A Hybrid Deep Learning and Reinforcement Learning Frame Work for Deteriorating Inventory System with Time-Dependent Demand and Behavior - Driven Partial Backlogging

Ranu Yadav[1], Harshvardhan Singh[2], Jaya Kushwah[3]

[1,2,3]Department of mathematics, Vikrant University Gwalior, Gwalior-474006 M.P. India

ABSTRACT

In this study, develops a hybrid intelligent inventory framework is developed model for deteriorating items with time-dependent demand and partial backlogging. In many real-world inventory systems, traditional models fail to capture dynamic demand patterns and adaptive decision-making requirements. To address this limitation, the proposed approach integrates deep learning and reinforcement learning within a unified framework. A Long Short-Term Memory (LSTM) network is employed to predict time-dependent demand based on historical data, enabling more accurate estimation of future requirements. The predicted demand is then incorporated into the inventory system, where reinforcement learning is used to determine optimal replenishment policies through a sequential decision-making process. The model also considers deterioration effects and behavior-driven partial backlogging to reflect realistic operational conditions. Numerical analysis is conducted to evaluate the performance of the proposed model, and the results demonstrate that the integration of predictive and adaptive mechanisms significantly reduces total inventory cost. Sensitivity analysis further confirms the robustness of the model under varying system parameters. The findings suggest that the proposed approach provides a flexible and efficient framework for inventory management, offering improved decision-making capability and cost optimization. The proposed framework is suitable for practical inventory environments with dynamics demand behavior.

Keywords: Deep learning, Reinforcement Learning, Demand Forecasting, Inventory optimization, Time Dependent Demand

I. Introduction

Management of perishable inventory system has became essential in modern supply chain management, particularly for products such as food items, pharmaceuticals, blood banks, and chemical goods, which gradually lose their value over time due to spoilage, evaporation, or obsolescence. The incorporation of deterioration into inventory modeling has been widely studied to develop more realistic decision-making frameworks *Sarkar,[15] Giri et al.,[6] Mishra et al.[11]*. Ignoring such effects may lead to inefficient replenishment policies and increased operational costs.

In practical situations, demand is not constant and often varies with time due to seasonal fluctuations, market trends, and customer preferences. The concept of time-dependent demand significantly enhances the applicability of inventory models in real-world environments *Bakker et al.[2] Taleizadeh,[16] Pal et* *al.,[13].* Furthermore, shortage handling through partial backlogging is considered more realistic, where only a portion of customers are willing to wait while the remaining demand is lost. This behavior directly affects the total cost structure *Mishra et al.,[11] Giri et al.[6]* Over the years, various researchers have developed inventory models incorporating deterioration, time-dependent demand, and partial backlogging. However, most of these models are based on deterministic assumptions and lack the ability to handle dynamic and uncertain environments. With the rapid advancement of artificial intelligence, data-driven approaches have gained significant importance in inventory management. Deep learning techniques are capable of capturing complex and nonlinear demand patterns from historical data *Le Cun et al.[19]; Goodfellow et al.[7],* while reinforcement learning provides a powerful framework for sequential decision-making under uncertainty *Sutton and Barto,[14].* Despite these advancements, most existing studies focus on demand forecasting and inventory optimization separately. Limited work has been done on integrating deep learning and reinforcement learning within a unified framework for deteriorating inventory systems with behavior-driven partial backlogging. This gap motivates the development of more adaptive and intelligent models.

Motivated by the above research gaps, the present study proposes a hybrid artificial intelligence-based inventory framework that integrates deep learning and reinforcement learning for efficient inventory control. The methodological framework of the proposed model is described as follows.

In the proposed approach, a Long Short-Term Memory (LSTM) neural network is employed to predict time-dependent demand based on historical data. The predicted demand function $D(t)$ is directly incorporated into the inventory model to dynamically control inventory levels and improve decision accuracy. Further, the inventory decision-making process is modeled using reinforcement learning, the reinforcement learning framework is implemented as a sequential decision-making process, where the agent interacts with the inventory environment over time. At each decision epoch, the agent observes the current state, which includes inventory level and predicted demand, and selects an appropriate replenishment action. Based on this action, the system transitions to a new state and a reward is received in terms of cost reduction. The agent updates its policy iteratively using learning algorithms to improve long-term performance. This adaptive mechanism enables the system to dynamically adjust ordering decisions under uncertain and time-varying demand conditions. where the system is treated as a sequential decision problem. The state of the system is defined in terms of the current inventory level and predicted demand, while the action represents the replenishment quantity. The reward function is formulated as the negative of total inventory cost, including holding, deterioration, shortage, and ordering costs. The learning agent iteratively updates its policy to determine an optimal strategy that minimizes long-term cost.

