Available online at www.alphanumericjournal.com

alphanumeric journal
The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 11, Issue 2, 2023

Received: December 10, 2023
Accepted: December 31, 2023
Published Online: December 31, 2023

AJ ID: 2023.11.02.MIS.02
DOI: 10.17093/alphanumeric.1402897
Research Article

Time Series Prediction with Digital Twins in Public Transportation Systems

Mehmet Ali Ertürk, Ph.D.*

Assist. Prof., Department of Computer Engineering, Faculty of Computer and Information Technologies, Istanbul University, Istanbul, Turkiye, mehmetali.erturk@istanbul.edu.tr

* İstanbul Üniversitesi, Bilgisayar ve Bilişim Teknolojileri Fakültesi, Bozdoğan Kemeri Cad. No:1 Vezneciler / Fatih, İstanbul, Türkiye

ABSTRACT

Classical traffic and transportation control centers must be more robust with the rapid spread of electric, intelligent, autonomous, and software-defined vehicles. Existing traffic management strategies have significant drawbacks in public safety, predictive maintenance, tuning the core functionality of vehicles, and managing mobility. We can renovate this system with next-generation intelligent Digital Twin (DT) technologies. This research proposes a time-series prediction system through Digital Twins to manage the public transportation system with Facebook's Prophet. This study presents a model framework to build a Digital Twin application in Intelligent Public Transportation Systems and uses a public data set to validate the model with Facebook's Prophet library by forecasting metro line passenger flows. According to the results, the Mean Absolute Percentage Error (MAPE) is 0.017 for a 1-day horizon.

Keywords: Digital Twin, Intelligent Transportation Systems, IoT, Time Series Prediction

1. Introduction

The roads are growing with electric, intelligent, self-driving, and software-controlled cars daily. Classic traffic management strategies can lead to inefficiency and problems when managing technologically equipped intelligent connected vehicles and infrastructure. Several challenges need to be addressed for public transportation systems. These include managing mobility, proactively protecting against cyber security threats, ensuring efficient over-the-air (OTA) software updates, tuning vehicle functions, predictive maintenance for vehicles and traffic infrastructures, effectively managing vehicle malfunctions, planning disaster scenarios, and handling secure communication between vehicles, infrastructures, and pedestrians.

These issues can be addressed by renewing classic traffic and transportation management centers with next-generation intelligent Digital Twin technologies for public safety (Wang et al., 2022). A Digital Twin is a virtual representation of a physical system or process to perform simulation tests and artificial intelligence applications with real-time data. This study proposes a Digital Twin scheme to predict metro usage within rush hours in Istanbul. This framework enables the development of algorithms and policies to respond to events in real-time to overcome the disadvantages addressed. The presented model demonstrates a three-layer architecture consisting of data, virtualization, and service. In this architecture, an open data set is used to validate the proof-of-concept study to illustrate how DT can solve problems in Intelligent Transportation Systems (ITS) from a public transportation point of view.

The study's main contribution is to illustrate how the Digital Twin approach can be adopted by the public—transportation system to handle the planning of the metro infrastructures. In Section 2 related work about Digital Twins and ITS are discussed, in Section 3 the presented model and the framework is presented. Section 4 discusses the findings in this study and Section 5 concludes the paper.

2. Related Work

The importance of the Intelligent Transportation System has increased in recent years and is becoming an essential part of today's digital world. There have been numerous ITS studies, and all address different aspects of the system. However, Digital Twin has yet to be studied in detail yet. There is an open gap in the literature regarding a proper introduction to DT in ITS from a public transportation perspective. This section will discuss the literature studies on DT systems.

Micro mobility is essential in urban traffic; vehicles such as e-scooters, e-bicycles, and human-powered micro-vehicles have become vital when managing urban transportation. Scooter and bicycle sharing systems attract the city for quick and clean transportation (Oeschger et al., 2020).

Electricity generation is essential in ITS since the recent movement in electric vehicles and aspects of green energy generation with DT have a high potential. Monitoring, configuring, maintenance, and power supply of wind turbines are important in electricity production. Modernizing the management of wind turbines using new generation with DT can be part of the ITS for electricity supply needs. Designing energy generation systems through Digital Twins can be addressed to optimize this flow, increase production efficiency, and control turbines. A digital twin is a virtual version of a physical asset/process. Real-time data communication between the physical entity and the virtual copy is provided. DT approach is used in wind turbines to analyze, monitor, and predict power generation through deep-learning networks (Fahim et al., 2022).

