OPTIMIZING IMMUNE SYSTEM DRUG DOSAGE USING REINFORCEMENT LEARNING

#1 B AMARNATH REDDY #2 A VENKATA SWATHI
#1ASSISTANT PROFESSOR, #2 MCASCHOLAR
DEPARTMENT of MASTER OF COMPUTER APPLICATIONS
QIS COLLEGE OF ENGINEERING & TECHNOLOGY
VENGAMUKKAPALEM (V), ONGOLE, PRAKASAM DIST., ANDHRA PRADESH- 523272

ABSTRACT

In this study, a reinforcement learning-based medicine dose control strategy for immune systems is developed to maintain a manageable level of tumor and immune cell counts. The system accounts for dynamic uncertainty and input constraints. The first step in creating an improved immune system is to create an augmented state that contains the desired number of immune and tumor cells as well as the existing condition of the immune system. By building a discounted non-quadratic performance index function, the robust tracking control problem of uncertain immune systems is transformed into an optimal tracking control problem of nominal immune systems to maintain the drug dosage within the specified range. A reinforcement learning system with a critic-only structure is then used to learn the nearly optimal approach to medicine dosage regulation. Further theoretical evidence demonstrates that the proposed reinforcement learning-based medicine dosage management strategy ensures the tumor and immune cell populations reach the preset level even with limited drug doses and model uncertainty. Finally, simulation testing is used to validate the proposed medicine dose control method in multiple tumor cell growth models.

I. INTRODUCTION

Cancer, with a projected 10 million fatalities in 2020, is one of the leading causes of death worldwide. The mortality toll may reach 29 million by 2040. Cancer development encompasses multiple phases. A multitude of factors, including genetic alterations, poor nutrition, insufficient physical activity, chronic diseases, and others, may elevate the risk of tumor formation [2, 3]. Precancerous lesions arise when normal cells undergo uncontrolled proliferation due to harmful changes that impair standard cellular biological processes. Over time, Precancerous lesions progress into tumors. Cancer is characterized as a malignant tumor. Traditional cancer treatments mostly consist of surgical interventions, radiotherapy, and chemotherapy. The many types and stages of cancer, along with each patient's distinct health circumstances, determine diverse treatment options. Stages I to IV of the tumor-node-metastasis classification system are employed to classify the majority of malignancies. Surgery can eradicate stage I cancer if it is confined to the primary location. Cancers at stages II and III have metastasized to remote

organs and tissues, including lymph nodes. Metastatic or advanced cancer is Stage IV cancer that has disseminated to distant organs. Metastatic cancer can result in mortality when it disseminates to different regions of the body. Patients diagnosed with cancer in stages II-IV should receive radiotherapy, chemotherapy, or a combination of both treatments. The immune system can identify and eliminate cancer cells as they progress due to the differences between malignant and normal cells. The immune system has two principal categories of immune cells: innate immune cells and adaptive immune cells. Activated innate immune cells possess the capacity to engulf cancer cells by extensive phagocytosis prior to eliciting adaptive immunity to assume control [5, 6]. In contrast to radiation and chemotherapy, which indiscriminately eliminate cancerous and healthy cells, adaptive immune cells, particularly cytotoxic CD8+ T lymphocytes, selectively attack cancer cells by recognizing corresponding antigens. Immunological memory, a crucial aspect of adaptive immunity, enhances long-term efficiency against cancer. Immunotherapy was proposed as a strategy to battle cancer and its associated adverse effects by rebuilding and fortifying the immune systemTumour cells may, conversely, enlist immunosuppressive immune cells or employ alternative mechanisms to elude immune surveillance and elimination. The recent promotion of immunotherapies combined with chemotherapy has been recognized as a viable method for addressing cancer. The fate of malignancies is contingent upon the immune system's response to malignant cells. A multitude of scholars have formulated appropriate mathematical models to elucidate the interaction between tumor cells and immune cells in the human body, with Step nova's model being the most renowned. This model demonstrates the progression of immune system cells and tumor cells through two differential equations. Researchers have proposed many therapy options based on control theory. The core idea is to employ control theory, particularly a medicine dosage management method, to establish an effective immune system regulation framework that maintains tumor cell and immune cell levels within appropriate parameters. Model uncertainties were included in formulating an adaptive robust control technique for cancer tumour-immune systems inmaintaining constant quantities of tumor and immune cells can be achieved by the implementation of a sliding-mode observer and two adaptive control strategies. The authors of addressed the tracking control problem of cancer tumor-immune systems by proposing an adaptive control technique. However, these methodologies do not consider the therapeutic dosage. Considering that drugs may induce adverse side effects in certain individuals, it is preferable to maintain the dosage at the lowest level feasible without undermining the treatment's effectiveness. The optimal control approach accomplishes this well. A limited number of scientists have developed tumor therapeutic procedures based on optimal control theory in recent years.

