

Bulud əsaslı elektron dövlət xidmətlərinin səmərəliliyinin artırılması metodu

Fərqanə Abdullayeva

AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan

a_farqana@mail.ru

Xülasə— Məqalədə bulud infrastrukturunda virtual maşınların (VM) iş yükünün proqnozlaşdırılması üsulu təklif edilmişdir. Proqnozlaşdırmanı həyata keçirmək üçün variasional avtoenkoderlərin rekonstruksiya ehtimalları istifadə edilmişdir. Rekonstruksiya ehtimalı dəyişənlərin paylanması dəyişkən olduğunu nəzərə alan ehtimal meyardır. İşdə variasional avtoenkoderin rekonstruksiya ehtimallarının qiymətləri VM-nin yüklənmə dərəcələrini göstərir. Eksperimentlərin nəticəsi göstərmişdir ki, variasional avtoenkoderlər sadə dərin neyron şəbəkələri ilə müqayisədə VM-ların iş yükünün proqnozlaşdırılmasında daha yaxşı nəticələr vermişdir. Variasional avtoenkoderlərin generativ xarakteristikaları verilənləri rekonstruksiya edərək yüklənmənin növlərini (səviyyələrini) müəyyən etməyə imkan verir.

Açar sözlər— *Bulud əsaslı elektron dövlət; iş yükünün proqnozlaşdırılması; virtual maşın; dərin neyron şəbəkələri; dərin təlim; variasional avtoenkoder.*

I. GİRİŞ

Son illər elektron dövlət (e-dövlət) bir çox ölkələrə geniş istifadə olunur. E-dövlət vətəndaşlar və dövlət üçün olduqca faydalıdır. Lakin e-dövlətin effektiv xidmət göstərməsinə neqativ təsir edən amillər vardır [1]. E-dövlətin problemlərini aradan qaldırmaq məqsədi ilə və dövlət xidmətlərinin göstərilməsini yaxşılaşdırmaq üçün hazırda dünya ölkələrində bulud texnologiyaları geniş tətbiq olunur. Bulud texnologiyalarının tətbiqi e-dövlətin müxtəlif aspektlərinin yaxşılaşdırılmasına ciddi təsir edə bilər.

Bulud əsaslı e-dövlətin effektivliyi elektron dövlət istifadəçilərinin dəyişkən tələblərinə cavab verən virtual resursların effektiv idarə edilməsindən asılıdır.

Bulud infrastrukturunda resursların planlaşdırılmasını həyata keçirməzdən əvvəl VM-ların tutumunun qiymətləndirilməsi aparılmalıdır [2]. VM-nin resurs utilizasiyası səviyyəsinin qiymətləndirilməsi və müəyyən edilməsi daxil olan tələbat üçün resursların (məsələn, CPU, yaddaş) ayrılmasında bazis hesab olunur [3].

VM-ların iş yükünün proqnozlaşdırılması sahəsində olduqca çox sayda yanaşmalar təklif olunmuşdur. Bu yanaşmaların əksəriyyətində VM-ların iş yükünü və ya onların

bulud infrastrukturunda CPU sərfiyyatını proqnozlaşdırmaq üçün neyron şəbəkələr [4, 5, 6], obrazların tanınması [7], xətti reqressiya [8] kimi proqnozlaşdırma üsulları təklif edilmişdir. Bu metodlar VM-ların gələcək iş yükünü əvvəlki t zaman intervalındakı keçmiş iş yükü verilənlərinə əsasən proqnozlaşdırırlar [9]. Bu işlərdə resursların proqnozlaşdırılması SaaS və PaaS səviyyələrində aparılır. Bu səviyyələrdə bulud provayderləri infrastrukturlarında bu tip proqramların olmasından xəbərdar olurlar və onların davranışlarını asanlıqla izləyə bilirlər. Lakin IaaS səviyyəsində VM-da hansı proqramın icra olunacağı haqqında, ayrı-ayrı VM-larda yerinə yetiriləcək proqramlar (tətbiqlər) haqqında heç bir məlumat olmur. Belə olduqda isə VM-ların iş yükünün və CPU sərfiyyatının qiymətləndirilməsində proqramların davranışından istifadə etmək mümkün olmur. Bu mühitdə VM-ların iş yükünə istifadəçilərin davranışları və istifadəçilərin ani qəbul etdiyi qərarları təsir edir. Bu səbəbdən iş yükündə fluktuasiya əvvəlki iş yükündən və əvvəlki CPU sərfiyyatından asılı olmur və istifadəçilərin resursa olan dinamik, gözlənilməz, fluktuativ tələbatlarından asılı olaraq dramatik şəkildə dəyişə bilər.

