SYNOPSIS

ON

"Akashic: The Smart Resource Management System"

Submitted in

Partial Fulfillment of requirements for the Award of Degree

of

Bachelor of Technology

In

Computer Science and Engineering

(Artificial Intelligence & Machine Learning)

By

(Project Id: 22_AIML_2A_20)

Ayush Agnihotri (2101641530049)
Anish Singh (2101641530027)
Ashutosh Pandey (2101641530046)
Abhraneel Singh (2101641530009)

Under the supervision of
Yashi Rastogi
(Assistant Professor)

PSIT
Kanpur

Pranveer Singh Institute of Technology.

Kanpur - Agra - Delhi National Highway - 19
Bhauti -Kanpur - 209305.
(Affiliated to Dr. A.P.J. Abdul Kalam Technical University)

1. Introduction

As we all know that, it won't be wrong to say that our life on digital media these days rely on passwords. Hence, while the purpose of passwords is to protect your account, we have reached that point in time where in return we are supposed to give the right protection to our passwords as well. But we are failing at that! So, our project **Akashic** will be working for its solution. As, according to the research conducted by HYPR in the last two and a half years all around the United States, India and Canada, the statistics show that internet users have been absolutely careless in managing their passwords; both in personal life and at work. Diving deep into the study, 78% of the 500 respondents accepted that they had to reset the passwords for one personal account in the last 90 days of their calendars. 57%, on the other hand, claimed that they had to go for password reset of their work account.

Another worrying statistic revealed that 65% of the users rely on digital and physical lists when it comes to managing their password and despite the heavy advertisements, only 30% had password managers installed on their devices that also can be harmful for their Data Privacy. When users were obliged to reset the password, 49% of the respondents said that they used the same phrase with little changes here and there so that the password should remain easy to remember. Going even worst, 35% of the users keep their passwords in unprotected files such as Excel spreadsheets or text documents stored onto their computers. So, our Smart Login System will be a smart solution for this smart world.

So, in this smart era of Fast-Moving world, everyone needs a smart login system to prevent their hesitations of typing credentials long and hectic password which is sometimes tough to remember and there are chances of breaching of ID and Passwords. We are proposing a solution in the form of new, advance and secure facial recognition login system. That will determine the face of the account owner and will provide a secure login to the user panel by simply asking a security pin for more security purposes. In case if the server is unable to determine the face due to some issues, then the account owner will be having a secondary method to recover his/her account.

2. Project Objective

The main aim of this project is to resolve the issues of remembering hectic passwords and protect it from data breaching. And our system will also protect the user's panel from being getting login by any unauthorized person. As the main objective of our project is to provide such a protected ML Based Login System that is unhackable as, it will require the face of the account owner to gain access to their accounts.

We have applied efficient pre-processing techniques including label-encoding and normalization that improve the accuracy of the models. Further, using various feature selection approaches, we have identified and prioritized a number of risk factors. Extensive experiments have been conducted to analyze the performance of the model using two different datasets. Our model is compared with some recent study and the results show that the proposed model can provide better accuracy of 2.71% to 13.13% depending on the dataset and the adopted ML algorithm. Finally, a machine learning algorithm showing the highest accuracy is selected for further development. We integrate this model in a web application using python flask web development framework. The results of this study suggest that an appropriate preprocessing pipeline on clinical data and applying ML-based classification may predict diabetes accurately and efficiently.

3. Feasibility Study:

- **Technical Feasibility**
  I. Our login system is ML based so, there will be very less chances of Login Failures (Login Failures can only occur under the circumstances of usage of old versions of browsers).
  II. None type of data can be breached while using our login system.
  III. There is no way for any type of Hacker to login through our system as it requires the person to be there.

IV.   We will be using ML to recognize whether the real person is there or its image is being shown to the camera.

V.   The whole project will be completely responsive for all frames devices.

VI.   Interactive UI.

VII.   No lag in mobile devices.

VIII.   Also works in low internet speed.

- **Legal feasibility**

  Our Login System is fully Legalized as it doesn't violate any Illegal, Unauthorized or Irrelevant Activity. It doesn't violate any copyright and doesn't intend to harm any authorities, person or government.