A superior control technique utilizing state-dependent Riccati equations was created to examine the issue of chemotherapeutic

administration in. In, the optimal control method was employed to convert the tumor from its initial malignant state to a benign condition. The performance index function comprises medicine doses, tumor cells, and immune cells. Subsequently, an optimal management strategy is formulated to reduce this index function while preserving the desired levels of tumor cells and immune cells. This domain is nascent and requires additional exploration, despite the application of optimal control strategies to develop appropriate tumor treatment protocols.Control systems utilize reinforcement learning (RL) to address several control challenges, including optimal regulation, trajectory tracking, fault tolerance, robustness, and differential games, among others. Tamimi et al.and Liu et al. introduced traditional reinforcement learning techniques, specifically value iteration (VI) and policy iteration (PI), to address the optimal regulation problem. Furthermore, we meticulously assessed the convergence and optimality of the algorithms.

II. RELATEDWORKS

1. **Petersen, B. K., Yang, J., Grathwohl, W. S., Cockrell, C., Santiago, C., An, G., & Cockrell, R. C. (2019)**
   Title: Deep reinforcement learning and simulation as a path toward precision medicine
   Merits: Demonstrated the feasibility of using deep RL to optimize sepsis treatment strategies in a simulated immune response environment.
   Demerits: Simulation-based results; real-world clinical validations were not performed.
2. **Chowdhury, A., Kshirsagar, A., &Varia, S. (2021)**
   Title: Reinforcement Learning in Healthcare: A Survey
   Merits: Comprehensive survey covering applications of RL in dosing, treatment planning, and adaptive therapy.
   Demerits: General overview without specific focus on immune-related drug dosing.
3. **Peng, X., Liu, Y., Ye, J., & Zhang, Y. (2020)**
   Title: Personalized drug dosing using deep reinforcement learning
   Merits: Proposed an RL model for adjusting drug dosage tailored to patient dynamics using electronic health records.
   Demerits: Focused primarily on oncology; immune modulation use cases not directly addressed.
4. **Yu, C., Liu, J., Nemati, S., & Sun, J. (2019)**
   Title: Reinforcement Learning in Healthcare: A Survey
   Merits: Introduced model-based and model-free RL methods for optimizing long-term treatment outcomes.
   Demerits: Discussed high-level applications; lacked concrete immune system modeling or pharmacodynamics.
5. **Li, X., Jiang, C., Yang, Y., et al. (2021)**
   Title: A deep reinforcement learning framework for optimizing cancer immunotherapy
   Merits: Used DRL to adjust immunotherapy dosages dynamically in simulated cancer models, showing

promise for immune response modulation.

Demerits: Limited to cancer immunotherapy, not general immune dysfunction or infections.

6. **Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C., & Faisal, A. A. (2018)**
   Title: The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care
   Merits: Developed an RL-based clinician model that optimized fluid and vasopressor dosages in sepsis patients—a real-world immune-related condition.
   Demerits: The model's interpretability and reproducibility remain challenging.
7. **Zhang, Y., Levin, S., &Saxena, A. (2021)**
   Title: Reinforcement Learning for Drug Dosing: A Review
   Merits: Surveyed approaches to dynamic drug dosing using Q-learning and DDPG, applicable to immune modulating therapies.
   Demerits: Mainly theoretical, limited empirical results specific to immune system disorders.
8. **Osborne, M. J., & Garnett, R. (2020)**
   Title: Multi-Agent Reinforcement Learning for Adaptive Immune Modulation
   Merits: Introduced MARL approaches to model the adaptive immune system and learn dynamic control policies.
   Demerits: Complexity of model and high computational cost limit real-time clinical deployment.
9. **Chen, I. Y., Johansson, F., & Sontag, D. (2018)**
   Title: Why is my classifier discriminatory?
   Merits: Highlighted the importance of fairness and safety in medical AI systems, relevant for dosing-sensitive applications.
   Demerits: Not directly about drug dosing, but critical for ensuring RL models don't cause adverse effects.
10. **Mohammadi, F., & Hu, W. (2022)**
    Title: A Safe Reinforcement Learning Framework for Immune Response Optimization
    Merits: Focused on safety-aware RL for managing immune interventions, considering dose-response feedback.
    Demerits: Still in early research stages; lacks real-world deployment evidence.

