



EvalUMAP Workshop

UMAP2019

Generation & Evaluation of Personalised Push-Notifications

Kieran Fraser, Bilal Yousuf, Owen Conlan
ADAPT Centre, Trinity College Dublin

- ❖ Proposed Challenge
- ❖ Gym-Push
- ❖ Evaluation Metrics
- ❖ Challenge Entry
- ❖ Results & Discussion
- ❖ Limitations
- ❖ Final Thoughts

Currently no established or standardized means for **repeatable** and **comparative evaluation** of algorithms and systems in the UMAP space.

Goal

Shared Task

“focuses on user model generation using logged mobile phone data, with an assumed purpose of supporting mobile phone notification suggestion.” ¹

1. Proposal for a Shared Challenge in the UMAP Space, EvalUMAP Whitepaper 2019



“the challenge is to create an approach to generate personalized notifications on individuals’ mobile phones, whereby such personalization would consist of deciding what events (SMS received, etc.) to show to the individual and when to show them.”¹

Challenge 1

- Given 3 months historical notification data (for training)
- Develop a user model which generates a personalized notification given a context
- Using Gym-Push, user model is evaluated using test data and evaluation metrics

Challenge 2

- Given small sample of notification data (no training)
- Develop an adaptive user model which generates a personalized notification given a context
- Using Gym-Push, user model is evaluated, in simulated “real-time”, using test data and evaluation metrics

1. Proposal for a Shared Challenge in the UMAP Space, EvalUMAP Whitepaper 2019



OpenAI Gym

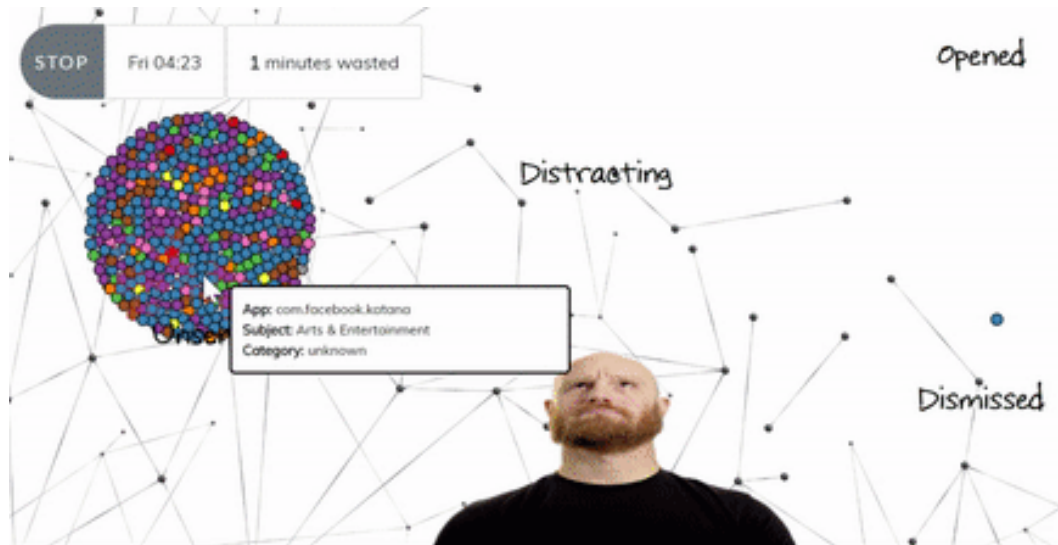
Open source toolkit for “developing and comparing reinforcement learning algorithms” ¹



Gym-Push

Custom OpenAI Gym environment simulating push-notification overload on mobile device users