- **Scheduling Feasibility**

  To complete the project on time we have divided our project work in multiple small milestones, and our all team members are assigned with the particular task which they are used to do so.

	3 - 7 Oct	10 - 14 Oct	17 - 21 Oct	24 - 28 Oct	31 Oct - 4 Nov	7 - 11 Nov	14 - 18 Nov	21 - 25 Nov	28 - 30 Nov
	M T W T F	M T W T F	M T W T F	M T W T F	M T W T F	M T W T F	M T W T F	M T W T F	M T W
Resource Planing and Work Division									
Assignment of Work									
Web-Frontend Development									
Skeleton Code									
Web-Page Desgining									
Adding Responsivness									
Web-Backend Development									
Routing the Path									
Integrating ML									
Establising the Connection frontend									
Web Security									
Data Collection									
Raw Data Collection									
Backend Development									
Module Training									
Increasing Accuracy									
Setup Connection with Web Backend									

- **Operational Feasibility**

  High degree of any accuracy will be obtained for the facial recognition as per the data made and used to train the machine.

4. Methodology/ Planning of work

- **Dataset collection** – It includes data collection and understanding the data to study the hidden patterns and trends which helps to predict and evaluating the results. Dataset carries more than 1500 rows i.e., total number of data and 10 columns i.e., total number of features. Features include Facial recognition, Gender determination, age recognition etc.
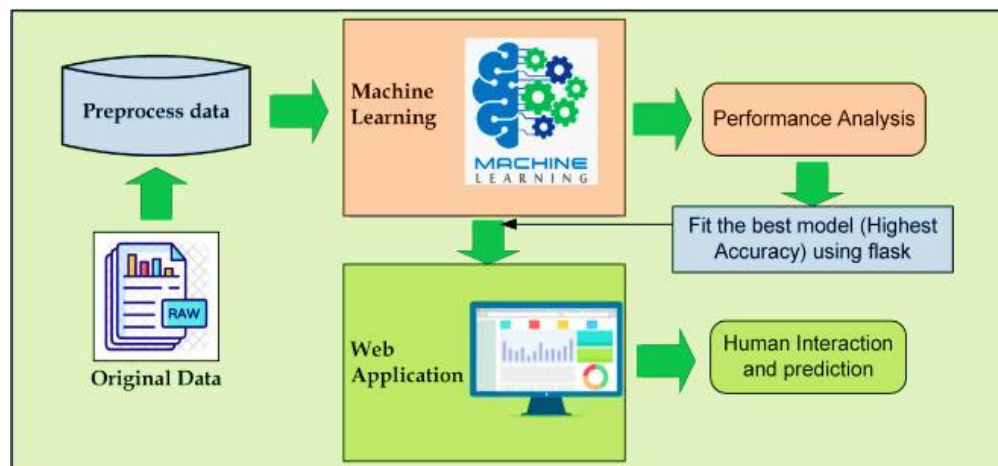
- **Data Pre-Processing** – This phase of model deals with models training and handling of inconsistent data in order to get more accurate and precise results like in this. Dataset ID is inconsistent so we dropped the feature. This dataset doesn't contain missing values. And we pretrained all our datasets to gain maximum accuracy and persistency**.**

- **Design and implementation of classification model** – In this research work, comprehensive studies are done by applying different ML classification techniques like DT, KNN, RF, NB, LR, SVM

- **Feature selection** – Pearson's correlation method is a popular method to find the most relevant attributes/features and resemblance. The correlation coefficient is calculated in this method, which correlates with the output and input attributes. The coefficient value remains in the range by between −1 and 1. The value above 0.5 and below −0.5 indicates a notable correlation, and the zero value means no correlation.

- **Splitting of data** – After data cleaning and pre-processing, the dataset becomes ready to train and test. In the train/split method, we split the dataset randomly into the training and testing set. For Training we took more than 2000 samples and for testing we took 500+ samples.

