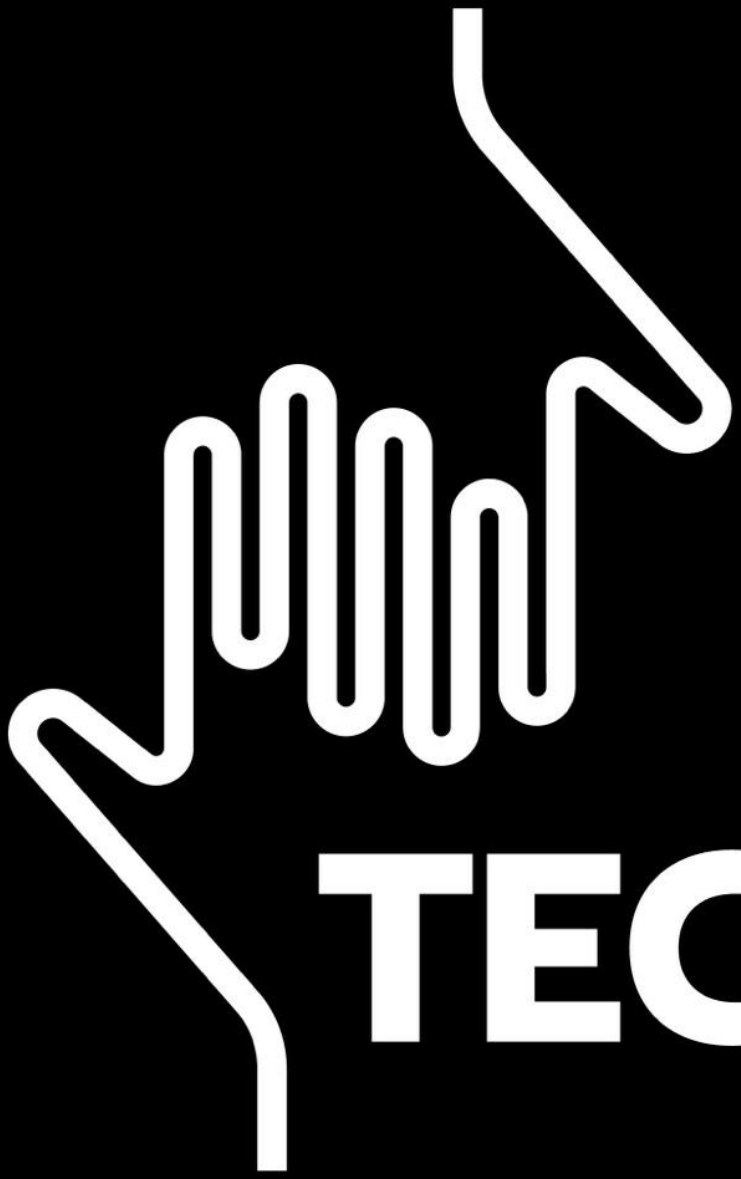


TECHJAM

data

TRACK

TECHJAM 2017

TECHJAM 2017

Safety & Facility at KBTG Building

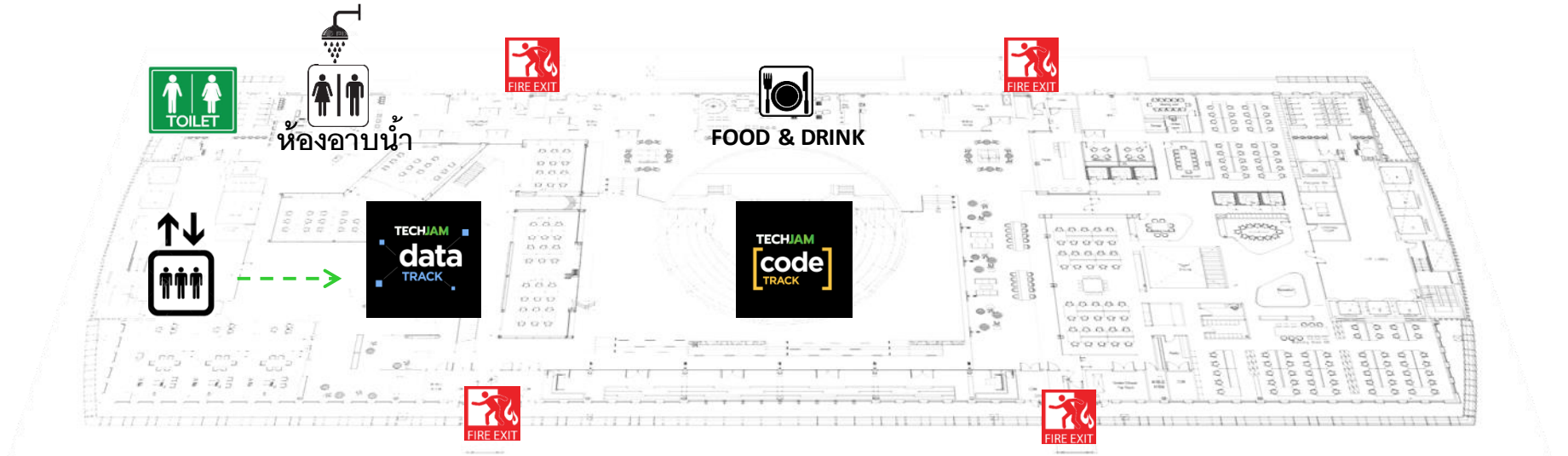
Fl. G

ธนาคารกสิกรไทย KASIKORNBANK · KBTG · LINE · Google Developers · GDG Thailand · SoA+D

TECHJAM 2017

Safety & Facility at KBTG Building

Fl.11



ธนาคารกสิกรไทย KASIKORNBANK · KBTG · LINE · Google Developers · GDG Thailand · SoA+D

TECHJAM 2017

Safety & Facility at KBTG Building

Don't

กรุณาแต่งกายสุภาพ

ห้ามสูบบุหรี่ทุกพื้นที่

กรุณาติดบัตรผู้ติดต่อ

เข้าพื้นที่ที่กำหนดให้เท่านั้น

ทิ้งขยะในที่เตรียมจัดไว้ให้

เบอร์ติดต่อด่วน 02-008-1314 (ส่วนงานความปลอดภัยอาคาร KBTG)

ธนาคารกสิกรไทย KASIKORNBANK | KBTG | LINE | Google Developers | GDG Thailand | SoA+D

TECHJAM 2017

Schedule

data TRACK

Time	Activity
	Saturday
08.30 - 09.00	Registration
09.00 - 10.00	Problem Briefing
10.00	Submission Open
11.30 - 13.00	Lunch
13.00 - 15.00	Mentoring Session #1 (5 minutes each)
16.00 - 18.00	Mentoring Session #2 (5 minutes each)
22.00	Submission Closed
	Sunday
08.30 - 11.30	Model Discussion with Judges (5 minutes each)
11.30 - 12.00	Certificate Handout & Photos
12.00 - 13.30	Lunch
13.30 - 14.00	Welcoming Speech and Keynote Speech
14.00 - 16.00	Award Ceremony & Mini Party

Disclaimer

All data and labels given in this problem are synthesized and not related to any customer or any real person.

TECHJAM 2017