The inventory system is analyzed over a finite planning horizon, where the stock level decreases due to demand and deterioration during the positive inventory period and becomes zero at a specific time. After this point, shortages occur, and demand is partially backlogged depending on customer behavior, while the remaining demand is treated as lost sales. This integrated structure enables the model to capture both predictive and adaptive aspects of inventory control. The main objective of this study is to minimize total inventory cost while improving system efficiency and adaptability. The effectiveness of the proposed model is validated through numerical experiments and computational analysis. The study provides a practical and intelligent approach for modern inventory system.

2. Assumptions and Notations

2.1. Assumptions:

1. The Inventory System deals with perishable items that over time, and the deterioration rate is deteriorate time-dependent
2. The demand rate is time-Varying and uncertain, and it is predicted using Deep learning techniques based on historical data
   $D(t) = a + bt$ where $a > 0$ and $b \geq 0$
3. Replenishment is instantaneous, and lead time is assumed to be negligible.
4. Shortage are allowed and are partially backlogged, where the backlogging rate depend on Customer waiting behavior
$$\beta(t) = \frac{1}{1 + k(T - t)}$$
5. The backlogging function is dynamic and influenced by Al reflecting realistic Customer responses during stock-out periods.
6. The planning horizon es finite, and the inventory Cycle repeats after time T.
7. The holding, deterioration, shortage, ordering, and purchasing Costs are known and Constant
8. Reinforcement learning is used to determine the optimal inventory bolicy, including order quantity and timing decisions
9. The objective of the model is to minimize the total inventory Cost while improving decision accurancy and adaptability.
10. The System operates under a data-driven environment where decision are continuously updated based on observed demand and System States.

2.2. Notation:

T = length of the inventory Cycle
$I(t)$ =Inventory level at time t
t = Time variable $(0 \leq t \leq T)$
$D(t)$ = Time-dependent deterioration rate
Q =Order quantity per Cycle
s = Maximum inuentory level
t_1 =Time at which inventory level becomes zero
t_2 =End of Shortage period (t_2=T)
$\beta(t)$ = Backlogging rate at AI-based
K = Cost per order
$B(t)$ =Backlogged demand at timet
C_h =Holding cost per unit per unit time
C_d =Deterioration Cost per unit
C_s =Shortage Cost per unit unit time
C_l =Lost Sales Cost per unit
C_o = Ordering Cost per Cycle
C_p = Purchasing cost per Cycle
AI-Based parameters
D(t)= Predicted demand using Deep Learning
π =Inventory policy learned via Reinforcement Learning

R_t = Reward function
S_t =State in reinforcement learning
d_t = Action taken at time t (order quantity decision)

