第15回テキストアナリティクスシンポジウム

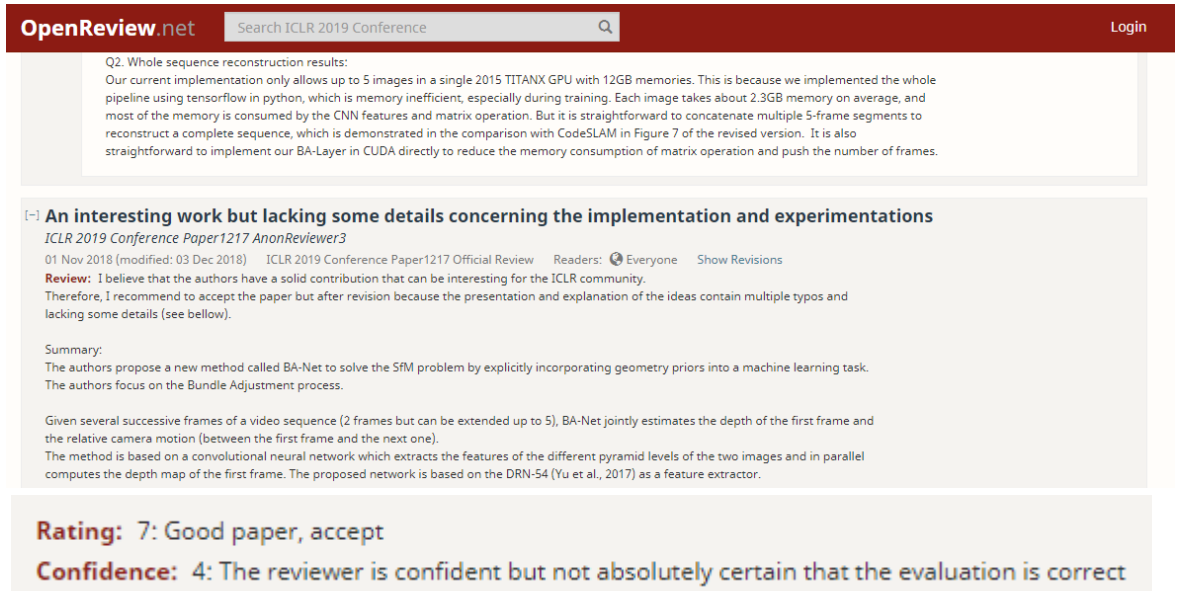
ICLR2019参加報告

フューチャー株式会社　貞光 九月　田中 駿

ICLRの概要

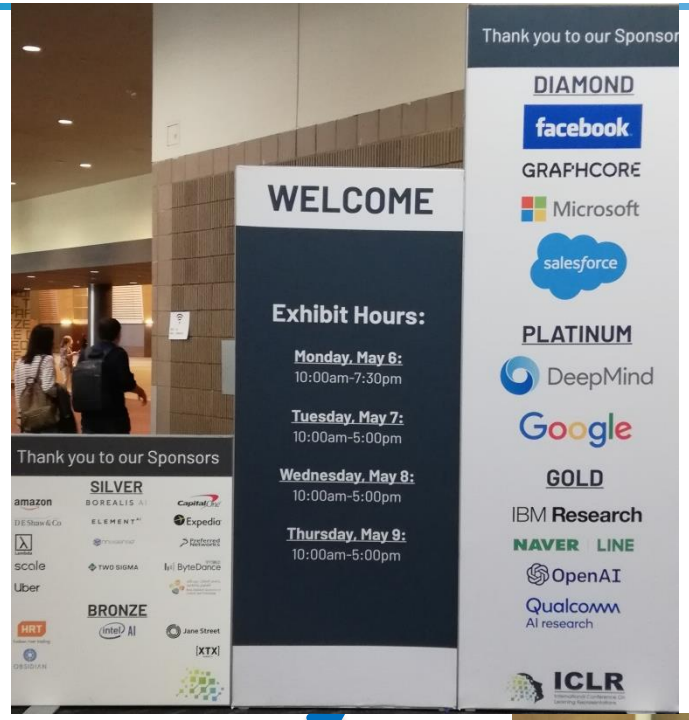
- 7回目と歴史は浅いが、DNN系の質の高い論文が集まる。類似会議として、NeurIPS(NIPS), ICML
- Statistics:
  - Submit 1590件, 1400 reviewers
  - Oral 24件 (accept rate=1.5%)
    - single session. Posterでも発表するのが特徴
  - Poster 454件 (accept rate=28.6%)
    - 国別ではアメリカ、中国が圧倒。ヨーロッパ,韓国も多い。
      - 日本人の発表は5件程度。PFN,リコー,東大。日本人ではないがNEC labの存在感あり
- Invited talk
- Workshop (accept rate > 50-70% ?)
- Sponsor booth

Openreviewの例(oral)

