Time series in sensor data using state of the art deep learning approaches: A systematic literature review

Luis-Roberto Jácome-Galarza[1[0000-0002-2886-3372]], Miguel-Andrés Realpe-Robalino[1[0000-0001-8711-5596]], Jonathan Paillacho-Corredores[1[0000-0001-6489-067X]], José-Leonardo Benavides-Maldonado[2[]]

[1] Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ingeniería en Electricidad y Computación, CIDIS. Gustavo Galindo Km. 30.5 Vía Perimetral, Guayaquil, Ecuador.
[2] Universidad Nacional de Loja, La Argelia, Loja-Ecuador
lrjacome@espol.edu.ec, jspailla@espol.edu.ec, mrealpe@fiec.espol.edu.ec, jose.benavides@unl.edu.ec

Abstract. IoT (Internet of Things) and AI (Artificial Intelligence) are becoming support tools for several current technological solutions due to significant advancements of these areas. The development of the IoT in various technological fields has contributed to predicting the behavior of various systems such as mechanical, electronic, and control using sensor networks. On the other hand, deep learning architectures have achieved excellent results in complex tasks, where patterns have been extracted in time series. This study has reviewed the most efficient deep learning architectures for forecasting and obtaining trends over time, together with data produced by IoT sensors. In this way, it is proposed to contribute to applications in fields in which IoT is contributing a technological advance such as smart cities, industry 4.0, sustainable agriculture, or robotics. Among the architectures studied in this article related to the process of time series data we have: LSTM (Long Short-Term Memory) for its high precision in prediction and the ability to automatically process input sequences; CNN (Convolutional Neural Networks) mainly in human activity recognition; hybrid architectures in which there is a convolutional layer for data pre-processing and RNN (Recurrent Neural Networks) for data fusion from different sensors and their subsequent classification; and stacked LSTM Autoencoders that extract the variables from time series in an unsupervised way without the need of manual data pre-processing.Finally, well-known technologies in natural language processing are also used in time series data prediction, such as the attention mechanism and embeddings obtaining promising results.

Keywords: time series, deep learning, recurrent networks, sensor data, IoT.

1. Introduction

The IoT (Internet of things) is formed by many connected devices and transmit vast amounts of data [1]. It is currently being applied in fields like agriculture [2, 3], smart cities [4], smart homes [5, 6], health care [7, 8], and human activity recognition [9, 10]. As time goes by, these IoT sensors are used, more insights can be obtained from them that may help to predict the behavior of systems which is useful for systems maintenance [11, 12], yield performance [13, 14], resource allocation [15], or business planning [16]. Deep neural networks have achieved state-of-the-art results in complex tasks like image recognition [17] using CNN [18], human language understanding [19] using LSTM [20], or gaming [21] with reinforcement learning [22]. It is relevant to understand how well deep learning algorithms can extract patterns in time series and what trends those algorithms can find. When designing a deep learning architecture for time series forecasting, RNN and LSTM networks are considered, but CNN networks are also used to report good results [23]. Accordingly, this paper aims to analyze what deep learning architectures are being utilized to forecast time series, their applications, efficiency, challenges, and trends in time.

The rest of this paper is structured as follows: Section 2 describes the research methodological framework; section 3 shows the results obtained; section 4 presents the limitations and future work, and section 5 concludes the article.

2. Materials and Methods

We used Barbara Kitchenham's methodology [24], which includes: formulation of research questions, search process, inclusion and exclusion criteria, data extraction, and data analysis and classification. The following research questions were posed to attain this purpose:

- RQ1: What deep learning architectures are being used in projects involving time series in sensor data, and what are their application and efficiency?
- RQ2: What challenges are found in projects with time series in sensor data?
- RQ3: What trends are found in the applied methodologies in projects with sensor data and time series?

A manual search on the Science Direct, Google Scholar, and Springer databases was conducted. The search was performed in June 2020. The search string was the following: “deep learning” AND “sensor data” AND “time series”. Additionally, only scientific papers written in English and later than 2016 were considered.

From the first search process, 8033 articles were extracted (532 Science Direct, 61 Springer, 7440 Google Scholar). After that, the extraction process resulted in 40 articles with the following year distribution: 2016 (3 articles), 2017 (11 articles), 2018 (10 articles), 2019 (10 articles), and 2020 (6 articles).

The data analysis and classification steps are explained below.

