International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

The Significant Of Biases In Learning Algorithms Generalization

Apoorva Ganapathy[1], Md. Shakawat Hossain[2],Md. Mahofuzur Rahman[3], ABM Asadullah[4], Ruhul Amin[5]

[1]*Senior Developer, Adobe Systems, San Jose, California, United States*
[2]*Associate Professor, Department of Accounting & Information System, Jagannath University, Dhaka 1100, Bangladesh*
[3]*Assistant Professor & Chairman, Department of Business Administration, Port City International University,Bangladesh*
[4]*Department of Economics and Management Sciences, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
[5]*Senior Data Entry Control Operator (IT), ED-Maintenance Office, Bangladesh Bank (Head Office), Dhaka, Bangladesh*

Abstract: In machine learning, presumed a limited set of examples there is typically numerous explanation that can flawlessly appropriate working out data; nonetheless, the 'inductive bias of learning algorithm' chooses and place in order those answer that comes to an understanding with its statement as mentioned above. However, when there are no understandings in the analytic procedure, a likely approach to examine this bias is to investigate the feedback/outcome performance of the learning algorithm. The problem with this method is that both feedback and outcomes are in elevated height, for example spreading over images or overlapping, making it hard to distinguish the feedback-outcome relationship thoroughly. An approach for investigating "high-dimensional" devices is to task them against a lesser dimensional cosmos where the investigation is possible. It was to this gap that we find it interested in investigating the feedback-outcome relationship of system, with the help of biases generalization of learning intelligence. This article will analyze its performance by sticking out the image interplanetary onto a prudently selected low dimensional property of interplanetary. Motivated by investigational approaches from cognitive psychology, we investigate respective learning algorithms with prudently planned working out datasets to illustrate when and how the prevailing models produce new characteristics and their blends. We classify resemblances to human psychology and confirm that these patterns are reliable and steady across generally utilized prototypes and structural designs.

Keywords: Learning algorithm, deep learning, generalization, biases, Cognitive psychology

1. INTRODUCTION

The purpose of a concentration prediction algorithm is to understand a distribution from working-out data. Moreover, consistent and unbiased concentration prediction is unrealistic (Rosenblatt, 1956; Efromovich, 2010). The same thing applies to distinct scenery, where the volume of likely distributions measures even more exponentially concerning dimensionality (Arora and Zhang, 2017), recommending tremendously high data demands. Because of this, the statement developed by a 'learning algorithm', or its inductive bias, is crucial when there is the involvement of performing data management. In place of mere concentration prediction algorithms, including interpolating a Gaussian distribution through maximum probability, which can describe the distribution that is generated assuming some working-out data. Moreover, for composite algorithms consisting of deep generative patterns including 'Generative Adversarial Networks (GAN)' and 'variation automatic coders (VAE)' (Kingma and Welling, 2013; Goodfellow et al., 2014; Rezende et al., 2014; Ho and Ermon, 2016; Zhao et al., 2018; Ahmed et al., 2021), the pattern of the inductive bias is very challenging to distinguish.

When there are no understandings in the analytic procedure, a viable approach to examine this bias is to investigate the feedback/outcome performance of the learning algorithm. The problem with this method is that both feedback and outcomes are in elevated height, for example spreading over images or overlapping, making it hard to distinguish the feedback-outcome relationship thoroughly. An approach for investigating "high-dimensional" devices is to task them against a lesser dimensional cosmos where the investigation is possible. Also, the same problems are what has affected 'cognitive psychologists'. As graphical perception works are very complex, 'cognitive psychologists' and 'neuroscientists' have created well-ordered experiments to examine the optical structure. For instance, research on cognitive and demonstration of color, shape, etc., has resulted in crucial innovations that include ensemble representation (Alvarez, 2011; Ganapathy, 2020), model improvement effect (Minda and Smith, 2011), and Weber's hypothesis (Stevens, 2017). Owing to the challenges as mentioned above, we plan to take on experimental approaches from perception psychology to illustrate the significance of biases in learning algorithm generalization.

The Objective of the Study

The objective of this study is to illustrate the feedback-outcome relationship of the system with the help of biases generalization of learning intelligence. This article will investigate its performance by sticking out the image interplanetary onto a prudently selected low dimensional property of interplanetary.

