UPRIGHT PUBLICATIONS
https://upright.pub/index.php/ijrstp/

Original Contribution

Quantum Vision Investigations Frame Worked after Long Short-Term Typed Memory

Takudzwa Fadziso[1]

Keywords: Quantum Vision, LSTM, Network Architecture

International Journal of Reciprocal Symmetry and Theoretical Physics

Vol. 5, Issue 1, 2018 [Pages 1-8]

In this paper, we show the manner in which machine learning models experiments when it comes to quantum physics. The cornerstone of the newfangling quantum technologies like quantum cryptography and quantum computation is quantum entanglement. The ones that of greater interest are complicated quantum conditions having over two particles as well as a significant count of entangled quantum stages. Considering a high-dimensional and multi-particle state like this, it is not often possible to reframe an experimental premise that will be able to generate the same results. As such, in order to discover interesting experiments, one needs to, at random, formulate millions of premises or setups on a computer, after which one will compute the output states respectively. This work is used to demonstrate that machine learning models are capable of providing more substantial development compared to random searches. The paper shows how a Long Short-Term Memory Network (also called LSTM) is capable of effectively learning how to handle modelling for quantum experiments through the accurate prediction of the output state attributes for particular setups while not having to make computing states an essential consideration. With this approach, one is not only able to conduct faster searches but also be able to take a critical step towards the modelling of high-dimensional quantum tests with the use of generative machine learning algorithms.

INTRODUCTION

In the last 10 years, man-made neural networks have, with outstanding performances, been implemented in a cornucopia of scientific fields, with commercial use cases and daily tasks. For example, they have been applied to medical diagnosis, board games and self-driving vehicles (Esteva et al., 2017; Silver et al., 2017). Dissimilar to conventional neural systems, Long Short-Term Memory networks have frameworks with recurring links (Hochreiter, 1991; Ahmed, 2009; Hochreiter & Schmidhuber, 1997). With these recurring connections, the architectures are enabled to process sequential pieces of information like speech and text (Sutskever et al.,2014). This capability to process sequences can be specifically critical for the design of complicated quantum experiments because the final condition of quantum particles is quite reliant on the elements as well as experimental premises that the considered particles undergo. Case in point, in the field of quantum-optic experiment, photons can traverse a series of beam splitters, wave plates and holographic-type plates.

Quantum states that are high-dimensional are essential for multiparticle as well as multi setting disregard for native realist models and for implementations in upcoming quantum technologies such as error correction and quantum comms for quantum-type computers (Shor, 2000;

[1]Institute of Lifelong Learning and Development Studies, Chinhoyi University of Technology, ZIMBABWE
Corresponding email: takudzwafadziso@gmail.com

Kaszlikowski et al., 2000; Bynagari, 2015; Donepudi, 2014b; Manavalan, 2016; Manavalan & Donepudi, 2016; Donepudi, 2016; Neogy and Ahmed, 2015; Begum et al., 2012; Ahmed & Dey, 2009). Already, for only a few quantum stages and a trio of photons, it becomes generally infeasible for people to determine the setup that is needed for the final stage of a quantum state. An instance of this kind of automatic process is the MELVIN algorithm (Krenn et al., 2016), which employs an entire toolbox of optical components, the indefinitely contrived sequences of the said components, computes the outcome quantum stage and finds out if the state is of interest (that is, whether it involves a reasonable number of quantum levels or is maximally entangled.

The premise that was introduced by MELVIN has been discovered during lab-based experiments (Malik et al., 2016; Erhardet al., 2018b). Recently, another reinforcement learning method was applied to the formulation of new experiments (Melnikov et al., 2018). These advances have inspired our investigation into the way LSTM systems can train quanto-optical setups and forecast the features of the quantum states resulting from the learning. This task's level of perceived human difficulty is beyond that which exists in other deep learning assignments such as text generation and object recognition. We learn the neural systems with the use of millions of MELVIN-generated setups. The humongous volume of data makes deep learning the method of first choice. To assess the models, we cluster cross validation (Mayr et al., 2016).