1. Reading the abstracts and conclusions.
2. Searching for each criterion within the complete content of the articles.
3. Reading the whole article when necessary.
4. Classifying articles by criteria.

3. Results

3.1 First query results

Table 1 is presented to answer the RQ1. This table includes the list of analyzed papers, their application, the deep learning architecture, and efficiency.

Table 1. List of analyzed papers

Paper id	Application	Year	Deep learning architecture	Efficiency
P001	Water quality monitoring network	2019	LSTM	69.2%, 70.3%, 98.3%, and 76% improvement in RMSE, MAE, MAPE, SMAPE
P002	Generate sensory data for a deep learning-based discriminator model	2017	LSTM - MDN	Discriminator distinguishes 50%.
P003	Indoor Air Quality Analysis	2017	GRU LSTM	84.69% accuracy
P004	Heterogeneous Human activity recognition, user identification with motion analysis, car tracking with motion sensors	2017	CNN GRU	99.7% accuracy
P005	Real-time activity classification.	2016	CNN spectrogram	85% accuracy
P006	Human emotion classification	2019	CNN LSTM	95% accuracy
P007	Machinery fault diagnosis	2017	DNN	100% accuracy
P008	Predictive monitoring of production processes	2017	LSTM FFNN	87% accuracy
P009	Human activity recognition	2016	CNN	94.79% accuracy
P010	Human activity recognition based on mobile sensor data	2020	DRN CNN GAF	96 % accuracy
P011	Driving behavior identification	2019	CNN, LSTM, GRU	97% accuracy
P012	Marine sensor data prediction	2020	DBEN	91% accuracy
P013	Predicts if participants would attend to a notification within 10 minutes	2017	RNN	40% improvement respect to random baseline
P014	Forecasting the future use of energy for the home appliances	2017	GMDH	63.23 RMSE
P015	Machine health monitoring (MHM)	2018	CONV LSTM	8.39 RMSE
P016	Atrial fibrillation detection using wearables.	2018	CONV LSTM	False positive and false negative rates below $2 \times 10{-}3$
P017	Aircraft hard-landing prediction	2018	LSTM	$1.1\times10{-}3$ MSE
P018	Multimodal sensor fusion for human activity recognition	2019	LSTM	94.46%, accuracy
P019	Anomaly detection	2017	LSTM	97%, 95%, 97% accuracy
P020	Predicting remaining useful life of machines	2017	GRU RNN	466 MSE
P021	Monitor the performance degradation of commercial aircraft air conditioning	2020	LSTM	0.0466 error

	systems (ACS)			
P022	Construction equipment activity recognition	2019	LSTM	79.9%, 96.7% accuracy
P023	Aircraft landing speed prediction	2018	LSTM	3.5 RMSE
P024	Data augmentation using synthetic data for time series classification	2018	CNN	97% accuracy
P025	Remaining useful life RUL prediction model	2019	CNN	2.5, 6.41, 3.25E-03, 10.72, 6.26, 3.12, 19.02 RMSE
P026	Time series classification with multivariable data.	2018	CNN	97.4%, 97.7% accuracy
P027	Human activity recognition with inertial sensors	2016	CNN	97.01% accuracy
P028	Human activity recognition from inertial sensor time series	2018	LSTM	92% accuracy
P029	Multi time series anomaly detection	2019	CNN RNN	Anomalies reduced by 3%
P030	Landslide hazard prediction	2018	LSTM, BPNN	81.2%, 62% accuracy
P031	Data-driven anomaly detection for UAV sensors	2019	LSTM	99% accuracy
P032	Recovering missing air quality data	2018	LSTM	0.63, 2.90, 1.87 MAE
P033	Missing sensor data prediction in IoT	2020	LSTM	0.07, 0.08, 0.17 RMSE
P034	Time series data for equipment reliability analysis.	2020	FFNN	75%, 82%, 85%, 79%, 80% accuracy
P035	Recognizing and forecasting the underlying high-level states from raw sensory data for activity recognition	2017	CNN LSTM	87.47%, 80.21% accuracy
P036	Sleep apnea severity detection	2017	LSTM-RNN	100%, 99.9% accuracy
P037	Activity recognition via multichannel sensor data.	2020	CNN LSTM	0.755, 0.933, 0.918 F1 score
P038	Mobile application usage prediction	2019	LSTM	89.55%, 95.53% accuracy
P039	Wearable-based Parkinson’s disease severity monitoring	2019	FCN	84.3% accuracy
P040	Classify the dangerous levels of methane concentration.	2019	RBM DBN	97.6%, 99.2%, 96.3% accuracy

