## โดรงการ CO-OP SHOWCASE : ก้าวแรกสู่อาชีพ ประจำปี 2567



# I-BIT CWIE DAY 2024



## FIND USER'S BEHAVIOR WICHAYAPHAN TRAITHIPTHOMRONGCHOKE 6407012662048





• SmyleLand is a game that can be played on various applications and websites. It presents the idea of creating a platform that can be used as a SmyleLand reward to claim various rights in real life as Play to Earn Privilege. User can play the minigame in Partner Landmark and earn special rights range from movie tickets, dining hall tickets, special privileges to receive and discounts from restaurants and coffee shops, etc.

**Pattern Identification in Privilege Assignment & Redemption:** 



#### **User Retention & Engagement Analysis (RFM Analysis)**:

• Pain Point:

Company did not know the current active user's retention and engagement status result.

• Objective:

To understand user retention and engagement, we applied the RFM (Recency, Frequency, Monetary) analysis. This method categorizes users based on how recently and frequently they engage with the game, as well as the value they generate through in-game privileges.

• Process:

User data was queried from the database, including fields such as user ID, last login date (recency), play frequency, and privilege value (monetary). Each of these metrics was then scored from 1 to 5 using Python, with score ranges separated by the 20th percentile. The users were grouped into segments based on their RFM scores.

• Results:

The results showed segmented groups such as "Champions" (users with high scores in all categories) and "Lost" users (low scores across the board). This segmentation made it easy to identify which users are highly engaged and which are at risk of leaving. We visualized these findings in a table showing the number of users based on their recency and FM scores, with segments highlighted for clarity.

#### • Benefits:

Company can create personalized experiences that keep users interested, improve retention, and encourage more activity. This approach helps apply the right strategies for different user groups, leading to business growth and long-term success.

• Pain Point:

Company didn't know specific odds patterns that can lead to assigned and redeemed privileges.

• Objective:

To detect patterns in user behavior that influence privilege assignment and redemption.

• Process:

Data from SQL and MongoDB (fields included User ID, Play Time, Partner, Reward, Privilege Status) was combined and processed using Python. Machine learning techniques were then applied:

- Isolation Forest was used to detect anomalies in user behavior, helping to flag unusual patterns, like repetitive or automated activity.
- PCA (Principal Component Analysis) was applied to reduce the dataset's dimensionality, making pattern identification easier.
- K-means clustering grouped users into four clusters based on their assignment and redemption behavior. Transactions were labeled as either "o" (redeemed) or "x" (assigned).
- Results:

This analysis identified clusters of users based on their behavior. For example, bot-like patterns were flagged through the Isolation Forest method, and K-means clustering showed groups of users with similar privilege engagement. The results were exported into an Excel file for further analysis, with the cluster numbers and anomaly flags added as extra columns.

• Benefits:

Company can see data visualized in different ways but pattern still remains unclear.





	UserId	Recency	Frequency	Monetary	R	F	M	FM	RFM
0	0186f5fd-4	404	1	69	1	1	1	1	111
1	01ee1778-	163	16	1780	5	5	5	5	555
2	01fe5159-	203	23	2868	3	5	5	5	355
3	03ad8f56-	134	24	2908	5	5	5	5	555
4	03e6c2bc-	287	10	1419	1	5	5	5	155
5	05d9326f-	502	6	1069	1	4	4	4	144
6	081afd99-	160	10	1520	5	5	5	5	555
7	0980bc69-	499	3	128	1	4	1	2.5	141
8	0a48d73d-	196	15	1828	4	5	5	5	455
9	Of6c50ee-	365	1	19	1	1	1	1	111
10	Ofb8e66b-	424	2	88	1	2	1	1.5	121
11	11a4fb21-	135	27	3529	5	5	5	5	555
12	11fbea27-	259	15	2200	2	5	5	5	255
13	133d80f5-	420	5	869	1	4	3	3.5	143
14	14cd7b2e-	499	3	128	1	4	1	2.5	141
15	156ec074-	152	26	3508	5	5	5	5	555
16	16c80daf-	364	8	1288	1	4	5	4.5	145



### **Optimization of Privilege Assignment Timing:**

• Pain Point:

Company did not know which hour was the best to let users come and get assigned privileges.

• Objective:

To discover the optimal times for assigning in-game privileges by analyzing user activity patterns.

#### • Process:

Data was queried for user ID, play datetime (hour, day, month, year), assigned datetime, minigame type, and partner category. Power BI was then used to visualize the data, focusing on the separation of play and assigned activity counts.

- Visualization:
  - Bar graphs displayed daily and monthly play patterns to reveal trends over time.
  - Line graphs were used to show hourly trends, identifying peak user activity hours.
  - Box and whisker plots provided an overview of datetime distribution, offering insights into when users are most active.

#### • Results:

From these visualizations, we found that users primarily play games and are assigned privileges between 10 PM and midnight, with certain minigames (like Claw Machine, Lucky Card, and Lucky Wheel) being the most popular. This information enables the company to strategically assign privileges during peak times, maximizing engagement.