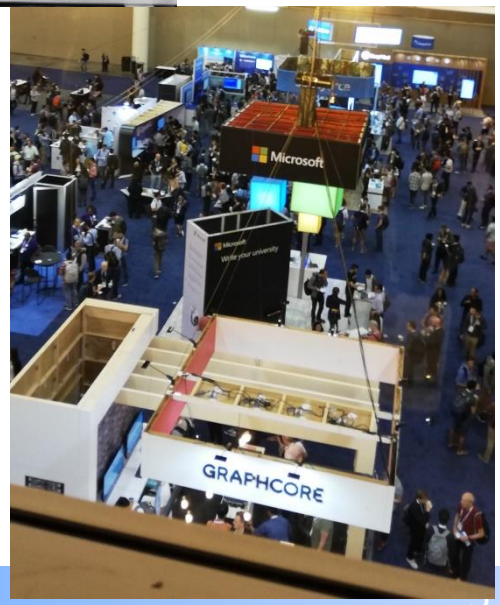


Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

会場の様子とSponsors



いたるところに
スポンサー名が掲示



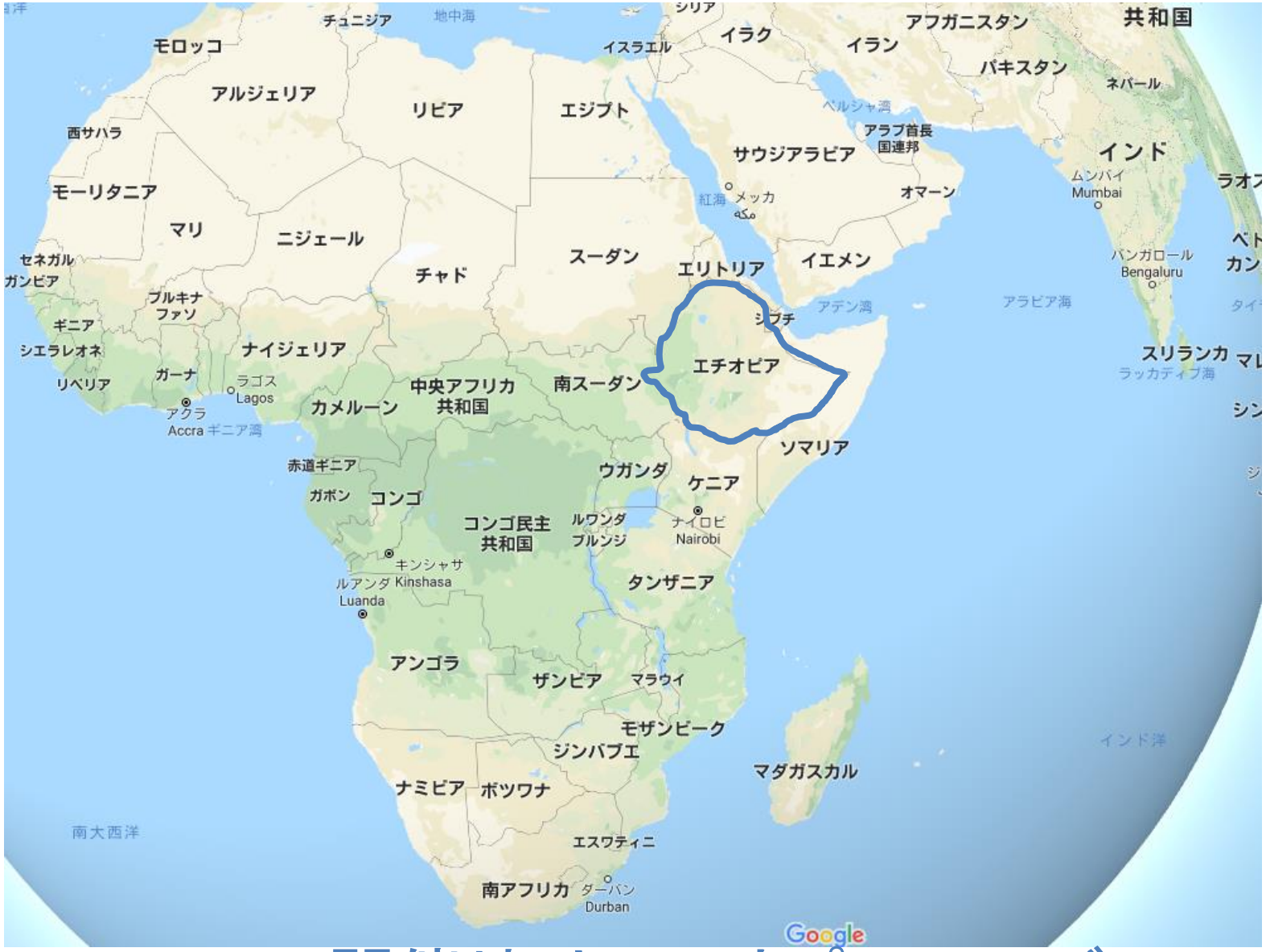
Diamond $80K
Graphcore, MS, salesforce, Facebook

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

Active Steps for ICLR 2019

- Events by **Women in ML, Black in AI, LatinX in AI, Queer in AI,** and others.
- 100% gender parity in program team, invited speakers, and session chairs.
- 50 additional student travel awards given to underrepresented students.
- Workshops each have at least two female organizers.
- Introduction of ICLR code of conduct (https://iclr.co/public/CodeOfConduct)
- Area Chair minor steps from 2018 → 2019
  - 13% → 25% Female
  - ~2% → ~5% Underrepresented minorities
  - 0% → ~5% Geographically diverse institutions

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential



ICLR2020の開催地はエチオピアのアディスアベバ

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

Future → アフリカ・ルワンダ

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

Invited Talk

- ***Highlights of Recent Developments in Algorithmic Fairness***
- ***Can Machine Learning Help to Conduct a Planetary Healthcheck?***

