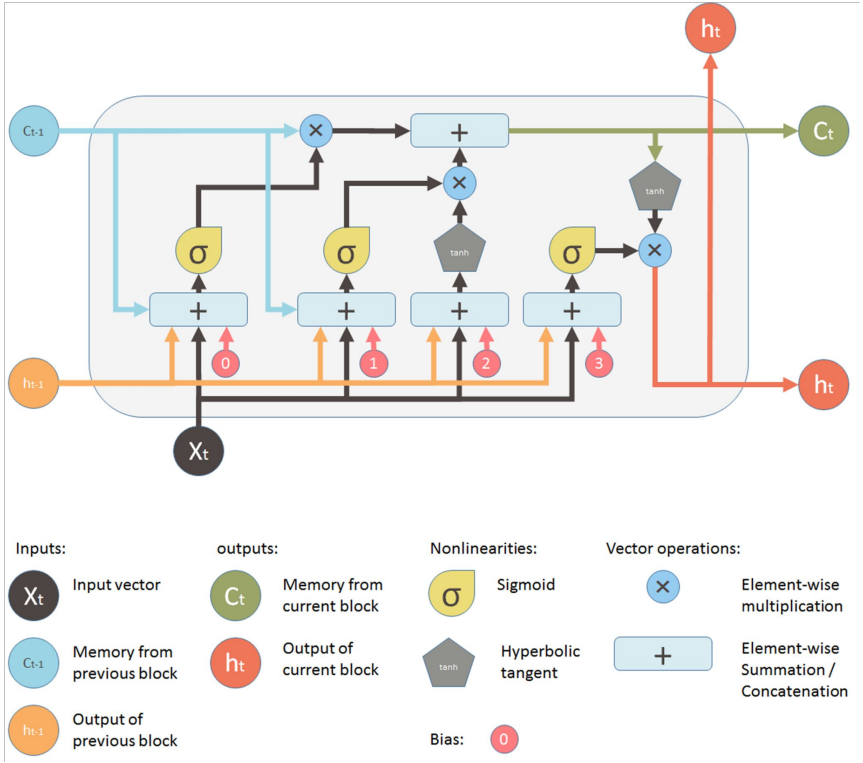
A web application framework written in Python