#### • Benefits:

Company can track when users play to earn and redeem privileges. This information can help decide the best times to offer more or fewer privileges.



### **Prediction of Privilege Redemption Behavior:**

• Pain Point:

Company doesn't know if users are going to redeem the assigned privileges they got in the future or not.

- Objective:
- To predict whether users will redeem their assigned privileges using a machine learning model.
- Process:

The combined dataset was used to build a predictive model using the XGBoost algorithm. The dataset was split 70-30 for training and testing, ensuring robustness:

- Non-numerical data, such as partner category and privilege type, were transformed using onehot encoding.
- SMOTE (Synthetic Minority Over-sampling Technique) was applied to balance the dataset due to the disproportionate number of assigned privileges compared to redeemed ones.
- A confusion matrix was used to evaluate the model's performance on metrics such as accuracy, precision, recall, and F1 score.
- Results:

The predictive model performed well after applying SMOTE, showing high accuracy in predicting true positives, true negatives, and false negatives. Although false positives (cases predicted to redeem but actually assigned) were less accurate, this issue was minor since users could still redeem privileges later. The model's accuracy was confirmed after testing it on new data two months later, and the results were exported into Excel with additional columns for predicted redemption status and probability of redemption.

• Benefits:

Company get usable model that can predict if user will redeem privileges in the future or not. This enables companies to manage resources more effectively, enhance user engagement and retention,