2. PROBLEM STATEMENT

Employing a biased framework has enabled a systematic approach to evaluate generalization models of the sophistication patterns, including 'Generative Adversarial Networks' and 'variation automatic coders' (Kingma and Welling, 2013; Azad et al., 2021). It is interesting to know that this pattern is dependable through models, datasets, and hyper-variables selections. Additionally, hitherto conveyed experiments on perception psychology show that most of these designs have remarkable resemblances (Ganapathy, 2021a). For instance, as soon as offered with a working out fixed where all images comprise precisely three matters, in cooperation Generative Adversarial Networks' and 'variation automatic coders usually create 2-5 device, with log-normal designed dissemination (Figure 1). If the working out set comprises multiple approaches (e.g., all pictures contain one or the other 2 or 10 devices) the performance is observed to be the same as that of a lines strainer — the algorithm performances as if it is worked out distinctly on 2 and 10 devices and then means the two distributions. An exemption is as soon as the modes are close to each other (e.g., 2 and 4 devices) where the prototype is observed for improvement (Minda and Smith, 2010).

International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

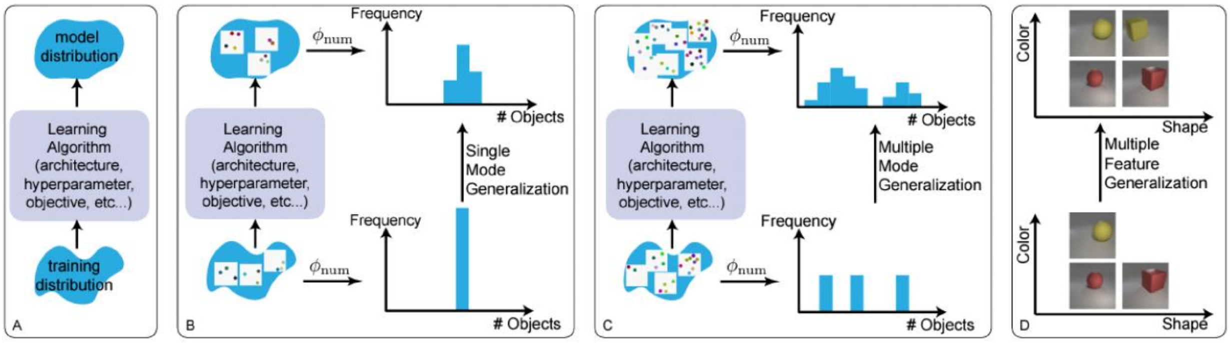

Figure 1: Design Dissemination

Studying Generalization via Probing Structures

Cognitive psychology was the source of inspiration and offered a new structure to investigate the inductive bias of procreative systems through a set of analytical frameworks. The probing was concentrated on images; nonetheless the methods can also be functional to additional domains.

The support set of *p(x)* is given as $\mathcal{S} \subset \mathcal{X}$. A set of probing structure as a tuple of functions is defined as $\phi = (\phi_1\ldots, \phi_k)$ where respective ϕ_1 plots an input in *S* to a value. For instance, one of the structures ϕ_i: $S \rightarrow N$ possibly willplot the input graphic to the number of entities (numerosity) in that graphic(Figure 2). The feature space, Z is denoted as ϕ range. For any selection of *p(x)*, with a minormishandling of symbolization, we stand forp(z) as the (induced) distribution on Z by ϕ(x) when x ~ p(x). Instinctivelyp(z) is the forecast of p(x) onto the structureplanetary Z.

When a learning algorithm *A* gives a learned distribution q(x), it isa scheme to feature planetaryutilizing ϕ. Our objective is to explorein what way p(z) varies from q(z), that is, the generalization performance of the learning algorithm constrained to the feature planetary Z. In the input planetary X, even estimating the spaceamid p(x) and q(x) is problematic. However in feature planetary Z whichcan notmerelyselect if q(z) is diverse from p(z) nonetheless also describein what way they are dissimilar. For instance, if p(z) is a distribution over graphics with blue and redthree-way relationship (,) and red circles ( ),we can explore whether q(z) takes a broad view to blue circles ( ). The number of colors for the respective circle is investigated that is; the color that essentiallyis in the working out data so that q(z) produces circles of all colors. Such inquiries are significant to describe the inductive bias of prevailingprocreative modeling algorithms.