Architecture

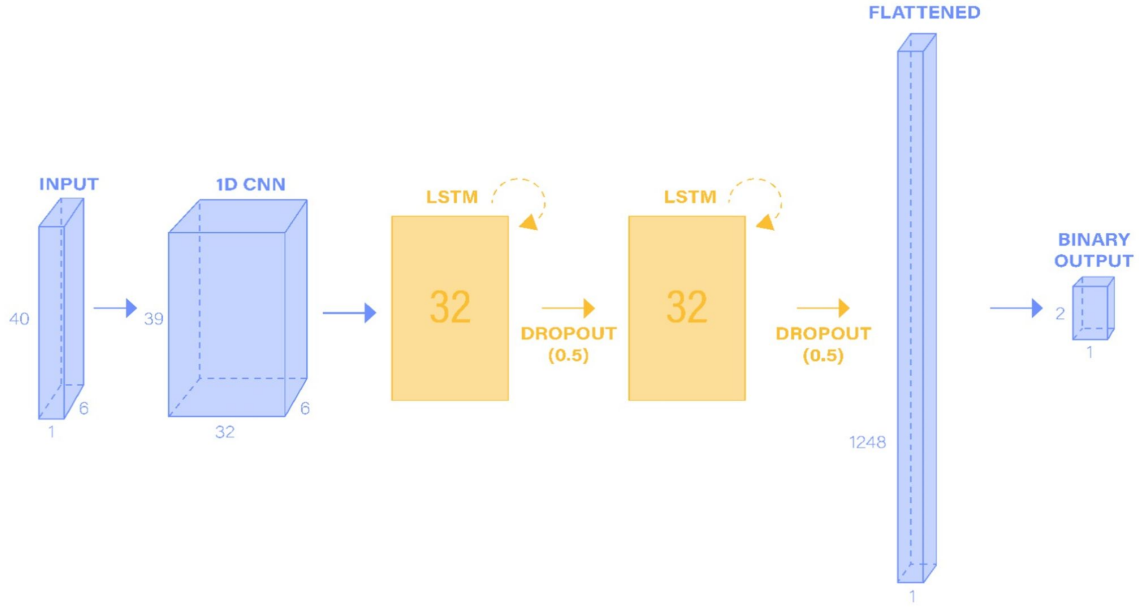


ML model - LSTM

- LSTM networks are type of Recurrent neural networks that can learn order dependence.
- It is composed of multiple gates.
- Each gate has a single layered Neural Network connected to the next last cell.
- It addresses the issue of RNN long-term dependency.
- RNN does not provide an efficient performance as the gap length rises. The LSTM may keep information for a long time by default.
- Used for time-series data processing prediction, and classification.

