**REINFORCEMENT LEARNING**

**PROJECT REPORT**

****

**Faculty name: Ms. Neetu Singla**

**STUDENT NAME & ROLL NO :**

Keerti (21CSU260)

Aditya Sindhu(21CSU278)

Sowmya Sri(21CSU326)

Snigdha (21CSU283)

Abhishek Dubey(21CSU323)

**Semester:** 5th

**Group:** AIML-B (A3)

**Department of Computer Science and Engineering**

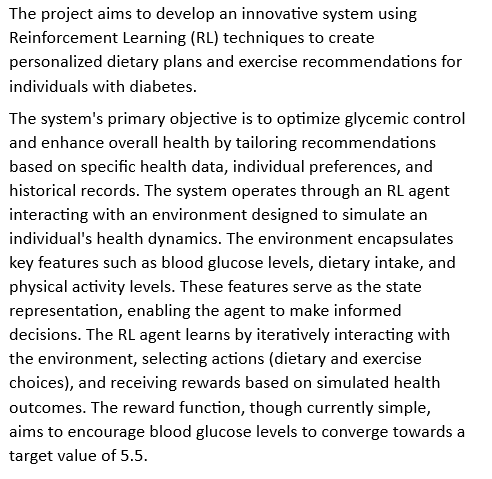
**The NorthCap University, Gurugram- 122001, India**

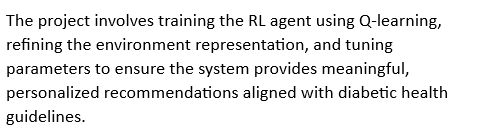
**Session 2023-2024**

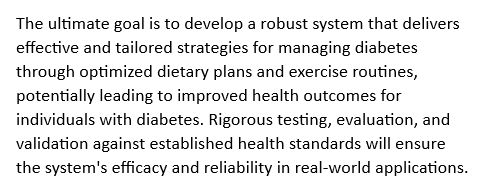
**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **S.No** |  | **Page No.** |
| **1.** | **Project Description** | **3** |
| **2.** | **Problem Statement** | **5** |
| **3.** | **Research Paper Summaries** | **6** |
| **4.** | **References** | **14** |
| **5.** | **Review Of Papers In Tabular Form** | **15** |
| **6.** | **Code Link** | **16** |

**PROJECT DESCRIPTION**







**PROBLEM STATEMENT**

Design an RL-based system for personalized dietary planning and exercise recommendations ." The objective is to develop an RL agent that tailors dietary plans and exercise routines for individuals with diabetes, taking into account their specific health data, dietary preferences, and historical records, to optimize glycaemic control and overall health.

**RESEARCH PAPER SUMMARIES**

**[1]**

**Offline reinforcement learning for safer blood glucose control in people with type 1**

**Diabetes (ABHISHEK DUBEY)**

SUMMARY

Glucose control is a big challenge for people with type 1 diabetes. They need to

carefully monitor their blood sugar levels and take insulin injections to keep their blood

sugar in a healthy range.

Reinforcement learning (RL) is a type of machine learning that can be used to teach

computers to control systems, like insulin pumps. RL algorithms learn by trial and error,

interacting with the environment to learn what actions lead to the desired results.

Online RL algorithms learn by interacting with the environment in real-time. This can be

dangerous for people with type 1 diabetes, because if the algorithm makes a mistake, it

could lead to a dangerous blood sugar spike or drop.

Offline RL algorithms can learn from data that has already been collected. This means

that offline RL algorithms can be trained without interacting with the environment in real

time, making them safer for use in glucose control.

In this paper, the authors evaluate three offline RL algorithms: batch-constrained deep

Q-learning (BCQ), conservative Q-learning (CQL), and twin delayed deep deterministic

policy gradient with behavioral cloning (TD3-BC). They train and test these algorithms

on a cohort of 30 virtual patients and compare their performance to the strongest online

RL and control baselines.

The results show that offline RL can yield more effective and safer insulin dosing

policies than online RL, without the need for patient interaction during training. Offline

RL also requires significantly smaller samples of data, making it more applicable for use

in real patients. Offline RL in the presence of practical limitations, such as missing CGM data and

sub-optimally set PID parameters. They show that offline RL is robust to these limitations

and can still learn effective control policies.