Acronyms: Bi-LSTM Bidirectional Long Short-Term Memory; GRU Gated Recurrent Units; FCNN Fully Connected Neural Network; GAN Generative Adversarial Networks; DFNN Deep Feedforward Neural Networks; LSTM Long Short-Term Memory; DBEN Deep Belief Echo State Network; RNN Recurrent Neural Networks; BPNN Back Propagation Neural Networks; DRN Deep Residual Networks; CNN Convolutional Neural Networks; DAE Deep Autoencoder; FCN Fully Convolutional Net; DBN Deep Belief Networks.

Deep learning architectures. It can be seen in Fig. 1 that LSTM is the most used architecture. LSTM networks can automatically learn high-level representative features containing the long-range temporal relationships. Stacked LSTM autoencoders extract the features from time series in an unsupervised manner avoiding the manual feature engineering. CNN are preferred for human activity recognition. Human activities are hierarchical because complex activities are composed of basic actions. Human activities are also translation-invariant because different people perform the same activity differently. A fragment of an activity can be performed at different points in time. CNN can address raw temporal signals in machine health monitoring (MHM) tasks without any time-frequency transformation. Convolutional kernels can be thought of as signal filters and remove the need for manual signal processing; in more complex architectures, CNN is used for feature extracting while RNN is used for time series classification.

Embeddings (numerical vectors that represent features of a system) can be used for prediction. The embeddings of two machines (P020) with similar operational behavior are close to each other. Heart rhythm embedding (P016) for both Afib (Atrial Fibrillation) and NSR (Normal Sinus Rhythm) vary between patients, so unsupervised clustering applied to heart rhythm embeddings can improve the model's accuracy. Furthermore, paper P015 explains that fully connected structures of DAE and DBN may lead to high computation costs and overfitting problems caused by huge model parameters. To avoid those disadvantages, they use a hybrid Conv LSTM architecture for machine health monitoring.

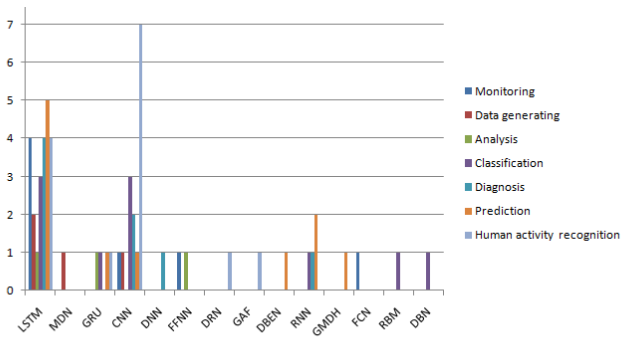

Fig.1. Deep learning architectures and their applications

Multi-modal data fusion architectures. In paper P004, the authors present a multi-modal data fusion model for car tracking with motion sensors (accelerometer, gyroscope, magnetometer, and GPS for the ground truth), heterogeneous human activity recognition (accelerometer, gyroscope), and user identification with biometric motion analysis (accelerometer, gyroscope). The Fourier transform is used to obtain the frequency dimension. A CNN extracts local features within each sensor modality and merges the local features into global features hierarchically; a stacked GRU extracts temporal dependency. In paper P006, they achieve motion detection fusing various sensor modalities: on-body, environment, and location data. Two convolutional layers analyze the concatenated data. After that, a fully convolutional layer is used, then a LSTM layer processes the temporal information, and a softmax layer is used for classification. The model gets better performance than single CNN and LSTM. In paper P010, they use multi-modal data fusion for human activity recognition. The model encodes sensor data as images using GAF (Gramian Angular Fields), representing time series data in polar coordinates instead of the Cartesian coordinates. A CNN analyzes the image representation of each sensor's data. The model performs better than random forests, SVM, restricted Boltzmann machines, FCNN and LSTM. In paper P018, the authors use multi-modal data fusion for human activity prediction. The data from on-body IMU devices (tri-axial accelerometer, gyroscope, and magnetometer) are analyzed individually by an LSTM classifier; the results from each sensor modality are used as features for a model that integrates the results, applying simple soft-voting based on the weighted average of class probabilities (soft-voting gets higher overall performance than hard-voting because it can give more weight to highly confident votes). This model uses random forests, kNN, and SVM for the final classification. In P029, they present an industrial case where three types of anomalies detection are considered: point, context-specific, and collective.

