**Chapter 1**

# Introduction

In this era of digitization, technology and the way of living is changing fast. Technology is now becoming a new usual way of living. About 5 years ago, in India, there was less awareness of the internet and its uses, but today every other person has access to it. The problem is not about using this technology but having enough awareness about its security.

Security and privacy are two major concerns in our day and age. As most of the things are moving online, the vulnerability increases with the ease of access and usage that it provides. We have online banking, study programs, online food ordering apps, government services, online payments. Moreover, the use of social media has increased with the usage of internet services. There is hardly any person who does not hold a social media account. After all these advances in technology, internet services and applications, we still need more security and privacy. Even though security is also kept in mind while launching new services or products, the question still stands. Is that security sufficient?

We have all of our information, including our financial, personal photographs, personal text messages, location information and more, stored on our phones. Whenever we connect to the internet, all such information is at stake. This information can be used for many malicious activities such as accessing our bank accounts, tracing our location, tracking our user history. They get to know what we have purchased and give personalized advertisements whenever you search for some product online or on any social media site. Hence, in this advancing age of smartphones, it is essential to know how safe your phone is along with your laptop.

These all are the vulnerabilities of the information stored on the internet. We all know that once something is up on the internet, it may always stay up there. There are very few people in the world who understand how the online systems work and how vulnerable it could be. What type of data can prove to be very harmful if put on the internet? There are people from this small percentage too who do not care about how other people use the data. This shows that even the aware users ignore the privacy terms and conditions and blindly click accept whenever they require to use a service. So, leave alone the unaware users who just follow these trends that are going on in the world.

These are the legal ways your private information can be taken and maybe even sold to the data scientists who require such data or to the advertisements seller to get the target audience. There are many ways in which your privacy and security can be compromised.

Hacking is one type of breach into your system from remote. It depends on how a loophole in the code of service can be used to the advantage and steal any information that the hacker seems fit to benefit from. This is an illegal manner of taking the information. This area itself has many types.

But what if someone can do it the legal manner. Just like how any person simply clicks “OK” while installing an application on the phone and gives all the permission it asks for blindly. What if someone takes your media storage completely online without you noticing. Ask yourself this question that how many times have you read the complete terms and condition before clicking on “Accept” and give permissions to the camera, microphone, location and storage to any app that you have installed.

Let’s say that a person has not given any permission to the application, but by coming into physical contact of the phone he/she can unlock the phone to transfer any data he/she wants. This can be possible if the person knows the PIN code of the phone. Unlocking can be very tough if the person does not know the actual PIN as the PIN can be of different lengths. It may lock the phone for a specific amount of time until it allows the user to retry. But using all the possible PINs to unlock the phone is time-consuming, and the probability that you get the right PIN without permanently locking the phone is very low.

With the advancing in the virtual reality (VR) and the importance given to user experience with the phone we normally see the functionalities like auto-rotate screen, pedometer being used by almost every other user and if noticed, these functions are used without any permissions asked the user. When you enable auto-rotate screen, then it keeps track of the orientation of the phone in 3SD space and rotates it in either landscape or portrait mode. Even while playing games like car racing, they too use these motion sensors which requires no permissions to be asked and enhances the gaming experience. A smartphone which gives no permissions to any app can still be vulnerable to this type of attack that takes data from motion sensors of the phone.

Focusing specifically in the direction of exploiting the motion and position sensors of the android smartphones in the background, we aim to record the sensor data, using a gyroscope (angular velocity), accelerometer (linear acceleration) and gravity, from the phone and predict the numeric PIN used to lock the phone. These sensors require no permissions when accessed and can easily leak information to the background service. This background service can upload such sensitive data from the background itself to remote storage or can be taken in real-time database form. Internet connection to a mobile phone is a very common thing nowadays, and any of these remote access will not be a limitation to the attacker.

It is observed that whenever the screen is touched on specific portions, a spike is created in the motion and position sensors. Position and motion sensors tell us about the orientation and movement of the phone in 3D space. Through extracting such information of the motion of the phone when the screen is touched and released these patterns can be analyzed and used to know which part of the screen was touched, we can crack the PIN of a phone.

To do so, we log the data from sensors such as accelerometer, gyroscope and gravity, to know the orientation and movement of the phone in 3D space. This raw data is first processed so that it is noise-free and contains all the required features to recognize the pattern. This processed file is then used as an input dataset to feed into a supervised multiclass machine learning (ML) model. As the numeric PINs contain numbers from (0-9), we will be using multiclass classification for 10 classes which represent the keys pressed, or in other words, tell which part of the screen was touched where that particular number key will be present when the user is entering the PIN.

Our method will not be limited to the length of the PIN and will be used to predict individual digit from the complete sensor data so that it is flexible to any PIN length. It can also be seen that the PIN is not only used for unlocking the phone but also for online transactions. There are PINs to be set for online payment apps and can be scaled up to be used for this kind of exploitation also.

We are going to introduce two methods in which we will be explaining how the data is collected and how the data is being processed before feeding it to the machine learning model as input. We have tried different configurations for each model used and have compared each processing method and sensor type data to eliminate the models that do not give accurate results or the sensor type which will not be having the data which is required to get the correct predictions.

For the actual exploitation of the sensor data, we will be building a background service to allow remote access to all such data in such a way that the user is not aware of the background process that is running. The app will store the data locally and also upload on cloud storage. These files are further processed in a similar manner in which the training data is processed. It is then given as input to the trained machine learning model and give out results.

We will be considering the results for 50 4-digit PINs and see by what probability we can predict individual digits of the PIN. We will also compare our results to the related work on this particular topic. 50 PINs contain 200 individual digits, so the probability of predicting on single digit can be calculated and compared with the brute force method of exploitation.

If it is possible to train machine learning models to predict which part of the screen is touched by the user to an extent, then it would mean that it could be extended further to predict the keyboard input also and in turn increase the vulnerability associated to it. This can increase the concerns related to security and can become a security risk level same as the keylogger in a device.

With the increasing usage of smartphones for net banking and other secured services, this exploitation if scaled to a level can then become a high risk to all such applications as no password or OTP will then ever be hidden from the attacker. Even the two-step verification will stand as useless and ineffective against this type of attack.