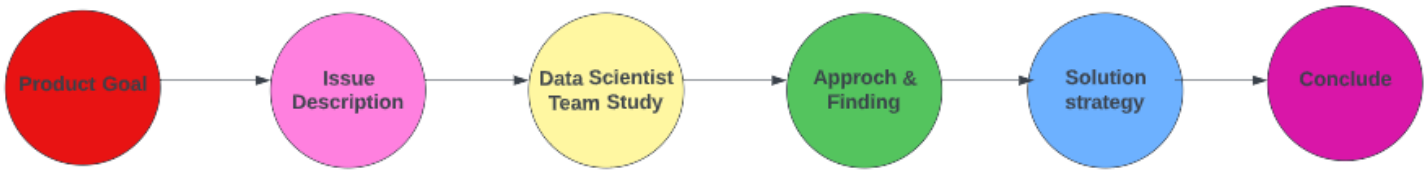
Fraudulent Users Case Study – Rohit Sahu (Freelancing trainee)

Case Study

#X real money gaming APP:

#GameLogo



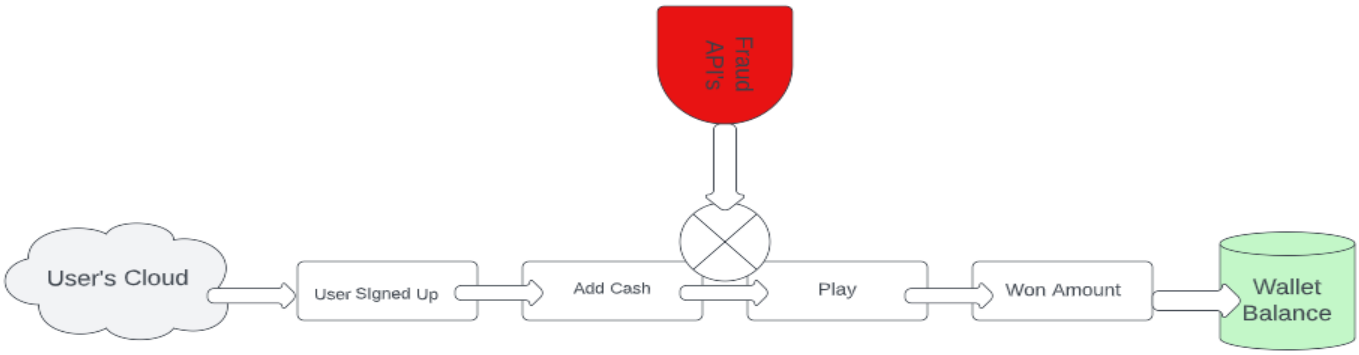
FUNCTIONAL BLOCK DIAGRAM

Product Goal:

Our product major seeks to improve product security and app seal from fraudulent user base as the product seeks to protect genuine users who play the most on the platform and is good for our system and product revenue.

Description of the situation:

WitZeal Corporation– is a gaming brand and has an own-created listed game and product gaming App for android and Ios. In starting of production, OEM wants to sell our product from the Fraud users base and fraud API. when fraud users will land on our product and will create the best high score with the help of other fraud APIs alter apps which are used in the market. When the product will understand what is going on the product anomaly listed users will withdraw payment in our case wallet.



Fraudulent Users Case Study – Rohit Sahu (Freelancing trainee)

- ☐ CEO key scenario-based point:
    - Understand the behaviour of Users to find the characteristics.
    - What types of factors would you want to explore to understand how product anomaly will capture automatically and block the same?
    - X% amount of users landing on our product and out of these users 5% probabilities to found the anomaly users.

Data Scientist Team Study:

Product completed 5-6 years with the Gaming relation and anomaly detection activity are doing manually till date.
For our understanding, we connected with all Data Analysts and collected all the observations on what is going on with the product which is related to the Anamoly users.

According to our analysis and historical data set, 0.39% of users were detected as fraud and other users were put on the whitelisted characteristics.