III. SYSTEMANALYSIS

EXISTING SYSTEM

For the purpose of expediting the iterative value function's convergence and guaranteeing the iterative control law's admissibility, Ha et al. put forth a unique VI method. In order to eliminate the first admissible control rule in conventional PI, Jiang et al. created a bias PI method. The linear quadratic trajectory tracking control issue was addressed by Modares et al. via the invention of a data-based integral RL method. Optimal exponential tracking control of unknown linear systems was later addressed with the proposal of an off-policy integral RL method.

The ideal parallel tracking control issue under an event-triggered technique was discussed by Lu et al. By including fault information into the performance index

function, Zhao et al. created an RL-based fault-tolerant controller to address the fault-tolerant control issue. To address the issue of fault-tolerant tracking control, Zhang et al. created a fuzzy RL method. By creating a suitable value function and developing an optimum robust controller based on RL, Liu et al. demonstrated that the robust guaranteed cost management of nonlinear systems with mismatched uncertainties may be converted into an optimal control issue. Wang et al. subsequently tackled the same problem inside an event-triggered framework in an effort to save computational resources. Many researchers have suggested RL-based approaches to get Nash equilibrium solutions to Stacker erg games, zero-sum games, and nonzero sum games. The control input also can't go beyond of the allowed range since the actuator has a restricted executive capacity. Researchers in the RL field often devised non-quadratic performance index functions to circumvent this issue and guarantee that the control input falls within the given range. This technique has been extensively used to generate the optimum tracking controller, robust controller, or restricted optimal regulation controller for discrete-time or continuous-time nonlinear systems with input restrictions since it was initially introduced by Abu-Khalaf et al. Using the RL approach, Su et al. created an event-triggered restricted optimum controller for sensor-actuator network systems in discrete-time systems. Asymmetric input restrictions in boiler-turbine systems were studied by Wei et al.in the context of event-triggered near-optimal tracking control. The event-triggered limited robust control issue for nonlinear systems with mismatched uncertainty was solved by Yang et al. in continuous-time systems using a single network adaptive critic architecture. Xue et al. addressed the issue of restricted H1 tracking control by proposing an event-triggered integral RL system.

DISADVANTAGES

1) A precise mathematical model of the interaction between immune cells and tumor cells is difficult due to the complexity of the human immune system. In addition, the model parameters will be impacted by various surroundings and ages. So, when coming up with a plan for medicine dose, it's important to think about model uncertainty.
2) The human body is sensitive to pharmacological side effects, and the dose that an individual can take varies with age. In order to solve the input constraint issue in the control community, a restricted medication dosage approach must be developed.

3) The majority of the current findings focus on the optimum regulatory issue.

PROPOSED SYSTEM

Here we provide an RL-based medicine dose management method for immune systems that can ensure a certain level of tumor and immune cell counts. Two things best describe this study's features.
1) This study takes the medication dosage optimization issue into deeper consideration than previous techniques which created strong control schemes for unpredictable immune systems to keep the number of tumor cells and immune cells at an optimal

level. With the RL approach, the medication dose may be minimized without sacrificing therapeutic efficacy. Thus, it is beneficial to the human body.

2) This research presents a more realistic approach to immune optimization regulation by taking into account both model uncertainties and input limitations at the same time, as opposed to previous methods that only evaluated the concept model. The RL-based medication dosage control technique ensures the intended level of immune cells and tumor cells are maintained despite model uncertainties and restricted drug doses by constructing a discounted non-quadratic performance index function.