TARGET VALUES

We consider a quantum optical test with a trio of photons possessing orbital angular momentum (also known as OAM) (Yao & Padgett, 2011; Erhard et al., 2018a). A photon's OAM comprises an integer whose sign and size are a respective representation of the handedness and shape of the photon's wave front. Case in point, succeeding a series of optical components, a quantum state with three particles may develop this form: |Ψi=12(|0,0,0i+|1,0,1i+|2,1,0i+|3,1,1i). This state is, in turn, a representation of a physical condition where there is ¼ possibility the entire three photons will have OAM values of 0 for the initial term. Meanwhile, there is a ¼ possibility that photons number 1 and 3 will have OAM values of 1. Then, photon 2 will have 0 OAM value in the second term, and this order will continue for the remaining terms. In generality, we have interest in the two major attributes of the quantum states: are they high-dimensional and maximally entangled?

A state's dimensionality is represented via its Schmidtrank vector (also called SRV) (Huber & de Vicente, 2013; Huber et al., 2013). The state |Ψi is undeniably maximally entangled due to the fact that all the terms on the right hand corner possess similar values in amplitude. 4,2,2 is the state/s SRV because the photon has entanglements with the rest of the network, in four dimensions. Meanwhile, photons number 2 and 3 are entangled with the other two photons 2-dimensionally. If a setup's output proves maximally entangled and agrees with additional restrictions (that is, when it behaves normally despite multi-pair emissions and negative labelling (yE=0). The targeted label which captures the dimensionality of the state is referred to as the SRV SRV=(n, m, k)>. We learn LSTMs to directly forecast the attributes of these states (SRV and entanglement) from a specific experimental setup while not needing to forecast the quantum state.

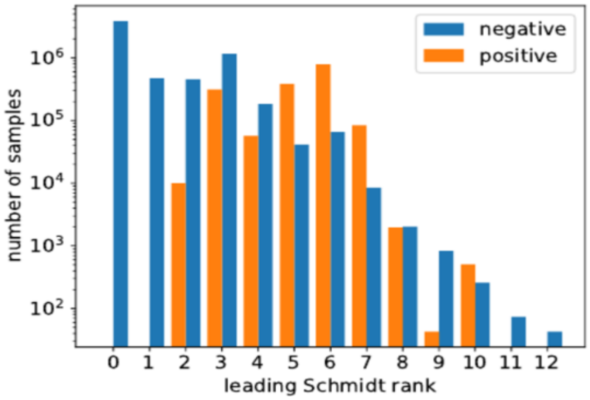
Network Architecture

In the sequence we used to depict the processing model, we learned two networks; one for SRV regression (targetSRV) and the other for entanglement classification (target). Why did we sidestep multi-task training in this context? It's because we do not want to include correlations between the SRV and entanglement for our models. Case in point, the SRV (6,6,6) was only noticed for the samples that were not maximally entangled, which is an ideal correlation. It would make a multi-task system to automated mark such samples as negatives just because they possess SRV. By learning separate networks, we will be lowering the risk of incorporation of such correlation types. An arrangement of N elements was fed into a system via its sequence of single optical elements x= (x1, x2, ..., xN)>. In our data, N ranges from 6 to 15. The LSTM we use deliberately has 2048 hidden units and comes with a component embedding area that has 64 different dimensions. Component embedding, as a technique, is similar to embedding words (Mikolov et al., 2013).

EXPERIMENTS

Datasets

The MELVIN-generated dataset comprises 7,853,853 different arrangements, 1,638,233 of which are labelled as positive samples. Every arrangement is made up of optical element sequences and a pair of target values: yE and ySRV. Our interest is in if the learned model is capable of extrapolation to unnoticed SRVs. As such, we put the data in cluster mode by leading the Schmidt ranking method. The figure below is an illustration of the number of negative as well as positive samples in the data set meant for every n.