Anomaly Users Characteristics:

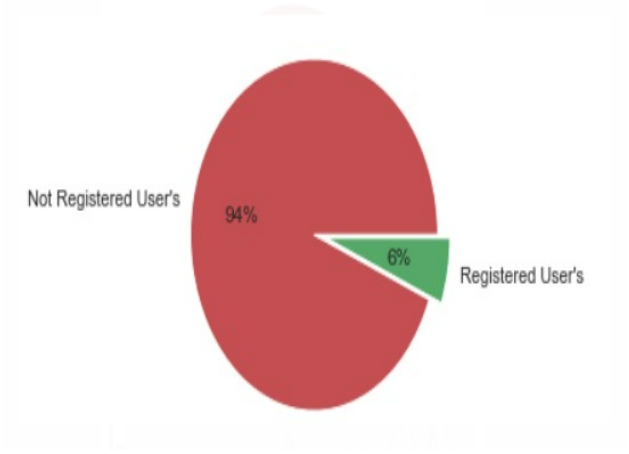
1. A user signed up with temp mail Ids
2. Number not registered
3. Max redemption within 4-5 Days.
4. Didn't add much money to play with compared to Withdraw Wallet money.
5. Focus on the specific game type

We will discuss only 0.39% of users

Out of these all user amount of X% users signed up with a temp mail id and added their own account and played well without capturing the product.

Click to generate the temp Mail Ids

https://temp-mail.org/en/



Mobile Registration– A few X% users are not registered their phone numbers, but they link accounts and withdraw the money when he will win some amount. The very high percentage of users listed in the 3rd base key point means when users were landing on our product than a

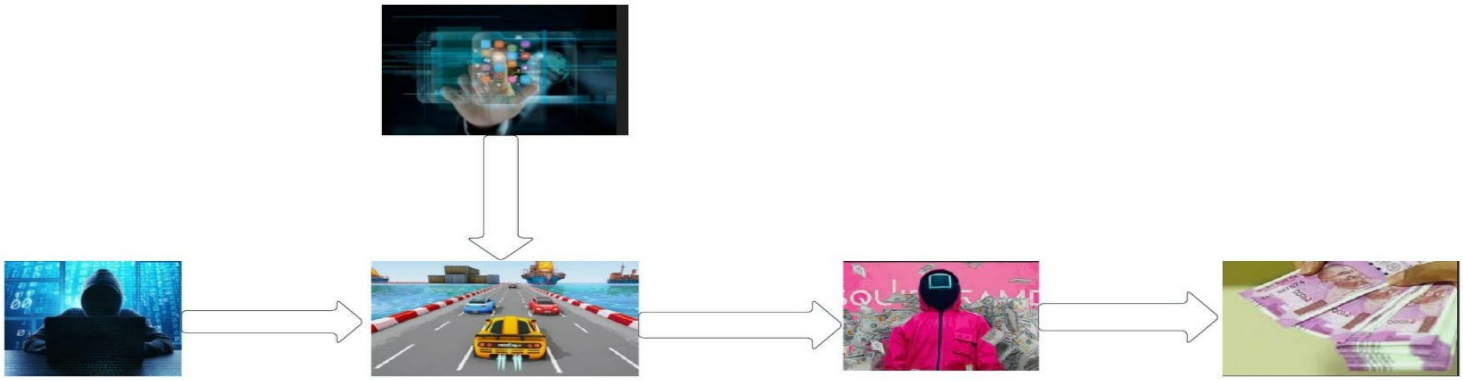
percentage of days retention is not much better than the after 4-5 days (user acquisition)

Redemption– Most users are thinking we will hit the score with the help of some product (product help to override the score call API to change the real score to a fake one).

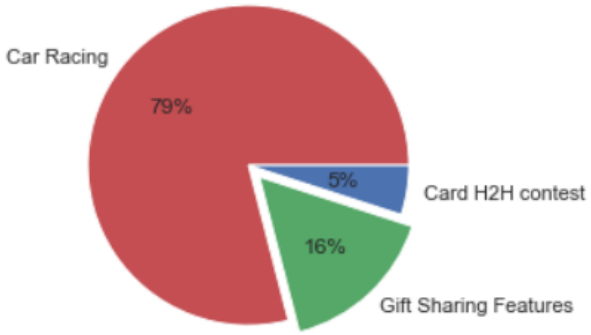
Add Money– Who users are coming on the product but already mind makeup with doing something fraud, the same users will not add much money, if we will compare it to the redemption, all users will 1K-10K% more money which he has been added on our system.