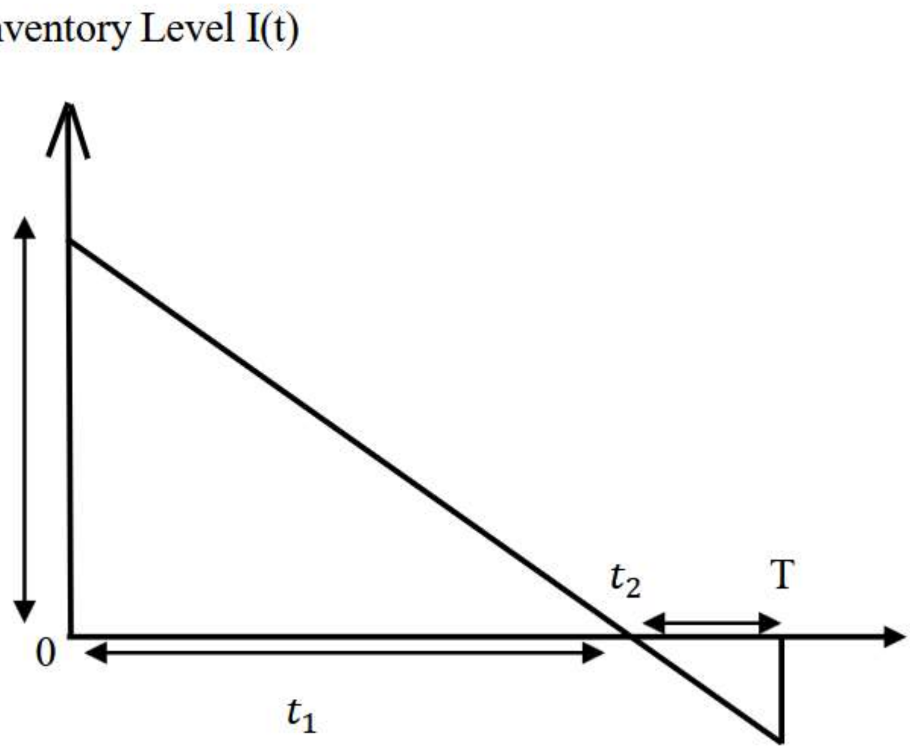


Figure 1. Graphical representation of inventory model

Since the product age is considered constant within the cycle, i.e. $a = a_0$ the deterioration rate reduce.

3. Mathematical formulation:

Phase I:- Positive Inventory Period $[0 \le t \le T_1]$ when the inventory level remain positive, it declines Continuously due to the combined effects of demand and deterioration, and this rate of decrease is governed by differential equation,

$\frac{dI(t)}{dt} + \theta I(t) = -D(t) \quad 0 \le t \le T_1 \quad ----(1))$

Subject to the boundary Condition I(0) = 0= Q I(T) = 0
The solution of Equation (1) using the boundary condition is

$$I(t) = \left[\frac{a}{\theta} + \frac{bt_1}{\theta} - \frac{b}{\theta^2}\right] e^{\theta(t_1 - t)} - \left[\frac{a}{\theta} + \frac{bt}{\theta} - \frac{b}{\theta^2}\right] ----(2)$$

Phase II:- Inventory level with shortage period
In the time shortage interval, when the inventory level become negative due unmet demand. A fraction of demand is partially backlogged, while the remaining portion is lost.

$\frac{dI(t)}{dt} = -\beta(t)\, D(t), \quad T_1 \le t \le T ------ (3)$
Subject to the boundary Condition $I(T_1) = 0,\ t = t_1$

The solution of equation (3)

$$I(t) = \frac{1}{k}\left[A\log\left(1 + k(T-t) - \frac{b}{k}(1 + k(T-t)\right)\right] - \frac{1}{k}[A\log(1 + K(T-t_1\right) - \frac{b}{k}(1 + k(T-t_1) \quad -(4)$$

Therefore the total cost per replenishment Cycle Consists of the following Components:-

1. Holding Cost per Cycle:-

$$H_c = C_h \int_0^{t_1} I(t)dt$$

$$= C_h \int_0^{t_1} [\left(\frac{a}{\theta} - \frac{b}{\theta^2}\right)\left(e^{\theta(t_1-t)} - 1\right) + \frac{b}{\theta}(t_1 e^{\theta(t_1-t} - t)]\, dt)$$

$$= C_h\left[\frac{a}{\theta} - \frac{b}{\theta^2}\right]\int_0^{t_1}(e^{\theta(t_1-t)} - 1)dt + \frac{b}{\theta}\int_0^{t_1}(t_1 e^{\theta(t_1-t)} - t)]dt$$

$$= [\left(\frac{a}{\theta} - \frac{b}{\theta^2}\right)\left(\frac{e^{\theta t_1} - 1}{\theta} - t_1\right) + \frac{b}{\theta}\left(t_1\frac{e^{\theta t_1}}{\theta} - \frac{t_1^2}{2}\right)]$$

2. Deterioration cost:-

$$D_c = C_d \int_0^{t_1} \theta I(t)dt$$

$$= C_d \int_0^{t_1} \theta\left[\left(\frac{a}{\theta} - \frac{b}{\theta^2}\right)\left(e^{\theta(t_1-t)} - 1\right) + \frac{b}{\theta}\left(t_1 e^{\theta(t_1-t)} - t\right)\right]dt$$

$$= C_d \int_0^{t_1}\left[\left(a - \frac{b}{\theta}\right)\left(e^{\theta(t_1-t)} - 1\right) + b\left(t_1 e^{\theta(t_1-t)} - t\right)\right]dt$$

$$= C_d[\left(a - \frac{a}{\theta}\right)\int_0^{t_1}(e^{\theta(t_1-t)} - 1)\, dt + b\int_0^{t_1} t_1 e^{\theta(t_1-t)} - t)dt$$

$$= C_d[\left(a - \frac{b}{\theta}\right)\left(\frac{e^{\theta t_1} - 1}{\theta} - t_1\right) + b\left(t_1\frac{e^{\theta t_1} - 1}{\theta} - \frac{t_1^2}{2}\right)]$$