Problem Statement

data TRACK

You, as a data scientist, are tasked to utilize credit card information and transactions to determine buying behavior of all the customers. The given data dates from January 2016 to June 2017. The goal is to predict amount of spending and buying frequency of five designated categories in July 2017.

TECHJAM 2017

Problem Statement

data TRACK

You, as a data scientist, are tasked to utilize credit card information and transactions to determine buying behavior of all the customers. The given data dates from January 2016 to June 2017. The goal is to predict amount of spending and buying frequency of five designated categories in July 2017.

Info	month1	month2	.....	month18	month19
card1			....		???
card2			....		???
.	.	.		.	.
.	.	.		.	.
cardN			....		???

ธนาคารกสิกรไทย KASIKORNBANK | KBTG | LINE | Google Developers | GDG Thailand | SoA+D

Input Data

techjam_final_data.zip contain 3 files

- cc_log.csv
- cc_info.csv,
- final_categories.csv

Input Data

cc_log.csv

card_no	txn_dt	txn_tm	bill_amt	card_acpt_cty	mrch_tp_cd	card_type
4410123456000001	2016-07-12 00:00:00	11:25:03	6220	TH	6300	visa
4410123456000001	2017-05-13 00:00:00	12:58:04	760	TH	5719	visa
4410123456000001	2017-04-09 00:00:00	18:12:42	1440	TH	5651	visa
4410123456000001	2017-04-07 00:00:00	18:03:37	1120	TH	5651	visa
4410123456000001	2017-05-02 00:00:00	13:03:34	1180	TH	5331	visa