Total Number of Examples

Training Set	Test Set
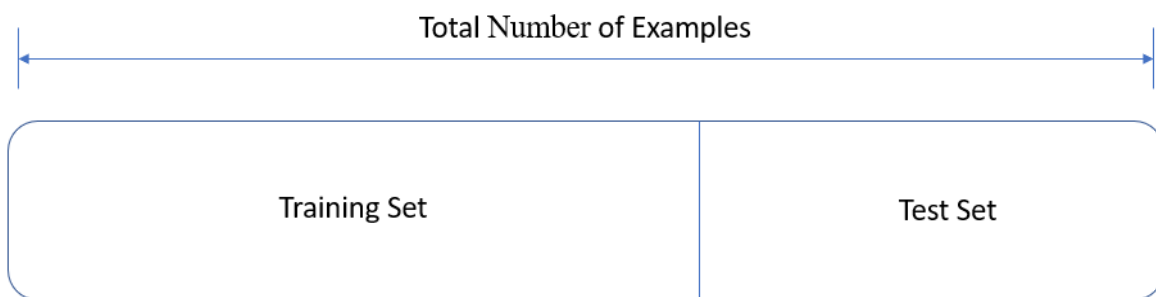
- **Scaling and Normalization** – We performed feature scaling by normalizing the data from 0 to 1 range, which boosted the algorithm's calculation speed. Scaling means that you're transforming your data so that it fits within a specific scale, like 0-100 or 0-1. You want to scale data when you're using methods based on measures of how far apart data points are, like support vector machines (SVM) or k-nearest neighbors (KNN). With these algorithms, a change of "1" in any numeric feature is given the same importance.

- **Machine learning classifier** – We have developed a model using Machine learning Technique. Used different classifier and ensemble techniques to predict diabetes dataset. We have applied SVM, LR, DT and RF Machine learning classifier to analyze the performance by finding accuracy of each classifier All the classifiers are implemented using scikit learn libraries in python. The implemented classification algorithms are described in next section.

5. Tools/Technology Used:

5.1 Hardware Requirements

- Camera
- A stable internet connection
- Central Processing Unit (CPU) — Intel Core i5 6th Generation processor or higher. An AMD equivalent processor will also be optimal.
- Graphics Processing Unit (GPU) — NVIDIA GeForce GTX 960 or higher.
- SSD
- RAM — 8 GB minimum, 16 GB or higher is recommended
- Operating System — Ubuntu or Microsoft Windows 10 or higher is recommended.

5.2 Software Requirements

- Python 3.10.7
- NodeJS 18.9.0
- Anaconda
- CUDA Toolkit & cuDNN
- VPS
- OpenCV
- Database
- Cryptography
- TensorFlow
- Keras
- Pandas

6. References:

- "Study", by Daniyal Malik, December 23, 2019 [Online] Available: https://www.digitalinformationworld.com/2019/12/new-password-study-finds-78-of-people-had-to-reset-a-password-they-forgot-in-past-90-days.html

- "Content moderation using machine learning: a dual approach", by Jen Person, Developer Advocate, August 19, 2022 [Online] Available: https://blog.tensorflow.org/2022/08/content-moderation-using-machine-learning-a-dual-approach.html

- "Portrait Depth API: Turning a Single Image into a 3D Photo with TensorFlow.js", by Ruofei Du, Yinda Zhang, Ahmed Sabie, Jason Mayes, Google on May 10, 2022 [Online] Available: https://blog.tensorflow.org/2022/05/portrait-depth-api-turning-single-image.html

- "Body Segmentation with MediaPipe and TensorFlow.js", by Ivan Grishchenko, Valentin Bazarevsky, Ahmed Sabie, Jason Mayes, Google, on January 31, 2022 [Online] Available: https://blog.tensorflow.org/2022/01/body-segmentation.html

- "Preparing Your Dataset for Machine Learning" on March, 2021 [Online] Available: https://www.altexsoft.com/blog/datascience/preparing-your-dataset-for-machine-learning-8-basic-techniques-that-make-your-data-better/