DT applications are not limited to the energy sector; they can be applied in any field today, and this study aims to show how DT can be applied to public transportation systems. An overview of city transportation and its components is given in Figure 1. It comprises buses, trams, metros, emergency service vehicles, bicycles, and other automobiles. It is a challenging task to deal with such an extensive network.

Wang et al. (2022) introduce a digital twin framework to mode mobility with case studies. The authors present a digital win model that covers different aspects of mobility management with cloud and artificial intelligence (AI) based approaches. In the study, three physical systems, humans, vehicles, and traffic, are covered, and cloud-based architecture uses the Amazon Web Services (AWS) platform.

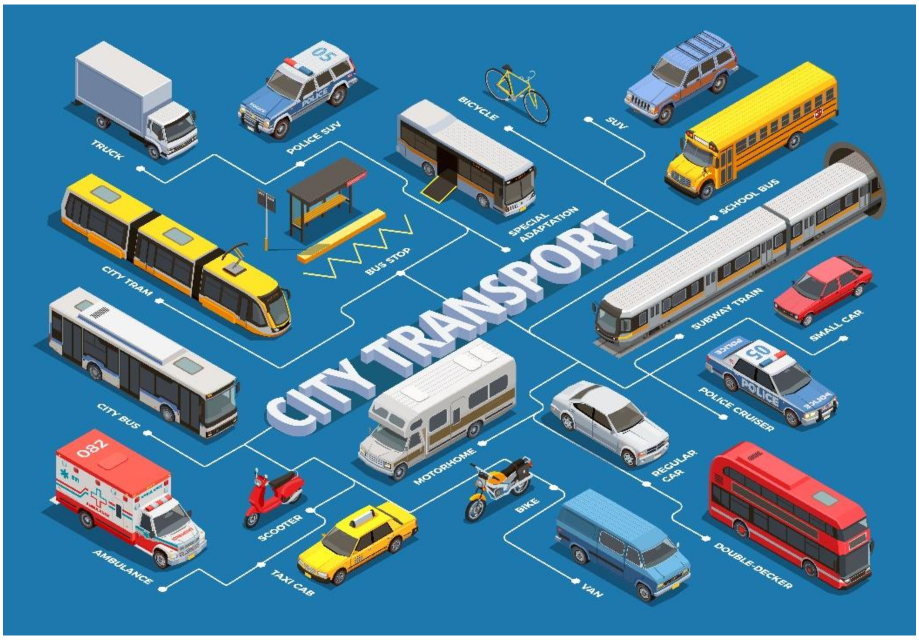

Figure 1. Components of a City Transportation

Digital Twins are used for sustainable development interms of efficient resource allocation, green technologies, and monitoring progress in sustainable development. DT can be used in different efficient energy allocation applications, from managing water supplies to urban planning. Big data and AI applications can model traffic demands and demand planning. (Tzachor et al., 2022).

Digital Twins can be used to for smart cities to model complex infrastructures using existing tools. Ivanov et al. (2020) uses traffic flow information, urban status with different sensors, outdoor surveillance cameras, and other data sources to present a conceptual view of Digital Twins for cities. By its nature, the integration of Digital Twin with Intelligent Transportation Systems will require a celebrator to allocate necessary resources depending on the severity level of the system. To address this problem, a Blockchain-based digital twin for smart cities is presented (Liao et al., 2022). DT is served as a service to meet different requirements of the different ITSs, and blockchain architecture solves the permission requirements of different ITSs.

The Internet of Things (IoT) and Digital Twin technologies rely heavily on data, particularly time series data that records events with timestamps. Time series data can be analyzed using statistical methods, machine learning approaches, or any data point that can be considered a signal. Signal processing techniques can be applied to analyze such data (Erturk & Vollero, 2020; Sabatini & Vollero, 2022).

Facebook's Prophet is a library designed to forecast time series data. It initially written with R language and adopted for Python to analyze and predict future events (Taylor Sj, 2017). The library can deal with seasonality in the time series data and is used in different fields, such as cloud resource predictions. (Daraghmeh et al., 2021), sales prediction in supermarkets (Jha & Pande, 2021) or high accuracy air temperature forecasting (Toharudin et al., 2023) .

Long Short Term Memory (LSTM) is a recurrent neural network (RNN) based deep learning method that has considerable performance in language and machine translation modeling. The LSTM layers take input from the previous layer, which enables accounting for short historical data on learning (Jalali et al., 2022). LSTM is a possible alternative to time-series forecasting. However, nonlinear behavior can increase resource usage even for simple tasks. For this reason, in this study, only Facebook's Prophet is considered.