Input Data

cc_info.csv

card_no	card_type	opn_dt	exp_dt	cr_lmt_amt	prev_cr_lmt_amt	main_zip_cd	cr_line_amt	incm_amt	brth_estb_yr
4410123456000001	visa	1997-10-14	1017	50000	0	73120.0	50000	21000	1953.0
4410123456000002	master	2010-06-29	620	146000	0	43000.0	146000	72000	1965.0
4410123456000003	visa	2014-06-18	619	22000	0	23000.0	22000	20000	1978.0
4410123456000004	visa	2014-11-03	1119	1000000	0	10200.0	2141000	1800000	1955.0
4410123456000005	visa	2014-04-10	419	80000	0	57160.0	80000	32000	1969.0
4410123456000006	visa	2015-12-18	1220	100000	0	10150.0	100000	66000	1991.0

Input Data

final_categories.csv

Categories	MCC
Contracted Services	742
Contracted Services	763
Contracted Services	780
Contracted Services	1520
Contracted Services	1711
Professional Services and Membership Organizat...	8931
Professional Services and Membership Organizat...	8999
Government Services	9211
Government Services	9222
Government Services	9223

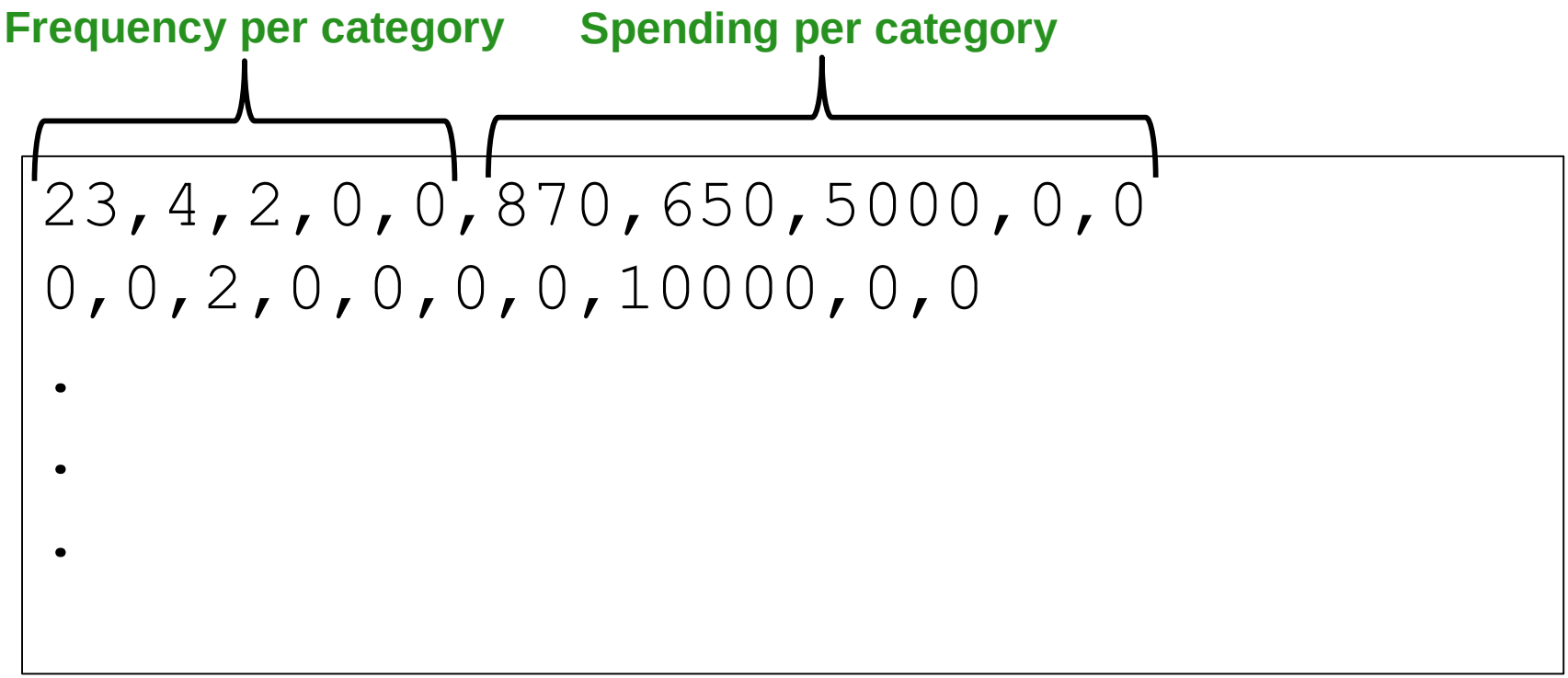
Output Data

team_NN.csv

```
23,4,2,0,0,870,650,5000,0,0
0,0,2,0,0,0,0,10000,0,0
.
.
.
```

Output Data

team_NN.csv



Deliverables

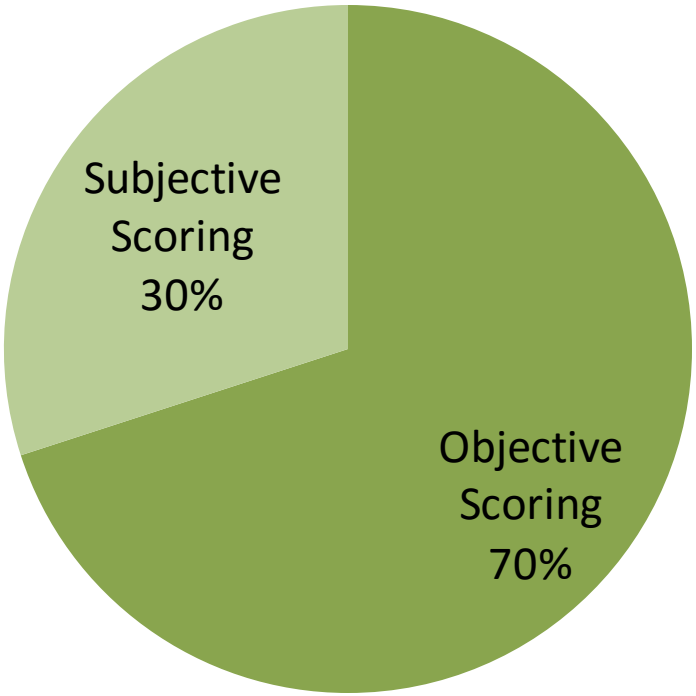
Saturday

- Output file (only the last submission before 10pm counts)
- Zip file containing source files and README file (only the last submission before 10pm counts)

Sunday

- 5-minute Presentation + 1-minute Q&A
  - How did you build your model? Why so?
  - How else can you leverage this data to generate business value?

Scoring Criteria



Objective Score (70 points)