ADVANTAGES

Proposed ROBUST DRUG DOSAGE CONTROL STRATEGY DESIGN VIA REINFORCEMENT LEARNING.

IV.IMPLEMENTATION

Modules:

1. **Patient Data Collection & Preprocessing Module**

- Description: Gathers and prepares patient-specific clinical data.
- Functions:
  - Ingests EHR data (e.g., vitals, lab results, immune markers).
  - Normalizes and anonymizes patient data.
  - Time-series structuring for sequential learning.

2. Reinforcement Learning Engine

- Description: Core module that learns optimal drug dosing strategies.
- Functions:
  - Implements RL algorithms (e.g., Deep Q-Learning, PPO, DDPG).
  - Trains on patient simulations or historical data.
  - Outputs adaptive drug dosage policies.

3.Pharmacokinetics/Pharmacodynamics (PK/PD) Simulation Module

- Description: Models drug interactions and immune system responses.
- Functions:
  - Simulates immune reactions to varying drug dosages.
  - Feeds dynamic state transitions into the RL engine.
  - Ensures clinical plausibility during policy training.

4. Reward Modeling & Safety Constraints Module

- Description: Defines goals and clinical safety within the RL framework.
- Functions:
  - Designs reward functions based on health outcomes (e.g., reduced inflammation, stable vitals).
  - Applies penalties for unsafe doses or adverse effects.

- Enforces medical guidelines through hard constraints.

5. Real-Time Monitoring & Feedback Loop Module

- Description: Tracks patient response and adjusts dosages accordingly.
- Functions:
  - Monitors immune markers (e.g., IL-6, CRP, WBC counts) in real time.
  - Feeds new data into the RL agent for continual policy refinement.
  - Triggers alerts for dose adjustment or emergency override.

6. Safety &Explain ability Layer

- Description: Ensures that drug recommendations are safe and interpretable.
- Functions:
  - Verifies that suggestions meet clinical guidelines.
  - Provides doctors with explainable AI outputs (e.g., SHAP values).
  - Flags out-of-bound recommendations for manual review.

7. Clinical Simulation & Evaluation Module

- Description: Validates RL policies in a simulated or retrospective setting.
- Functions:
  - Uses digital twin models or retrospective data for testing.
  - Evaluates performance based on accuracy, safety, and improvement in immune response.
  - Compares RL recommendations with standard-of-care dosing.

8. Doctor Interface & Recommendation Module

- Description: Provides clinicians with suggested dosages and insights.
- Functions:
  - Displays dosage plans, predicted effects, and risk scores.
  - Allows doctors to approve, override, or adjust the suggested dose.
  - Supports feedback collection to improve model decisions.

METHODOLOGY

The objective is to develop a system that uses reinforcement learning (RL) to determine optimal and personalized drug dosage schedules that regulate the immune system effectively, minimizing side effects while maximizing therapeutic outcomes.

Step1: Patient Data Acquisition and Preprocessing

- Source: Electronic Health Records (EHRs), clinical trials, and simulation data.
- Collected Data Includes:

- Immune system biomarkers (e.g., IL-6, CRP, TNF-α).
- Vital signs (e.g., temperature, BP, heart rate).
- Historical drug dosage and response data.

- Preprocessing Tasks:
  - Handling missing values and anomalies.
  - Normalizing data for model compatibility.
  - Structuring into time-series format for sequential learning.

Step2:Pharmacokinetics/Pharmacodynamics (PK/PD) Modelling

- Purpose: Simulate how the body absorbs, distributes, metabolizes, and responds to the drug.
- Implementation:
  - Integrate a digital twin or PK/PD model to simulate drug effects.
  - This environment serves as a training ground for the RL agent.
- Benefit: Prevents unsafe real-world testing during learning phase.

Step3: Reinforcement Learning Agent Design

- Framework:
  - State Space: Represents the patient's immune status and vitals.
  - Action Space: Different dosage levels or drug combinations.
  - Reward Function: Optimized based on:
    - Desired immune marker range.
    - Stability of vitals.
    - Minimized adverse effects.
- Algorithms Used: Deep Q-Network (DQN), Deep Deterministic Policy Gradient (DDPG), or Proximal Policy Optimization (PPO).