倫理

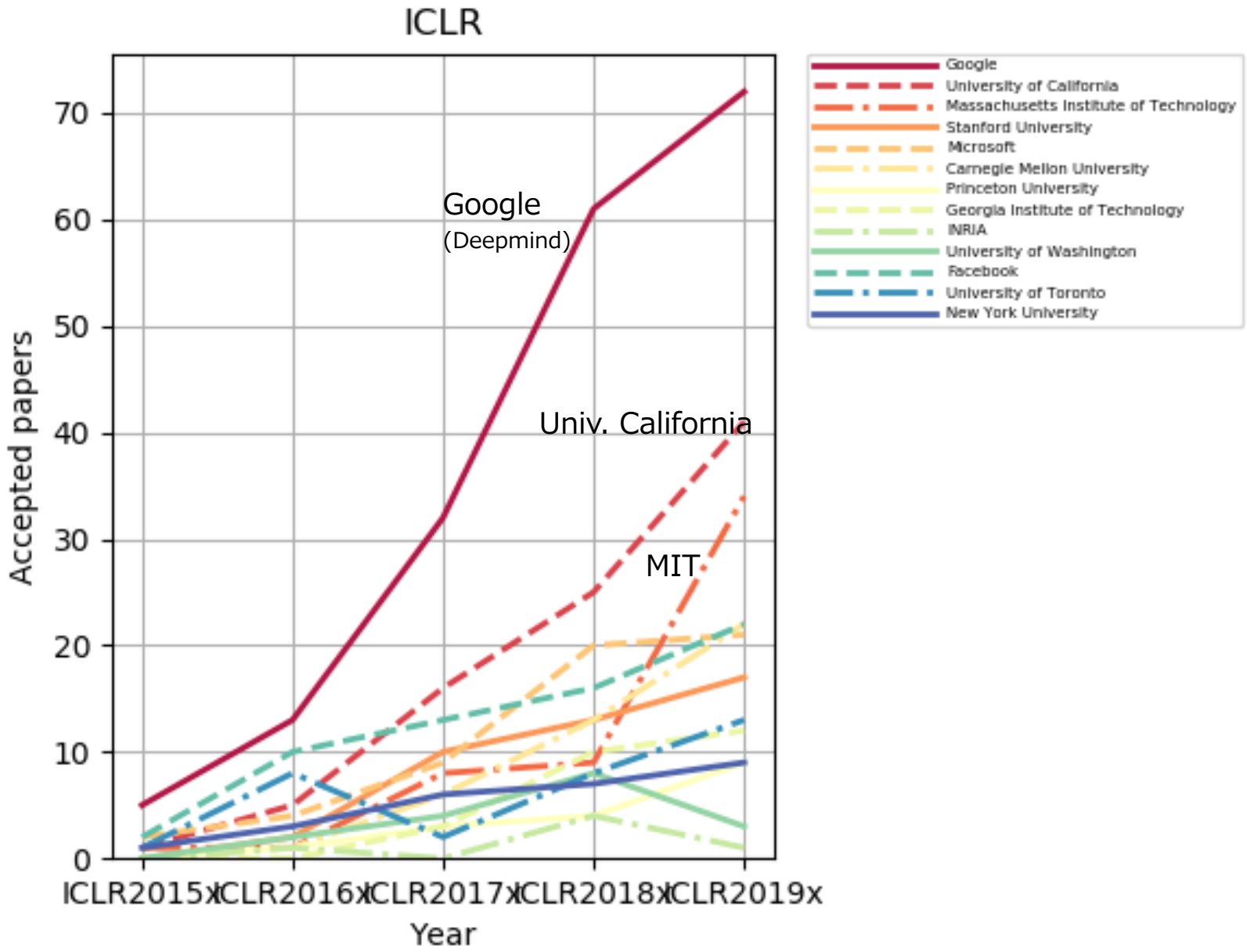
- ***Learning Natural Language Interfaces with Neural Models (by Mirella Lapata)***
- ***Learning (from) language in context***

言語

- ***Adversarial Machine Learning (by Ian Goodfellow)***
- ***While We're All Worried about Failures of Machine Learning, What Dangers Lurk If It (Mostly) Works?***
- ***Learning Representations Using Causal Invariance***
- ***Developmental Autonomous Learning: AI, Cognitive Sciences and Educational Technology***

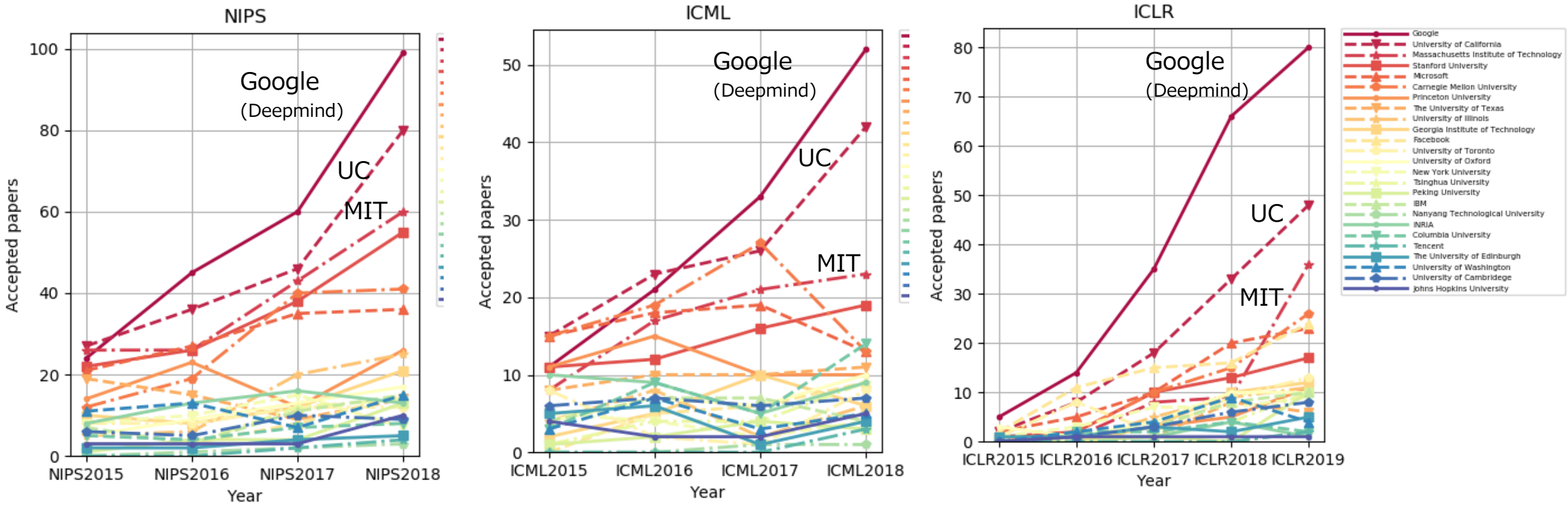
Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