How to calculate objective score:

- Each card's prediction will yield a ***card error*** between [0, 1].
- Your team's objective score is equal to your 1 minus average card error (over all cards), and then multiplied by 70 (i.e., ***obj_score*** = (1 - ***avg_card_err***) * 70)

How to calculate card error for each card's prediction:

- For each prediction type (frequency or spending) and category, calculate an error (absolute difference between the predicted number and the actual number)
- For each prediction type (frequency or spending) and category, normalize your error to [0, 1] range using minimum error and maximum error across all teams.
- Your card error is an average of normalized errors across all prediction types and all categories.

Objective Score Examples (2 categories, 2 cards, 3 teams)

		actual	team1	team2	team3
card1	cat1_freq	5	1	3	4
	cat1_spend	250	100	200	80
	cat2_freq	7	3	5	1
	cat2_spend	1000	50	500	800
card2	cat1_freq	3	10	8	1
	cat1_spend	30	250	800	50
	cat2_freq	4	20	5	7
	cat2_spend	800	100	50	450

Objective Score Examples (2 categories, 2 cards, 3 teams)

		actual	team1	team2	team3
card1	cat1_freq	5	1 (**4**)	3 (**2**)	4 (**1**)
	cat1_spend	250	100 (**150**)	200 (**50**)	80 (**170**)
	cat2_freq	7	3 (**4**)	5 (**2**)	1 (**6**)
	cat2_spend	1000	50 (**950**)	500 (**500**)	800 (**200**)
card2	cat1_freq	3	10 (**7**)	8 (**5**)	1 (**2**)
	cat1_spend	30	250 (**220**)	800 (**770**)	50 (**20**)
	cat2_freq	4	20 (**16**)	5 (**1**)	7 (**3**)
	cat2_spend	800	100 (**700**)	50 (**750**)	450 (**350**)