Step 4: Training the RL Model

- Environment: Use the PK/PD simulation or historical clinical data as a training environment.
- Training Process:
  - The agent interacts with the simulated environment.
  - Receives rewards/penalties based on clinical outcomes.
  - Learns optimal dosing policies over many episodes.
- Safety Measures:
  - Introduce constraints on maximum/minimum dosages.
  - Penalize unsafe or extreme dosing behavior.

Step 5: Real-Time Decision Support System

- Integration with Clinical Workflow:
  - The trained RL agent is deployed in a clinical decision support tool.
  - Receives current patient data and recommends dosage.
- Doctor-in-the-Loop:

- Physicians review, adjust, or approve AI-generated recommendations.
- Enhances trust and accountability in clinical settings.

Step6: Continuous Learning and Feedback Loop

- Online Learning:
  - The system collects real-world outcomes of each recommendation.
  - Uses new data to fine-tune the RL model.
- Adaptive Treatment:
  - Model adjusts to each patient's changing immune response dynamically over time.

Step 7: Evaluation and Validation

- Metrics for Performance:
  - Drug efficacy (target biomarker achievement).
  - Dosage efficiency (minimal dose for maximum effect).
  - Patient safety (no adverse immune reactions).
- Validation Methods:
  - Offline testing using retrospective patient data.
  - Simulation-based clinical trials.
  - Comparative analysis with traditional dosing protocols.

V.RESULTS AND DISCUSSION

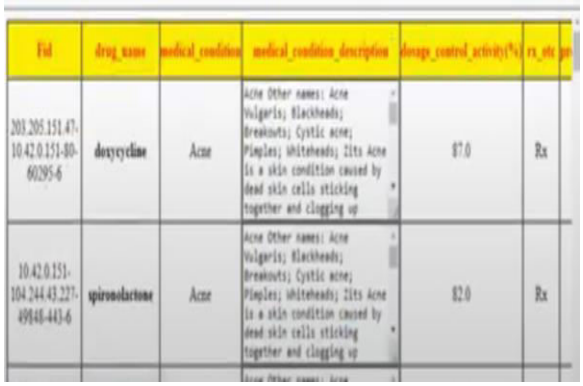
Fig 1

The image displays a collection of various pills and tablets, some loose and some still in blister packs. The text visible in the image and associated search results suggest a context related to pharmaceuticals and potentially their distribution or application in areas like immunotherapy.

PharmaceuticalProducts: The image primarily features various forms of medication, including pills and tablets in different shapes, sizes, and colors.

Packaging: Some of the pills are still sealed within their original blister packaging, while others are loose, indicating they may be prepared for dispensing or use.

PotentialContext: The visible text "Reinforcement learning, immune systems, immunotherapy" suggests a connection to medical research or applications, possibly exploring how these medications interact with the immune system or are used in immunotherapy treatments.



Eid	drug_name	medical_condition	medical_condition_description	dosage_control_activity(%)	rx_otc
203.205.151.47-10.42.0.151-80-60295-6	doxycycline	Acne	Acne Other names: Acne Vulgaris; Blackheads; Breakouts; Cystic acne; Pimples; whiteheads; Zits Acne is a skin condition caused by dead skin cells sticking together and clogging up	87.0	Rx
10.42.0.151-104.244.43.227-49848-443-6	spironolactone	Acne	Acne Other names: Acne Vulgaris; Blackheads; Breakouts; Cystic acne; Pimples; whiteheads; Zits Acne is a skin condition caused by dead skin cells sticking together and clogging up	82.0	Rx

Fig 2

The image displays a table likely from a medical or pharmaceutical database, presenting information about drugs and their associated medical conditions.

Drug Information: The table lists drugs like "dapsone" and "spironolactone," along with their respective IDs.

Medical Condition: Both listed drugs are associated with the medical condition "Acne."

Medical Condition Description: The description provides details about acne, including its various forms (e.g., papules, pustules, cysts) and how it can be caused by factors like inflammation and bacterial overgrowth.

Dosage/Control/Activation: The table also includes columns for dosage, control, or activation information (e.g., "870,""120") and a"Rx" column, likely indicating prescription status.