#Image Pic 2.1.1

Focus– I observed, that users were doing fraud, the majority of these users focus on a single game, which means data was completely biased, meaning users wanted to earn more without skill used and withdraw money as soon as possible in the account.
The most user who was fraud-focused on the Car race game and because of this game we can change the score update APIs and update high score within a second.



Descriptive Approach & Finding– During our 1 month of team Study about product data, I found something interesting about the product latency, lacks and loopholes. 1 game is very interesting for the fraudulent users and something interesting data parameter values cached of the real fraud users.
Car Racing - 78.6% UU, Gift Sharing features - 15.9% UU, Card H2H contest - 5.12% UU



– Users A1– Some app exists to do fraud in the gaming environment and because with the help of this app directly, we can change the score APIs.
Suppose, users are playing in the H2H contest, A users are related to real users and B is a fraudulent category user, now B users before

submitting the score will hit score APIs to change the real score to a fake score and then he will win the contest.

- Users A2– Every time features are not helpful for the product growth because in this case product aim to retain the users who are coming through the referral affiliate channel because the PDAU of referral users was very less like 1.39% only.

  But users were following a very smart strategy.

  Feature- of the product - users can be signed up with an email id, but on the same device, he can create more users (Device Ids = 1).

  Users are creating new or unique user ids using Temp mail IDs and sharing the joining bonus amount on the single users like the user collective bank.

  Example - A users have 100 user Ids and the joining of each user is 2$. If users will share their money with a friend of 0.12$ so A will get 0.12*99$ in their wallet, then we can earn more than without investing the money in the product.

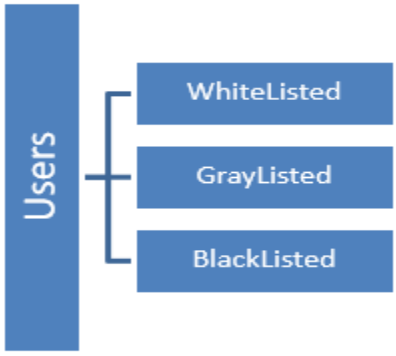
- Users A3– Card Game is a very high revenue game of the product. OEM wanted the solution on higher priority. Because some users are doing fake with the game. Live to the game if users have 3 different cards and not even 0 pairs in poker then fraudulent category users will upload higher category card pair with the help of third party app and win the game with use set of skills.

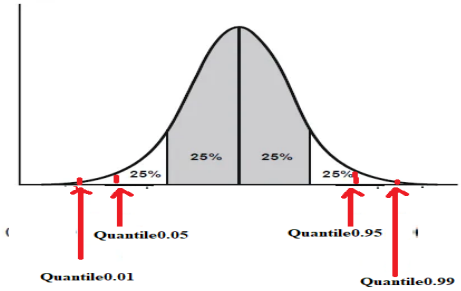
Data Scientist Call:

Approach to solve Problems:

The Data Scientist team are collecting historical data from the warehouse, to know about the user behaviour on the platform and to know the ratio of users onboard to redeem money in the wallet.

The team converted the user's statics values into 3 different categories.





C-1 WhiteListed — if users' data values in under the 1st quantile then can be put in this category.

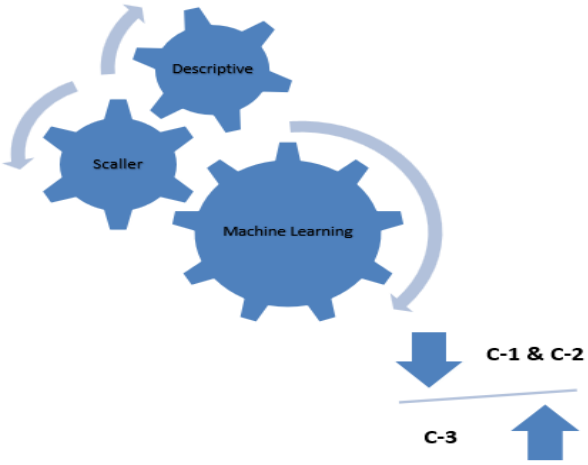
C-2 GrayListed Users – if Users' data values under the 2nd and 3rd quantile can be put in this category.