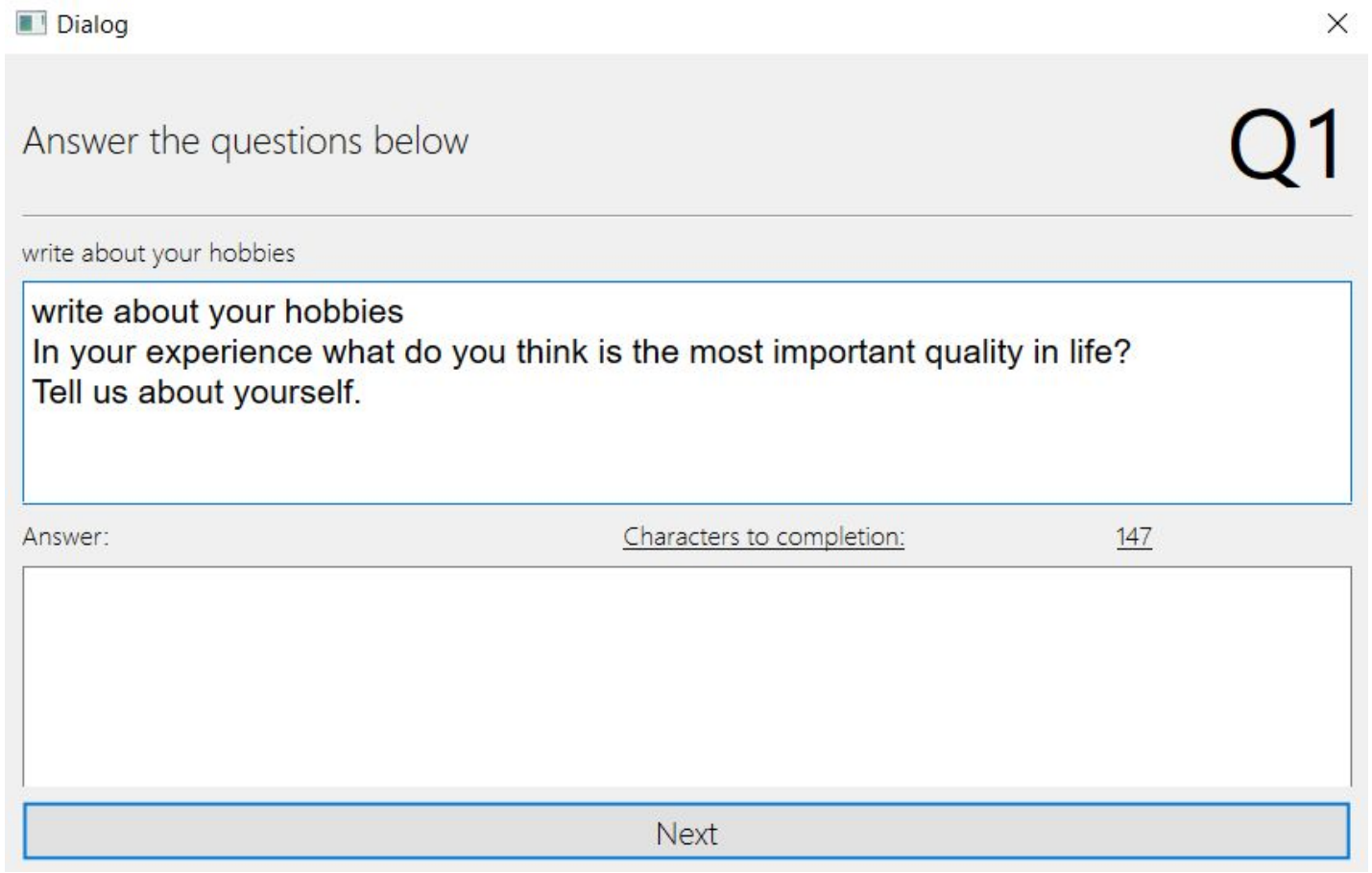


1D 2LSTM



Window size	Loss	Accuracy	Validation loss	Validation accuracy
20	0.2507	0.8927	0.2335	0.9001
30	0.1092	0.9558	0.1043	0.9561
40	0.0757	0.9718	0.0824	0.9707
50	0.0211	0.9927	0.0188	0.9957