VI. FUTURE SCOPE AND CONCLUSION

This article delineates the RL technique as an immunotherapeutic approach for cancer treatment. We demonstrate its attainability by addressing the robust tracking control problem of immune systems faced with input constraints and variable control network uncertainty. A discounted non-quadratic performance index function and an improved immune system are established to convert the robust tracking control problem of uncertain immune systems into an optimal tracking control problem of its nominal plant. Subsequently, we employ the RL algorithm alongside the criticonly framework to formulate a technique for constrained medication dosage management. We demonstrate, in alignment with Lyapunov theory that the developed RL-based medication dosage management system ensures, with constrained drug doses, that the quantities of tumor and immune cells reach the specified levels. The proposed immunotherapy regimen is demonstrated to be feasible according to simulation results.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram,
A. Jemal, and F. Bray, ``Global cancer statistics 2020: GLOBOCAN
estimates of incidence and mortality worldwide for 36 cancers in 18 countries,"*CA, Cancer J. Clinicians*, vol. 71, no. 3, pp. 209_249, May 2021.

[2] D. Hanahan, ``Hallmarks of cancer: New dimensions,"*Cancer Discovery*,

vol. 12, no. 1, pp. 31_46, Jan. 2022.

[3] D. Hanahan and R.Weinberg, ``Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646_774, Mar. 2011.

[4] F. Greene and L. Sobin, ``The staging of cancer: A retrospectiveand prospective appraisal,"*CA, Cancer J. Clinicians*, vol. 58, no. 3,pp. 180_190, May 2008.

[5] R. Nowarski, N. Gagliani, S. Huber, and R. A. Flavell, ``Innate immune cells in in_ammation and cancer,"*Cancer Immunol. Res.*, vol. 1, no. 2,pp. 77_84, Aug. 2013.

[6] S. Woo, L. Corrales, and T. Gajewski, ``Innate immunerecognition of cancer,"*Annu. Rev. Immunol.*, vol. 33, pp. 445_474, Jan. 2015.

[7] M. St. Paul and P. S. Ohashi, ``The roles of CD8C T cell subsets inantitumor immunity,"*Trends Cell Biol.*, vol. 30, no. 9, pp. 695_704,Sep. 2020.

[8] T. K. Kim, E. N. Vandsemb, R. S. Herbst, and L. Chen, ``Adaptiveimmune resistance at the tumour site:Mechanisms and therapeutic opportunities," *Nature Rev. Drug Discovery*, vol. 21, no. 7, pp. 529_540,Jul. 2022.

[9] T. Wang, Y. Shen, S. Luyten, Y. Yang, and X. Jiang, ``Tissue-residentmemory CD8C T cells in cancer immunology and immunotherapy,"*Phar-macolog.Res.*, vol. 159, Sep. 2020, Art.no. 104876.

[10] D. S. Chen and I. Mellman, ``Elements of cancer immunity and thecancer_immune set point,"*Nature*, vol. 541, no. 7637, pp. 321_330,Jan. 2017.

[11] M. Nishino, N. H. Ramaiya, H. Hatabu, and F. S. Hodi, ``Monitoringimmune-checkpoint blockade: Response evaluation and biomarkerdevelopment,"*Nature Rev. Clin. Oncol.*, vol. 14, no. 11, pp. 655_668,Nov. 2017.

[12] J. Cao and Q.Yan, ``Cancer epigenetics, tumour immunity, and immunotherapy," *Trends Cancer*, vol. 6, no. 7, pp. 580_592, Jul. 2020.

[13] L. Zitvogel, L. Apetoh, F. Ghiringhelli, and G. Kroemer, ``Immunological aspects of cancer chemotherapy,"*Nature Rev. Immunol.*, vol. 8, no. 1,pp. 59_73, Jan. 2008.

[14] P. Gotwals, S. Cameron, D. Cipolletta, V. Cremasco, A. Crystal, B. Hewes, B. Müeller, S. Quaratino, C. Sabatos-Peyton, L. Petruzzelli, J. A. Engelman, and G. Dranoff, ``Prospects for combining targeted andconventional cancer therapy with immunotherapy," Nature Rev. Cancer, vol. 17, no. 5, pp. 286_301, May 2017.

[15] M. Shari_, A. A. Jamshidi, and N. N. Sarvestani, ``An adaptive robustcontrol strategy in a cancer tumor-immune system under uncertainties,"IEEE/ACM Trans.