経年変化 × 研究組織ごとの採択数



Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

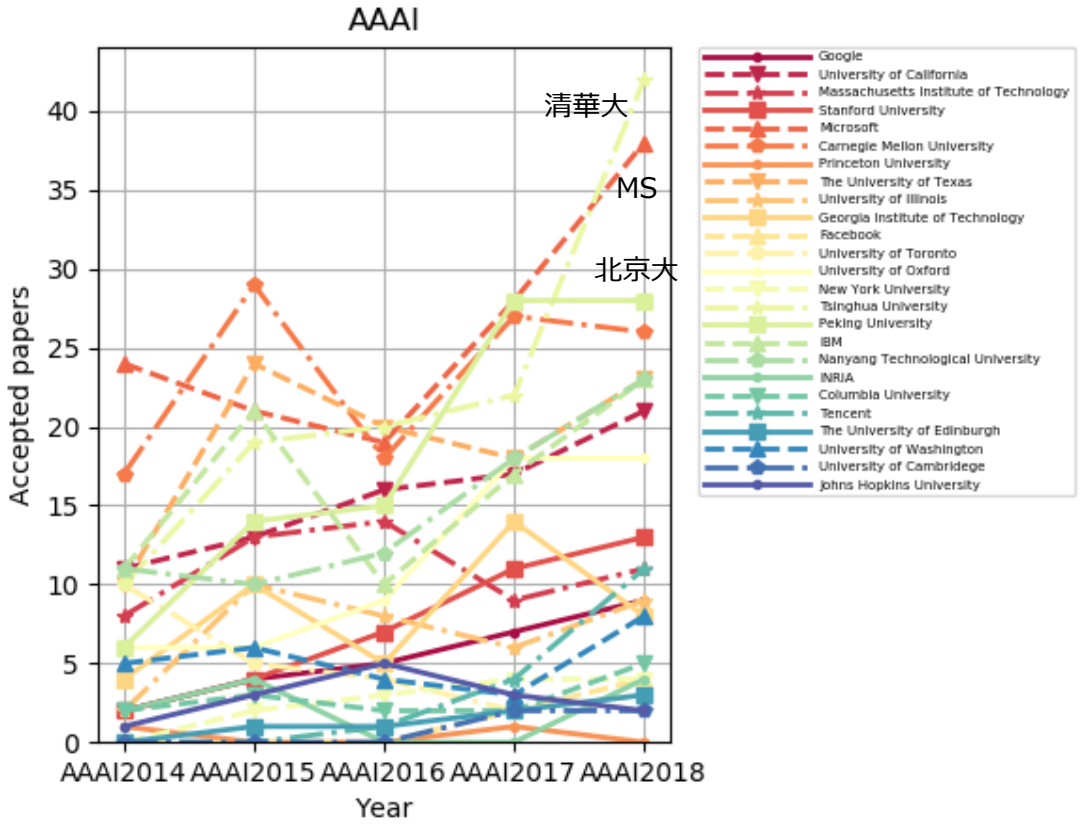
- 機械学習主要3会議のTOP3は同じ

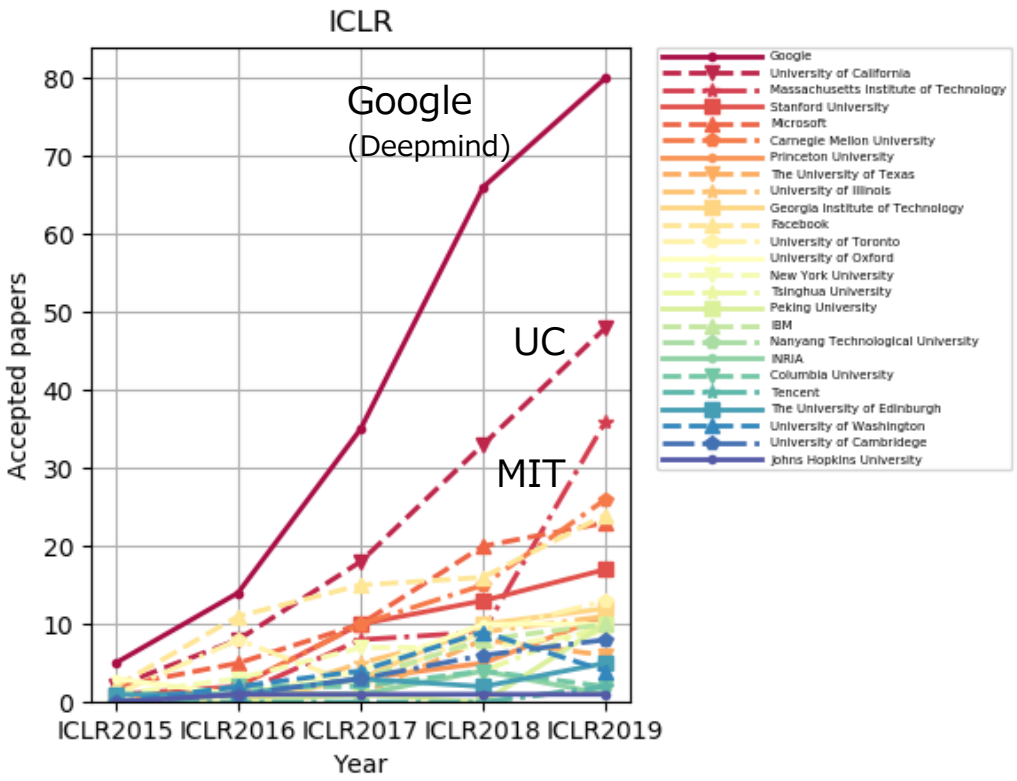


Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

他会議との比較（AAAI）

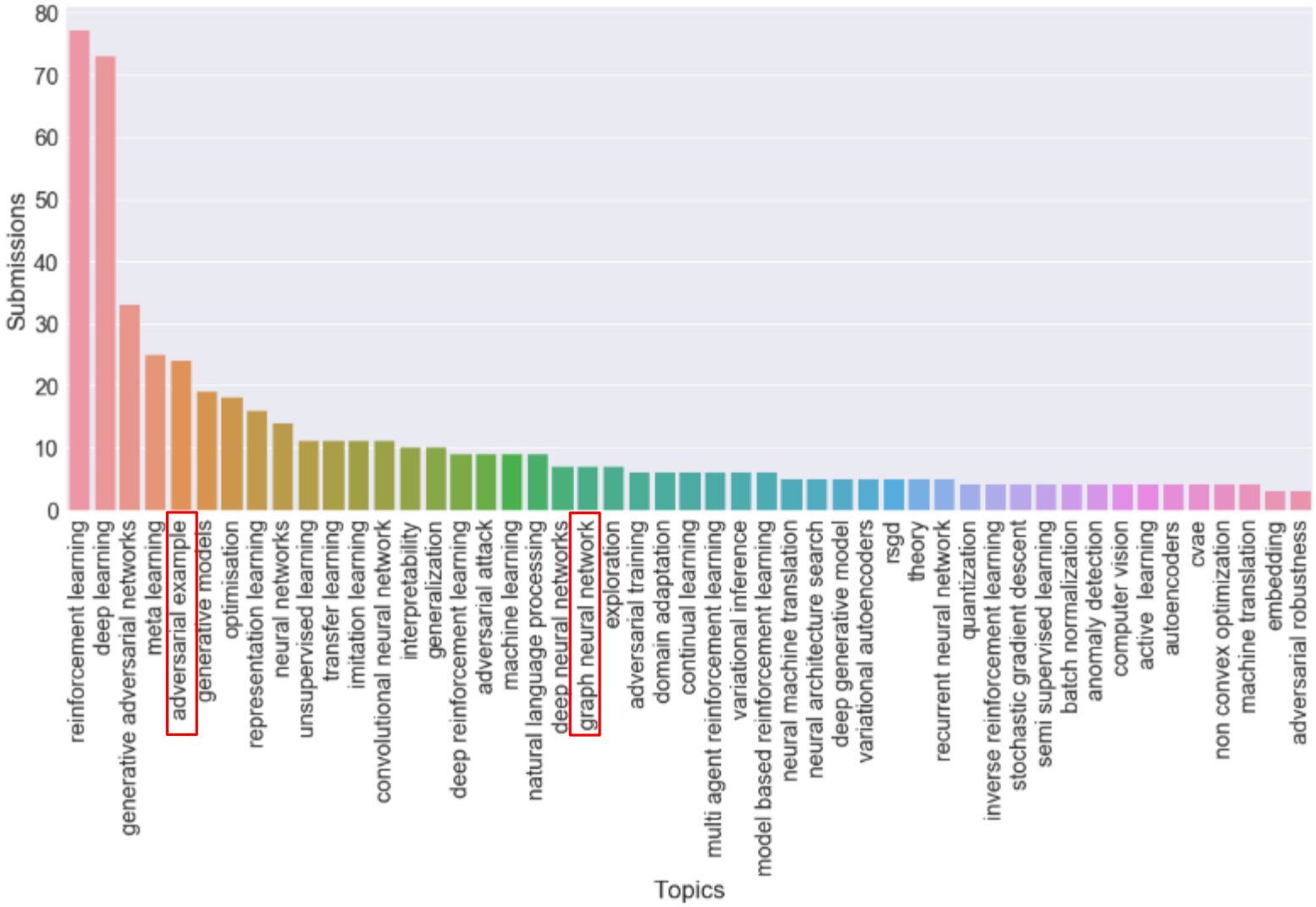