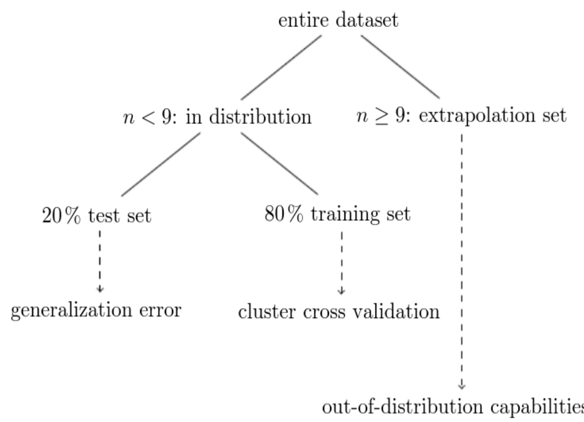


Workflow

All the samples that are with the ≥9 range are relocated to a specific extrapolation group comprising just 1,754 setups. The leftover data (that is, the samples possessing < 9) are grouped into a conventional test set and training set while 20 percent of the obtained data is drawn randomly. As evident in the workflow, the test set is used for the estimation of the traditional generalization inaccuracy. Meanwhile, the extrapolation set is used as a light shedder on the capacity of the trained model to deliver results higher than that of the Schmidt ranking figures. Should the model be extrapolated successfully, it is possible for us to hope for the discovery of experimental arrangements that culminate in new and interesting quantum states.

The cross-validation of clusters (also called CCV) is an assessment into the technique employed to conventional cross validation. Rather than classifying the folds, CCV categorizes them with respect to their clustering behaviors. As such, the CCV eliminates the semblances between the training and validation set, as well as stimulates the conditions in which the hoarded foldshave have not yet been ascertained. This will, therefore, allow us to conduct an investigation to the capacity of the system to discover the setups that have been withheld.

For this paper, we use the CVV alongside nine folds. A total of 7 in the fold align with the leading Schmidt number. The samples that have n=1 (not entangled) and n=0 (which isn't even a valid three-photon condition) are definitively negative. The samples are a representation of special cases of yE=0 arrangements, plus it is not critical for these cases to be generalized unless they are adequately trained. So, the 4,300,268 samples possessing < 2 are randomly split into two different folds in a way that the model will continually see some of these specific samples when learning is ongoing.



Outcomes of the Experiment

We observed to find out whether the LSTM network has been trained about something related to quantum physics. A model of ideal nature will be able to correctly recognize the positive steps and as well discard as much negative arrangements as possible. Such a behavior is mirrored in the true positive rate of the metrics, which is TPR=TP/(TP+FN), as well as in the true negative rate expressed as TNR=TN/(TN+FP). Herein, TP is the true positive, TN is the true negative, FP is the false positive and FN is the false negative.

The precision value or positive prediction value is the metric tasked with the responsibility of quantifying the success rate inside the positive forecasts. This is defined as PPV=TP+FP. For the sake of every CCV fold, we attribute an arrangement as interest should it meet two criteria. Firstly, it needs to be classified as positive (expressed as ˆyE> τ) where τ is the threshold for categorizing the activation of the sigmoid output. The second criterion is that the SRV forecast (expressed as ˆySRV = (ˆn, ˆm, ˆk)>) is as such that eSRV = (n, m, k)> exists alongside kySRV −ˆySRVk2< r. The SRV we call "radius".

Samples that are grouped as interesting (and uninteresting) and truly positive (and negative) are denoted. We also denote the samples that are categorized as interesting (uninteresting) and truly negative (and positive) as the false positives (and false negatives). The stochastic gradient is used as the descent for learning the LSTM with a momentum of 0.5 and a batch size of 128. The mini batches are samples in such a manner that the negatives and positives among the samples appear on a matched frequency when learning is ongoing. To ascertain a balanced SRV regression, the Schmidt rank vector figure in the lead is misused in terms of a class mark.