Data collection interface

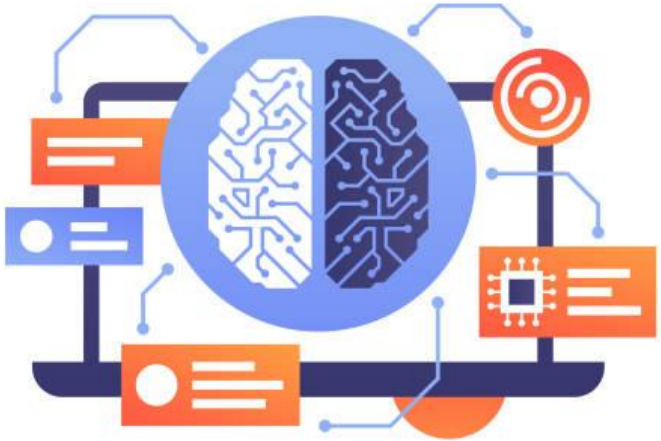


Stored data

email	input
54667865775	[(20, 72, 0.1733405590057373, 0.10733509063720703, 0.3622567653656006, 0.18891620635986328), (72, 20, 0.10733509063720703, 0.14170074462890625, 0.08950090408325195, -0.017834186553955078), (20, 69, 0....
54667865775	[(20, 72, 0.18693280220031738, 0.09209227561950684, 0.27198028564453125, 0.08504748344421387), (72, 20, 0.09209227561950684, 0.18463969230651855, 0.07603240013122559, -0.01605987548828125), (20, 69, 0...
54667865775	[(20, 69, 0.171600341796875, 0.13174104690551758, 0.40543103218078613, 0.23383069038391113), (69, 20, 0.13174104690551758, 0.16679000854492188, 0.14616727828979492, 0.014426231384277344), (20, 86, 0.1...
54667865775	[(20, 87, 0.1540677547454834, 0.09438347816467285, 0.14645695686340332, -0.007610797882080078), (87, 20, 0.09438347816467285, 0.14756226539611816, 0.16701841354370117, 0.07263493537902832), (20, 72, 0...
54667865775	[(20, 73, 0.2108173370361328, 0.08055734634399414, 0.3004474639892578, 0.089630126953125), (73, 8, 0.08055734634399414, 0.13869738578796387, 0.4104583263397217, 0.32990097999572754), (8, 79, 0.1386973... Original length 15879
54667865776	[(20, 84, 0.19189667701721191, 0.1067957878112793, 0.3057372570037842, 0.11384057998657227), (84, 20, 0.1067957878112793, 0.15377068519592285, 0.13530850410461426, 0.02851271629333496), (20, 72, 0.153...
54667865776	[(20, 87, 0.14860796928405762, 0.10635924339294434, 0.19494318962097168, 0.04633522033691406), (87, 20, 0.10635924339294434, 0.16165375709533691, 0.09315657615661621, -0.013202667236328125), (20, 72, ...
54667865776	[(20, 89, 0.2004392147064209, 0.10097575187683105, 0.49045705795288086, 0.29001784324645996), (89, 20, 0.10097575187683105, 0.16101956367492676, 0.12842226028442383, 0.027446508407592773), (20, 79, 0....
54667865776	[(20, 89, 0.18639612197875977, 0.11510133743286133, 0.4722898006439209, 0.28589367866516113), (89, 20, 0.11510133743286133, 0.16070890426635742, 0.10483384132385254, -0.010267496109008789), (20, 79, 0...
54667865776	[(20, 84, 0.18563532829284668, 0.0849001407623291, 0.25797224044799805, 0.07233691215515137), (84, 20, 0.0849001407623291, 0.15724468231201172, 0.10068798065185547, 0.015787839889526367), (20, 72, 0.1...
54667865777	[(20, 83, 0.20749640464782715, 0.10120415687561035, 0.4076526165008545, 0.20015621185302734), (83, 20, 0.10120415687561035, 0.15362930297851562, 0.11365437507629395, 0.012450218200683594), (20, 79, 0....
54667865777	[(20, 73, 0.21127867698669434, 0.09474921226501465, 0.2636141777038574, 0.052335500717163086), (73, 20, 0.09474921226501465, 0.15621423721313477, 0.13132262229919434, 0.03657341003417969), (20, 32, 0....
54667865777	[(20, 72, 0.17711949348449707, 0.08550047874450684, 0.2137773036956787, 0.03665781021118164), (72, 20, 0.08550047874450684, 0.14603519439697266, 0.09757471084594727, 0.01207423210144043), (20, 69, 0.1...
54667865777	[(20, 73, 0.19798707962036133, 0.08600020408630371, 0.1907331943511963, -0.007253885269165039), (73, 20, 0.08600020408630371, 0.1426098346710205, 0.14249086380004883, 0.05649065971374512), (20, 32, 0....