In this study, Facebook's Prophet is used to present how a time-series prediction can be applied to the digital twin system by forecasting the number of passengers on public transportation within rush hours.

3. Digital Twin Model

This study uses a Digital Twin system architecture to address the problems. The digital twin technology is a virtual representation of a physical system, entity, or process. Digital twin systems benefit real-time (or almost real-time) communication with two-way data transfer. Data received from a physical twin can be represented virtually in a real-time manner, and the virtual twin is fed with real-time data. This enables the simulation and analysis of physical entities in a virtual world that cannot be achieved via classic simulation systems. An overview of the digital twin system is given in Figure 1. A digital twin system has several components and layers, in this study a three-layer approach is selected. Each layer is discussed in the following section.

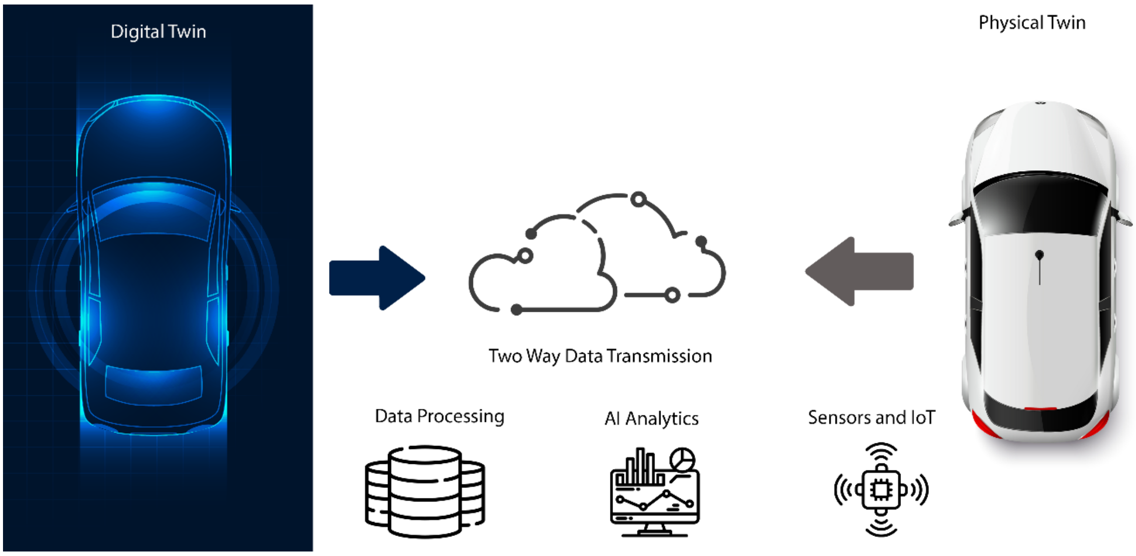

Figure 2. An overview of a Digital Twin System

3.1. Data Layer

A digital twin system needs real-time data to effectively model a physical entity. As shown in Figure 2, a vehicle infrastructure can be represented in a digital world virtually by modelling its data and behavior. The data layer is responsible for collecting data from various sources such as Internet of Things (IoT) external sensors, data APIs (i.e. weather) and vehicle data logs that has been processed through the operations. All the information flow must be wisely managed and handled.

The first layer is the data, which handles data collection from different physical assets and preprocessing to deliver other layers (Barricelli et al., 2019). Software-defined and autonomous vehicles employ various sensors and network connections to collect large amounts of data at different severity levels. These sensor data are vital since autonomous actions depend on environmental states and perceptions through these sensors. Therefore, the proposed data layer will solve data collection issues to achieve real-time data without losing information. Urban traffic consists of vehicles, traffic signs, infrastructures, and pedestrians, and handling a large data flow for decision-making is challenging. The model will be generic; however, only public transportation scenarios are considered for validation.

The design of Intelligent Transportation System (ITS) focused architecture benefits from Artificial Intelligence applications. Representing vehicle infrastructure with a generic data model enables data storing and processing in machine learning tasks more efficiently. In this study, time-series data is stored on a graph-based database system. Also, complex data and instant data access can be handled with graph query languages (GQL). Using adequate graph stores also reduces unnecessary repeated tasks by self-optimized data governance. In addition, the data aggregation will include other data sources, such as weather or environmental data or API calls from different systems, such as Vehicle-to-Everything (V2X), IoT sensors, and third-party apps.