A similar model described by Gretton et al. (2007) has been employed earlier to assess the space amid p(x) and q(x). Specifically, the FID grade (Heusel et al., 2017), the mode grade (Che et al., 2016), and the inception grade (Salimans et al., 2016) utilize hidden labels/features of a pre-worked-out CNN classifier as ϕ and determine the behavior of generative patterning algorithms by likening p(z) and q(z) beneath this estimation. In disparity, for the reason that the focus is to investigate the accurate variance between q(z) and p(z), we select ϕ to be interpretable high-level labels or features stimulated by experimental work in perception psychology, example includes color, numerosity, etc. employing low dimensional estimation utility ϕ has added advantage. The reason is that Z is low discrete and dimensional in the synthetic datasets and isessential in the immeasurable data system. In all the experiments carried out, the support of p(z) is set within the range of 500, so the exact calculation of p(z) (Rosenblatt, 1956) with the rationally sized dataset that is between hundreds of thousands to million samples in the experiments). The most important remark is that although D is an actual or precise calculation of p(z), the learned dissemination q(z) is not, so this abridged setting is adequate to disclosenumerousfascinatinginductive biases of the patterning algorithms.

Evaluation and feature assortment

Feature ϕ that fulfill 2 requirements which include:

- They are essential to human cognition, and have been investigated in perception psychology, and
- They are stress-free to examine one or the other by human ruling or consistent algorithms.

The investigated features consist of shape, size, color, numerosity, and location of the respective object. For shape and numerosity, independent assessments are employed via 3 human assessors, whereas other studied features are simple to assess by automated algorithms.

Model

To ascertain that the outcome is not influenced by the selection of hyper variables and model structure were adopted, 2 dissimilar model families were adopted, are Generative Adversarial Networks (GAN) – WGAN-GP (Gulrajani et al., 2017) and variation automatic coders (Kingma and Welling, 2013). Dissimilar network structures and hypervariable selections, including both fully connected networks and conventional networks.

Distinguishing Generalization on a Separate Feature

In this subsection, generalization is investigated when projecting the input distance X to a single label that is p(z) is a single-dimensional distribution, the learning algorithm's output is p(z) is scrutinized first when label or feature and is then operated to comprise an only value that is p(z) is delta function or unit impulse. The question such as stated below is asked; while all pictures in the working out set portray 5 objects, in what way numerous objects will the generative classic yield? It is expected by one that since the labeltakings a single value, and then has hundreds of thousands of discrete instances, the learning algorithm would snapprecisely this stable feature value (Ganapathy, 2021b). Though this is not factual, representation ofrobust inductive bias.

When learned dissemination q(z) is called, the input dissemination has a single-mode impulse response of the exhibiting algorithm. This terminology is lent from signal doling out principle for the reason by finding out the performance is the same to that of a linear filter, if p(z) is buttressed on many values, the principle's outcome q(z) are close by together. In this instance, findingmodelimprovementresults and the learning algorithm createsdissemination that "combines" the 2 modes. In conclusion, we validate our method of learning respective single labelsseparately by presenting that the learning algorithm's performance on the respective label is frequentlyself-regulating of additional features we learning.

Convolution Effect and Prototype Enhancement Effect

Probing the algorithm's performance when p(z) is unimodal, we study its performance when p(z) is multi-modal. We observe that in label distance the outcome distribution can be very well categorized by convolving the feedback dissemination with the learning algorithm's outcome on the respective individual manner (impulse reply) if the feedbackapproaches are far from respective other (the same to a linear filter in signal dealing out). Nevertheless, we discover that this no extensiveclutches when the impulses are nearby to each other, anywhere we detect that the classicproduces unimodal and more focuseddissemination than involvedness would forecast. We request this consequencepatterndevelopment in similarity with the designimprovementoutcome in understanding psychology (Smith and Minda, 2000; Minda and Smith, 2011).

Individuality of Structures

Here wedisplay that respective features we cogitate can be investigatedautonomously of the other. We discover that the generalization performancebeside a specific feature measurement is justly stable as we modify the distribution in other dimensions. As a result, we can decompose the examinationcrosswaysscopes.