TECHNOLOGIES THAT ARE USED IN THIS PAPER:

* Batch-constrained deep Q-learning (BCQ)
* Conservative Q-learning (CQL)
* Twin delayed deep deterministic policy gradient with behavioral cloning (TD3-BC)

**[2]**

**Reinforcement learning application in diabetes blood glucose control: A systematic review (SNIGDHA SHRIVASTAVA)**

SUMMARY

In the Research paper the method and technologies using online databases, covering publications from 1990 to 2019. This involved searching for relevant articles, likely using search terms related to RL, BG control, and DM. Then in the initial stage got set of selection criteria was established to select the most relevant papers according to the title, keywords, and abstract. then Questions were established and answered in the second stage.

Initially, selection criteria were established to filter relevant papers based on titles, keywords, and abstracts. Research questions were subsequently formulated and answered using information extracted from the selected articles. The search yielded 404 articles, which were reduced to 347 after eliminating duplicates. A rigorous screening process, by the predefined inclusion and exclusion criteria, led to the removal of 296 articles, leaving 51 that were considered relevant. A comprehensive assessment of the full-text content resulted in 29 articles that underwent critical analysis. Inter-rater agreement was measured using the Cohen Kappa test, with disagreements resolved through discussion.

The review concludes that advances in health technologies and mobile devices have made it increasingly feasible to implement RL algorithms for optimizing glycemic control in diabetes. However, there remains a scarcity of literature specifically addressing the application of these algorithms to BG regulation. The paper also notes that RL algorithms are being designed for BG adjustment, and their utilization in diabetes research is on the rise, suggesting a growing trend in their use for BG control in the future.

Additionally, the review highlights a lack of focus in the literature on factors influencing BG levels, such as meal intake and physical activity, which should be incorporated into the control problem.

Finally, it emphasizes the need for clinical validation of these algorithms, indicating an avenue for further research and development in this field.

**[3]**

**Reinforcement Learning for Multiple Daily Injection (MDI) Therapy in Type 1 Diabetes (T1D ) (SOWMYA )**

SUMMARY

This research paper presents an innovative approach to managing type 1 diabetes using reinforcement learning (RL) techniques. It introduces a closed-loop insulin administration framework specifically designed for multiple daily injection (MDI) therapy. The RL agent, powered by the soft actor-critic (SAC) algorithm, dynamically adjusts insulin dosages based on real-time glucose levels, meal consumption, and prior actions.

KEY FINDINGS AND HIGHLIGHTS:

* Improved Glucose Control: The proposed closed-loop control strategy effectively reduces glucose fluctuations and significantly extends the time during which blood glucose levels stay within the desired target range of 70-180 mg/dL. Weekly average glucose levels exhibit notable reductions across various meal scenarios.
* Increased Time in Target Range: The study demonstrates a substantial enhancement in the percentage of time spent within the optimal blood glucose range (70-180 mg/dL), indicating improved glycemic management through the RL-based approach.
* Robustness: The RL-based model displays resilience against meal disruptions and variations in insulin sensitivity, effectively maintaining blood glucose levels within the target range and minimizing the risk of hypoglycemia.
* Personalized Insulin Delivery: The model tailors insulin doses in real-time, offering customized treatment plans tailored to the unique needs of individuals with type 1 diabetes.

CONCLUSION

This research paper introduces a groundbreaking approach to addressing type 1 diabetes using RL techniques. It presents a closed-loop insulin administration framework tailored for MDI therapy, driven by the SAC algorithm. The results underscore the potential of RL-based closed-loop insulin administration models to significantly enhance glycemic control in individuals with type 1 diabetes. This approach holds promise for delivering personalized and effective insulin solutions to those reliant on multiple daily injections, thereby improving their overall well-being.

[4]

**Reinforcement Learning Models and Algorithms for Diabetes Management**

**(ADITYA SINDHU)**

SUMMARY

Diabetes management can be formulated as a Markov decision process (MDP) problem, which can then be solved using reinforcement learning (RL). In an MDP, an agent makes sequential decisions in a dynamic and noisy environment. Traditional RL algorithms solve the MDP problem without using the transition probability matrix and have been applied to diabetes management. This model-free approach allows the agent to learn which actions are appropriate under different states through trial and error.