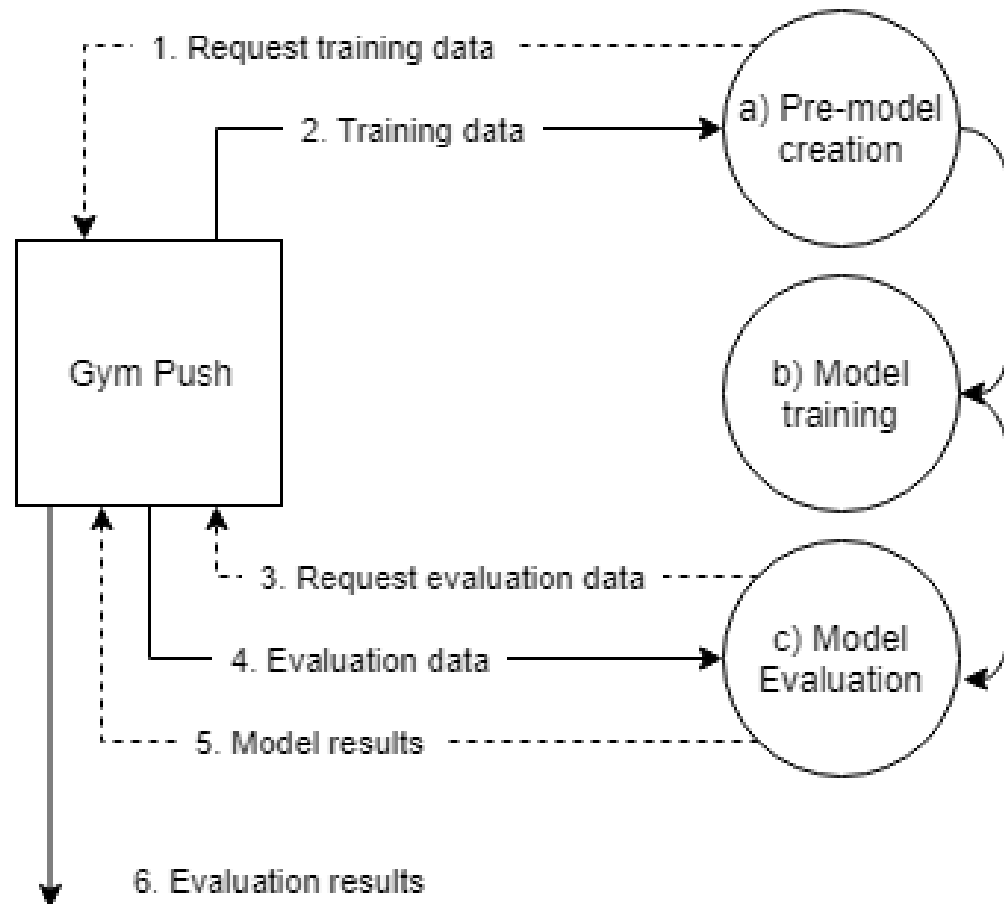
1. <https://gym.openai.com/>



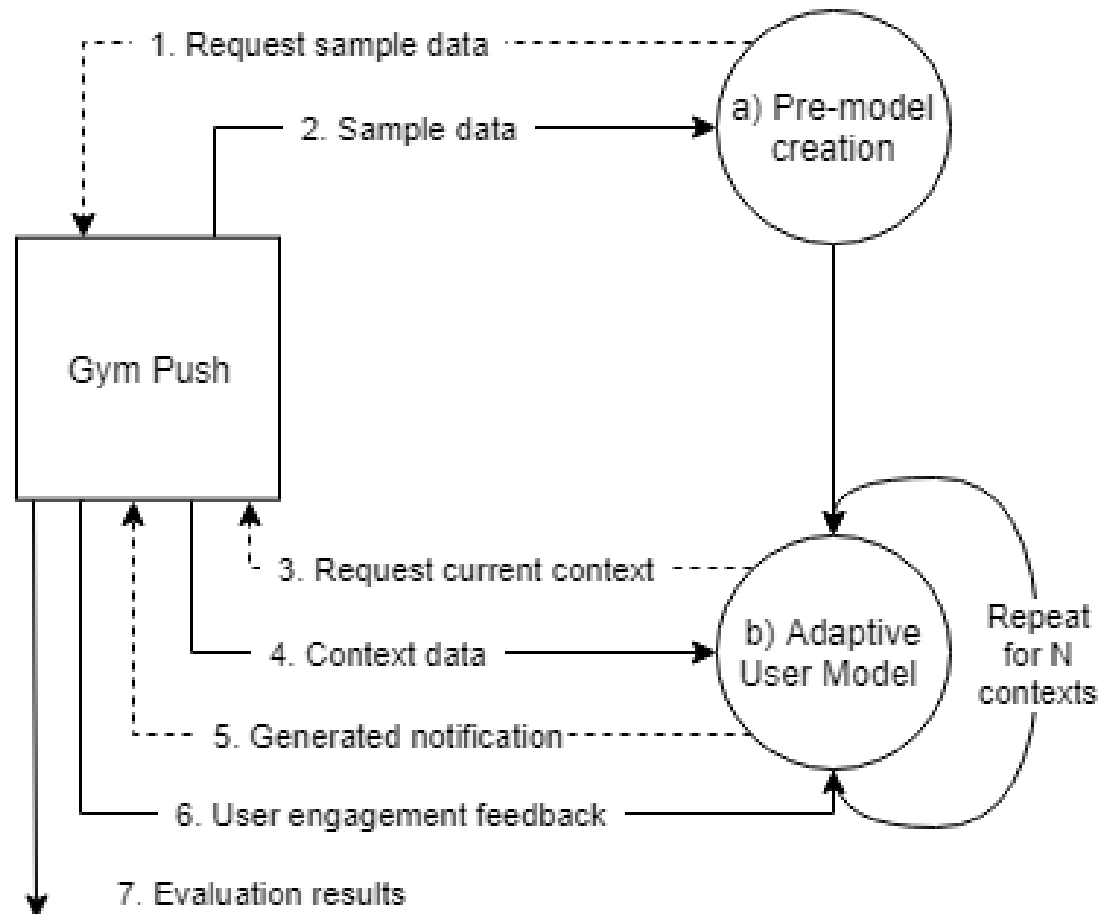
Gym-Push

- Ease of installation – pip, docker, hosted
- Multiple communities – RL, UMAP, HCI
- End-user interface
- Established Online Leaderboard