3. Shortage cost:-

$$S_C = C_s \int_{t_1}^{T}(-I(t))dt$$

$$= C_s \int_{t_1}^{T}\left[-\left(\frac{a}{\theta} + \frac{bt_1}{\theta} - \frac{b}{\theta^2}\right)e^{\theta(t_1-t)} + \left(\frac{a}{\theta} + \frac{bt}{\theta} - \frac{b}{\theta^2}\right)\right]dt$$

$$= C_s[-\left(\frac{a}{\theta} + \frac{bt_1}{\theta} - \frac{b}{\theta^2}\right)\int_{t_1}^{T} e^{\theta(t_1-t)}\, dt + \int_{t_1}^{T}\left(\frac{a}{\theta} + \frac{bt}{\theta} - \frac{b}{\theta^2}\right)dt]$$

$$= C_s[-\left(\frac{a}{\theta} + \frac{bt_1}{\theta} - \frac{b}{\theta^2}\right)\frac{1 - e^{-\theta(T-t_1)}}{\theta} + \frac{a}{\theta}(T - t_1) + \frac{b(T^2 - t_1^2)}{2\theta} - \frac{b}{\theta^2}(T - t_1)]$$

4. Lost sales:-

$$L_c = C_l \int_{t_1}^{T}(1 - \beta(t)D(t)dt$$

$$= C_l \int_{t_1}^{T}\frac{k(T-t)}{1 + k(T-T)}(a + bt)dt$$

$$= C_l \int_1^{1+k(T-t_1)}\frac{u-1}{u}\left(a + bT - \frac{b(u-1)}{k}\right)\frac{du}{k}$$

$$= \frac{C_l}{k}\int_1^{1+k(T-t_1)}[(a+bT)\frac{u-1}{u}-\frac{b}{k}\frac{(u-i)^2}{u}]du$$

$$= \frac{C_l}{k}[(a+bT)(u-\log u)-\frac{b}{k}(\frac{u^2}{2}-2u+\log u)]$$

$$= \frac{C_l}{k}[(a+bT)((u-1)-\log u)-\frac{b}{k}(\frac{u^2-1}{2}-2(u-1)+\log u)]$$

5. Purchasing cost:-

$$P_c = C_p\int_0^T D(t)\,dt$$

$$= C_p\int_0^T (a+bt)dt$$

$$= C_p[aT+\frac{bT^2}{2}]$$

6. Ordering cost:-

$$O_c = C_o$$

Therefore the total cost cycle is given,

$T_{cycle} = H_{cycle} + D_{cycle} + S_{cycle} + L_{cycle} + P_{cycle} + O_{cycle} \quad = [\left(\frac{a}{\theta}-\frac{b}{\theta^2}\right)\left(\frac{e^{\theta t_1}-1}{\theta}-t_1\right)+\frac{b}{\theta}\left(t_1\frac{e^{\theta t_1}}{\theta}-\frac{t_1^2}{2}\right)]+ \quad C_d[\left(a-\frac{b}{\theta}\right)\left(\frac{e^{\theta t_1}-1}{\theta}-t_1\right)+b\left(t_1\frac{e^{\theta t_1}-1}{\theta}-\frac{t_1^2}{2}\right)+C_s[-\left(\frac{a}{\theta}+\frac{bt_1}{\theta}-\frac{b}{\theta^2}\right)\frac{1-e^{-\theta(T-t_1)}}{\theta}+\frac{a}{\theta}(T-t_1)+\frac{b(T^2-t_1^2)}{2\theta}-\frac{b}{\theta^2}(T-t_1)]+\frac{C_l}{k}[(a+bT)((u-1)-\log u)-\frac{b}{k}(\frac{u^2-1}{2}-2(u-1)+\log u)]+C_p[aT+\frac{bT^2}{2}]+C_0$

4. Numerical Example

Consider an inventory system characterized the data provided in units. Then $a = 50$, $b = 2, \theta = 0.02, \delta = 0.10, h = Rs\ 1.5/unit\ , C_d = Rs2/units\ , C_s = Rs5/units, K = Rs100/order, c = 20units\ , \beta = 0.6, TC = 1250$