3. EXPERIMENTAL METHODS

Numerosity

Two different datasets are used for this study, a toy dataset anywhere are k non-spreading or non-overlapping dots that is with random location and color in the picture, as in the numerosity, prediction task in perception psychology (Nieder et al., 2002; Piazza et al., 2004; Ganapathy, 2021c), and the CLEVR dataset where there are k objects(with random shape, location, color, and size) in the section (Johnson et al., 2017). Instanceworking out and produced images are presented in Figure 2.

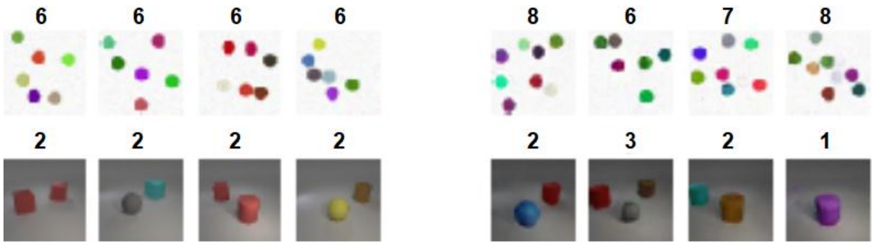

Figure 2:An instance of working out and producing pictures with numerosity annotated

Color ratio

For this label, we employ the dataset presented in Figure 3. Respective pie has numerous features: ratio of size zsize, red color zred, and location zloc. In these tests, we selectthe ratio of red to be 15%, 35%, 55%, 75%, 95% respectively, although the other features (location and size) are designatedhomogeneously at random in the thoroughgoing range endorsed in the dataset.

Figure 3: Pictures from working out and created a set

Difficulty Consequence and Pattern Improvement Consequence

For these tests, we use the color ratio feature of the pie dataset in Figure 4. We work out the classic with 2 bimodal disseminations, one with 30% or 40% red (2 closemodes), and the other with 30% or 90% red (2 distant modes). We also discover numerous otherselections of feature/modes, and they demonstrate undistinguishable designs.

Individuality of Structures

For these tests, we use the pie dataset in Figure 3. We study allthree features: ratio of size,zsize; red color, zred, and location,zloc, and display that the learningalgorithmsreact to respectiveself-determining of the other features. For respective feature, we choose 3fixed standards (0.3, 0.4, 0.9 for fraction of red color, 0.5 0.55 0.8 for size, and -0.05 0.0 0.2 forlocation). For the respective fixed importance of the feature beneathtraining, the other sorts can take 1-50 randomstandards. For instance, when learning if generalization on the fraction of red color,zred is self-determiningof other sorts, the training dissemination p(zred; zsize; zloc) is selected such that the peripheral onp(zred) is unvarying on 0:3; 0:4; 0:9g. If zred is self-determining on of the other features, the learneddisseminationq(zred) would only rest on this peripheral p(zred) then not p(zsize; zlocjzred). Tosearchfor dissimilarchoices for p(zsize; zlocjzred) we choose 1-50 random standards as the backing of thisrestricteddissemination (Vadlamudi et al., 2021). This covers a very wide range of connectionsamidzred and the other 2features from intenselyconnected (1 value) to very faintlyconnected (50 values).

Describing Generalization on Numerous Features

This subsection focuses on the combined distribution over numerous features. However, we investigate when a learning algorithm worked out on a small number of amalgamations can generalize to new ones. We observe that if the working out distribution only consists of a few number of blends that are ten to twenty in a feature distance Z, the learned dissemination learns them virtually correctly. Moreover, as there are more blends in the working out set, the pattern commences creating new ones (Ganapathy et al., 2021). We find this performance to be very reliable in a transversely dissimilar setting.

Three dissimilar datasets were employed to study generalization on numerous features. These include:

- Pie Dataset –the datasets as presented in Figure 3 is used, which include 4 features namely; size (five potential values), y location (nine possible values), x location (nine values), a fraction of red color (five values). There are two thousand approximately potential blends, and randomly choose from ten to four hundred blends as p(z) to create our working out set.
- 3 MNIST – Pictures that consist of 3 MNIST digits were used. For the respective working out sample, we first arbitrarily sample a 3 number amid 000 to 999, but for the respective digit, we sample arbitrarily MNIST fitting to that group. There are one thousand blends, and we arbitrarily choose from ten to hundred and sixty of them to produce our working-out set.
- 2 CLEVER object – CLEVER dataset where the respective object has 22 features; its geometric shape and it is color. One shape is selected from this dataset which takes only a sector of the likely colors, and one color that allocated to only a section of the likely shapes.