For validation purposes a public transportation data source is selected to show as proof of concept with open-source tools.

3.2. Virtualization Layer

The second layer of the presented framework is the virtualization layer. In this layer, the physical entities are virtually represented to model in digital space. This design represents the public transportation system as a virtual entity using a custom Python application. It is possible to represent the state of the transportation line and its state on a dashboard. Also, the model can be extended for further virtualization needs.

3.3. Service Layer

The last and the third layer of the proposed system is the service layer, which has a traffic control dashboard with an artificial intelligence application to monitor and manage traffic. The service layer will orchestrate the system and running processes from a central user interface. This layer monitors ITS vehicles in a selected region. The service layer can monitor each vehicle's status and the traffic network's overall status; maintenance prediction on trains and what-if simulation analysis can be performed through the control dashboard.

To validate presented model, this study uses Facebook's Prophet library to predict metro lines in Istanbul. The service layer can be extended to use additional libraries or algorithms.

4. Experiments

A digital twin system is developed using python[1] to test modeled framework. The designed system uses data set from Istanbul Metropolitan Municipality Open Data Portal[2] to model public transportation as Digital Twins.

The DT models a physical entity on virtual space, for this purpose "ÇEKMEKÖY-ÜSKÜDAR" metro line is selected to model on virtual space. A Python client is developed to simulate physical metro behavior using the data set. It sends metro passenger pass information to DT's data layer. Through this, the virtualization layer represents the virtual state of the Metro Line, and time forecasting is implemented on the service layer for better management of the transportation lines.

4.1. The Data Set

The Data Set is the Hourly Public Transport Data Set of Istanbul City which is obtained from IMM Open Data Portal[3] and it contains all transportation data from January 2023 to October 2023. The description of the data set is given in Table 1.

Attribute	Type	Description
transition_date	Text	Holds transition date of the transportation vehicle.
transition_hour	Text	Hour value of the transition.
transport_type_id	Text	1: Highway 2: Rail 3: Marine Transportation.
road_type	Text	Transpiration area
transfer_type	Text	Passenger transition type: Transfer or Normal.
number_of_passage	Text	Count of number of trips per hour
number_of_passenger	Text	Individual passenger using transportation vehicle given hour

Table 1. Data Description

[1] Python programming lagnague, https://www.python.org/
[2] IMM Open Data Portal, https://data.ibb.gov.tr/en/
[3] Hourly Public Transport Data Set, https://data.ibb.gov.tr/tr/dataset/hourly-public-transport-data-set

This study only uses "USKUDAR-CEKMEKOY" metro line, containing 47.418 records. Only May to September 2023 is selected, and holidays and weekend data are removed. Hourly usage of the metro line in rush hours (7-9 a.m.) is given in Figure 3.