CNN blocks individually process sensors; then, the resulting feature maps are fed to an RNN. A multi-channel CNN architecture is used to process multiple time series. A single feature map containing the main features of all the time series is obtained. The number of extracted features per window is much higher for the Multi-head convolution, which performs better than the multi-channel CNN architecture. In paper P030, they use multi-modal data fusion for landslide hazard prediction. They combine DEM (digital elevation model), HR-RS (high-resolution remote sensing) images, 1:50,000 GM (geologic maps), and meteorological data. A CNN processes the images, the build-up index, and the stream power index. The DEM, HR-RS, GM, and meteorological data are fed into a LSTM for classification. The model performs better than SVM, decision trees, and backpropagation neural networks.

3.2 Second query results

In order to answer the RQ2, Fig. 2 shows that complexity is the main challenge reported in manufacturing projects since the data has high-dimensional characteristics. Missing or noisy data is also found as a challenge in all kinds of projects. In projects with wireless sensor data, low bandwidth can be a problem. The lack of computation power and energy consumption are common problems in projects that use sensors in mobile devices. Imbalanced classes can be an issue since more errors are found on samples from minority classes. Data privacy is a limitation in some driving datasets. The lack of performance metrics is a limitation in unsupervised learning and generative adversarial networks. Marine data (P012) has fluctuations, outliers, noise, and complex characteristics. The intermittent nature of heart conditions such as AFib (P016) is a challenge for data collection. Human movements are encoded in a sequence of time, and the activities are not defined by one small window of data alone, and multiple sensor devices are required (P028). There is much more data related to normal behavior in anomaly detection than anomalous data (P029). Sensor-generated data is expected to be noisy, uncertain, erroneous, and missing due to the lack of battery power, communication errors, and malfunctioning devices (P033). Healthcare projects report high costs for acquiring medical datasets.

Regarding the mitigation of the impacts of the detected challenges, the following measures are explained:

Sampling frequency. If all sensors have the same sampling rate, 2D spatio-temporal sequence data is formed. Down-sampling reduces the number of dimensions in the original time series and makes it easier to learn patterns; in contrast, up-sampling can be done by the Lagrange interpolation method.

Time-series data augmentation. Some techniques are slicing window, jittering, scaling, time warping, and permutation (in paper P024, it is reported that data augmentation improved the performance of the CNN model considerably). Some algorithms incur a heavier loss for errors on samples from minority classes; putting more focus on the minority classes during training is needed. Over-sampling and under-sampling techniques are usually used to balance the number of instances of both classes.
Sliding window. Time series data is processed in a window-based manner. The sliding window length affects the number of parameters of each layer. Allowing overlap between segments avoids creating gaps in the output signal.
Other tools and techniques. Other techniques such as batch normalization are commonly used, transfer learning in time series data is proposed, but it needs further research. Likewise, InfluxDB, an open-source distributed database of time series events and metrics, is popular for storing real-time IoT data; the Ray v0.7.6 distributed application framework is used for hyperparameter search. In paper P037, they evaluate the training and architectural parameters on the mentioned framework.

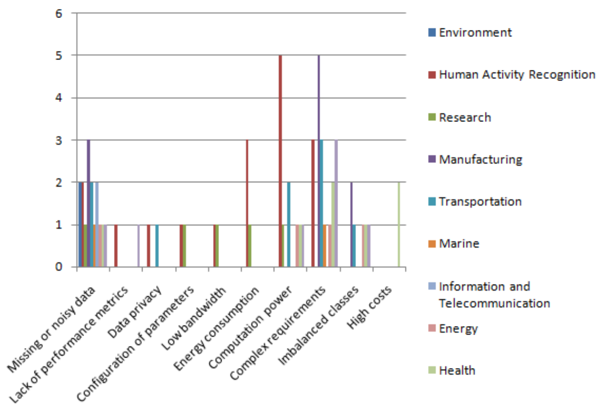

Fig.2. Challenges detected in the projects of time-series and sensor data

3.3 Third query results

In order to answer RQ3, Fig. 3 shows that the use of a single architecture is still the most utilized approach for time series forecasting and LSTM networks as the most popular choice. LSTM is proven to be the most stable and robust RNN model to learn long-range temporal patterns in practical applications.