4. RESULT AND DISCUSSION

Numerosity

As presented in Figure 2 and quantitatively investigated in Figure 4 – comprises of 3 images, the first image shows the amount of dots the learning algorithm is workedout on, and the solid line is the dissemination over the amount of dots the patterns produces, while the second image in the center shows the dissemination over the amount of CLEVER objects the system produced. Producing CLEVER is harder so we investigate few numerosities, then the generalization model is the same as dots. The last image in Figure 4 shows the numerosity cognitive of monkeys according to Nieder and Miller (2003), respective solid interpolations show the probability a monkey judges an incentive to have the same numerosity as an orientation incentive. In cooperation with colored dots and CLEVER tests, the learned dissemination does not yield a similar number of items as in the dataset on which it was workedout (Khan et al., 2021). The dissemination of the numerosity in the produced pictures is centered at the numerosity from the dataset, with a slight bias towards over-prediction. For instance, when workedout on the pictures with 6 dots as shown in Figure 2, the created pictures consist of anywhere from 4 to 9 dots. So far, studies have shown neurons that answer to numerosity in human and monkey brains (Nieder et al., 2002; Piazza et al., 2004; Ganapathy, 2021d). According to both performance data and neural data, 2 noticeable features about these neurons were reported by Nieder and Merten (2007). These are;

- The larger differences forlarger numerosity, and
- Asymmetric reply with more moderate slopes for larger numerosities likened to fewer one

It is interesting to know that deep generative patterns generalize in the same manner with respect to the numerosity property.

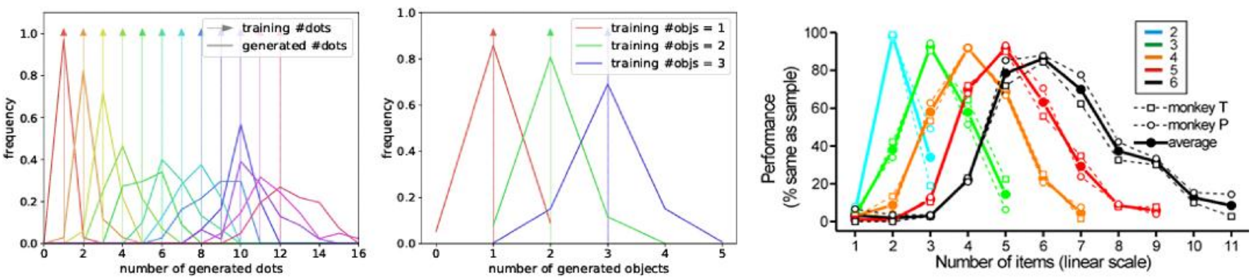

Figure 4: Quantitative investigation of a number of dots in the images

Color Ratio

Figure 5 presented the result of the color ratio. It was observed that the learned property dissemination q(z) is well verging on by a Gaussian centered at the value the system is worked out on. The shape for the small ratio at ten percent and the larger ratio at ninety percent is what was observed. This was dependable for Weber's law (Stevens, 2017) that states humans are better sensitive to comparative alteration than final adjustment. Not like numerosity, generalization in this field is symmetric.