#### optimize marketing strategies, and ultimately increase revenue.

Co	onfusion Matrix													
			A 8 C 0 E F	6	н	1	,	ĸ	L	м	N	0		Q
33679			1 Userid Partner Reward Privilegeid Time StartLandma	rk FinishLandmark	TravelDuration	StartTravelTime	FinishTravelTime	StaminaCost	Status	CreatedTime	RedeemedTime	Category	RedeemPrediction	RedeemProbabili
		20000	380 34532 Koala Th Discourt66c865ae8 2024-08-23 17:34:22							1 2024-08-23 17:34:23	2024-08-23 17:34:44	IT & Gadg	1	1
		- 30000	381 31631 Koula Th Discourt66e 7681c8(2024-08-22 23:32:28							6 2024-08-32 23:32:29	00.00,00	IT & Gade	1	1
			382 37100 Koala Th Discourt66b4/579e/ 2024-08-08 23:42:33							0 2024-08-08 21:42:33	00:00:00	IT & Gadg	1	1
			383 30566 Koela Th Discourt66ed96/7e82024-08-03 09:33:27							0 2024-08-03 09:33:28	00:00:00	IT & Gadg	1	1
			384 39715 GAGA Free Top/66c534dtci 2024-08-21 07:29:05 SmyleLand	GAGA	300	2034-08-21 07:19:20	2024-06-21 07:24:20	0 25		1 2024-08-21 07:29:05	2024-09-14 21:47:40	Food & Be	1	
	1026		385 27774 GAGA Free Topi66cb0c068(2024-08-25 17/48:29							0 2024-08-35 17:48:39	00,00,00	Food & Ba	1	1
	1920	- 25000	386 27820 8AN CH/Ame Cite66cb8b1c8(2024-08-25 12:44:44							1 2024-08-25 12:44:45	2024-08-25 17:85:50	Food & Br	1	1
			387 28056 JAN CH/ Free Cris66c1db83d 2024-08-18 17:22:59							1 2024-08-18 17-22-59	2024-09-08 16:52:04	Food & Ba	1	1
			388 31337 Karl Kerl Discourt/66c3d2b9c12024-08-18 12:53:45							0 2024-08-18 12:53:45	00:00:00	Food & Br	1	1
			389 27814 JVN CW/ free Cric66c65c818/2024-08-22 00:06:09 BarBQ Plaza	JIMN CHA	240	2024-08-21 23:50:50	2024-08-21 23:54:50	0 24		1 2024-08-22 00:06:10	2024-09-05 11:03:34	Food & Bu	1	1
			190 26640 Keala Th Discourt66/9/14a8(2024-08-24 18:17:30							0 2024-08-24 18:12:31	00-00-00	IT & Carle	1	1
		- 20000	391 27752 JUNI CH/ Free Wit 660/07622v 2024 08-05 13:50:11 Curle BAGA	JAM CHA	300	2024-08-05 13:31:09	2024/08/05 13:36:05	9 24		0 2024 08 05 13 50 11	00.00.00	Food & Bc	1	1
		20000	392 31337 Yak Ka is Discountilitiesheb109 2024-08-22 14540-17 Karl Korl	Yak Kallon	300	2024-08-18 17:5450	2024-06-18 17:5/642	7 25		0 2024-08-32 14:40:48	00110000	Food & Be	1	
			593 59655 Cale DACVoucher 66c446b3ci 2024 08 20 14 33:07 Koala Thailand	Cale BAGA	300	2034-08-20 07:31:51	2024 08 20 07:36:50	1 25		0 2024 08 20 14:33:08	00.00.00	Food & Bc	1	
			398 31716 CACA Tree Topition/blottel/20/406-21 11:3707/Text And Kins	CLACIA	300	2020-08-21 12:47:09	2026-06-21 126-20	0 2%		0 2014-06-21 12:07:08	OLIEBYCE	Lood & He	1	
			195 27789 JAN CIT/ Free Wig 65oc7700ci 2024 08 02 13:08:48 Everyday Thai	TO IVAN CITA	300	2026 08 02 12 59:33	2026 08 02 13 08 33	1 25		0 2024 08 02 12 04 49	00-00-00	Food & Dc	1	
		15000	576 S/115 JAN CR4 (new Pecititedbe /ew 2024-08-05 12:22:06 twender   hei	IwIJAN CHA	500	2004-08-03 12:16:55	2024-06-08 12:21:45	25		1 2024-08-08 12:22:08	2004-09-07 15:52:14	Lood & Be	1	
		- 15000	397 39664 Cate DACVoucher 66d15d2e8 2024-08 30 12:48:30 Koala Thailand	Cate BAGA	300	2024-08-30 12:38:31	2024/08/30 12:43:30	1 25		0 2024/08/30 12:48:30	00:00:00	Food & Bc	1	
			208 20884 Cufe BACVoucher 66bee/42cl 2024-08-56 12:18:42 Everyday That	Te Cale BAGA	300	2004-08-16 13:11:09	2024-08-16 13:16:00	25		0 2024-08-16 10:18-43	00.00.00	Food & Ba	1	
5			199 77570 Kosla Th Discourt@LafedOtel 2024-08-05 02:56-56	Koola Thailand						0 2024-08-05 02:56:57	00-00-00	IT & Gade	1	
			400 27724 Keyly Th Discourt66c5653c8(2024-08-26-00-09-16	Keyla Thailand						0 2024-08-36 00-09-16	00-00-00	IT & Garde	1	
			401 31602 Koala Th Discourt66a/df12e82024-08-05 03:05/251	Koala Thailand						0 2024-08-05 01:05-03	001100:000	IT & Gade	1	
		- 10000	402 26601 Kari Kari Thai toa 66c46799ci 2024 08-20 16:53:29 Yosurtland	Kari Kori	16	2034-08-20 16:52:36	2024-08-20 16:52:53	2 9		0 2024-08-20 16:53:29	00.00.00	Food & Ba	1	
	474		401 27022 Karl Keel Thai tea 64carl6/28 2024-08-25 05/50/26							1 2024-08-25 09:00/27	2028-09-14 22:06:63	Food & Br	1	
	4/4		404 37145 JUN CH/ Free Per 66cee8e7e 2024-08-04-09-35-19							1 2024/08/04 09 35:19	2024-08-09 13:10:57	Food & Ba	1	
			405 37138 Sw/SOPh E-cash y 66c63c038(2024-08-22 02:12:04 JAN CHA	BartiOPlana	300	2024-08-22 02:06:40	2026-06-22 02:11:40	24		1 2020-06-32 02:12:04	2021-09-04 13:27:53	Food & Be	1	
			406 36754 JAN CH/ Free Cris66ac68c8el 2024-08-02 12:04:08 BarBOPlana	JIAN CHA	300	2024-08-02 11:54:13	2024-08-02 11:59:11	3 25		0 2024-08-02 12:04:08	00-00-00	Food & Bc	1	
		- 5000	407 27814 JAN CR/ Inex City Mca25a68 2024-08-25 01:25:12 BarthOPara	JIAN CRA	240	2028-08-25 01:21:15	2026-06-25 01:21:11	24		1 2026-08-35 01:25-63	2024-09-05 11:02:51	Frond & No.	1	
		5000	408 24712 Yak Ka k Discourt66pe/Wiled 2024-08-04 10:19:49							0 2024-08-04 10:19:50	00:00:00	Ford & Bc		
			409 31716 JAN CW/ Iron Pag 66g665098( 2024-08-22 05-35-11 Keala Thailan	JAN CHA	250	2024-08-21 21:50-40	2024-08-21 21-54-50	0 21		0 2024-08-22 05-35-11	00-00-00	Food & Br		
			410 39760 GAGA Free Wh 66ct6e1580 2024-08-29 01:36:05 THAI THAI Shi	peGAGA	250	2024-08-29 01:25:49	2024-08-29-01:29:54	9 21		1 2024 08 29 01:36:05	2024-08-29 14:56:58	Food & Be	1	
			411 34106 Koula Th Discourt 66bea89ac 2024-08-16 08:17:14							1 2024-08-16 08:17:15	2004-08-16 08:17:34	IT & Gade	1	
			12 34532 Koala Th Discourt 64bea/342ci 2024-08-16 08-20:02							1 2024-08-16 08:20:03	2024-08-16 08:20-21	IT & Gade		
! .			The state of the second s	Calls Red to	200	3034 OR 18 12/22/34	1014-04-14-18-01-14			A 1014 04 10 00 44 10		Course of the low		
Assigned	Redeemed													
	Predicted													

# Data Analytics