We trained the models with the aid of early halting following 40,000 weight update stages for the entanglement categorization system. Meanwhile, for the SRV regression system, we used 14,000 stages. We then performed a hyperparameter search in advance for a dataset in semblance to the learning set. The discovery ratio is the amount of different SRVs for which no less than 20 percent of the samples are discovered again under the auspices of our approach. That is, it is recognized as "interesting" and can be divided by the amount of different SRVs existing in their respective clusters. 0.9996 is the TNR for fold and 0.659 is the accuracy of the extrapolating set 9-12. The error bars are 95 percent of binomial proportion intervals of confidence. The performance of the model is highly contingent upon the τ and r parameters.

Lastly, we conduct an investigation into the standard distribution-related generalization inaccuracy with the test set, which is about 20 percent of the raw information involved. 10.2 is the value of the entanglement learning BCE loss. The true positive rate (or TPR) is 0.9469 ±0.00041, while the true negative rate (or TNR) is 0.9271. The test error that is corresponding is a TNR of 10.4 and a TPR of 0.9261, which is, respectively, ±0.00038and0.9427 ±0.00065. As for SRV regression, the value of the SRV's learning loss, according to the equation, is 2.247. The precision (which is r=3) stands at 93.82 percent and the real distance between the forecast and the labeller is 1.3943.

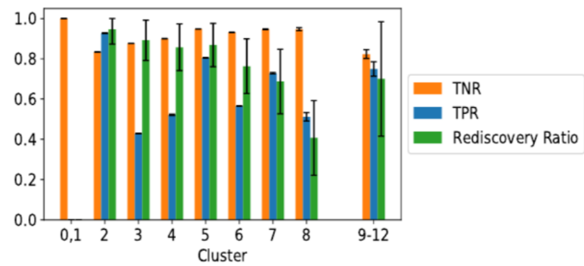
Specifically, the Long Short-Term Memory Network is (Hochreiter and Schmidhuber, 1997) is a unique kind of RNN, one that has the capacity to train long-term sequential reliances as well as surmount the disappearing gradient or detonate the gradient challenges categorically inherent to the mere RNNs (Pascanu et al., 2013). For this, we consider a pair of fresh equalizer models based on the RNNs of the LSTM. Then, we make use of LSTM realization obtained from Kera, thus using a forget gate (Gers et al., 2000).

Also, like we did prior, we fed the NS received into the equalizer, considering several neighbors from the two sides of the "interesting" NS element. Nevertheless, dissimilar to the FFNN equalizer case scenario, we simultaneously fed the real as well as imaginative aspects of the spectra as a two-variable multivariate sequence. The major aspect of the equalizer is the gated RNN that is stacked with a pair of bi-directional LSTM or BLSTM layers possessing 96 nodes in every strati. For the BLSTM, we use the optimization algorithm known as NADAM to bring the loss to its barest minimum.

While equalizing the spectra, we noticed that the RNN has the ability to extract sequential characteristics from a single tap onto the following (and then to the previous one in the forward trajectory of the BLSTM). As a result, we were expecting that this kind of equalizer would be able to perform significantly better than the FFNN with operations in a data stack without having to comprehend the relationship as well as the order of data points. Thereafter, with the goal of improving the equalizer performance of the BLSTM, we add a one-dimensional convolutional strategy before the layers of the BLSTM. The one-dimensional CNN serves as a filter that extracts attributes from the NS.

In an actual sense, the hybrid NN that has been ascertained by adjoining the LSTM and the CNN is known for efficient work in the areas of visual recognition, visual description, entity recognition when it comes to processing language (J. P. Chiu and E. Nichols, 2016) and the detection of sound events (Lim and Park, 2017). It is as such that, naturally, we expect to realize a more substantial performance because we added the one-dimensional CNN. We use just a single one-dimensional convolutional layer comprising 64 filters, in the absence of non-linear activation controls as well as max-pooling levels.