C-3 BlackListedUsers – if users' data values are outer from the 3rd quantile then we can put them in this category.

Definition:

In statistics and probability, **Quantiles** are cut points dividing the range of a probability distribution into continuous intervals with equal probabilities or

dividing the observations in a sample in the same way. There is one fewer quantile than the number of groups created.



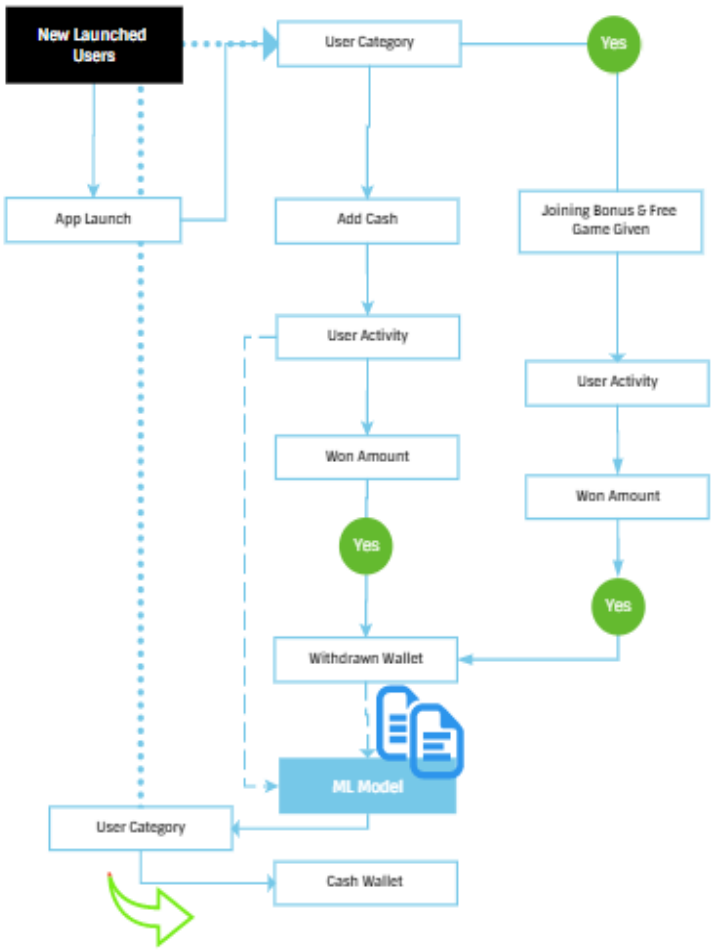
To know about the user's category, we were going the some user's features like

##Featured lists #Image Pic 2.1.1

Based on these all features, we will predict the user's behaviour statics values. Below mentioned some statics techniques we can judge the users' category.

Descriptive Analysis— Means, Mode, median, Avg /users Deposit and avg /users behaviour.

In the process to capture and updating status accordingly

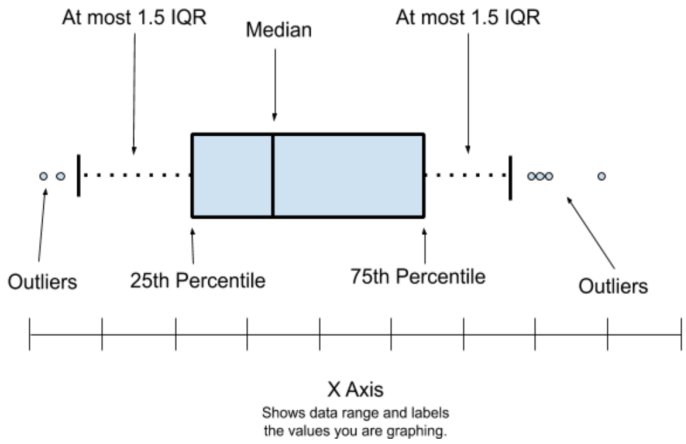


Machine learning Model structure, we have been involved to find the best **predictive users category** score.