Objective Score Examples (2 categories, 2 cards, 3 teams)

		actual	team1	team2	team3
card1	cat1_freq	5	1 (4 -> ***1***)	3 (2 -> ***0.33***)	4 (1 -> ***0***)
	cat1_spend	250	100 (**150**)	200 (**50**)	80 (**170**)
	cat2_freq	7	3 (**4**)	5 (**2**)	1 (**6**)
	cat2_spend	1000	50 (**950**)	500 (**500**)	800 (**200**)
card2	cat1_freq	3	10 (**7**)	8 (**5**)	1 (**2**)
	cat1_spend	30	250 (**220**)	800 (**770**)	50 (**20**)
	cat2_freq	4	20 (**16**)	5 (**1**)	7 (**3**)
	cat2_spend	800	100 (**700**)	50 (**750**)	450 (**350**)

Objective Score Examples (2 categories, 2 cards, 3 teams)

		actual	team1	team2	team3
card1	cat1_freq	5	1 (4 -> ***1***)	3 (2 -> ***0.33***)	4 (1 -> ***0***)
	cat1_spend	250	100 (150 -> ***0.83***)	200 (50 -> ***0***)	80 (170 -> ***1***)
	cat2_freq	7	3 (4 -> ***0.5***)	5 (2 -> ***0***)	1 (6 -> ***1***)
	cat2_spend	1000	50 (950 -> ***1***)	500 (500 -> ***0.4***)	800 (200 -> ***0***)
card2	cat1_freq	3	10 (7 -> ***1***)	8 (5 -> ***0.6***)	1 (2 -> ***0***)
	cat1_spend	30	250 (220 -> ***0.27***)	800 (770 -> ***1***)	50 (20 -> ***0***)
	cat2_freq	4	20 (16 -> ***1***)	5 (1 -> ***0***)	7 (3 -> ***0.13***)
	cat2_spend	800	100 (700 -> ***0.88***)	50 (750 -> ***1***)	450 (350 -> ***0***)

Objective Score Examples (2 categories, 2 cards, 3 teams)

		actual	team1	team2	team3
card1	cat1_freq	5	1 (4 -> ***1***)	3 (2 -> ***0.33***)	4 (1 -> ***0***)
	cat1_spend	250	100 (150 -> ***0.83***)	200 (50 -> ***0***)	80 (170 -> ***1***)
	cat2_freq	7	3 (4 -> ***0.5***)	5 (2 -> ***0***)	1 (6 -> ***1***)
	cat2_spend	1000	50 (950 -> ***1***)	500 (500 -> ***0.4***)	800 (200 -> ***0***)
	card1_error		**0.83**	**0.18**	**0.5**
card2	cat1_freq	3	10 (7 -> ***1***)	8 (5 -> ***0.6***)	1 (2 -> ***0***)
	cat1_spend	30	250 (220 -> ***0.27***)	800 (770 -> ***1***)	50 (20 -> ***0***)
	cat2_freq	4	20 (16 -> ***1***)	5 (1 -> ***0***)	7 (3 -> ***0.13***)
	cat2_spend	800	100 (700 -> ***0.88***)	50 (750 -> ***1***)	450 (350 -> ***0***)
	card2_error		**0.79**	**0.65**	**0.03**

Objective Score Examples (2 categories, 2 cards, 3 teams)

		actual	team1	team2	team3
card1	cat1_freq	5	1 (4 -> ***1***)	3 (2 -> ***0.33***)	4 (1 -> ***0***)
	cat1_spend	250	100 (150 -> ***0.83***)	200 (50 -> ***1***)	80 (170 -> ***0***)
	cat2_freq	7	3 (4 -> ***0.5***)	5 (2 -> ***0***)	1 (6 -> ***1***)
	cat2_spend	1000	50 (950 -> ***1***)	500 (500 -> ***0.4***)	800 (200 -> ***0***)
	card1_error		0.83	0.18	0.5
card2	cat1_freq	3	10 (7 -> ***1***)	8 (5 -> ***0.6***)	1 (2 -> ***0***)
	cat1_spend	30	250 (220 -> ***0.27***)	800 (770 -> ***1***)	50 (20 -> ***0***)
	cat2_freq	4	20 (16 -> ***1***)	5 (1 -> ***0***)	7 (3 -> ***0.13***)
	cat2_spend	800	100 (700 -> ***0.88***)	50 (750 -> ***1***)	450 (350 -> ***0***)
	card2_error		0.79	0.65	0.03
	avg_card_error		**0.81**	**0.42**	**0.26**