This can be represented by the following recursive relationship:

Vπt(st)=rt+1(st+1)+γVπt(st+1)

Where:

* Vπt(st) is the value of state st under policy π at time t
* rt+1(st+1) is the reward received for taking action at and transitioning to state st+1
* γ is the discount factor

In the context of diabetes management, the agent could be a blood glucose regulator that interacts with the operating environment (the human body and glucose variability). The agent would select actions (insulin dosages) based on the state (the patient's clinical condition, such as blood glucose level) and receive feedback in the form of the next state and delayed reward (e.g., time in range, during which the blood glucose level is within the normal range).

System Models

A system model is a representation of the operating environment and provides the agent with information about the glucose-insulin dynamics of an average virtual patient (AVP) in diabetes management. The system model must be realistic and complete, so the US Food and Drug Administration (FDA) encourages the use of its approved system models when running simulations. This helps to avoid using animals in preclinical tests.

Some examples of approved system models include:

* Ferdinando's system model
* The Lehmann-Deutsch physiological model
* Shifrin's system model
* Lee's system model
* Palumbo's system model
* Hovorka's glucoregulatory model

Reinforcement Learning Approaches

Several RL approaches have been proposed for diabetes management, including:

* Traditional RL
* Model-based RL
* Multi-agent RL
* Actor-critic RL
* Deep Q-networks
* Gaussian process approximation
* Proximal policy optimization

These RL techniques explore how diabetes can be effectively managed by establishing appropriate representations for states, actions, and rewards while leveraging advanced algorithms.

Conclusion

The use of RL for diabetes management is a promising area of research with the potential to significantly improve glycemic control and quality of life for patients with diabetes. Further research is needed to develop and validate RL-based systems in clinical settings.

**[5]**

**Optimized glycemic control of type 2 diabetes with reinforcement learning: A proof-of-concept trial (KEERTI)**

SUMMARY

Reinforcement Learning for Personalized Insulin Delivery in Type 2 Diabetes

The research paper "Optimized Glycemic Control of Type 2 Diabetes with Reinforcement Learning: A Proof-of-Concept Trial" introduces a novel approach to managing type 2 diabetes using a model-based reinforcement learning (RL) framework called RL-DITR. The goal of this study was to develop personalized insulin plans that effectively control blood sugar levels while minimizing the risk of hypoglycemia and other complications.

Key Findings and Highlights:

* The RL-DITR framework demonstrated consistent performance in terms of effectiveness, safety, and patient acceptability over time. Remarkably, results improved upon retesting, indicating the model's ability to learn and adapt.
* RL-DITR uses patient data such as blood sugar levels, insulin doses, and meal information to determine the most suitable insulin regimen for each individual. This personalized approach has the potential to achieve better blood sugar control than traditional fixed insulin plans.
* The study involved 40 participants with type 2 diabetes who were randomly divided into two groups: the RL-DITR group and the control group. The RL-DITR group received personalized insulin plans generated by the RL-DITR framework, while the control group received standard care.

The primary outcome measure was the amount of time spent within the target blood sugar range (70-180 mg/dL). The RL-DITR group showed a significant improvement in this measure compared to the control group.

Secondary outcome measures included the time spent in hypoglycemia and hyperglycemia states. The RL-DITR group had a lower risk of hypoglycemia while maintaining a similar risk of hyperglycemia compared to the control group.

The RL-DITR group had a lower average daily insulin dosage and reported higher satisfaction with their treatment, suggesting that the personalized insulin plans generated by the RL-DITR framework were both more effective and acceptable to the participants.

Conclusion:

The RL-DITR framework is a promising new approach to managing type 2 diabetes. It has the potential to improve blood sugar control, reduce the risk of hypoglycemia, and improve patient satisfaction. Further research is needed to validate RL-DITR in larger clinical trials and to implement it in real-world clinical practice.

**REFERENCES**

[1][Offline reinforcement learning for safer blood glucose control in people with type 1 diabetes - ScienceDirect](https://www.sciencedirect.com/science/article/pii/S1532046423000977)--Abhishek

[2] [Reinforcement learning application in diabetes blood glucose control: A systematic review - ScienceDirect](https://www.sciencedirect.com/science/article/abs/pii/S0933365718304548#:~:text=A%20Q%2Dlearning%2Dbased%20reinforcement,for%20the%20in%2Dsilico%20trials)—Snigdha

[3] [Reinforcement Learning for Multiple Daily Injection (MDI) Therapy in Type 1 Diabetes (T1D)](https://www.mdpi.com/2673-7426/3/2/28)--Sowmya