A total number of 18 features (the above-mentioned list of features) help to find this user category parameter.

Working model— based on the last 6 months' user's historical activity on the same platform. we are calculating score values and based on these values divided into the 3 quantiles (or 3 subgroups).
If users come through a new affiliate channel (New users)--*first we will set the lowest category* of the users then increase based on the user's behaviour. If the user's activity scores or some calculus values are under the *first quantile then we will put them in the **White Listed*** users category (Case 1st).
If the calculus values are *above the p values then put them in the grey list category.* This means now users are ready to further investigation and the model will generate the mode KPIs for

further investigation (Case 2nd) but not in high-priority tasks, Only we will put these users under observation.



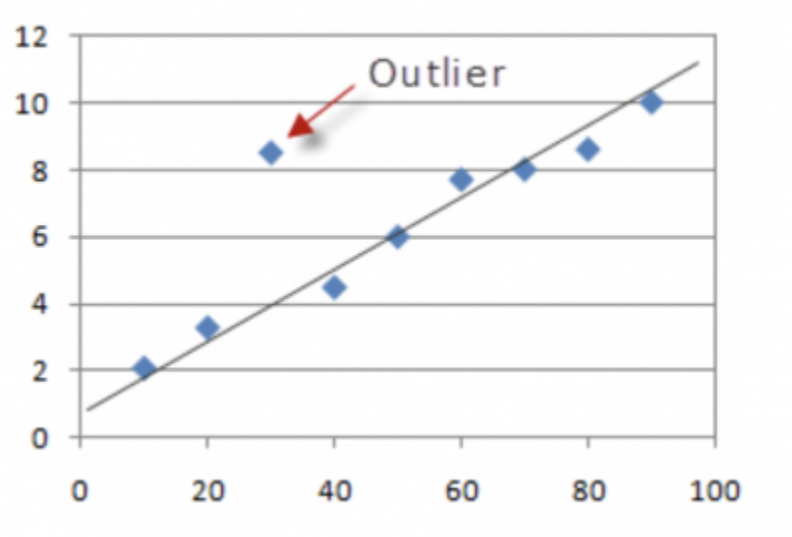
Who users exist in the Gray list, we will investigate these users for the next round category.

For this category, we will compare the ***user games-wise score and some aggregation parameters.***

If the user's values and scores are out of the box or quantile level then will we will put them in the *BlackList category and block the users for the product.*

Out of the Box User category Level

"Black List"



Conclusion:

Users would sign up for the product with some standard entry-level category list, and then the predictive model would store this user in our inquiry list or further investigation (user process flow in the fig above).

When will launch the App the model will collect all data and if will going to withdraw our amount in the case wallet then the model will decide, whether it can withdraw the money or not (in cash flow, the predictive model will interrupt to withdraw the case from the fraudulent category users).

Product Benefits:

1. The product can observe the black list users and we don't have to change the threshold model will handle it accordingly.
2. Proactively monitor for potentially fraudulent or high-risk events
3. Compute transactional risk factors to determine the legitimacy
4. Detect illegitimate transactional behaviours online
5. Provide alerts and analysis tools for administrators

6. Ensure compliance with data privacy and security regulations
7. Increased coverage to prevent fraud
8. Preventing Complex Fraud

Reference

https://venngage.com/blog/case-study-examples/
https://www.youtube.com/watch?v=nGzYzq3Wsos
https://www.altexsoft.com/blog/datascience/how-the-hospitality-industry-uses-performance-enhancing-artificial-intelligence-and-data-science/
https://productgym.io/product-manager-case-study-with-solution/
https://www.subex.com/article/what-are-the-benefits-of-using-a-fraud-management-system/