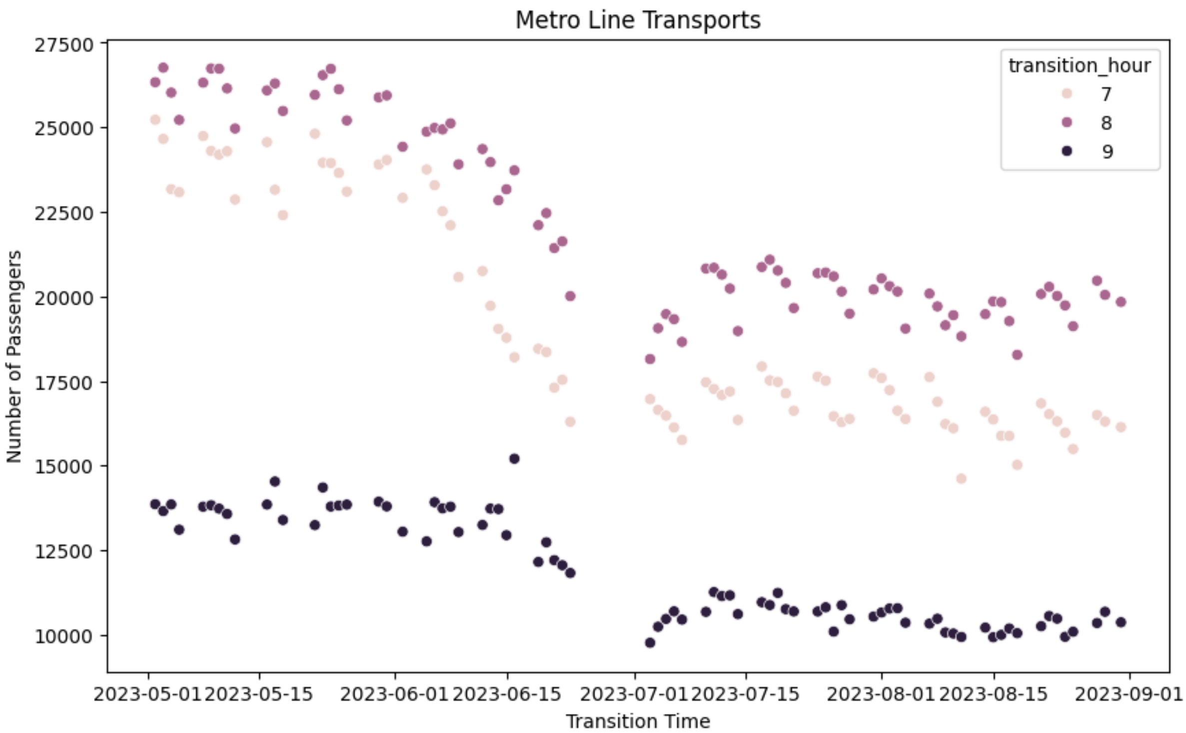

Figure 3. Metro (USKUDAR-CEKMEKOY) rush hours

4.2. Time-series Forecast of Metro Usage

The brief information about metro line is given previous subsection, by using Facebook's Prophet library, the metro usage for passenger is forecast by using past data. The dataset utilized in this experiment consists of passenger flow data gathered from January 2023 to October 2023 on the selected metro line. To validate the accuracy of the data, the dataset is split on 1st August 2023 to forecast daily data points, which are then compared with the actual data.

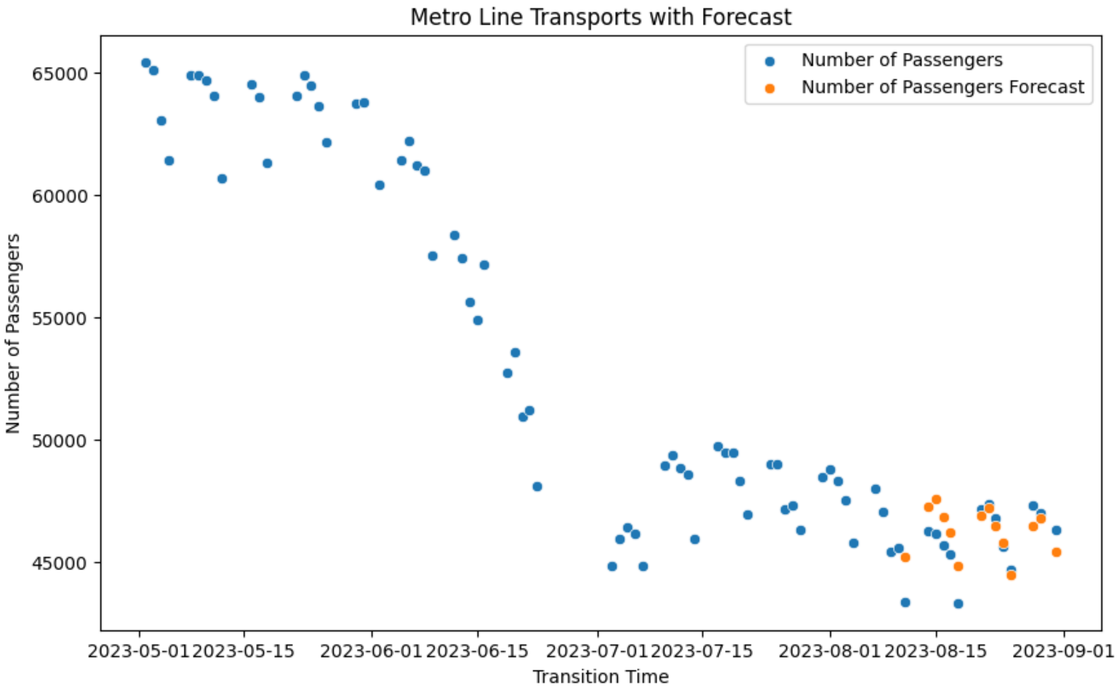

Figure 4. Metro (USKUDAR-CEKMEKOY) daily forecast

For performance measure MAPE (Mean Absolute Percentage Error) is used, and results shared on Table 2. According to the obtained results, it is shown that the forecasting of Facebook Prophet has an error of 0.017 for a 1-day horizon, which has excellent potential for predicting passenger flows in railway transits.