TECHJAM 2017

Objective Score Examples (2 categories, 2 cards, 3 teams)

data TRACK

		actual	team1	team2	team3
card1	cat1_freq	5	1 (4 -> ***1***)	3 (2 -> ***0.33***)	4 (1 -> ***0***)
	cat1_spend	250	100 (150 -> ***0.83***)	200 (50 -> ***1***)	80 (170 -> ***0***)
	cat2_freq	7	3 (4 -> ***0.5***)	5 (2 -> ***0***)	1 (6 -> ***1***)
	cat2_spend	1000	50 (950 -> ***1***)	500 (500 -> ***0.4***)	800 (200 -> ***0***)
	card1_error		0.83	0.18	0.5
card2	cat1_freq	3	10 (7 -> ***1***)	8 (5 -> ***0.6***)	1 (2 -> ***0***)
	cat1_spend	30	250 (220 -> ***0.27***)	800 (770 -> ***1***)	50 (20 -> ***0***)
	cat2_freq	4	20 (16 -> ***1***)	5 (1 -> ***0***)	7 (3 -> ***0.13***)
	cat2_spend	800	100 (700 -> ***0.88***)	50 (750 -> ***1***)	450 (350 -> ***0***)
	card2_error		0.79	0.65	0.03
	avg_card_error		0.81	0.42	0.26
	obj_score		(1-0.81)*70=**13.3**	(1-0.42)*70=**40.6**	(1-0.26)*70=**51.8**

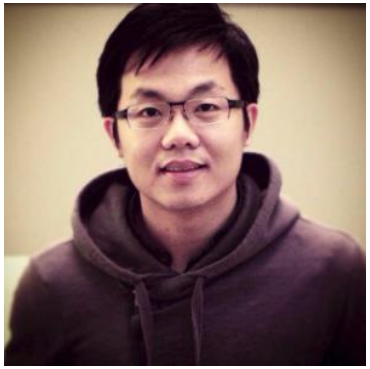
TECHJAM 2017

Subjective Score (30 points)

data TRACK

Scoring Criteria by Judges

- Communication (10 points)
- Technical skills (10 points)
- Business value generation (10 points)



Virot Chiraphadhanakul (GDE Machine Learning)

Sorawit Saengkyongam (GDE Machine Learning)

Tul Roteseree (Deputy Managing Director, KBTG)

Chetaphan Siridanupath (Senior Visionary Architect, KBTG)

Thadpong Pongthawornkamol (Senior Visionary Architect, KBTG)

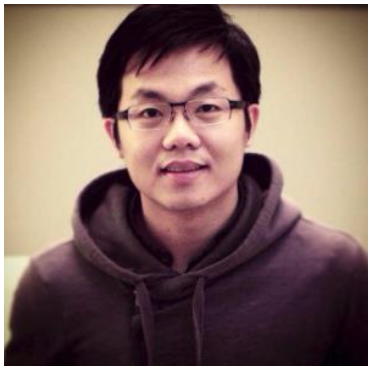
ธนาคารกสิกรไทย KASIKORNBANK | KBTG | LINE | Google Developers | GDG Thailand | SoA+D

TECHJAM 2017

Mentoring Session

data TRACK

- Meeting with 2 Google Developer Experts
- 2 periods (13.00 – 15.00, 16.00 – 18.00)
- 5 minutes per group
- Randomly ordered
- Except for asking for the solution, all questions are permitted



Virot Chiraphadhanakul

(GDE Machine Learning)

Sorawit Saengkyongam

(GDE Machine Learning)

ธนาคารกสิกรไทย KASIKORNBANK | KBTG | LINE | Google Developers | GDG Thailand | SoA+D

TECHJAM 2017

Rewards

data TRACK

Q&A


TECHJAM
2017
Human - centric AI
Better Living Through
Intelligent Technology