- AAAIとは傾向が大きく異なる。 AAAIは中国勢が特に強い





Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

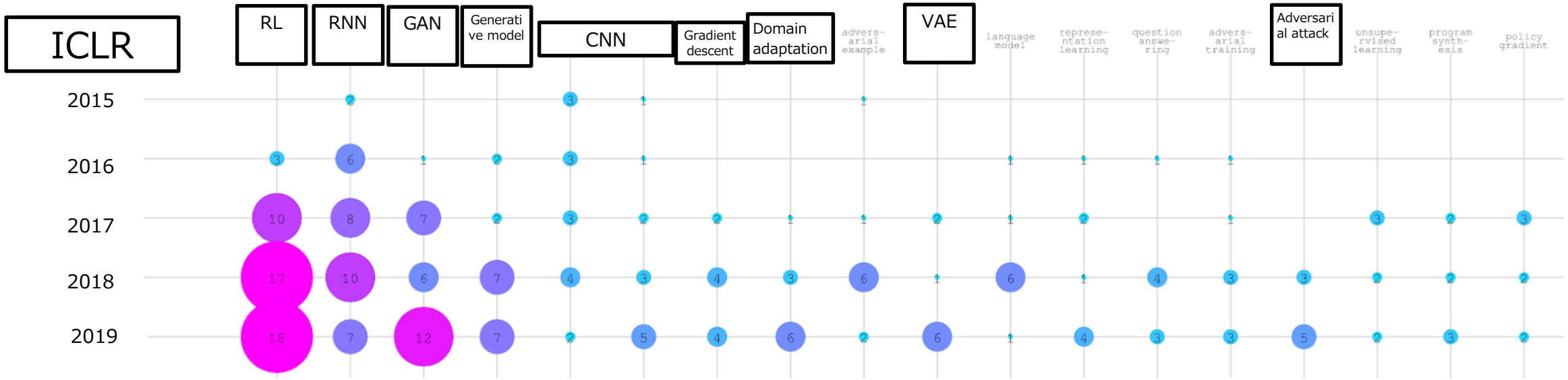
タイトルに含まれるキーワード



引用：https://ailab.criteo.com/iclr-2019-stats-trends-and-best-papers/

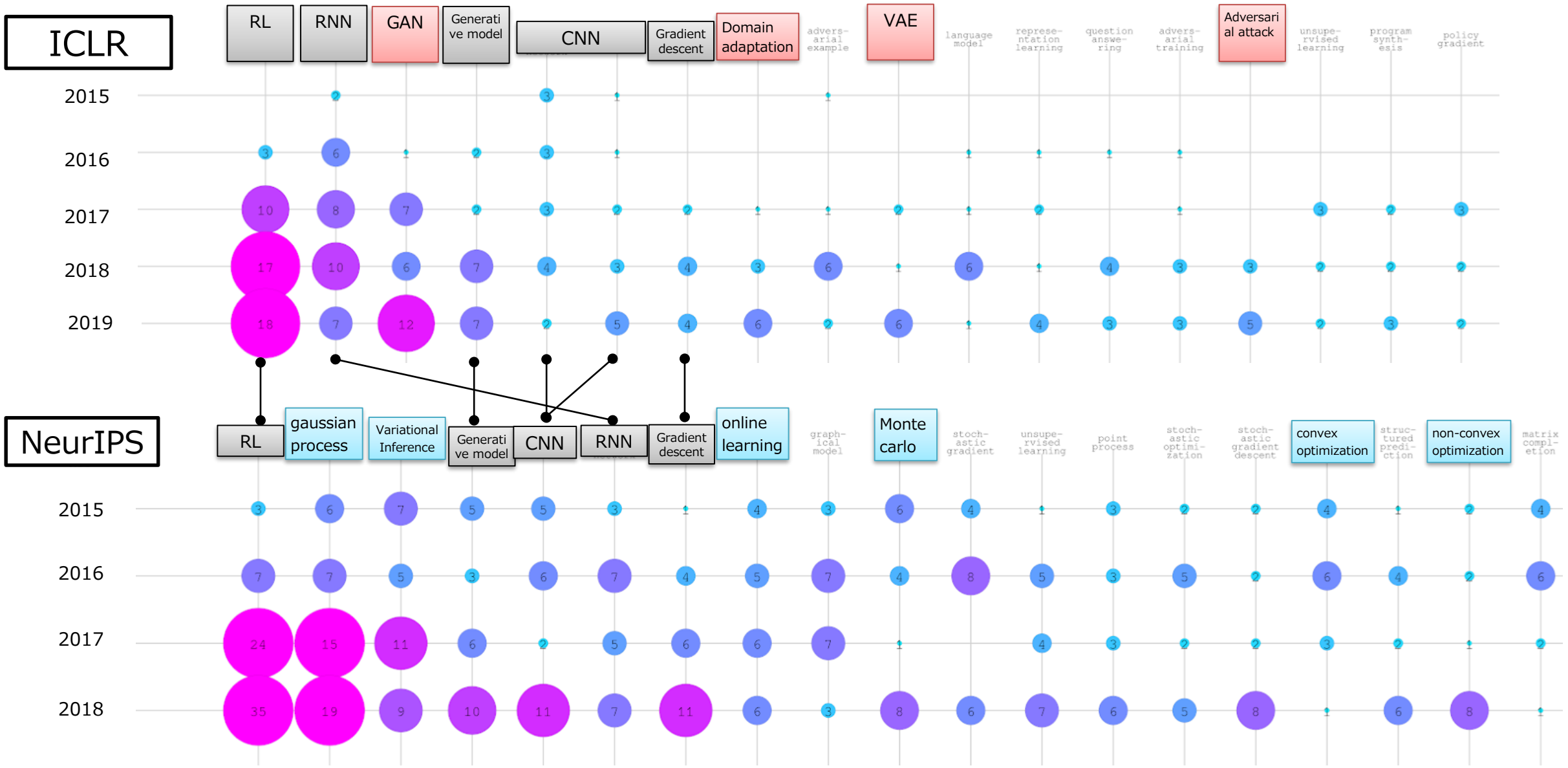
Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