[4] [(PDF) Reinforcement Learning Models and Algorithms for Diabetes Management (researchgate.net)](https://www.researchgate.net/publication/369390517_Reinforcement_Learning_Models_and_Algorithms_for_Diabetes_Management) – Aditya

[5] [Optimized glycemic control of type 2 diabetes with reinforcement learning: a proof-of-concept trial | Nature Medicine](https://www.nature.com/articles/s41591-023-02552-9)—Keerti

**REVIEW OF PAPERS IN TABULAR FORM**

|  |  |  |  |
| --- | --- | --- | --- |
| **Paper Title** | **Summary** | **Technologies/Algorithms** | **Challenges** |
| **[1]** Offline reinforcement learning for safer blood glucose control in people with type 1 Diabetes **(ABHISHEK DUBEY)** | - Addresses blood glucose control challenges in type 1 diabetes.  - Compares offline RL algorithms (BCQ, CQL, TD3-BC) to online RL and control methods.  - Offline RL is safer and requires less data. | - Batch-constrained deep Q-learning (BCQ)  - Conservative Q-learning (CQL)  - Twin delayed deep deterministic policy gradient with behavioral cloning (TD3-BC) | - Addressing sub-optimally set PID parameters and missing CGM data.  - Robustness to practical limitations. |
| **[2]** Reinforcement learning application in diabetes blood glucose control  **(SNIGDHA SHRIVASTAVA)** | - Conducts a systematic review of RL in diabetes control.  - Identifies a growing trend in RL applications for glycemic control.  - Highlights the need for incorporating meal intake and physical activity in control algorithms.  - Calls for clinical validation of RL algorithms. | - Online reinforcement learning (RL). | - Complexity of online RL algorithm design and implementation.  - Data requirements for online RL training.  - The need for clinical validation before patient use. |
| **[3]** Reinforcement Learning for Multiple Daily Injection (MDI) Therapy in Type 1 Diabetes (T1D)  **(SOWMYA SRI)** | - Introduces RL-based insulin management for type 1 diabetes using SAC.  - Shows improved glucose control, time in target range, and resilience to disruptions.  - Offers personalized insulin delivery. | - Soft Actor-Critic (SAC) Algorithm  - Mathematical Model  - In Silico Simulations  - Reinforcement Learning (RL)  - Data-driven and AI methods | - Real-time adaptation to meal consumption and glucose levels.  - Resilience to meal disruptions and insulin sensitivity changes. |
| **[4]** Reinforcement Learning Models and Algorithms for Diabetes Management **(ADITYA SINDHU)** | - Discusses RL as a solution to diabetes management.  - Mentions different system models for glucose-insulin dynamics.  - Explores various aspects of training RL models in diabetes management. - Mentions multiple RL approaches.  - Emphasizes patient data collection, personalization, and clinical evaluation. | -Markov Decision Processes (MDP)  - Q-learning, SARSA, Deep RL - Differential equations, RNNs, physiological models  - Contextual bandit algorithms - Reward design strategies  - Continuous monitoring and control systems  - Clinical trials and evaluation methodologies | Temporal Difference (TD) - The temporal difference is calculated as rt+1(st+1) + γ \* maxa∈A Qt(st+1, a) - Qt(st, at) in Q-learning. - Reflects the difference between the expected and actual rewards. |
| **[5]** Optimized glycemic control of type 2 diabetes with reinforcement learning: A proof-of-concept trial  **(KEERTI KOHLI)** | - Introduces RL-DITR for personalized insulin plans in type 2 diabetes.  - Shows consistent improvement in blood sugar control over time.  - Highlights reduced hypoglycemia risk and lower insulin dosage. | - RL-DITR framework for personalized insulin plan.  -Model-based reinforcement learning (RL)  - Patient model (RNN)  - Policy model (RNN)  - ClinicalBERT pre-trained model  - Natural language processing (NLP) pipeline | - Optimization of RL-DITR framework for long-term effectiveness and safety.  - Addressing individual patient variations and compliance. |

**CODE LINK:**

[RL-CSL348/RL\_PROJECT\_GROUP-4.ipynb at main · Sowmya0130/RL-CSL348 (github.com)](https://github.com/Sowmya0130/RL-CSL348/blob/main/RL_PROJECT_GROUP-4.ipynb)