Bu faktorları nəzərə alaraq təqdim olunan məqalədə Variasional Avtoenkoderlərə əsaslanan iş yükünün proqnozlaşdırılması üsulu təklif edilmişdir.

Variasional Avtoenkoderlər ehtimal proqnozlaşdırma üsuludur, variasional çıxarışla dərin təlimin birləşməsindən yaranmışdır. Variasional avtoenkoderlərinin üstünlüyü ondan ibarətdir ki, bu model ehtimal qiymətləndirmələr aparır. Bu ehtimal qiymətləri rekonstruksiya ehtimalları adlandırılır. Ehtimalların istifadəsi modeldə hər hansı bir spesifik sərhəd qiymətindən istifadə olunması problemini aradan qaldırır. Təqdim olunan məqalədə Variasional Avtoenkoderlər bulud mühitində virtual resursların həddindən artıq yüklənmə (overloading) ehtimalını qiymətləndirir.

Təklif olunan iş yükünün proqnozlaşdırılması üsulunun əsasında həddindən artıq yüklənmiş (zəif işləyən) VM-lar müəyyən olunur, bu maşınların ekstra iş yükünü yerinə yetirə bilən uyğun VM-lar çoxluğu proqnozlaşdırılır və tapşırıqları yerinə yetirmək üçün proqnozlaşdırılan bu VM-lar çoxluğu içindən tapşırıqları yerinə yetirəcək yeni VM-lər seçilir.

II. MƏSƏLƏNİN QOYULUŞU

Fərz edək ki, n sayda $D = \{(x_i, y_i)\}_{i=1}^n$ təlim verilənləri verilmişdir. Burada x_i D ölçülü vektor, $x_i \in R^D$, $y_i \in [1, K]$ isə uyğun siniflərin adlarıdır. Generativ modellərdən istifadə etməklə təlim verilənlərinə yaxın verilənlərin generasiyası tələb olunur.

A. Metodologiya

Variasional avtoenkoder variasional çıxarışı olan dərin generativ təlim modelidir. Variasional avtoenkoder sistemin davranışını öyrənmək və generasiya etmək üçün emal olunmamış verilənlərdən istifadə edir. Bu bölmədə fundamental anlayışlar haqqında məlumat verilir və VM-ların iş yükünün proqnozlaşdırılması metodologiyasının əsas ideologiyası şərh edilir.

B. Dərin təlim (Deep Learning)

Dərin təlim xam sistem verilənlərinin mürəkkəb konstruksiyalarını aşkarlamaq və öyrənmək üçün bir neçə sadə, qeyri-xətti emal laylarından ibarət olur.

Dərin təlim konsepsiyasının onilliklərdir mövcud olmasına baxmayaraq, o 2011-ci ildə ImageNet müsabiqəsinə təqdim olunmuş konvolyusional neyron şəbəkələrə həsr olunmuş layihənin möcüzəvi imkanlarından sonra, son bir neçə ildə məşhurlaşmışdır [10]. Həmin dövrlərdən başlayaraq dərin təlim “computer vision”, “genomics” kimi bir çox sahələrdəki nəaliyyətlərə təkan vermişdir. Dərin təlim həmçinin bulud infrastrukturunda anomaliyaların aşkarlanması [11], DDoS hücumların aşkarlanması [12] məsələlərində geniş tətbiq olunmuşdur. Dərin təlim verilənlərin qeyri-bircinsliyini və kütləvililiyini, sistemlərin mürəkkəbliyini nəzərə alaraq VM-ların yüklənmə dərəcələrinin proqnozlaşdırılması məsələsində də perspektiv imkanlara malikdir.