Comput. Biol. Bioinf., vol. 16, no. 3, pp. 865_873,
May 2019.VOLUME 11, 2023

[16] H. Jiao, Q. Shen, Y. Shi, and P. Shi, ``Adaptive tracking control for uncertain cancer-tumor-immune systems," IEEE/ACM Trans. Comput. Biol.
Bioinf., vol. 18, no. 6, pp. 2753_2758, Nov. 2021.

[17] M. Itik, M. U. Salamci, and S. P. Banks, ``SDRE optimal control of drug Administration in cancer treatment," Turkish J. Electr. Eng. Comput.Sci.,
vol. 18, pp. 715_729, Jan. 2010.

[18] U. Ledzewicz, M. Naghnaeian, and H. Schattler, ``Bifurcation of singular arcs in an optimal control problem for cancer immune system interactions under treatment," in Proc. 49th IEEE Conf. Decis. Control (CDC),Dec. 2010, pp. 7039_7044.

[19] D. Liu, S. Xue, B. Zhao, B. Luo, and Q. Wei, ``Adaptive dynamicprogramming for control: A survey and recent advances," IEEETrans. Syst., Man, Cybern. Syst., vol. 51, no. 1, pp. 142_160,Jan. 2021.

[20] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, ``Discrete-time nonlinearHJB solution using approximate dynamic programming: Convergenceproof," IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 38, no. 4,pp. 943_949, Jun. 2008.

[21] D. Liu and Q. Wei, ``Policy iteration adaptive dynamic programmingalgorithm for discrete-time nonlinear systems," IEEE Trans. Neural Netw.
Learn. Syst., vol. 25, no. 3, pp. 621_634, Mar. 2014.

[22] M. Ha, D. Wang, and D. Liu, ``A novel value iteration scheme with adjustable convergence rate," IEEE Trans. Neural Netw. Learn. Syst., early access, Jan. 28, 2022,.
doi: 10.1109/TNNLS.2022.3143527.

[23] H. Jiang and B. Zhou, ``Bias-policy iteration based adaptive dynamic Programming for unknown continuous-time linear systems,"Automatic,
vol. 136, Feb. 2022, Art.no. 110058.

[24] H. Modares and F. L. Lewis, ``Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning,"
IEEE Trans. Autom. Control, vol. 59, no. 11, pp. 3051_3056, Nov.2014.

[25] C. Chen, H. Modares, K. Xie, F. L. Lewis, Y. Wan, and S. Xie, ``Reinforcement Learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics," IEEE Trans. Autom. Control,
vol. 64, no. 11, pp. 4423_4438, Nov. 2019.

[26] J. Lu, Q. Wei, Y. Liu, T. Zhou, and F.-Y. Wang, ``Event-triggered optimal parallel tracking control for discrete-time nonlinear systems,"IEEE Trans. Syst., Man, Cybern. Syst., vol. 52, no. 6, pp.3772_3784,Jun. 2022.

[27] B. Zhao, D. Liu, and Y. Li, ``Observer based adaptive dynamic programming
for fault tolerant control of a class of nonlinear systems," Inf. Sci.,vol. 384, pp. 21_33, Apr. 2017.

[28] H. G. Zhang, K. Zhang, Y. Cai, and H. Jian, ``Adaptive fuzzy fault-tolerant Tracking control for partially unknown systems with actuator faults viaintegral reinforcement learning method," IEEE Trans. Fuzzy Syst., vol. 27,no. 10, pp. 1986_1998, Oct. 2019.

AUTHORS PROFILE

Mr**.** B**.**AMARNATH REDDY is an assistant professor in the Department of Master Of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. He earned his M.tech from Vellore Institute of Technology (VIT), Vellore. His research interests include machine learning, programming languages. He is committed to advancing research and Fostering innovation while mentoring studentsto excelin bothacademicandprofessionalpursuits.

MS**.**A**.**VENKATA SWATHI is a postgraduate student pursuing a Master of Computer Applications (MCA) in the Department of Computer Applications at QIS College of Engineering & Technology, Ongole an Autonomous college in Prakasam dist.She completed her undergraduate degree in BCA (Computers) from (ANU). Her academic interests include Cloud Computing, Artificial Intelligence, Cyber Security, and Data Structures.