Challenge 1



Challenge 2



15

Participants over
3 months

31,239

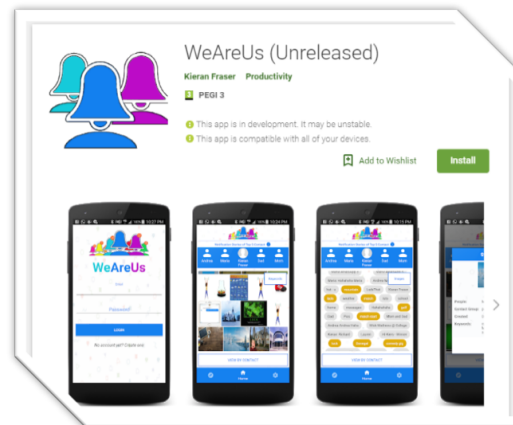
Notifications
Logged

291

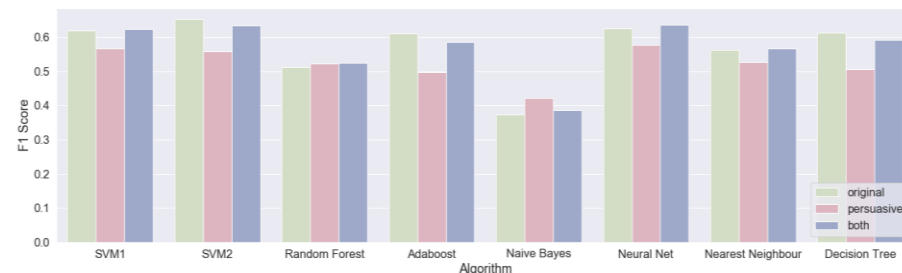
Questionnaire
Responses

4,940

Smartphone
Usage Logs

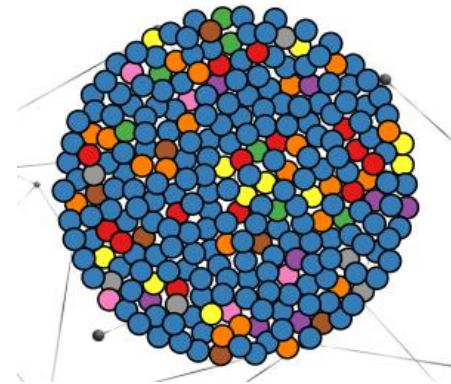
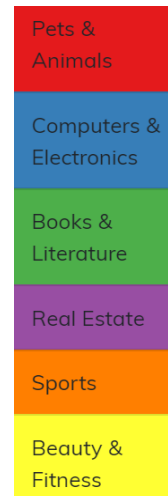
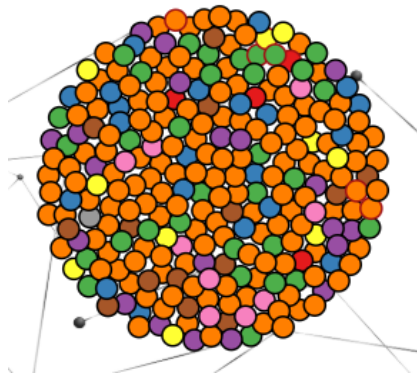
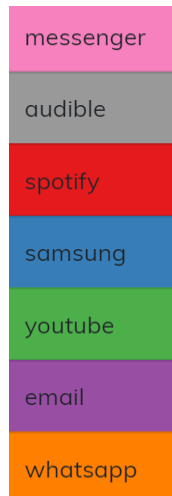
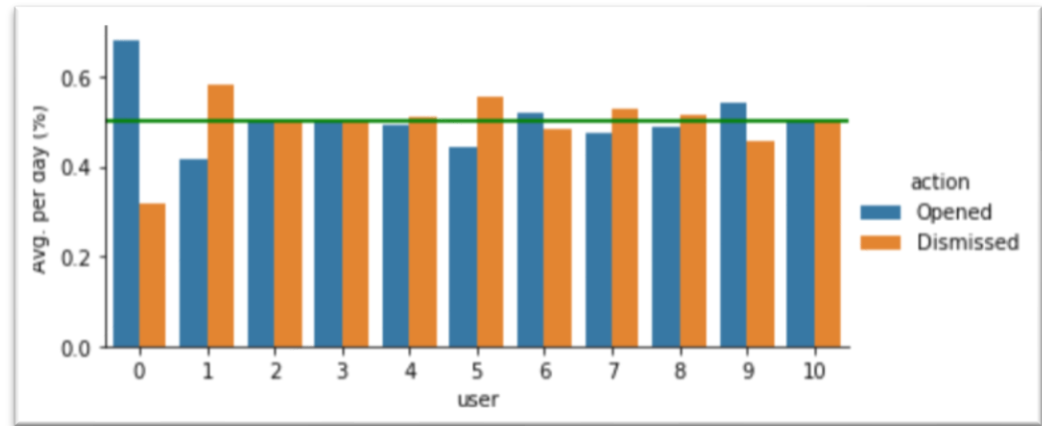
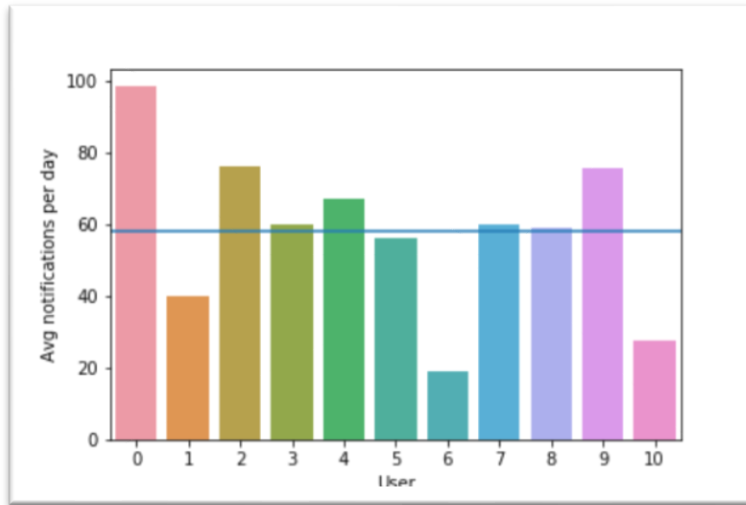


Notification accept%	33%	28%
Unique apps	140	35
Top Apps	WhatsApp, Gmail, SysUI, Vending	WhatsApp, Gmail, Vending, SysUI
App accept%	Whatsapp: 64% Gmail: 32% SytemUI: 7%	Whatsapp: 60% Gmail: 39% SystemUI: 7%
App Signific accept%	61%	55%
Category accept%	msg: 74% call: 66% alarm: 87% sys: 1% err: 6%	msg: 65% call: 50% alarm: 34% sys: 3.9% err: 5%
Unique subject	240	20
Subject accept%	Arts & Entertain: 67% Online Communities: 65% Education: 69% Mobile Apps: 88%	Arts & Entertain: 42% Online Communities: 29% Education: 73% Mobile Apps: 82%



*Train on Real, Test on Synthetic*¹ RMSE F1 scores differ in range 0.02 – 0.07 indicating synthetic data imitates real world data.