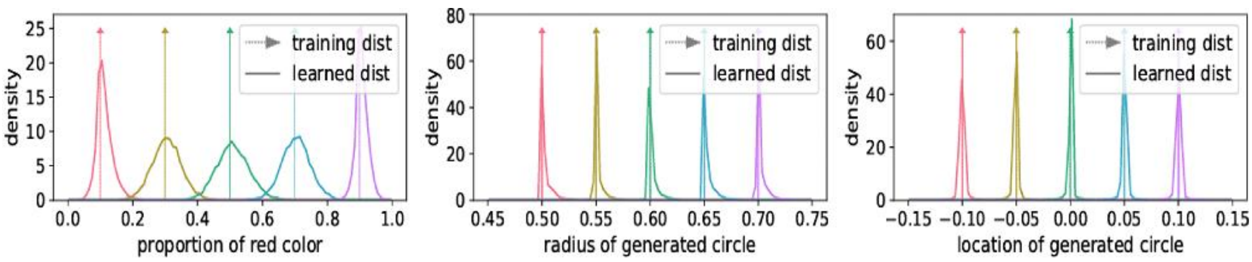

Figure 5: Produced Samples for respective Property

Difficulty Consequence and Pattern Improvement Consequence

Figure 6 demonstrated the convolution effect and prototype enhancement; when the working-out dissemination is sufficiently close, the modes capture together. The average of the two methods is allotted high likelihood. That is, objects with thirty-five percent red are the maximum possible to be produced; however, they never looked in the working-out set. When the modes are far from each other, convolution forecasts the performance of the system better (Vadlamudi, 2021). Again, these results depend on 'Generative Adversarial Networks (GAN)'/'variation automatic coders (VAE)' and diverse structures. Similar observations have been carried out in psychological tests.

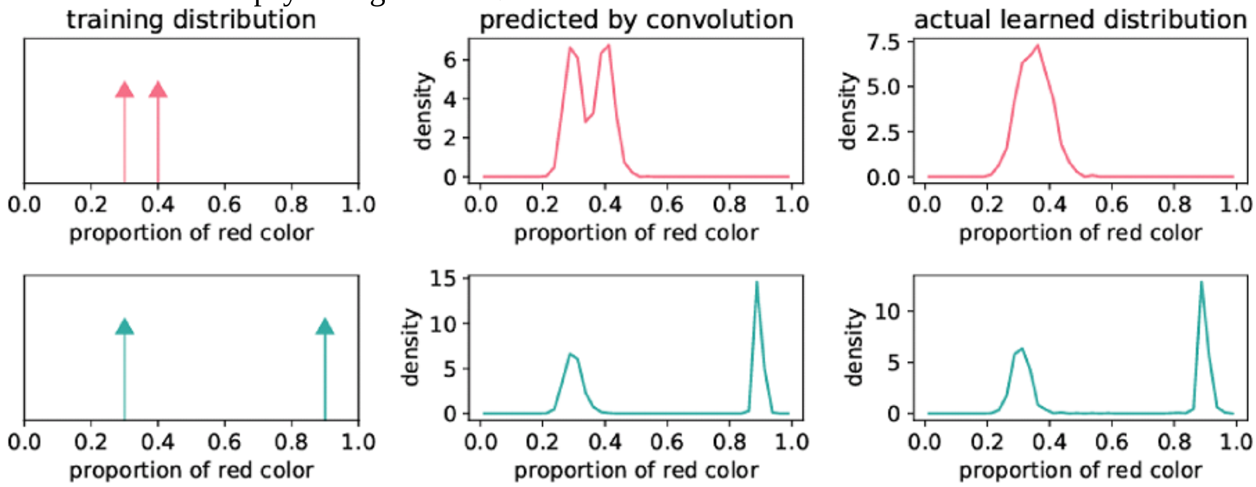

Figure 6: Demonstration of the Difficulty Consequence and Pattern Improvement Consequence

Individuality of Structures

Figure 7 shows the learned dissemination for respective structures as the other feature differ. It was observed that the learned dissemination for the separate part is justly self-determining of the different scopes. The only distinguished adjustment is an insignificant upsurge in modification if the other areas are arbitrary (Ahmed &Ganapathy, 2021). Interestingly, as the difference upsurges, modes that did not illustrate if the further pattern improvement instigated to merge.

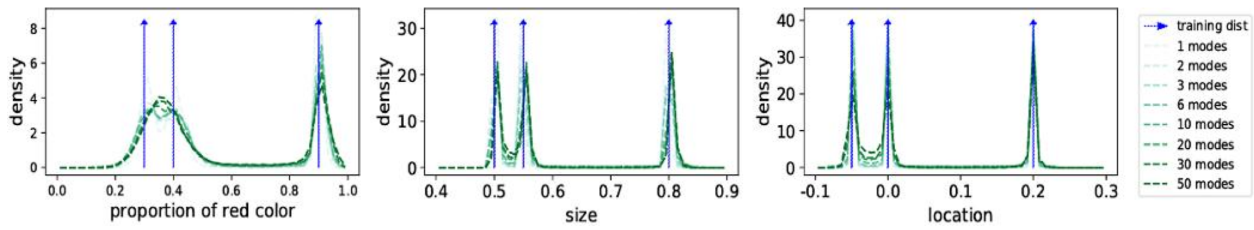

Figure 7: Learned Dissemination for respective Structure

Describing Generalization on Numerous Features

The pie and MNIST datasets we evaluated using exactness recall to compare the support. The recall addresses the ratio or fraction of blends in the backing of p(z) and q(x). A flawless memory denotes all blends that appear in the learned dissemination, while exactness is the ratio of combinations in the backings of both p(z) and q(x) (Figure 8). To pie and MNIST, generalization performance unfavorably is contingent on the number of available blends. For a CLEVER dataset, we precisely describe how q(z) varies from p(x). We employ a working-out set where a shape only takes a section of the likely colors and observe its possible generalization to other colors.