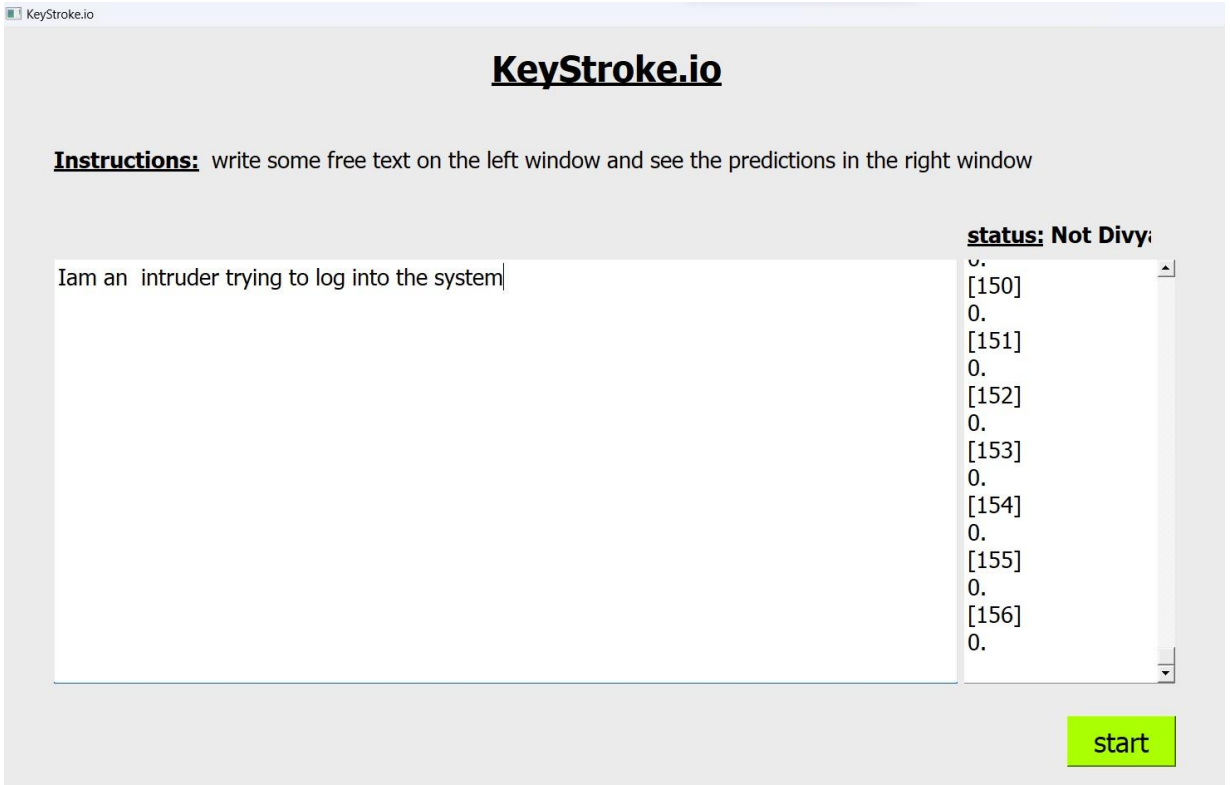
Training the model

- ➔ In order to gather data, the user responds to a sequence of 5 questions as shown in the previous slides.
- ➔ On the **Azure Cloud, MySQL** will be used to store the data.
- ➔ The LSTM model gets trained in around 20 minutes (on Colab with a **16GB GPU**), after which we can start validating the user as a legitimate user or as an intruder based on keystrokes.

