**REINFORCEMENT LEARNING PROJECT**

Logo

Description automatically generated

**GROUP-4**

**NAME & ROLL NO :**

Keerti (21CSU260)

Aditya Sindhu(21CSU278)

Snigdha (21CSU283)

Abhishek Dubey(21CSU323)

Sowmya Sri(21CSU326)

**FACULTY NAME:** Ms. Neetu Singla

**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

The research paper titled "Reinforcement Learning for Multiple Daily Injection (MDI) Therapy in Type 1 Diabetes (T1D)" introduces a fresh approach to managing type 1 diabetes using reinforcement learning (RL) techniques. It presents a closed-loop insulin administration framework tailored for multiple daily injection (MDI) therapy. The RL agent, powered by the soft actor-critic (SAC) algorithm, dynamically fine-tunes insulin dosages based on real-time glucose levels, meal consumption, and previous actions.

KEY FINDINGS:

1. Enhanced Glucose Control: The proposed closed-loop control strategy notably reduces glucose fluctuations and substantially increases the time during which glucose levels remain within the target range (70-180 mg/dL). Weekly average glucose levels witnessed significant reductions across various meal scenarios.
2. Time Spent in Target Range: The study reveals a substantial improvement in the percentage of time spent within the desired blood glucose range (70-180 mg/dL), signaling improved glycemic management through the RL-based approach.
3. Resilience: The RL-based model displays robustness against meal disruptions and changes in insulin sensitivity, maintaining glucose levels within the target range and minimizing the risk of hypoglycemia.
4. Personalized Insulin Delivery: The model tailors insulin doses in real-time, providing customized treatment plans for individuals with type 1 diabetes.

TECHNOLOGIES THAT ARE USED IN THIS PAPER:

1. Soft Actor-Critic (SAC) Algorithm: It serves as the foundation for making real-time, adaptive insulin dosage decisions based on data.
2. Mathematical Model: The research employs a mathematical model of the glucose-insulin metabolic system, incorporating a pharmacokinetics (PK) module for long-acting basal insulin.
3. In Silico Simulations: To circumvent limitations linked to real-world experimentation, the study harnesses in silico simulations. These simulations facilitate the creation of a comprehensive insulin delivery model and enable the manipulation of factors influencing blood glucose dynamics.
4. Reinforcement Learning (RL): RL serves as the central technology for automating insulin delivery. RL empowers the agent to learn the best insulin dosage decisions through iterative interactions with the environment.
5. Data-Driven and Artificial Intelligence (AI): The research mentions the use of machine learning (ML) and neural networks (NN) within the broader context of diabetes management. These methods find application in insulin dosage recommendations and future blood glucose predictions.

In conclusion, the research paper presents an inventive approach to address type 1 diabetes using reinforcement learning (RL) techniques. The study introduces a closed-loop insulin administration framework tailored for multiple daily injection (MDI) therapy. The RL agent, driven by the soft actor-critic (SAC) algorithm, dynamically fine-tunes insulin dosages based on real-time glucose levels, meal intake, and prior actions.

This study underscores the potential of RL-based closed-loop insulin administration models to substantially improve glycemic control among individuals with type 1 diabetes. It holds the potential to provide personalized and effective insulin delivery solutions for individuals who depend on multiple daily injections, enhancing their overall well-being.Top of Form

**[4]**

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

**(ADITYA SINDHU)**

SUMMARY

In this research paper, I learned that Diabetes management can be formulated as the MDP problem and subsequently solved using RL. In MDP, an autonomous agent makes sequential discrete-time decisions in the dynamic and noisy operating environment as time goes by. The traditional RL algorithm solves the MDP problem without using the transition probability matrix, and it has been applied in diabetes management. It is a model-free approach that allows the agent to discover through trial-and-error which actions are appropriate under different states which is equivalent to Equation :

 recursive relationship :Vπt(st)=rt+1(st+1)+γVπt(st+1)

The closed-loop agent e.g. blood glucose regulator interacts with the operating environment e.g., the human body and the glucose variability). The agent selects actions e.g., the insulin dosage based on the state e.g., the patient’s clinical condition, such as the blood glucose level, and receives feedback in the form of the next state and delayed reward (e.g., the performance of time in range during which the blood glucose level is within the normal range.

SYSTEM MODELS:

A system model, which provides the operating environment, represents 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, which helps to avoid using animals in preclinical tests. Like:

* Ferdinando’s system model
* The Lehmann-Deutsch physiological model comprises differential equations to represent the glucose-insulin dynamics of T1DM patients
* Shifrin’s system model, which is created based on experience, comprises a transition probability matrix to represent personal reactions to insulin dosage
* Lee’s system model, which is created based on experiments, comprises the gamma distribution 0(k, θ)representing the insulin time-action process
* Palumbo’s system model, which is created based on expert knowledge
* Hovorka’s glucoregulatory model, which is created based on experience, comprises parameters sampled from a prior log-normal distribution and parameter correlations sampled from healthy individual data.

The paper delves into various facets of training RL models in diabetes management, encompassing patient data, system models, simulators, simulation parameters, clinical studies, implementation strategies, experiment designs, and performance metrics. Several RL approaches have been proposed, including traditional RL, model-based RL, multi-agent RL, actor-critic RL, deep Q-networks, Gaussian process approximation, and proximal policy optimization. Through these RL techniques, the paper explores how diabetes can be effectively managed by establishing appropriate representations for states, actions, and rewards while leveraging advanced algorithms.

**[5]**

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

SUMMARY

The research paper titled "Optimized Glycemic Control of Type 2 Diabetes with Reinforcement Learning: A Proof-of-Concept Trial" introduces an innovative approach to managing type 2 diabetes using a specialized model-based reinforcement learning framework known as RL-DITR. The goal of this study is to create individualized insulin plans that effectively control blood sugar levels while minimizing the risk of low blood sugar events (hypoglycemia) and other complications.

KEY FINDINGS AND HIGHLIGHTS:

1. The RL-DITR framework displayed consistent performance in terms of effectiveness, safety, and patient acceptability over time. Remarkably, upon retesting, the results improved, indicating the model's capacity to learn and enhance its performance.
2. RL-DITR uses patient data such as blood sugar levels, insulin doses, and meal information to determine the most suitable insulin regimen for each person. This personalized approach demonstrates the potential for achieving better blood sugar control compared to traditional fixed insulin plans.
3. 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.
4. The primary measure of success 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.
5. Secondary outcomes included the time spent in low blood sugar (hypoglycemia) and high blood sugar (hyperglycemia) states. The RL-DITR group had a lower risk of hypoglycemia while maintaining a similar risk of hyperglycemia compared to the control group.
6. 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.

**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. | - Optimization of RL-DITR framework for long-term effectiveness and safety.  - Addressing individual patient variations and compliance. |