In correspondence to the imaginary and real aspects of the sampled NS, the dimensionality of the input multivariate process is modified from 2 all the way to 64, which is equivalent to the amount of convolutional filters. The layer of output is completely connected to the feed-forward level with a pair of nodes for the imaginary as well real parts of the sampled NS. For the network's final output, we have an NS sample of "equalized" nature, same as we had it in the preceding section.



RESEARCH OUTLOOK

According to the demonstrations obtained from our experiments, an LSTM-based neural system can be learned for the modelling of specific attributes on the part of complicated quantum systems. The approach we are using is not restricted to entanglement and the Schmidt ranking system but it can be generalized to carry out different objective functions, including fidelity, interference and high-dimensional, multi-particle quantum experiments during the exploitation of generative models. For this paper, we consider GANs (Generative Adversarial Networks) (Goodfellow et al., 2014) as well as the beam search system (Lowerre, 1976; Ahmed, 2016) as copyable techniques.

Formulating processes like text within adversarial environments has been carried out with the use of 1D CNNs and LSTMs (Gulrajaniet al., 2017; Yu et al., 2016). The approaches that are based on the LSTM are in the employ of reinforcement learning ideas for the alleviation of the challenge associated with the propagation of gradients via the soft maximum outputs of the system. Because the data is in structural similitude to text, the methods we consider are directly implementable to our setting. When it comes to beam search, there are two distinct ideas in consideration; a generative approach and a discriminative approach.

With the discriminative approach, one can incorporate the complete dataset, including the negative and positive samples (Ahmed & Dey, 2009; Bynagari, 2016; Donepudi, 2015; Azad et al., 2011; Ahmed et al., 2011; Abedin et al., 2012). Models learning for this exercise can be used in the discriminative methods in such a way that one is able to build fresh sequences through the maximization of the conviction that the results come out positive in arrangement. The generative method, on the other hand, is born of the idea to learn a model on the samples of positive nature, only to end up training their distribution through next element forecasts.

Inferentially, beam search is usable for the approximation of the most probable sequence considering a certain initial state (Bengio et al., 2015). A different option is the generation of fresh sequences from the distributed softmax of the system output at every sequence stance as is used for the models that generate text (Graves, 2013; Karpathy & Fei-Fei, 2015; Amin & Manavalan, 2017; Ahmed and Day, 2009b; Bynagari, 2014; Donepudi, 2014a; Fadziso & Manavalan, 2017; Manavalan, 2014; Manavalan & Bynagari, 2015; Manavalan & Ganapathy, 2014; Siddique & Ahmed, 2015; Ahmed and Neogy, 2009). On a general note, automated design processed in experiments has applications with broader scopes, beyond quantum optical arrangements. What's more, it can be of essence to several scientific fields that are not Physics.

OBSERVATIONS AND CONCLUSION

In this study, we have demonstrated that an LSTM-relianet neural system is capable of being learned to accurately forecast specific attributes without needing any explicit comprehension of quantum mechanics. When it comes to individuals, the complexity associated with analyzing quantum optical tests can go beyond those of other deep learning challenges such as the classification of images. The network exhibits optimum performances in spite of the unseen information beyond the capabilities in learning, distribution and extrapolation. This is somewhat a catalyst to the automated modelling of quantum-facing experiments with the generative machine learning techniques.

Our findings show that the proposed new kind of equalizers that are based on RNNS can also outperform the FFNN type. We noticed a 16 times BER delivery development with the equalizer contingent on the RUN environment, two bi-directional LSTM stratis and s 23 times BER improvement with an equalizer that is based on the hybrid 1D CNN - BLSTM for a 170 Gbit/s NFT communication system. That is, up to 1000 km of propagation distance. Nevertheless, we also noticed some drawbacks in learning the complex NN frameworks that involved the BLSTM. It is as such that we could not successfully get to the minimum BER as a function of the taps in close proximity. As such, it is our belief that the approach we used has the prospective capability of delivering even greater improvement when it is amalgamated with additional optimizations and regularizations.