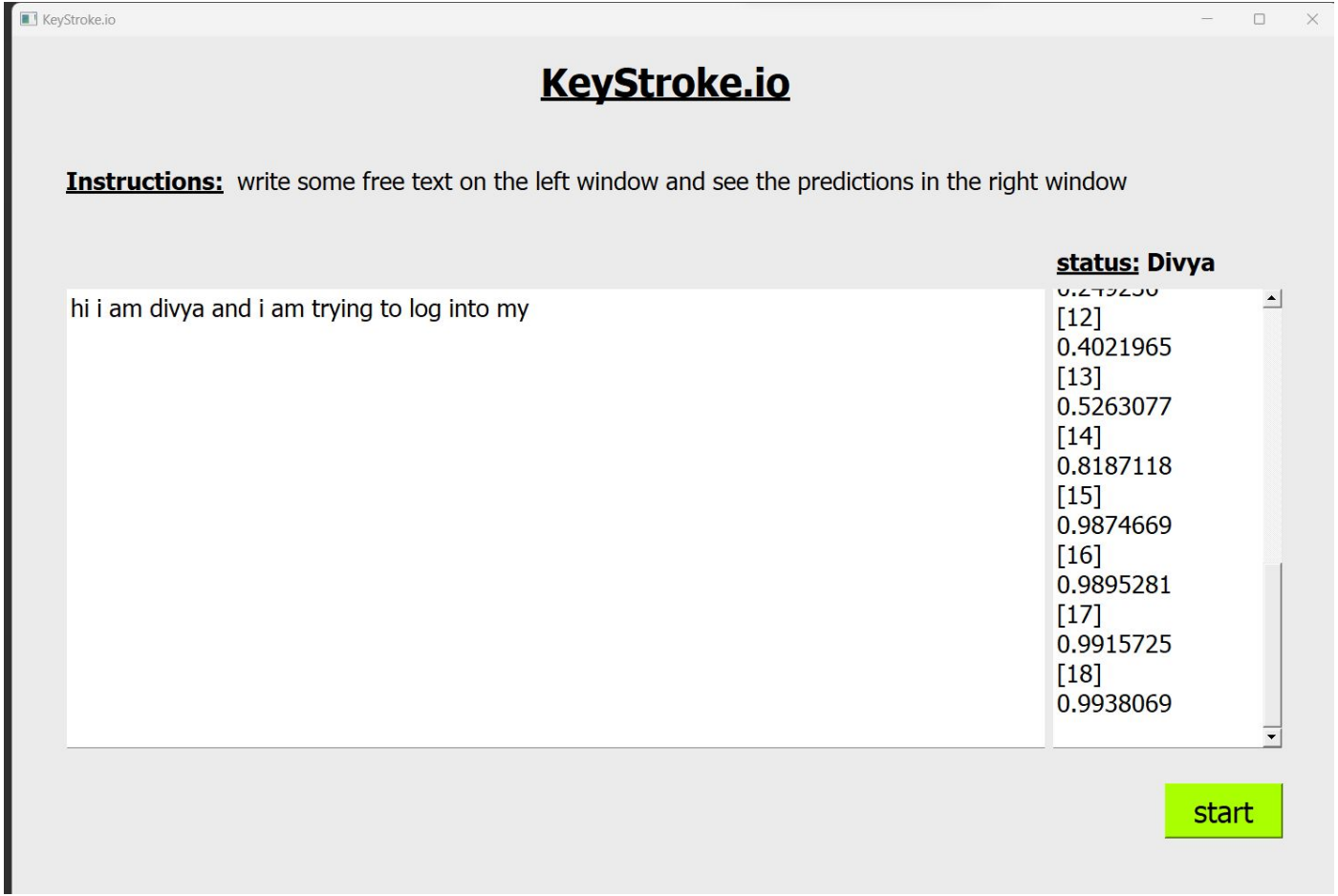


Validation

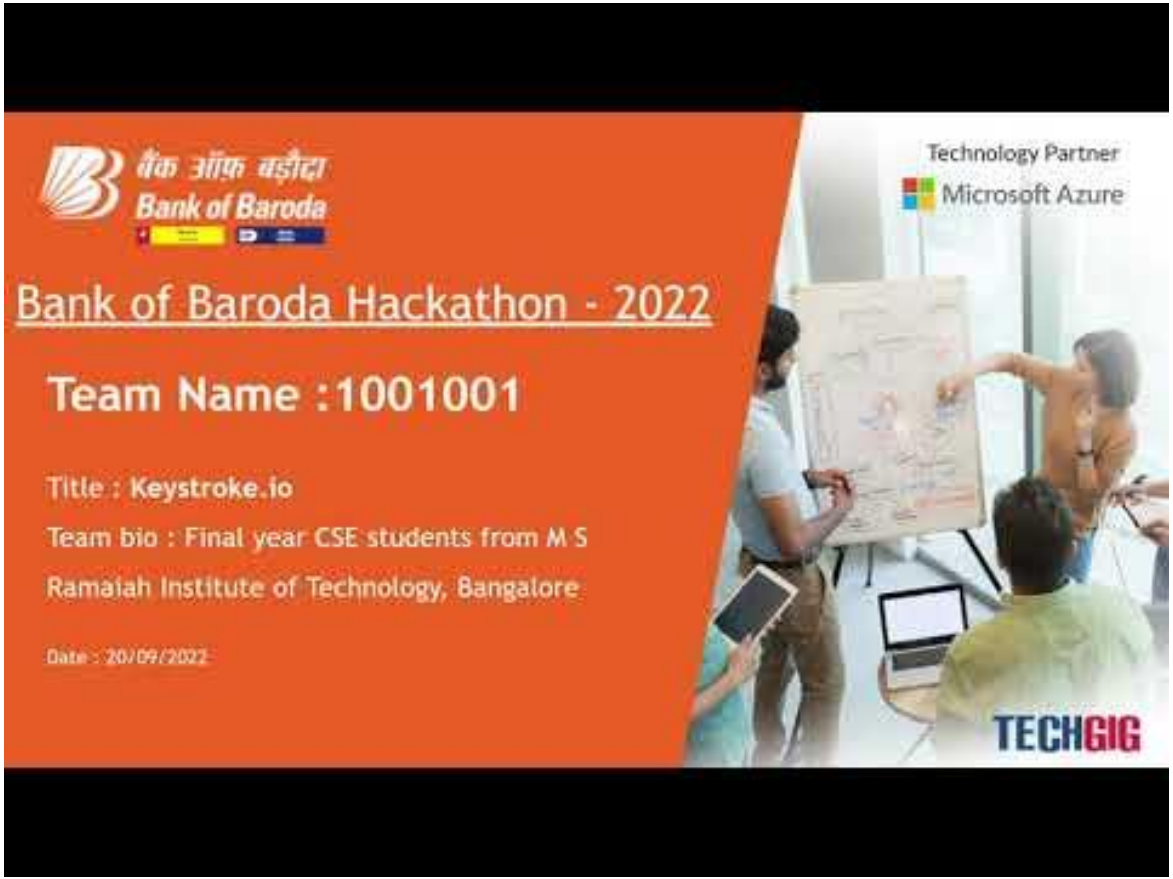
Results when typed by an intruder



Results when typed by a legitimate user



Demonstration



Why Keystroke.io?

- ➔ The behavioral biometric of Keystroke Dynamics uses the manner and rhythm in which an individual types the characters on a keyboard.
- ➔ The keystroke rhythms of a user are measured to develop a unique biometric template of the user's typing pattern for future authentication.
- ➔ It works passively in the background of a user web session to monitor the parameters and validate the user as either a legitimate user or an intruder.

Societal Impact-

- ➔ Stolen phones/e-gadgets cannot be used which reduces misuse of banking information.
- ➔ ATM frauds due to card skimming can be nullified.
- ➔ Stolen credentials become useless.
- ➔ Prevents User Substitution.
- ➔ Employee productivity and accountability.
- ➔ Personal Privacy Protection.
- ➔ Helps create a world of TRUST and EASE.

Scope of Work

Keystroke.io can be used by banks, surveillance facilities and anyone for their personal use.

- **Banks:** Banks can implement our idea in their web/mobile applications to track the behavioral patterns of their customers to continuously authenticate the identity of their users apart from the usual authentication.

- **Surveillance and Security Facilities:** To operate the security equipment and computers, a site under monitoring needs a specialized control room. To monitor the systems and make sure an unauthorised individual is not occupying the control room, it is crucial to constantly confirm the authorized personnels' identification.

- **Personal use:** Users who want to keep their data secured and ensure their privacy can use keystroke.io to trigger a response in case a third person accesses their phone or computer system.

Future Scope and scalability

- ➔ The suggested Proof-of-Concept is restricted to online authentication using desktops and laptops with physical keyboards, but it can potentially be extended to include mobile devices and other touch screens.
- ➔ The model may be taught a large number of patterns progressively using continuous learning while retaining its memory of knowledge acquired from earlier patterns.
- ➔ The recommended PoC tracks simply the user's keystroke dynamics for authentication, but it can be expanded to capture other behavioural patterns like mouse activity, device movement and touch screen behaviour.

TECHGIG

WELCOME to the new world of AUTHENTICATION

Thank You

Team member names -

1. K Divyasri
2. Kushagra Gupta
3. Lakshya Khandelwal
4. Praneeth Shetty