C. Generativ modellər (Generative Models)

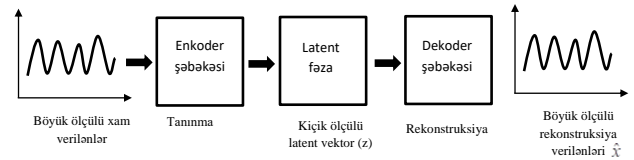
Generativ modellər dərin təlimin müəllimsiz təlim üsullarının bir bölməsidir, hadisələri verilənlərdən böyük ölçüdə cari nümunələri analiz edərək, verilənlərə yaxın yeni nümunələr generasiya etməklə öyrənir [13]. Generativ Adversarial şəbəkələr (Generative Adversarial Networks) [14], Dərin Avtoregressiv şəbəkələr (Deep Autoregressive Networks) [15] və Variasional Avtoenkoderlər (Variational Autoencoders) [16] ən məşhur generativ modelərdir. Məqalədə Variasional Avtoenkoderlərin bulud əsaslı elektron dövlət infrastrukturunda VM-ların keyfiyyət göstəricilərini xarakterizə edən verilənlərdən istifadə etməklə VM-ların yüklənmə dərəcələrinin proqnozlaşdırılması məsələsinə tətbiqinə baxılmışdır.

D. Variasional Avtoenkoder (Variational Autoencoder)

Variasional Avtoenkoder klassifikasiya prosesini verilənlərin ehtimalının variasional aşağı sərhəddini optimallaşdırmaqla həyata keçirir. Variasional Avtoenkoder variasional çıxarışla dərin təlimin hibridləşdirilməsi əsasında yaradılmışdır və posterior paylanmaların və böyük verilənlər bazalarının olduğu halda effektiv təlim həyata keçirə bilər. Variasional çıxarış ehtimal sıxlıqlarını optimallaşdırma yolu ilə aproksimasiya edir [17].

Variasional Avtoenkoder iki blokdən ibarətdir: enkoder və dekoder. Enkoder verilənlərin latent təsvirini yaradır, dekoder isə bu təsviri geri, orijinal verilənlərə konvertasiya edir.

Fərz edək ki, VM-ların yüklənmə dərəcəsinə göstərən verilənlər x , kodlaşdırılmış verilənlər isə z olsun. ϕ parametrlərinə malik enkoder şəbəkəsi verilmiş x giriş verilənlərini $q_\phi(z|x)$ şəklində verilmiş paylanmanın vasitəsi ilə z şəklində kodlaşdırır, θ parametrlərinə malik dekoder şəbəkəsi isə z verilənlərini $p_\theta(x|z)$ şəklində verilmiş paylanmanın vasitəsi ilə x verilənləri şəklində dekodlaşdırır. Bu proses şəkil 1-də təsvir edilmişdir.



Şəkil 1. Təlim verilənlərinə yaxın verilənlərin generasiyası prosesi

Fərz edək ki, verilənlərin loqarifmik ehtimalı aşağıdakı kimidir:

$$\log p(x) = \log p(x^{(1)}, \dots, x^{(n)}) = \sum_{i=1}^n \log p(x^{(i)}) \quad (1)$$

Onda

$$\log p_\theta(x) = D_{KL} \left(q_\phi(z|x^{(i)}) \parallel p_\theta(z|x^{(i)}) \right) + E_{q_\phi(z|x^{(i)})} \left[-\log q_\phi(z|x^{(i)}) + \log p_\theta(x, z) \right] \quad (2)$$

burada D_{KL} Kulbak-Leibler divergensiya funksiyasıdır (Kullback–Leibler divergence, KL Divergence). (2) ifadəsinin qiyməti hər zaman müsbət olduğundan aşağıdakı şərt ödənilir:

$$\log p_\theta(x) \geq E_{q_\phi(z|x^{(i)})} \left[-\log q_\phi(z|x^{(i)}) + \log p_\theta(x, z) \right] = L(\theta, \phi; x^{(i)}) \quad (3)$$

burada $L(\theta, \phi; x^{(i)})$ variasional aşağı sərhəddir və aşağıdakı kimi yazıla bilər:

$$L(\theta, \phi; x^{(i)}) = -D_{KL} \left(q_\phi(z|x^{(i)}) \parallel p_\theta(z) \right) + E_{q_\phi(z|x^{(i)})} \left[-\log p_\theta(x^{(i)}|z) \right] \quad (4)$$