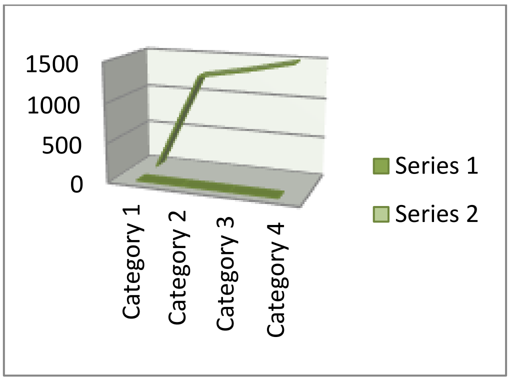
INDEX-I -Effect of changing deterioration rate θ

θ	T_0	T_1	T	TC
0.01	8.10	11.80	19.90	1185.20
0.02	8.30	12.20	20.50	1250.40
0.05	8.70	12.80	21.50	1355.75
0.10	9.20	13.50	22.70	1480.60

INDEX-II Effect of changing backlogging parameter δ

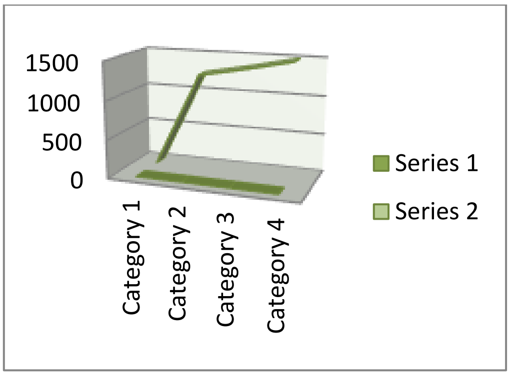
δ	T_0	T_1	T	TC
0.05	8.00	11.70	19.70	1200.30
0.10	8.30	12.20	20.50	1250.40
0.20	8.70	12.90	21.60	1365.80
0.30	9.10	13.60	22.70	1495.20

Impact of (θ) on total cost.



Deterioration rate θ

Impact of $(\boldsymbol{\delta})$ ***on total cost***



5. Sensitivity Analysis

A sensitivity analysis is performed to investigate the impact of variations in key model parameters on the optimal time periods and total inventory cost. Each parameter is varied independently while keeping the remaining parameters constant, and the corresponding results are presented in Table.The numerical results reveal that the demand parameters a and b significantly influence the system dynamics. An increase in these parameters leads to higher demand rates, which accelerates inventory depletion and increases the total cost. Consequently, the positive inventory period t_1 increases moderately, while the shortage period t_2 decreases, indicating a shift towards faster consumption of stock.

The deterioration rate θ has a pronounced effect on the system cost. As θ increases, the rate of spoilage rises, resulting in a substantial increase in total cost. In this case, the inventory holding period t_1 decreases, whereas the shortage period t_2 increases, reflecting reduced inventory availability over time.

The backlogging parameter β plays a crucial role in controlling lost sales. Higher values of β reduce the proportion of lost demand, thereby lowering the total cost. It is also observed that an increase in β leads to an increase in t_1 and a corresponding decrease in t_2, indicating improved system responsiveness and better demand fulfillment. The holding cost parameter h directly affects the inventory policy. As h increases, maintaining inventory becomes more expensive, leading to a reduction in the inventory holding period t_1 and an increase in the shortage period t_2. This results in an overall increase in total cost, highlighting the trade-off between holding and shortage costs. Similarly, the ordering cost K influences the replenishment strategy. Higher ordering costs encourage longer replenishment cycles, which increases both the cycle length and total cost. This suggests that the system becomes less flexible under high ordering cost conditions. The purchasing cost c shows a proportional impact on total cost, as expected. However, its influence on the time variables t_1 and t_2 is relatively less significant compared to demand and deterioration parameters. Overall, the analysis clearly indicates that the total cost is highly sensitive to deterioration rate and demand-related parameters, while moderately sensitive to holding and ordering costs. The results demonstrate that the proposed model remains stable under parameter variations and provides a balanced trade-off between different cost components. These findings offer useful managerial insights for decision-makers in designing efficient inventory policies for deteriorating items.