経年変化 × タイトルに含まれるキーワード



Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

経年変化 × タイトルに含まれるキーワード × 他会議との比較



Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

貞光リスト

- Keyword1: Adversarial example
  - ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness (oral)
    - University of Tubingen (ドイツ)
    - https://openreview.net/pdf?id=Bygh9j09KX

- Keyword2: GNN
  - How Powerful are Graph Neural Networks? (oral)

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness

Oral, University of Tubingen (ドイツ)
https://openreview.net/pdf?id=Bygh9j09KX

- 【WHAT/WHY】Imagenet(IN)で学習されたCNNがテキスチャに引きずられる問題を明らかにし、その改善策を示した
- 【HOW】IN画像に対し複数のtexture画像を生成し、IN pretrainモデルに対しfine-tuningするというシンプルなもの
- 【RESULT】IN単独で学習させた場合よりも高い精度

(a) Texture image
81.4% **Indian elephant**
10.3% indri
8.2% black swan

(b) Content image
71.1% **tabby cat**
17.3% grey fox
3.3% Siamese cat

(c) Texture-shape cue conflict
63.9% **Indian elephant**
26.4% indri
9.6% black swan

Figure 1: Classification of a standard ResNet-50 of **(a)** a texture image (elephant skin: only texture cues); **(b)** a normal image of a cat (with both shape and texture cues), and **(c)** an image with a texture-shape cue conflict, generated by style transfer between the first two images.

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

HOW: Stylized-ImageNet / Shape-ResNet

- 【データ】 Imagenet(IN)を色んなstyleに変換し、同じラベルとみなす　（Stylized-ImageNet (SIN) と呼ぶ）

Figure 3: Visualisation of Stylized-ImageNet (SIN), created by applying AdaIN style transfer to ImageNet images. Left: randomly selected ImageNet image of class `ring-tailed lemur`. Right: ten examples of images with content/shape of left image and style/texture from different paintings. After applying AdaIN style transfer, local texture cues are no longer highly predictive of the target class, while the global shape tends to be retained. Note that within SIN, every source image is stylized only once.

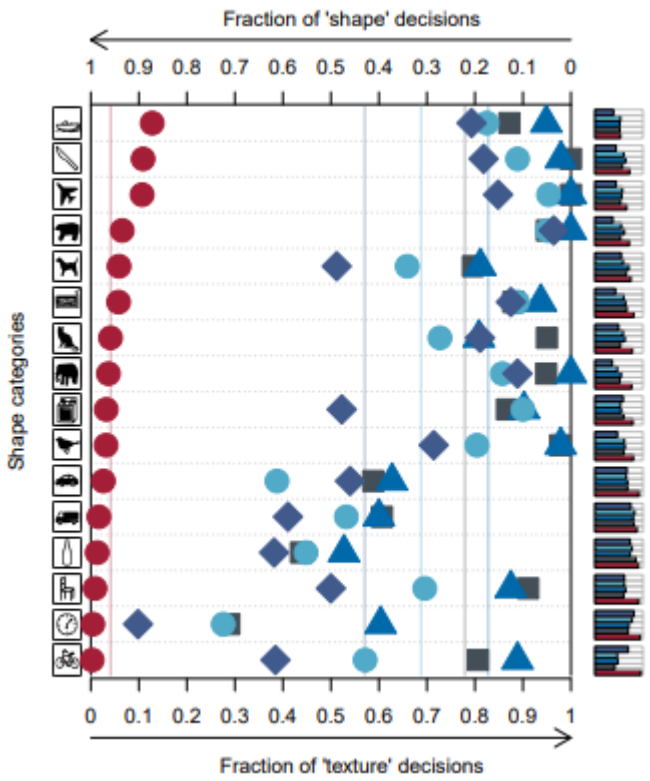
https://github.com/rgeirhos/Stylized-ImageNet