1. Esteban, C., Hyland, S.L., Rˆatsch, G.: Real-valued (medical) time series generation



Performance

Diversity

Response
Time

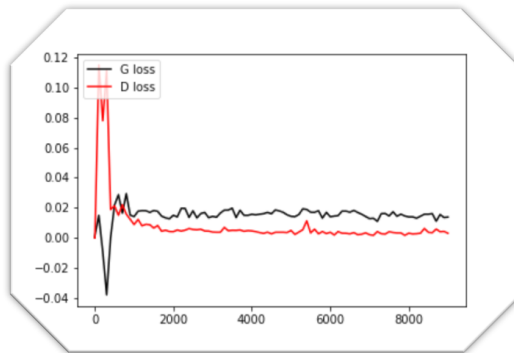
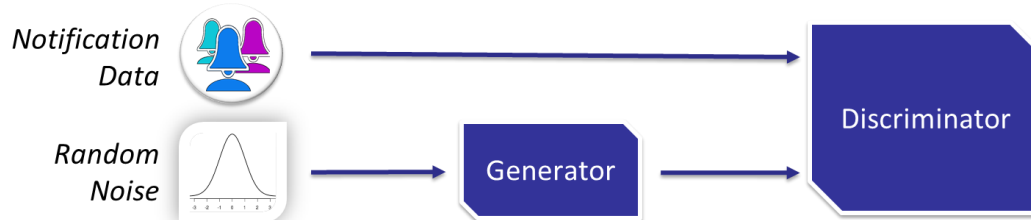
Learning
Rate

Simulated User

- *AdaBoost* Classifier chosen
- Trained on 3 months of historical user data
- Acc avg = 83.8%, F1 avg = 72.8%

Accuracy	Precision	Recall	F1
0.899713	0.929295	0.954237	0.941563
0.753049	0.713834	0.707143	0.709560
0.988000	0.988599	0.999377	0.993958
0.874490	0.912893	0.923490	0.917936
0.822222	0.674086	0.513115	0.581696
0.737730	0.715937	0.663380	0.687978
0.842683	0.801284	0.656000	0.717755
0.912500	0.716667	0.039394	0.073996
0.717647	0.717232	0.811111	0.760936
0.832824	0.858501	0.948734	0.901260