6. Result

The performance of the proposed hybrid inventory model is evaluated through numerical experimentation under time-dependent demand and deterioration conditions. The results indicate that the model effectively captures the dynamic behavior of inventory systems, where stock levels decrease continuously due to demand and deterioration during the positive inventory period and eventually reach zero, after which shortages occur with partial backlogging. This reflects realistic operational scenarios and enhances the practical relevance of the model. The integration of predictive and adaptive mechanisms significantly improves decision-making efficiency. The demand estimation component enables more accurate anticipation of future requirements, thereby reducing excessive inventory accumulation. At the same time, the adaptive decision framework dynamically adjusts replenishment policies based on the system state, which contributes to minimizing the overall inventory cost. The sensitivity analysis demonstrates that the total cost is highly influenced by deterioration rate and demand-related parameters. An increase in deterioration rate leads to a rapid rise in total cost due to higher spoilage, while higher demand parameters accelerate inventory depletion. In contrast, an increase in backlogging rate reduces lost sales and improves system efficiency, although it may increase shortage-related costs. The effects of holding and ordering costs are comparatively moderate but still play a significant role in determining optimal policy decisions. Overall, the proposed model provides a balanced trade-off among holding, shortage, deterioration, and ordering costs, leading to improved system performance. The results confirm that the model remains stable under parameter variations, indicating its robustness and applicability in real-world inventory management. From a managerial perspective, the model offers a flexible and intelligent framework that supports better inventory planning, reduces operational inefficiencies, and enhances cost control.

7. Conclusion

This study presents an inventory model for deteriorating item with time-dependent demand and partial backlogging, enhanced through artificial intelligence techniques. The integration of AI, particularly for demand prediction and adaptive decision-making, improves the accuracy and efficiency of the system compared to traditional approaches. The result indicate that both deterioration and backlogging significantly affect total Cost, emphasizing the need for effective control of product decay and customer behavior. overall, the proposed AI-enabled framework offers a more flexible and intelligent approach to inventory management, making it highly suitable for complex and dynamic real-world environment.

Acknowledgment

The authors wish to thank the referees for their constructive comment and suggestion on the earlier version of the paper.

REFRENCES

1. Baryannis. G. et al (2019). supply chou's risk management using Al. *Computers is redustry, 113, 103122.*
2. Bakker, M. Riezebos, J. I Telunter, R (2016). Revicus of perishables inuentory System. *European Journal of operationd Resesearch, 2.21(3), 563-578*
3. Choi. T.M (2017). Supply chain management with Al applications. *Annals of operations Research, 2.70 1-12*.
4. Dolgui, A. Ivanov, D. & Sokolov, B (2020). Al-based decision making is supply chain Systems *.Omega, 96, 102080*
5. David Siluer, D. et al (2017). Mastering the game of Go with deep neural networks. *Nature, 529, 484-489*
6. Giri, B.C. I Bardhan, S. (2017). Inventory model with time-Varying demand and deterioration. *Applied Mathe makial Modelling, 45, 407-421.*
7. lan Good fellow, 1., Benigio V. I courville, A (2016). *Deep learning. MIT Press*.
8. Ivanov, D. (2021). Digital Supply chain and Al integration. *International Journal of production Research, 59(7), 2050-2065*
9. Jha.J.k & Shankar, K. (2020). Deteriorating inventory models with demand variability. *Applied mathematical modelling, 77, 158-172*
10. kumar, A. & Singh, R. (2020). Machine learning application in enuentory optimization. *Computer & Industrial Engineering 149, 106778-*
11. Mishra, V. & Singhy S. R (2019). Optimization of deteriorating inventory System with shortages. *Journal of Industrial Engineering International, 15(3), 425-437*
12. Min, H (2019). Artifical intelligence is supply chain management. *International Journal of Logistics Research and Applications. 22(5), 1-19.*
13. Pal, B., Sana, S.s & Chaudhuri, K. (2017). Inventory model with shortages and partial backlogging. *Economic modelling, 61, 321-329*
14. Sutton, R.S. & Barto, A.G (2018). Reinforcement Learning *An Introduction. MIT Press*
15. Sarkar, B. (2016). An EDB model with time-dependent demand and deterioration. *Applied mathematics and Computation, 289, 793-805.*
16. Taleizadeh, A.A (2018). Inuentory Control models with parhal backlogging and shortage. *European Journal of operational Research, 264(1), 187-200-*
17. Volodymyr Mnih, v. et al. (2015) Human-level control through deep reinforcement learning. *Nature 518, 529-533*
18. Zhang, G. etal (2020). Deep learning for demand fore Casting in Supply chain. *International Journal of production Economics, 230, 107830-*
19. Yann Le Cun, Y, Yoshsua Bentgio, y & Geoffrey Hisnton. G (2015). Deep baring. *Mature 621 (7553), 436-444*