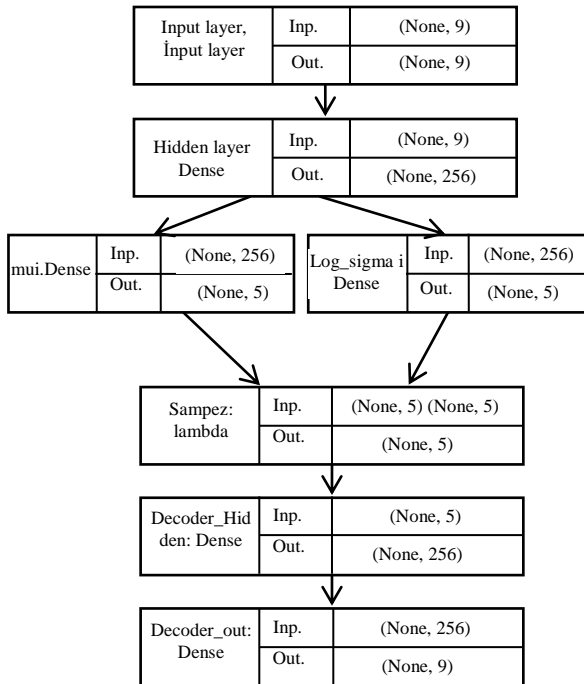
Bulud texnologiyaları infrastrukturunun mürəkkəbliyini nəzərə alaraq $q_\phi(z|x)$ və $p_\theta(x|z)$ parametrlərini

modelləşdirmək üçün dərin neyron şəbəkəsi istifadə edilmişdir. Enkoder və dekoderin öyrədilməsində Adam optimizasiya funksiyası istifadə edilmişdir. Optimallaşdırmanın məqsədi variasional aşağı sərhəddi maksimallaşdırmaqdan ibarətdir. Adam staxostik optimallaşdırma məsələləri üçün yaradılmış staxostik qradiyent eniş optimallaşdırılması üsuludur.

III. EKSPERİMENTLƏR

A. Variasional Avtoenkoder modelinin konfigurasiyası

Məqalədə dərin Variasional Avtoenkoder modeli *Theano* kitabxanasından istifadə etməklə qurulub və şəkil 2-də təsvir edilmişdir. Variasional Avtoenkoder modelinin arxitekturasının bir hissəsini enkoder, bir hissəsini isə dekoder təşkil edir. Şəbəkənin enkoder hissəsini 6 lay təşkil edir. Dekoder enkoderin əməliyyatına əks əməliyyatı yerinə yetirir və 6 laydan ibarətdir.



Şəkil 2. Dərin variasional avtoenkoder modeli

B. Verilənlər bazasının təsviri

İşdə 13 bulud qovşağından toplanmış 28147 nümunədən və 9 əlamətdən ibarət iş yükü verilənlər bazası istifadə edilmişdir [18]. 1 dəqiqədə tapşırıqların sayı, 5 dəqiqədə tapşırıqların sayı, 15 dəqiqədə tapşırıqların sayı, yaddaşın tutumu, diskin tutumu, CPU nüvələrinin sayı, hər bir nüvədə CPU-nun sürəti, şəbəkənin buraxma qabiliyyətinə görə ortalama qəbul edilən trafik kilobit həcmi (Kbps), şəbəkənin buraxma qabiliyyətinə görə ortalama ötürülən trafik kilobit həcmi (Kbps) istifadə

olunan verilənlər bazasının beş kateqoriyadan ibarət sinif nişanları təşkil edir. Bu kateqoriyalar çox az (very low, VL), az (low, L), orta (medium, M), yüksək (high, H), çox yüksək (very high, VH) kimi təsnif olunur.

Bulud resurslarında yükün proqnozlaşdırılması üsulunun test prosesi Azərbaycan Milli Elmlər Akademiyası, İnformasiya Texnologiyaları İnstitutunun verilənlər mərkəzində, Ubuntu 16.04.3 LTS amd64 system, 3312 GB memory, 2933.437 CPU MHz mühitində yerinə yetirilmişdir. Eksperimentlərin aparılması üçün Python proqram paketi və Tensorflow kitabxanası istifadə edilmişdir.

İş yükünün proqnozlaşdırılması variasional avtoenkoder, və dərin neyron şəbəkəsi əsasında aparılmışdır. Burada ilk öncə neyron şəbəkənin nişanlanmış verilənlər (labeled training data) əsasında 9 iş yükü atributları üzrə öyrədilməsi aparılır, daha sonra nişanlanmamış verilənlər (unlabeled testing data) əsasında iş yükünün proqnozlaşdırılması aparılır.