REFERENCES

A. A. Melnikov, H. Poulsen Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J.Briegel. Active learning machines learn to create new quantum experiments. PNAS, 115(1221),2018.

A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Sweater, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(115), 2017.

A. Graves. Generating sequences with recurrent neural networks. arXiv:1308.0850, 2013.

A. Karpathy and L. Fei-Fei. Deep visual-semantic alignments for generating image descriptions. InProceedings of the IEEE conference on computer vision and pattern recognition, pp. 3128–3137, 2015.

A. M. Yao and M. J. Padgett. Orbital angular momentum: origins, behavior and applications. Adv.Opt. Photon., 3(161), 2011.

A. Mayr, G. Klambauer, T. Unterthiner, and S. Hochreiter. DeepTox: Toxicity Prediction using DeepLearning. Frontiers in Environmental Science, 3(80), 2016.

Abedin, M. M. M., Ahmed, A. A. A., and Neogy, T. K. (2012). Mechanism of Accountability and Auditing: Public Sector Scenarios of Bangladesh. Journal of Business Studies, 4, 131-148.

Ahmed, A. A. A. & Dey, M. M. (2009). Timeliness attributes and the extent of accounting disclosure: a study of banking companies in Bangladesh. *Osmania Journal of International Business Studies, 4*(1).

Ahmed, A. A. A. (2009). The Effect of Timeliness Regulation of Corporate Financial Reporting: Evidence from Banking Sector of Bangladesh. *Accounting and Management Information Systems, 8*(2), 216 - 235. http://online-cig.ase.ro/jcig/art/8_2_4.pdf

Ahmed, A. A. A. (2016). Relationship between Foreign Direct Investment and Company Taxation: Case of Bangladesh. *American Journal of Trade and Policy, 3*(1), 11-14. https://doi.org/10.18034/ajtp.v3i1.394

Ahmed, A. A. A. and Day, M. M. (2009). Bank loan officers' perceptions of corporate financial disclosure in the banking sector of Bangladesh: An empirical analysis, Proceedings 2nd CBRC, Lahore, Pakistan, 1-12

Ahmed, A. A. A. and Neogy, T. K. (2009). Merger & Acquisitions (M&A) Goodwill Accounting: Principles and Practice. The Bangladesh Accountant, 65, 75-91.

Ahmed, A. A. A., & Dey, M. M. (2009b). Corporate Attribute and the Extent of Disclosure: A Study of Banking Companies in Bangladesh. Proceedings of the 5th International Management Accounting Conference (IMAC), OCT 19-21, 2009, UKM, Kuala Lumpur, MALAYSIA, Pages: 531-553. https://publons.com/publon/11427801/

Ahmed, A. A. A., Khan, W., & Hossain, M. S. (2011). Reporting Practice of Accounting Disclosure on Changes in Listed Companies of Bangladesh. ASA University Review, 5(1), 83-96. https://www.researchgate.net/publication/336664901

Amin, R., & Manavalan, M. (2017). Modeling Long Short-Term Memory in Quantum Optical Experiments. *International Journal of Reciprocal Symmetry and Physical Sciences*, *4*, 6–13. Retrieved from https://upright.pub/index.php/ijrsps/article/view/48

Azad, M. R., Khan, W., & Ahmed, A. A. A. (2011). HR Practices in Banking Sector on Perceived Employee Performance: A Case of Bangladesh. Eastern University Journal, 3(3), 30–39. https://doi.org/10.5281/zenodo.4043334

B. T. Lowerre. The Harpy speech recognition system. PhD Thesis, Carnegie Mellon University,Pittsburgh, 1976.

Begum, R., Ahmed, A. A. A., & Neogy. T. K. (2012). Management Decisions and Univariate Analysis: Effects on Corporate Governance in Bangladesh. Journal of Business Studies, 3, 87-115.