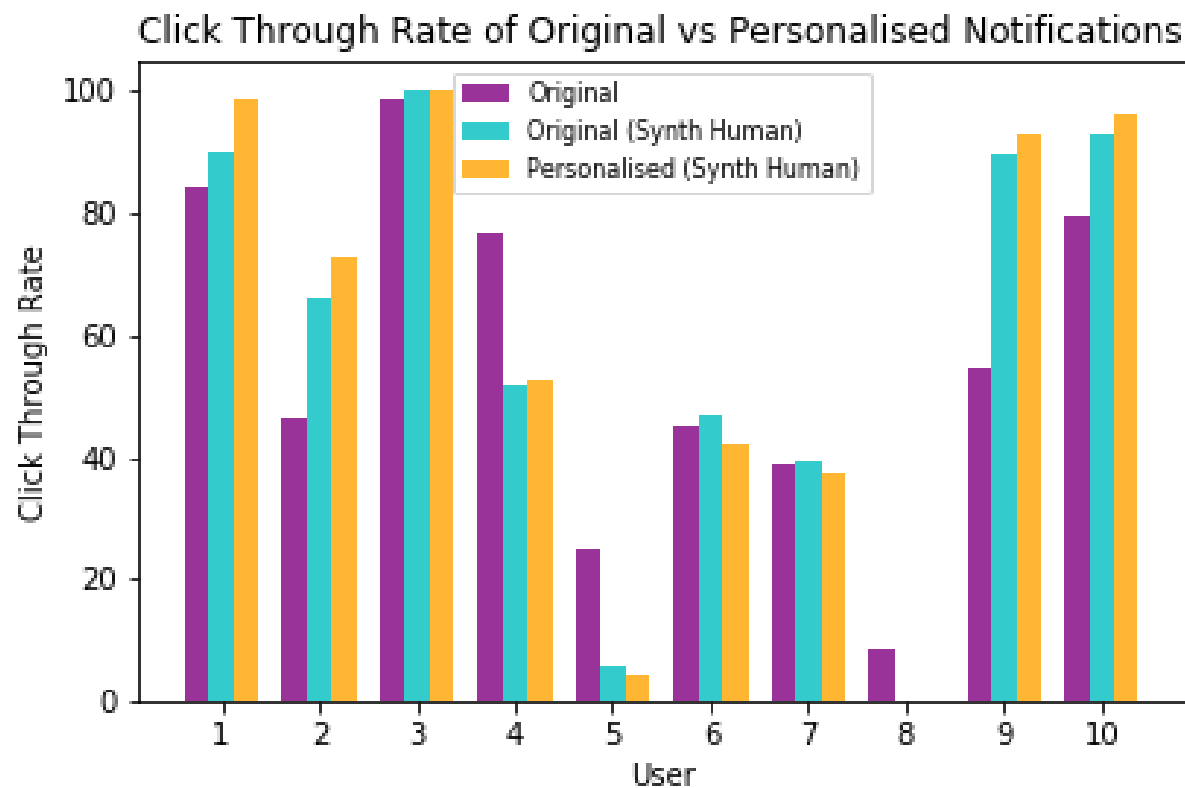
Challenge 1

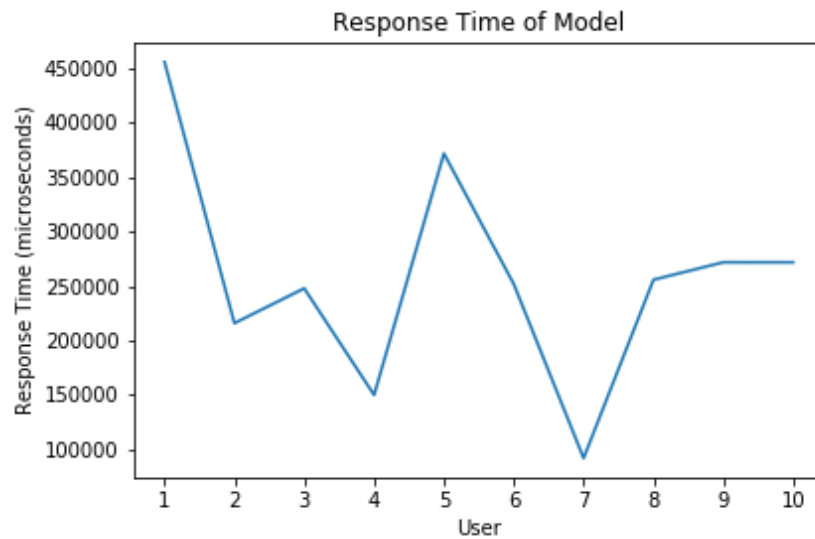


- MLP used for Generator & Discriminator
- Notifications OHE vector length 28
- Trained using RMSProp in 128 mini-batch chunks over 2000 epochs