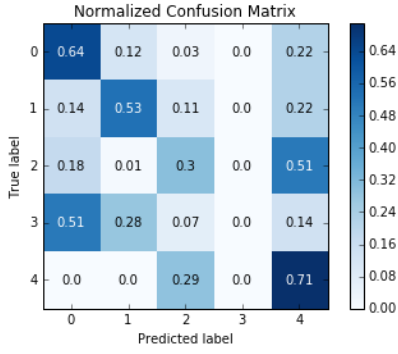
Variasional avtoenkoderdə verilənlər bazasının siniflərinə uyğun 5 gizli lay istifadə edilmişdir.

Modelin effektivliyinin qiymətləndirilməsi üçün precision, recall, F1-score, accuracy metrikaları istifadə edilmişdir. Bulud əsaslı elektron dövlət infrastrukturunda VM-ların iş yükünün proqnozlaşdırılması məsələsində Variasional Avtoenkoderin dəqiqliyi cədvəl 1-də verilmişdir. Cədvəl 1-dən görüldüyü kimi model 0, 1 və 4-cü siniflərin aşkarlanmasında yaxşı nəticələr vermişdir. Lakin 2 və 3-cü siniflərin aşkarlanmasında aşağı effektivlik nümayiş etdirmişdir. Bu onunla əlaqədar ki, model 2 və 3-cü siniflərdən olan verilənlərin strukturunu rekonstruksiya etməkdə çətinliklə qarşılaşmışdır.

Cədvəl 1. Variasional Avtoenkoder

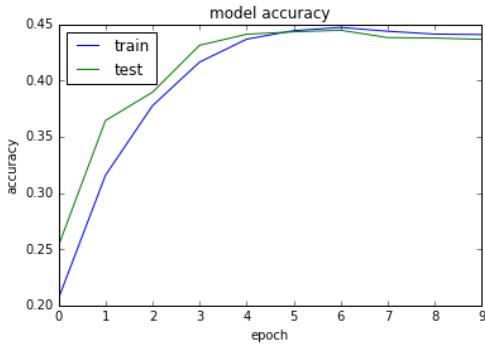
Sinif	Precision	Recall	F1-score	Accuracy
VL (0)	0.43	0.64	0.51	0.64
L (1)	0.57	0.53	0.55	0.53
M (2)	0.37	0.30	0.33	0.3
H (3)	0.00	0.00	0.00	0.0
VH (4)	0.40	0.71	0.51	0.71

Bu mənzərəni Variasional Avtoenkoderin xətlər matrisindən də (Şəkil 3) aşkar şəkildə görmək mümkündür.

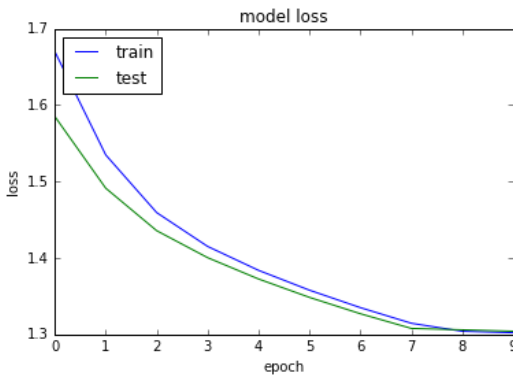


Şəkil 3. Variasional Avtoenkoderin xətalər matrisi

Variasional Avtoenkoderin iş yükünün proqnozlaşdırılması üzrə verilənlər bazasına tətbiqi zamanı itki funksiyasının dinamikasında azalma, dəqiqlik funksiyasının dinamikasında isə artım müşahidə edilmişdir (Şəkil 4, Şəkil 5).



Şəkil 4. Variasional Avtoenkoderin dəqiqlik funksiyasının artım dinamikası



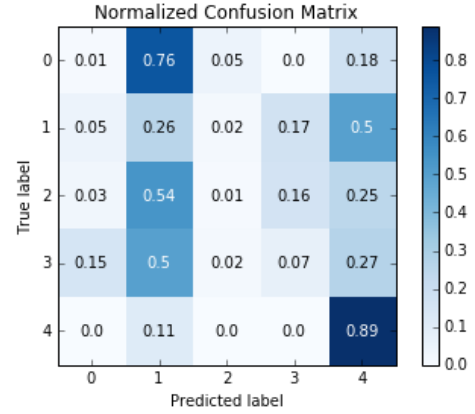
Şəkil 5. Variasional Avtoenkoderin itki funksiyasının azalma dinamikası

İş yükünün proqnozlaşdırılması üzrə verilənlər bazasına sadə dərin neyron şəbəkələrini tətbiq etdikdə isə (Cədvəl 2) model 4-cü sinifdən olan nümunələri yaxşı klassifikasiya edə bilmişdir, lakin qalan dörd sinifdən olan nümunələri aşağı effektivliklə tanıya bilmişdir.