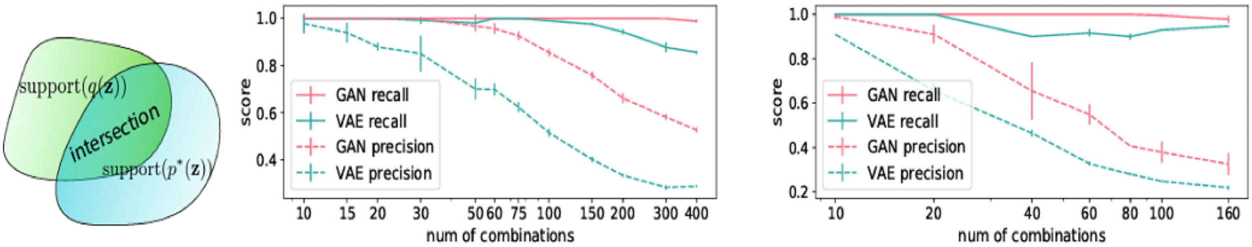

Figure 8: Exactness and Recall for GAN/VAE

5. CONCLUSION AND RECOMMENDATION

This study suggested methods and the significance of biases in learning algorithms generalization through prudently patterned working out sets. This was achieved by detecting the learned dissemination, and we accessed new understandings into the generalization performance of the patterns. We found distinct generalization patterns for each individual feature, some of which have similarities with perception tests on monkeys and humans. In addition, we visualized the learned dissemination and discovered new blends produced while conserving the peripheral on each feature. However, biases are very significant in learning generalization, which help predict or forecast the outcome of inputs that it has come across. We strongly believe that the framework and the mechanism we suggest will help stimulate further study into the empirical performance of this model.

6. REFERENCE

[1] Ahmed, A. A. A., &Ganapathy, A. (2021). Creation of Automated Content with Embedded Artificial Intelligence: A Study on Learning Management System for Educational Entrepreneurship. *Academy of Entrepreneurship Journal*, *27*(3), 1-10, https://doi.org/10.5281/zenodo.4973057

[2] Ahmed, A. A. A.; Paruchuri, H.; Vadlamudi, S.; &Ganapathy, A. (2021). Cryptography in Financial Markets: Potential Channels for Future Financial Stability. Academy of Accounting and Financial Studies Journal, 25(4), 1–9. https://doi.org/10.5281/zenodo.4774829

[3] Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in cognitive sciences*, *15*(3), 122–131.

[4] Arora, S. and Zhang, Y. (2017). Do gans actually learn the distribution? an empirical study. https://arxiv.org/abs/1706.08224

[5] Azad, M. M., Ganapathy, A., Vadlamudi, S., Paruchuri, H. (2021). Medical Diagnosis using Deep Learning Techniques: A Research Survey. *Annals of the Romanian Society for Cell Biology*, *25*(6), 5591–5600. Retrieved from https://www.annalsofrscb.ro/index.php/journal/article/view/6577

[6] Che, T., Li, Y., Jacob, A. P., Bengio, Y. and Li, W. (2016). Mode regularized generative adversarial networks. https://arxiv.org/abs/1612.02136

[7] Efromovich, S. (2010). Orthogonal series density estimation. WIREs Computational Statistics, 2(4), 467–476. https://doi.org/10.1002/wics.97

[8] Ganapathy, A. (2020). Everything-as-a-Service (XaaS) in the World of Technology and Trade. *American Journal of Trade and Policy*, *7*(3), 91-98. https://doi.org/10.18034/ajtp.v7i3.555

[9] Ganapathy, A. (2021a). Edge Computing: Utilization of the Internet of Things for Time-Sensitive Data Processing. *Asian Business Review*, *11*(2), 59-66. https://doi.org/10.18034/abr.v11i2.547