Bynagari, N. B. (2014). Integrated Reasoning Engine for Code Clone Detection. *ABC Journal of Advanced Research*, *3*(2), 143-152. https://doi.org/10.18034/abcjar.v3i2.575

Bynagari, N. B. (2015). Machine Learning and Artificial Intelligence in Online Fake Transaction Alerting. *Engineering International*, *3*(2), 115-126. https://doi.org/10.18034/ei.v3i2.566

Bynagari, N. B. (2016). Industrial Application of Internet of Things. *Asia Pacific Journal of Energy and Environment*, *3*(2), 75-82. https://doi.org/10.18034/apjee.v3i2.576

D. Kaszlikowski, P. Gnacínski, M. Zukowski, W. Miklaszewski, and A. Zeilinger. Violations of local realism by two entangled N-dimensional systems are stronger than for two qubits. Phys. Rev. Lett.,86(4418), 2000.

D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker,M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis. Mastering the game of Go without human knowledge. Nature, 550(354), 2017.

D.-A. Clevert, T. Unterthiner, and S. Hochreiter, “Fast and accurate deep network learning by exponential linear units (elus),” arXiv preprint arXiv:1511.07289 2 (2016).

Donepudi, P. K. (2014a). Technology Growth in Shipping Industry: An Overview. *American Journal of Trade and Policy*, *1*(3), 137-142. https://doi.org/10.18034/ajtp.v1i3.503

Donepudi, P. K. (2014b). Voice Search Technology: An Overview. *Engineering International*, *2*(2), 91-102. https://doi.org/10.18034/ei.v2i2.502

Donepudi, P. K. (2015). Crossing Point of Artificial Intelligence in Cybersecurity. American Journal of Trade and Policy, 2(3), 121-128. https://doi.org/10.18034/ajtp.v2i3.493

Donepudi, P. K. (2016). Influence of Cloud Computing in Business: Are They Robust?. Asian Journal of Applied Science and Engineering, 5(3), 193-196. Retrieved from https://journals.abc.us.org/index.php/ajase/article/view/1181

F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," Neural Comput. 12(10), 2451–2471 (2000). [CrossRef]

Fadziso, T., & Manavalan, M. (2017). Identical by Descent (IBD): Investigation of the Genetic Ties between Africans, Denisovans, and Neandertals. *Asian Journal of Humanity, Art and Literature*, *4*(2), 157-170. https://doi.org/10.18034/ajhal.v4i2.582

H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein, "Visualizing the loss landscape of neural nets," in Advances in Neural Information Processing Systems, vol. 31 (2018), pp. 6389–6399.

H. Lim and J. Park, "Rare sound event detection using 1d convolutional recurrent neural networks," in Detection and Classification of Acoustic Scenes and Events 2017, (2017), pp. 80–84.

H. Sak, A. Senior, and F. Beaufays, (2014). "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems 27, pp. 2672–2680. Curran Associates, Inc., 2014.

I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville. Improved training of wasserstein gans. In Advances in Neural Information Processing Systems 30, pp. 5767–5777.Curran Associates, Inc., 2017.

I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. InAdvances in neural information processing systems, pp. 3104–3112, 2014.

J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell, "Long-term recurrent convolutional networks for visual recognition and description," in Proceedings of the IEEE conference on computer vision and pattern recognition, (2015), pp. 2625–2634.

J. P. Chiu and E. Nichols, "Named entity recognition with bidirectional lstm-cnns," Trans. Assoc. Comput. Linguist. 4, 357–370 (2016). [CrossRef]

L. Yu, W. Zhang, J. Wang, and Y. Yu. Seqgan: Sequence generative adversarial nets with policy gradient. arxiv:1609.05473, 2016.