Bi-LSTMs were used mainly for recovering missing sequence data. Papers P018, P019, and P028 used stacked LSTM, which is built instead of using a higher number of memory blocks in a standard recurrent hidden layer. One more recurrent hidden layer is stacked, providing better generalization and robust temporal pattern learning. In P019, the model gets 97%, 95%, and 97% accuracy in anomaly detection performing better than traditional methods like cumulative sum and exponentially weighted moving average. In P028, the model gets 92% accuracy in human activity recognition having better results than SVM, deep convolutional LSTM, FCNN, and single LSTM. They state that batch normalization can obtain similar accuracy with fewer training epochs. Autoencoders explore the underlying structure in the dataset in an unsupervised manner; they can be categorized as simple or deep fully connected autoencoders, convolutional autoencoders, sequence-to-sequence autoencoders (LSTM autoencoders), contractive autoencoders, and variational autoencoders. The encoder can be used to read from the time-series data, and a decoder generates the missing data sequences. Hybrid approaches are also found in P006, P011, P015, P016, P035, and P037. CNN learn the spatial features and LSTMs learn the temporal features. The CNN can be treated independently, forming convolutional heads for individually feature extracting sensor data, having the advantage of easily adding new sensors. The ConvLSTM is described as an extension of LSTM, replacing the matrix multiplication in LSTM with a convolutional operation for performing spatio-temporal forecasting. In the paper, P035, the authors propose an LR-ConvLSTM (Long-term Recurrent Convolutional LSTM Network) model for time series classification and high-level state prediction. LRConvLSTM is a hybrid deep learning representation that has a stacked Convolutional LSTM component to learn complex temporal features and an LSTM component that learns external temporal features. Most of the projects use the “many-to-one” approach (the input is divided into fixed-length overlapping windows, then a model processes each window individually, generating a class prediction for each one, and then the predictions are concatenated). Researchers use the “many-to-many” approach in the paper P037. The entire output time series is generated with a single model evaluation;they propose a GPU-optimized (cuDNN) LSTM implementation for improving the model. In P029, scholars generate an already-trained model using transfer learning (Transfer Knowledge-Based - TKB) instead of training a model from scratch. Authors use transfer learning to improve their model for monitoring Parkinson's disease with wearable devices in paper P039, An optimal mechanism for time step searching is needed to improve learning models was found in P003.

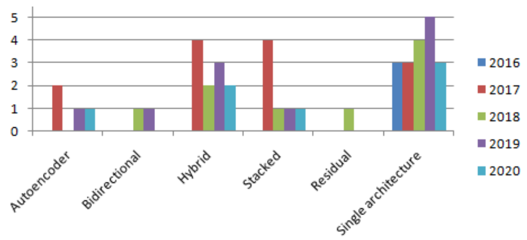

Fig.3. Deep learning architectures and their applications

4. Discussion

The current paper was limited by its main focus on deep learning techniques and restraint of classical approaches. As future work, we intend to extend this analysis to the projects that utilize the ConvLSTM and Temporal Convolutional Network architectures towards verifying their advantages. Additionally, tools for automatic hyper-parameter search need further research.

5. Conclusions

Deep learning has achieved impressive improvement in many research fields. This paper presents a systematic literature review about the use of DL in time series data produced by IoT sensors. Its insights may be beneficial for optimization in many industrial processes. The fusion of multi-modal sensor data is a challenging task suitable for deep learning networks since they can find patterns from raw and massive data. Findings revealed that multi-modal approaches increase the accuracy of the predicting models. Well-known technologies in natural language processing are also being used in time series prediction. Thus, the attention mechanism and embeddings disclosed promising results. The combination of CNN and LSTM networks is a good approach for improving time series data prediction. The CNNs are used for the pre-processing of the sensor data, and the RNNs are used for processing time patterns. CNN networks can also obtain spatial patterns. Moreover, Stacked LSTM Autoencoders are used for extracting variables from time series in an unsupervised way without performing manual data pre-processing.The challenges found in time series prediction are missing and noisy data, privacy restrictions of datasets, high costs of datasets, especially in the health area, high computational demand, and high wireless speed requirements for mobile devices. Applying transfer learning to time series prediction still needs research. Other relevant topics for time series prediction studies are Temporal Convolutional Networks and Sequence-to-Sequence learning.