Cədvəl 2. Sadə dərin neyron şəbəkəsi

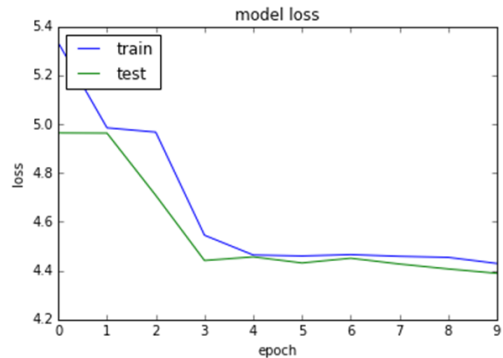
Sinif	Precision	Recall	F1-score	Accuracy
VL (0)	0.07	0.01	0.02	0.01
L (1)	0.01	0.26	0.01	0.26
M (2)	0.15	0.01	0.02	0.01
H (3)	0.29	0.07	0.11	0.07
VH (4)	0.55	0.89	0.68	0.89

Bu mənzərəni sadə dərin neyron şəbəkəsinin xətalər matrisindən də (Şəkil 6) aşkar şəkildə görmək mümkündür.

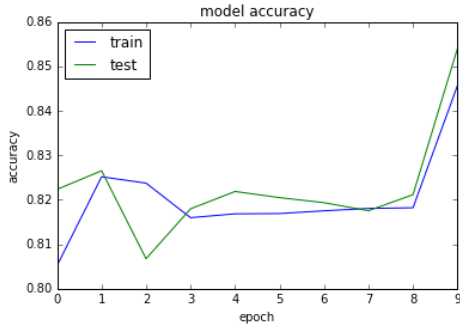


Şəkil 6. Sadə dərin neyron şəbəkəsinin xətalər matrisi

Sadə dərin neyron şəbəkəsinin iş yükünün proqnozlaşdırılması üzrə verilənlər bazasına tətbiqi zamanı itki funksiyasının dinamikasında azalma, dəqiqlik funksiyasının dinamikasında isə qismən artım müşahidə edilmişdir (Şəkil 7, Şəkil 8).



Şəkil 7. Sadə dərin neyron şəbəkəsinin itki funksiyasının azalma dinamikası



Şəkil 8. Sadə dərin neyron şəbəkəsi dəqiqlik funksiyasının artım dinamikası

NƏTİCƏ

Məqalədə variational avtoenkoderlərin generasiya etdiyi rekonstruksiya ehtimalları əsasında VM-ların iş yükünün proqnozlaşdırılması üsulu təklif edilmişdir. VM-ların yüklənmə dərəcələrinin xassələri və ehtimal qiymətləri təyin edilərək yüklənmə sinfini müəyyən etmək üçün Variational Avtoenkoder alqoritminin rekonstruksiya ehtimalları istifadə edilmişdir. Təklif edilmiş üsul VM-ların yüklənmə dərəcəsini xarakterizə edən verilənlər bazasında test edilmişdir. Nəticələr göstərmişdir ki, Variational Avtoenkoder alqoritmi VM-ların iş yükünün proqnozlaşdırılmasını sadə dərin neyron şəbəkələrinə nisbətən daha yaxşı həyata keçirir.

MİNNƏTDARLIQ

Bu iş Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondunun maliyyə yardımı ilə yerinə yetirilmişdir – Qrant № EİF-BGM-4-RFTF-1/2017-21/08/1