[10] Ganapathy, A. (2021b). Quantum Computing in High Frequency Trading and Fraud Detection. *Engineering International*, *9*(2), 61-72. https://doi.org/10.18034/ei.v9i2.549

[11] Ganapathy, A. (2021c). Robotic Process Automation: End from Entrepreneurial Content Editing to Go Live. *Academy of Entrepreneurship Journal*, *27*(3), 1-11. https://www.abacademies.org/articles/robotic-process-automation-end-from-entrepreneurial-content-editing-to-go-live-11331.html

[12] Ganapathy, A., Ahmed, A. A. A, Siddique, M. NEA., (2021). Easy URLs in the Content Management System with Crawlers for Added Security. Academy of Marketing Studies Journal, 25(4), 1-10.https://doi.org/10.5281/zenodo.5002945

[13] Ganapathy. A. (2021d). Pharmaceutical Company's Stock Leap through Machine Learning and AI Driven Input Injection. *Annals of the Romanian Society for Cell Biology*, *25*(6), 11923–11933. Retrieved from https://www.annalsofrscb.ro/index.php/journal/article/view/7783

[14] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems, 2672–2680.

[15] Gretton, A., Borgwardt, K. M., Rasch, M., Schölkopf, B. and Smola, A. J. (2007). A kernel method for the two-sample-problem. In Advances in neural information processing systems, 513–520.

[16] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Klambauer, G. and Hochreiter, S. (2017). Gan strained by a two time-scale update rule converge to a nash equilibrium. https://arxiv.org/abs/1706.08500

[17] Ho, J. and Ermon, S. (2016). Generative adversarial imitation learning. In Advances in Neural Information Processing Systems, 4565–4573, 2016.

[18] Johnson, J., Hariharan, B., van der Maaten, L., Fei-Fei, L., Zitnick, C. L. and R. Girshick, R. (2017). Clevr: A diagnostic dataset for compositional language and elementary visual reasoning. In CVPR.

[19] Khan, W., Ahmed, A. A. A., Vadlamudi, S., Paruchuri, H., Ganapathy, A. (2021). Machine Moderators in Content Management System Details: Essentials for IoT Entrepreneurs. Academy of Entrepreneurship Journa, 27(3), 1-11.https://doi.org/10.5281/zenodo.4972587

[20] Kingma, D. P. and Welling, M. (2013). Auto-encoding variationalbayes. https://arxiv.org/abs/1312.6114

[21] Minda, J. P. and Smith, J. D. (2011). Prototype models of categorization: Basic formulation, predictions, and limitations. *Formal approaches in categorization*, 40–64.

[22] Nieder, A. and Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *Journal of Neuroscience, 27*(22), 5986–5993.

[23] Nieder, A. and Miller, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. Neuron, 37(1), 149–157.

[24] Nieder, A., Freedman, D. J. and Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. Science, 297(5587), 1708–1711.

[25] Piazza, M., Izard, V., Pinel, P., Le Bihan, D. and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron, 44(3), 547–555.

[26] Rezende, D. J., Mohamed, S. and Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. https://arxiv.org/abs/1401.4082

[27] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. Proceedings of the National Academy of Sciences, 42(1), 43–47.

[28] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 832–837.

[29] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. and Chen, X. (2016). Improved techniques for training gans. In Advances in Neural Information Processing Systems, 2234–2242.

[30] Smith, D. J. and Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 3.

[31] Stevens, S. S. (2017). Psychophysics: Introduction to its perceptual, neural and social prospects. Routledge.

[32] Vadlamudi, S. (2021). The Internet of Things (IoT) and Social Interaction: Influence of Source Attribution and Human Specialization. *Engineering International*, *9*(1), 17-28. https://doi.org/10.18034/ei.v9i1.526

[33] Vadlamudi, S., Islam, M. A., Hossain, M. S., Ahmed, A. A. A., Asadullah, ABM. (2021). Watermarking Techniques for Royalty Accounts in Content Management Websites for IoT Image Association. Academy of Marketing Studies Journal, 25(4), 1-9. https://doi.org/10.5281/zenodo.5002840

[34] Zhao, S., Song, J. and Ermon, S. (2018). A lagrangian perspective on latent variable generative models. Proc. 34th Conference on Uncertainty in Artificial Intelligence.