References

1. Mahdavinejad, M., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., Sheth, A.: Machine learning for Internet of Things data analysis: A survey. Digital Communications and Networks, 4(3), 161-175 (2018).
2. Gondchawar, N., Kawitkar, R.: IoT based smart agriculture. International Journal of advanced research in Computer and Communication Engineering, 5(6), 838-842 (2016).
3. Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., Nillaor, P.: IoT and agriculture data analysis for smart farm. Computers and electronics in agriculture, 156, 467-474 (2019).
4. Kim, T., Ramos, C., Mohammed, S.: Smart city and IoT (2017).
5. Samuel, S.: A review of connectivity challenges in IoT-smart home. In 2016 3rd MEC International conference on big data and smart city (ICBDSC) pp. 1-4. IEEE (2016).
6. Kim, Y., Park, Y., Choi, J.: A study on the adoption of IoT smart home service: using Value-based Adoption Model. Total Quality Management & Business Excellence, 28(9-10), 1149-1165 (2017).
7. Ukil, A., Bandyoapdhyay, S., Puri, C., Pal, A.: IoT healthcare analytics: The importance of anomaly detection. In 2016 IEEE 30th international conference on advanced information networking and applications (AINA) pp. 994-997. IEEE (2016).
8. Tyagi, S., Agarwal, A., Maheshwari, P.: A conceptual framework for IoT-based healthcare system using cloud computing. In 2016 6th International Conference-Cloud System and Big Data Engineering (Confluence) pp. 503-507. IEEE (2016).

9. Rghioui, A., Sendra, S., Lloret, J., Oumnad, A.: Internet of things for measuring human activities in ambient assisted living and e-health. Network Protocols and Algorithms, 8(3), 15-28 (2016).
10. Shi, C., Liu, J., Liu, H., Chen, Y.: Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT. In Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing pp. 1-10 (2017).
11. Al-Douri, Y. K., Hamodi, H., Lundberg, J.: Time series forecasting using a two-level multi-objective genetic algorithm: a case study of maintenance cost data for tunnel fans. Algorithms, 11(8), 123 (2018).
12. Baptista, M., Sankararaman, S., de Medeiros, I., Nascimento Jr, C., Prendinger, H., Henriques, E.: Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. Computers & Industrial Engineering, 115, 41-53 (2018).
13. Kamir, E., Waldner, F., Hochman, Z.: Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. Isprs Journal of Photogrammetry and Remote Sensing, 160, 124-135 (2020).
14. Adeniyi, O. D., Szabo, A., Tamás, J., Nagy, A.: Wheat Yield Forecasting Based on Landsat NDVI and SAVI Time Series (2020).
15. Kadri, F., Harrou, F., Chaabane, S., Tahon, C.: Time series modelling and forecasting of emergency department overcrowding. Journal of medical systems, 38(9), 1-20. (2014).
16. Demir, E., Dincer, S.: Place and solution proposals of data mining in production planning and control processes: a business application. PressAcademia Procedia, 11(1), 189-193 (2020).
17. Wu, M., Chen, L.: Image recognition based on deep learning. In 2015 Chinese Automation Congress (CAC) pp. 542-546. IEEE (2015).
18. Qassim, H., Verma, A., Feinzimer, D.: Compressed residual-VGG16 CNN model for big data places image recognition. In 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC) pp. 169-175. IEEE (2018).
19. Tur, G., Celikyilmaz, A., He, X., Hakkani-Tür, D., Deng, L.: Deep learning in conversational language understanding. In Deep learning in natural language processing pp. 23-48. Springer, Singapore (2018).
20. Wen, T., Gasic, M., Mrksic, N., Su, P., Vandyke, D., Young, S.: Semantically conditioned lstm-based natural language generation for spoken dialogue systems. arXiv preprint arXiv:1508.01745 (2015).
21. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hassabis, D. Mastering the game of go without human knowledge. nature, 550(7676), 354-359 (2017).
22. Vinyals, O., Ewalds, T., Bartunov, S., Georgiev, P., Vezhnevets, A. Yeo, M., II, S.: A New Challenge for Reinforcement Learning. arXiv preprint ArXiv:1708.04782 (2017).
23. Liu, R., Meng, G., Yang, B., Sun, C., Chen, X.: Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine. IEEE Transactions on Industrial Informatics, 13(3), 1310-1320 (2016).
24. Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering–a systematic literature review. Information and software technology, 51(1), 7-15 (2009).

APPENDICES

The list of the analyzed papers is available in the Appendix available in:
https://www.dropbox.com/s/1hlpy6i1n0vb7rk/appendix.pdf?dl=0