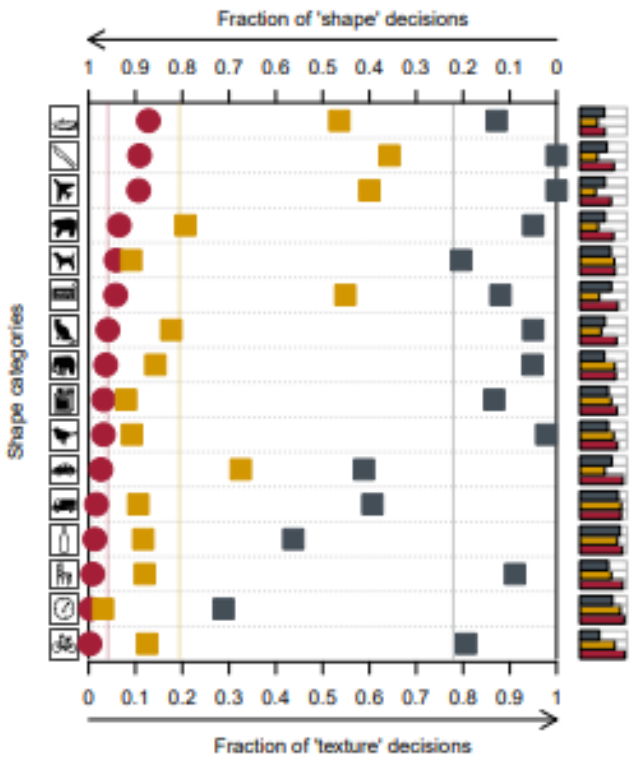
- SIN+INでpretrainしたあと、INのみでfine tunening　（“Shape-ResNet”と呼ぶ）

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

Result

人間が●
Alexnetが◆
VGG16が▲
GoogleNetが●
Resnetが■





人間が●
Resnetが■
Stylized-ImageNet(提案)が■
→SINが人間に近い判定

INでFine-tuning
しないと精度悪化

name	training	fine-tuning	top-1 IN accuracy (%)	top-5 IN accuracy (%)	Pascal VOC mAP50 (%)	MS COCO mAP50 (%)
vanilla ResNet	IN	-	76.13	92.86	70.7	52.3
	SIN	-	60.18	82.62	70.6	51.9
	SIN+IN	-	74.59	92.14	74.0	53.8
Shape-ResNet	SIN+IN	IN	**76.72**	**93.28**	**75.1**	**55.2**

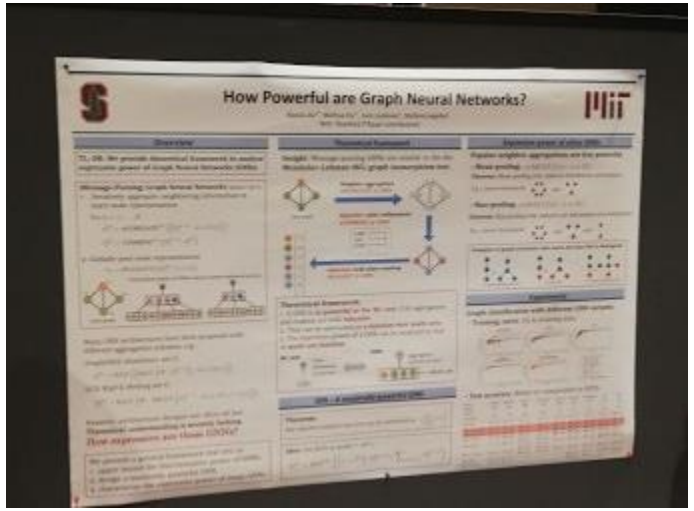
Table 2: Accuracy comparison on the ImageNet (IN) validation data set as well as object detection performance (mAP50) on PASCAL VOC 2007 and MS COCO. All models have an identical ResNet-50 architecture. Method details reported in the Appendix, where we also report similar results for ResNet-152 (Table 4).

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

How Powerful are Graph Neural Networks? (Oral, MIT+Stanford)

https://openreview.net/pdf?id=ryGs6iA5Km

- WHAT: GNNの理論的裏付けとそれに基づく精度改善
- WHY: GNNには色々あるが、理論的背景が不十分。もっと良くできるのでは？
  - GCN、GraphSAGE、DCNN
- HOW:
  - 理論的フレームワークの下(割愛)で、Graph Isomorphism Network（GIN)を提案
  - 加えて、従来ノードの表現は周辺ノードのmeanでやっていたのに対し、sumにすることで精度改善
- RESULT: 従来のSOTAと同等以上の精度を示す



Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

HOW : GRAPH ISOMORPHISM NETWORK (GIN)

GNNの一般的な形式による第k層の記述

$$a_v^{(k)} = \text{AGGREGATE}^{(k)}\left(\left\{h_u^{(k-1)} : u \in \underline{\mathcal{N}(v)}\right\}\right), \quad h_v^{(k)} = \text{COMBINE}^{(k)}\left(h_v^{(k-1)}, a_v^{(k)}\right), \tag{2.1}$$

ノードvの隣接ノード

h_v: vector of node

Graph SAGE (NIPS2017) の第k層の記述

AGGREGATE　　COMBINE

$$a_v^{(k)} = \underline{\text{MAX}}\left(\left\{\text{ReLU}\left(W \cdot h_u^{(k-1)}\right), \forall u \in \mathcal{N}(v)\right\}\right), \quad h_v^{(k)} = W \cdot \left[h_v^{(k-1)}, a_v^{(k)}\right]$$

GCN (ICLR2017) の第k層の記述

$$h_v^{(k)} = \text{ReLU}\left(W \cdot \underline{\text{MEAN}}\left\{h_u^{(k-1)}, \forall u \in \mathcal{N}(v) \cup \{v\}\right\}\right). \tag{2.3}$$

隣接ノード＋自ノード

GINの第k層の記述: 全部足す。

$$h_v^{(k)} = \underline{\text{MLP}^{(k)}}\left(\underline{\left(1+\epsilon^{(k)}\right) \cdot h_v^{(k-1)} + \sum_{u \in \mathcal{N}(v)} h_u^{(k-1)}}\right). \tag{4.1}$$

自ノード＋隣接ノード

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

HOW：GRAPH ISOMORPHISM NETWORK (GIN)