ƏDƏBİYYAT

- [1] K.E. Ali, S.A. Mazen, E.E. Hassanein, “A proposed hybrid model for adopting cloud computing in e-government,” *Future Computing and Informatics Journal*, 2018, In Press.
- [2] R.M. Alguliyev, Y.N. Imamverdiyev, F.J. Abdullayeva, “PSO based load balancing method in cloud computing,” *Automatic Control and Computer Sciences*, No.1, 2019, in press.
- [3] X. Meng, C. Isci, J. Kephart, L. Zhang, E. Bouillet, D. Pendarakis, “Efficient resource provisioning in compute clouds via VM multiplexing,” *Proceeding of the 7th International Conference on Autonomic Computing*, ACM, 2010, pp. 11-20.
- [4] J. Kumar, A.K. Singh, Workload prediction in cloud using artificial neural network and adaptive differential evolution, “*Future Generation Computer Systems*,” 2018, vol. 81, pp. 41-52.
- [5] Y. Cheng, W. Chen, Z. Wang, Z. Tang, Y. Xiang, Smart VM co-scheduling with the precise prediction of performance characteristics, “*Future Generation Computer Systems*,” 2016, In Press.
- [6] T. Mahdhi, H. Mezni, A prediction-Based VM consolidation approach in IaaS Cloud Data Centers, “*Journal of Systems and Software*,” vol., 146, 2018, pp. 263-285.
- [7] R. Hu, J. Jiang, G. Liu, L. Wang, Efficient Resources Provisioning Based on Load Forecasting in Cloud, “*The Scientific World Journal*,” Vol. 2014, 2014, pp. 1-12.
- [8] A. Abdelsamea, A.A. El-Moursy, E.E. Hemayed, H. Eldeeb, Virtual machine consolidation enhancement using hybrid regression algorithms, “*Egyptian Informatics Journal*,” Vol 18, no 3, 2017, pp. 161-170.

- [9] D. Yang, J. Cao, J. Fu, J. Wang, J. Guo, “A pattern fusion model for multi-step-ahead CPU load prediction,” *Journal of Systems and Software*, vol. 86, no. 5, 2013, pp. 1257-1266.
- [10] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet classification with deep convolutional neural networks,” 25th International Conference on Neural Information Processing Systems (NIPS 2012), Granada, Spain, Vol. 1, 2012, pp. 1097-1105.
- [11] M. Abdelsalam, R. Krishnan, Y. Huang, and R. Sandhu, “Malware detection in cloud infrastructures using Convolutional Neural Networks,” *IEEE 11th international conference on Cloud Computing (CLOUD)*, 2018, pp. 162-169.
- [12] Y. Imamverdiyev, F. Abdullayeva, “Deep Learning Method for Denial of Service Attack Detection Based on Restricted Boltzmann Machine,” *Big Data*, vol. 6, no. 2, 2018, pp. 159-169.
- [13] G.E. Hinton, and Z. Ghahramani, “Generative models for discovering sparse distributed representations,” *Philosophical transactions of the royal society of London, Series B: Biological sciences*, vol. 352, no. 1358, 1997, pp. 1177-1190.
- [14] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2014, pp. 2672-2680.
- [15] K. Gregor, I. Danihelka, A. Mnih, C. Blundell, and D. Wierstra, “Deep AutoRegressive Networks” In *International Conference on Machine Learning*, 2014, pp. 1242-1250.
- [16] D. Kingma, and M. Welling, “Auto-encoding variational Bayes,” 2nd International Conference on Learning Representations (ICLR 2014), Banff, Canada, 2014, pp. 1-14.
- [17] D.M. Blei, A. Kucukelbir, J.D. McAuliffe, Variational inference: A review for statisticians, *Journal of the American Statistical Association*, vol. 112, no. 518, 2017, pp. 859-877.
- [18] A.M. Al-Faifi, B. Song, M.M. Hassan, A. Alamri, A. Gumaei, “Performance prediction model for cloud service selection from smart data, *Future Generation Computer Systems*,” 2018, vol. 85, pp. 97-106.

EFFICIENCY IMPROVEMENT METHOD OF THE CLOUD-BASED E- GOVERNMENT SERVICES

Fargana Abdulayeva

Institute of Information Technology of ANAS, Baku,
Azerbaijan

a_farqana@mail.ru

Abstract- In this paper, the virtual machine (VM) workload prediction method in the cloud infrastructure is proposed. To make predictions the reconstruction probabilities of the Variational Autoencoders are used. The reconstruction probability is the probability measure that considers the variabilities of the variable distributions. In this paper, the reconstruction probability values of the variational autoencoders show the load degrees of the VMs. The results of the experiments are shown that variational autoencoder is produced better results in predicting VM workloads than simple deep neuronal networks. The generative characteristics of the variational autoencoders, by reconstructing the data allow identifying the types (levels) of the load.

Keywords- Cloud-based electronic government; workload prediction; a virtual machine; deep neural networks; deep learning; variational autoencoder.