Horizon	MAPE
1 days	0.017
2 days	0.016
3 days	0.019
4 days	0.030
7 days	0.029
8 days	0.063
9 days	0.090
10 days	0.072

Table 2. Performance evaluation

5. Conclusions and Future Work

The number of electric, intelligent, self-driving, and software-controlled cars on the roads is increasing daily. This poses a challenge for traditional traffic management methods that may need to be more efficient and effective when dealing with technologically advanced intelligent connected vehicles and infrastructure. To address this issue, there is a need to find better ways to manage public transportation systems. A study has presented a forecasting framework using a Digital Twin system for public transportation. The system uses Facebook's Prophet library to forecast passenger flows in metro lines, which can help plan mobility and transportation infrastructures. The results show that the forecasting algorithm is usable for predicting passenger flows in public transportation. As a future work, it is planned to extend the presented digital twin approach for different forecasting algorithms and compare and evaluate the results of different forecasting algorithms besides Facebook Prophets.

References

Daraghmeh, M., Agarwal, A., Manzano, R., & Zaman, M. (2021, 14-23 June 2021). Time Series Forecasting using Facebook Prophet for Cloud Resource Management. 2021 IEEE International Conference on Communications Workshops (ICC Workshops),

Erturk, M. A., & Vollero, L. (2020, 13-17 July 2020). GSP for Virtual Sensors in eHealth Applications. 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC),

Fahim, M., Sharma, V., Cao, T. V., Canberk, B., & Duong, T. Q. (2022). Machine Learning-Based Digital Twin for Predictive Modeling in Wind Turbines. IEEE Access, 10, 14184-14194. https://doi.org/10.1109/ACCESS.2022.3147602

Ivanov, S., Nikolskaya, K., Radchenko, G., Sokolinsky, L., & Zymbler, M. (2020, 17-19 Nov. 2020). Digital Twin of City: Concept Overview. 2020 Global Smart Industry Conference (GloSIC),

Jalali, S. M. J., Ahmadian, S., Kavousi-Fard, A., Khosravi, A., & Nahavandi, S. (2022). Automated Deep CNN-LSTM Architecture Design for Solar Irradiance Forecasting. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52(1), 54-65. https://doi.org/10.1109/TSMC.2021.3093519

Jha, B. K., & Pande, S. (2021, 8-10 April 2021). Time Series Forecasting Model for Supermarket Sales using FB-Prophet. 2021 5th International Conference on Computing Methodologies and Communication (ICCMC),

Liao, S., Wu, J., Bashir, A. K., Yang, W., Li, J., & Tariq, U. (2022). Digital Twin Consensus for Blockchain-Enabled Intelligent Transportation Systems in Smart Cities. IEEE Transactions on Intelligent Transportation Systems, 23(11), 22619-22629. https://doi.org/10.1109/TITS.2021.3134002

Oeschger, G., Carroll, P., & Caulfield, B. (2020). Micromobility and public transport integration: The current state of knowledge. Transportation Research Part D: Transport and Environment, 89, 102628. https://doi.org/https://doi.org/10.1016/j.trd.2020.102628

Sabatini, A., & Vollero, L. (2022, 27 June-1 July 2022). Graph Signal Processing for IoT Sensor Networks. 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC),

Taylor Sj, L. B. (2017). Forecasting at scale. PeerJ Preprints 5:e3190v2. https://doi.org/10.7287/peerj.preprints.3190v2

Toharudin, T., Pontoh, R. S., Caraka, R. E., Zahroh, S., Lee, Y., & Chen, R. C. (2023). Employing long short-term memory and Facebook prophet model in air temperature forecasting. Communications in Statistics - Simulation and Computation, 52(2), 279-290. https://doi.org/10.1080/03610918.2020.1854302

Tzachor, A., Sabri, S., Richards, C. E., Rajabifard, A., & Acuto, M. (2022). Potential and limitations of digital twins to achieve the Sustainable Development Goals. Nature Sustainability, 5(10), 822-829. https://doi.org/10.1038/s41893-022-00923-7

Wang, Z., Gupta, R., Han, K., Wang, H., Ganlath, A., Ammar, N., & Tiwari, P. (2022). Mobility Digital Twin: Concept, Architecture, Case Study, and Future Challenges. IEEE Internet of Things Journal, 9(18), 17452-17467. https://doi.org/10.1109/JIOT.2022.3156028