- 表現能力としては、sumがmean,maxより偉いはずで、それを理論的にも裏付けられたモデルが偉い＝GIN！

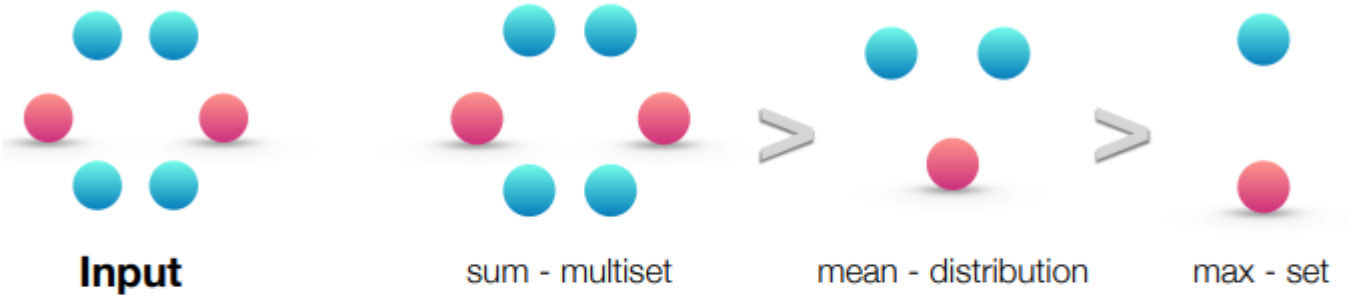
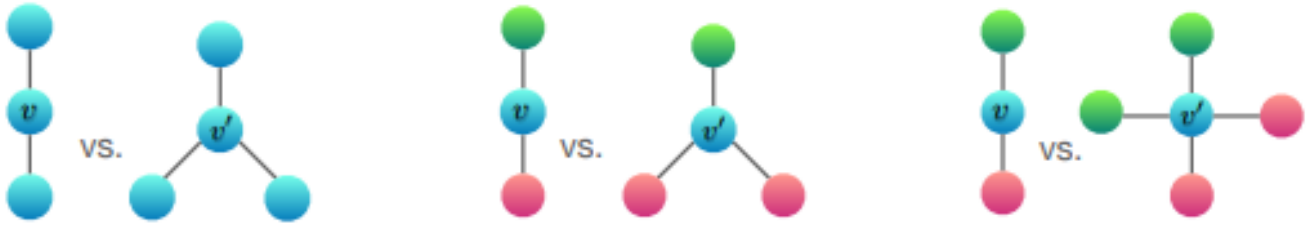


Figure 2: **Ranking by expressive power for sum, mean and max aggregators over a multiset.** Left panel shows the input multiset, *i.e.*, the network neighborhood to be aggregated. The next three panels illustrate the aspects of the multiset a given aggregator is able to capture: sum captures the full multiset, mean captures the proportion/distribution of elements of a given type, and the max aggregator ignores multiplicities (reduces the multiset to a simple set).



(a) Mean and Max both fail (b) Max fails (c) Mean and Max both fail

Figure 3: **Examples of graph structures that mean and max aggregators fail to distinguish.** Between the two graphs, nodes v and v' get the same embedding even though their corresponding graph structures differ. Figure 2 gives reasoning about how different aggregators "compress" different multisets and thus fail to distinguish them.

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential

Result

- シンプルなのに、他手法と比べて良い感じ。

	Datasets	IMDB-B	IMDB-M	RDT-B	RDT-M5K	COLLAB	MUTAG	PROTEINS	PTC	NCI1
Datasets	# graphs	1000	1500	2000	5000	5000	188	1113	344	4110
	# classes	2	3	2	5	3	2	2	2	2
	Avg # nodes	19.8	13.0	429.6	508.5	74.5	17.9	39.1	25.5	29.8
Baselines	WL subtree	73.8 ± 3.9	50.9 ± 3.8	81.0 ± 3.1	52.5 ± 2.1	78.9 ± 1.9	90.4 ± 5.7	75.0 ± 3.1	59.9 ± 4.3	**86.0 ± 1.8** *
	DCNN	49.1	33.5	–	–	52.1	67.0	61.3	56.6	62.6
	PATCHYSAN	71.0 ± 2.2	45.2 ± 2.8	86.3 ± 1.6	49.1 ± 0.7	72.6 ± 2.2	**92.6 ± 4.2** *	75.9 ± 2.8	60.0 ± 4.8	78.6 ± 1.9
	DGCNN	70.0	47.8	–	–	73.7	85.8	75.5	58.6	74.4
	AWL	74.5 ± 5.9	51.5 ± 3.6	87.9 ± 2.5	54.7 ± 2.9	73.9 ± 1.9	87.9 ± 9.8	–	–	–
GNN variants	SUM–MLP (**GIN-0**)	**75.1 ± 5.1**	**52.3 ± 2.8**	**92.4 ± 2.5**	**57.5 ± 1.5**	**80.2 ± 1.9**	**89.4 ± 5.6**	**76.2 ± 2.8**	**64.6 ± 7.0**	**82.7 ± 1.7**
	SUM–MLP (**GIN-ε**)	**74.3 ± 5.1**	**52.1 ± 3.6**	**92.2 ± 2.3**	**57.0 ± 1.7**	**80.1 ± 1.9**	**89.0 ± 6.0**	**75.9 ± 3.8**	63.7 ± 8.2	**82.7 ± 1.6**
	SUM–1-LAYER	74.1 ± 5.0	**52.2 ± 2.4**	90.0 ± 2.7	55.1 ± 1.6	**80.6 ± 1.9**	**90.0 ± 8.8**	**76.2 ± 2.6**	63.1 ± 5.7	82.0 ± 1.5
	MEAN–MLP	73.7 ± 3.7	**52.3 ± 3.1**	50.0 ± 0.0	20.0 ± 0.0	79.2 ± 2.3	83.5 ± 6.3	75.5 ± 3.4	**66.6 ± 6.9**	80.9 ± 1.8
	MEAN–1-LAYER (GCN)	74.0 ± 3.4	51.9 ± 3.8	50.0 ± 0.0	20.0 ± 0.0	79.0 ± 1.8	85.6 ± 5.8	76.0 ± 3.2	64.2 ± 4.3	80.2 ± 2.0
	MAX–MLP	73.2 ± 5.8	51.1 ± 3.6	–	–	–	84.0 ± 6.1	76.0 ± 3.2	64.6 ± 10.2	77.8 ± 1.3
	MAX–1-LAYER (GraphSAGE)	72.3 ± 5.3	50.9 ± 2.2	–	–	–	85.1 ± 7.6	75.9 ± 3.2	63.9 ± 7.7	77.7 ± 1.5

Table 1: **Test set classification accuracies (%).** The best-performing GNNs are highlighted with boldface. On datasets where GINs' accuracy is not strictly the highest among GNN variants, we see that GINs are still comparable to the best GNN because a paired t-test at significance level 10% does not distinguish GINs from the best; thus, GINs are also highlighted with boldface. If a baseline performs significantly better than all GNNs, we highlight it with boldface and asterisk.

Copyright 2019 by Future Corporation
Future Corporation Proprietary and Confidential