M. Erhard, M. Malik, M. Krenn, and A. Zeilinger. Experimental GHZ entanglement beyond qubits.Nature Photonics, 12(759), 2018b.

M. Erhard, R. Fickler, M. Krenn, and A. Zeilinger. Twisted photons: new quantum perspectives in high dimensions. Light: Science & Applications, 7(3):17146, 2018a.

M. Huber and J. I. de Vicente. Structure of multidimensional entanglement in multipartite systems.Physical Review Letters, 110(030501), 2013.

M. Huber, M. Perarnau-Llobet, and J. I. de Vicente. Entropy vector formalism and the structure of multidimensional entanglement in multipartite systems. Physical Review A, 88(4):042328, 2013.

M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger. Automated Search for newQuantum Experiments. Phys. Rev. Lett., 116(090405), 2016.

M. Malik, M. Erhard, M. Huber, M. Krenn, R. Fickler, and A. Zeilinger. Multiphoton entanglement in high dimensions. Nature Photonics, 10(248), 2016.

Manavalan, M. (2014). Fast Model-based Protein Homology Discovery without Alignment. *Asia Pacific Journal of Energy and Environment*, *1*(2), 169-184. https://doi.org/10.18034/apjee.v1i2.580

Manavalan, M. (2016). Biclustering of Omics Data using Rectified Factor Networks. *International Journal of Reciprocal Symmetry and Physical Sciences*, *3*, 1–10. Retrieved from https://upright.pub/index.php/ijrsps/article/view/40

Manavalan, M., & Bynagari, N. B. (2015). A Single Long Short-Term Memory Network can Predict Rainfall-Runoff at Multiple Timescales. *International Journal of Reciprocal Symmetry and Physical Sciences*, *2*, 1–7. Retrieved from https://upright.pub/index.php/ijrsps/article/view/39

Manavalan, M., & Donepudi, P. K. (2016). A Sample-based Criterion for Unsupervised Learning of Complex Models beyond Maximum Likelihood and Density Estimation. *ABC Journal of Advanced Research*, *5*(2), 123-130. https://doi.org/10.18034/abcjar.v5i2.581

Manavalan, M., & Ganapathy, A. (2014). Reinforcement Learning in Robotics. *Engineering International*, *2*(2), 113-124. https://doi.org/10.18034/ei.v2i2.572

Neogy, T. K. and Ahmed, A. A. A. (2015). The Extent of Disclosure of Different Components of Disclosure Index: A Study on Commercial Banks in Bangladesh. *Global Disclosure of Economics and Business, 4*(2), 100-110. https://doi.org/10.18034/gdeb.v4i2.139

P. W. Shor. Scheme for reducing decoherence in quantum computer memory. Phys. Rev. A, 52(R2493), 2000.

R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in

International conference on machine learning, (2013), pp. 1310–1318.

S. Bengio, O. Vinyals, N. Jaitly, and N. Shazeer. Scheduled sampling for sequence prediction with recurrent neural networks. In Advances in Neural Information Processing Systems 28, pp.1171–1179. Curran Associates, Inc., 2015.

S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput. 9(8), 1735–1780 (1997). [CrossRef]

S. Hochreiter and J. Schmidhuber. Long short-term memory. Neural Computation, 9(1735), 1997.

S. Hochreiter. Untersuchungen zu dynamischen neuronalen Netzen. Diploma Thesis, TU München,1991.

Siddique, M. N. & Ahmed, A. A. A. (2015). Congruence of Competitive Advantage and Transfer Pricing: A Study on Selected MNCs Operating in Bangladesh. *Asian Accounting & Auditing Advancement,* *5*(2), 119-126. https://www.researchgate.net/publication/354712086

T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. ICLR Workshop, arXiv:1301.3781, 2013.

--0--