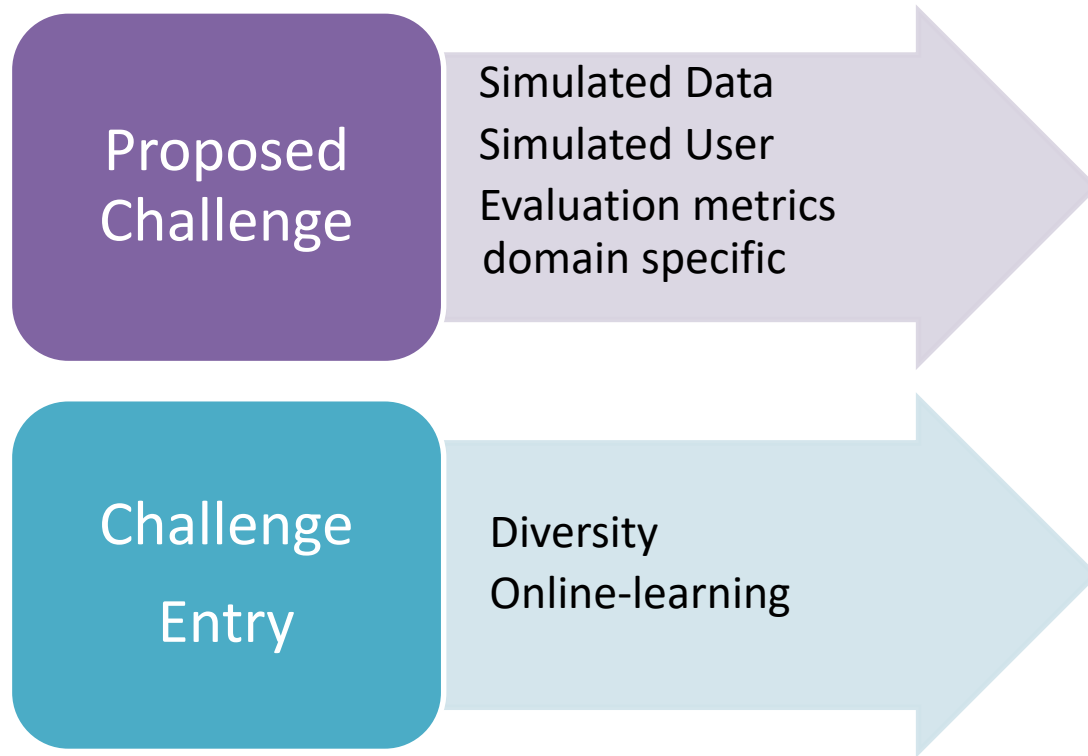


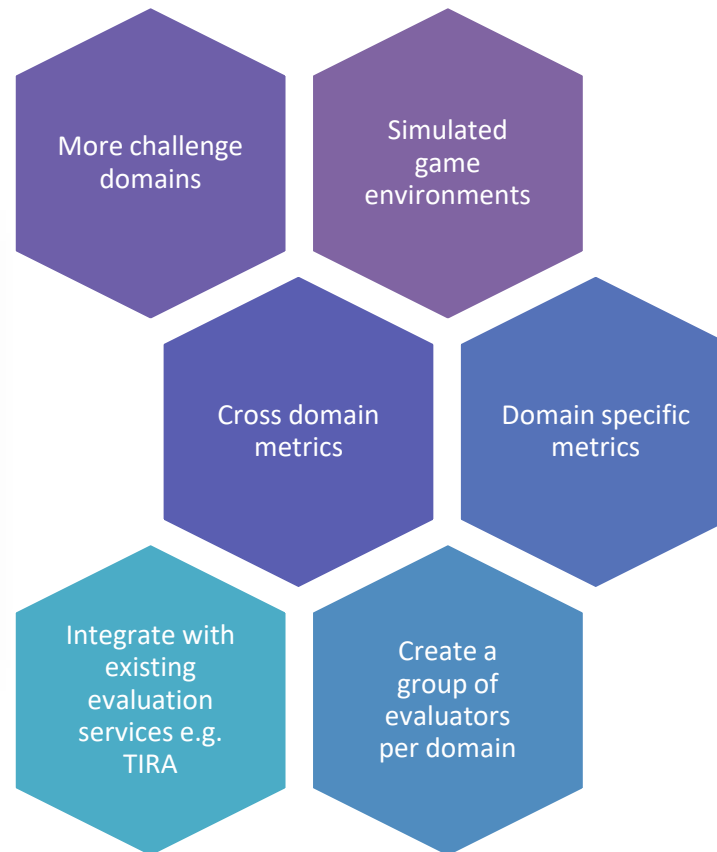
Context Features	Notification Features
activity	appPackage
noise	category
batteryLevel	numberUpdates
charging	subject
headphonesIn	priority
lightIntensity	ongoing
musicActive	visibility
place	
proximity	
ringerMode	
day	
time	
action	





Unique Apps		Unique Subjects	
Real	Synth	Real	Synth
23	8	12	1
8	3	6	2
3	1	4	1
8	4	2	1
12	4	17	2





Thank you.

Questions?

Email:

kieran.fraser@adaptcentre.ie

