IJBPAS, November, Special Issue, 2021, 10(11): 1192-1203

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

IMPROVED REINFORCEMENT LEARNING TO IDENTIFY THE DIABETICS TYPES BASED ON HEALTHCARE OPTIMUM POLICY

LAVANYA.R[1*], MANIKANDAN.S[2], AJITH KUMAR.S[3] AND HARIPRASANTH.S[4]

1. Assistant Professor in Computer Science at Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, India
2. II M.Sc. in Computer Science at Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, India
3. II M.Sc. in Computer Science at Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, India
4. II M.Sc. in Computer Science at Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, India

***Corresponding Author: Lavanya.R; E Mail: lavanyaprabhakaran07@gmail.com**

Received 24th July 2021; Revised 26th Aug. 2021; Accepted 30th Sept. 2021; Available online 1st Nov. 2021

https://doi.org/10.31032/IJBPAS/2021/10.11.1102

ABSTRACT

Insulin deficiency would be a feature of hyperglycemia Diabetics Type (DT), which was among the most frequent metabolic disorders. The inability of glucagon to function properly creates unmanageable blood sugar levels in the body that could lead to the existing conditions. As a consequence, early intervention of DT was critical to saving many lives. This paper provides a computer learning-based forecasting approach for detecting DT to achieve this goal. The PIMA Indian Women Mellitus database was used to create the classification models, which used the Q-learning technique from the Reinforcement Algorithm (RL) methodology. A model generates an over predicated RL and instructs that learning operative discover to best strategy of three criteria's (such as weight, insulin level, and patient age) to diagnose patients with DT. A subject's data could be in any of 330 multiple states. The reliability, specificity, memory, F-measure, and AUC values of the suggested RL study were compared to those of government approaches like K Nearest Neighbor (KNN) suggested in terms of effectiveness.

Keywords: Hyperglycemia; Diabetics Types; Reinforcement Learning; Decision Tree; K nearest Neighbor

INTRODUCTION

Hyperglycemia would be a life-altering disease that predominantly affects each year. DT accounts for 90 percent or 95 percent of all diagnosed cases, according to the 2017 national Mellitus latest statistics. DT would be a significant disease that affects a large number of people from around the world and is on the rise **[1]**. DT would be a leading cause of illness and death, and it increases the risk of chronic disease by two to four occasions. As a consequence, early identification of DT is critical for improving patients' quality of living and extending overall life expectancy **[2]**. To comprehend the patterns in data, Machine Learning (ML) methods are used. ML is being used in a range of applications these days, including biomedical information retrieval, feature extraction, investment analysis, fault diagnosis, illness diagnosis, computational linguistics, assay surveillance, and dynamic game running **[3]**.

Associations derived via machine learning could be used to enhance our understanding or forecast potential outcomes. ML researches a wide range of topics, including analytics, management, and computer programming. Almost all ML challenges are documented or judged as efficiency issues when it comes to data sets **[4]**. Electronic Health Records (EHRs) became more commonly accessible in current history, piquing enthusiasm in healthcare monitoring recommendations to improve health care decisions and customer monitoring **[5]**. As an outcome, the treatment recommendation research shifts from a knowledge-based to an informal approach. A goal in such a situation is to develop a design that accurately understands the information **[6]**. Most machine learning techniques fall into three classifications: monitored, unregulated, or RL focused specifically Q-learning, which would be a court hearing technique that an operator learns of the challenging situation using feedback to its activities and interactions.

The individual learns a translating mechanism from a series of facts to the eventual outcome after observing an insufficient sequence of attendees **[7]**. The training dataset is used to complete this learning experience. When the agent was trained, the resultant mapping model can be used to new training dataset for pattern recognition or forecasting. Using EHRs containing actual information such as Body Mass Index (BMI), hyperglycemia, and age, the proposed RL model can detect DT dynamically and constructively. In this study, the suggested Q-learning approach is consistent with basic yet effective controlled ML algorithms such as KNN to forecast whether or not an individual has DT **[8-11]**. The information has been taken from the PIMA Indian Women Hyperglycemia dataset.

Numerous researchers focus on mellitus information or developed Mellitus modeling techniques. Numerous predictors were used as foundation predictors for AdaBoost computation, including Multinomial Naive Bayes (MNB), Support Vector Machine (SVM) **[12-13]** AdaBoost computation efficiency with classification model. PIMA Indian data set, multiple classification algorithms such as J48 DT, MNB, Random Forest (RF), and Multilayer Perspective should be used to compare or forecast hyperglycemia. In aspects of accuracy and consistency, Multilayer Interpretation performed the best. When contrasted to K-Mean, and KNN, the SVM classifier had greater forecasting accuracy DT patients **[14]**. The effectiveness of RL approaches for blood sugar monitoring automation. In a series of studies, they evaluated the contribution of different deep RL algorithms versus non-RL methods **[15]**. Based on this finding, RL approaches are comparable with benchmarks and are better equipped to handle concealed behavior responses, which suggests that RL could be effective to enhance blood sugar management techniques. RL-based methodology for treating hyperglycemia in a patient with Type 1 diabetes (T1D) who takes insulin injections, in which RL agent learns to choose the acceptable behavior through exploratory and strategic learning. In this strategy, the authors used physical exercise, BMI, and glycosylated hemoglobin (HbA1C) level as state variables of diabetes patients. Hyperglycemia control using a Q-iteration structure **[16]**. For T1D patients, the RL agent used the best insulin strategy. An individual personalization was achieved by modifying data restricted as in form of a hyperglycemia curve using a quasi regression technique using functional properties.

A model-free Q-learning methodology was used in an operative off-line simulation to restrict the blood sugar level of a T1D patient. A work's main characteristics were consistency, resistance to interruption, or ambiguity, as well as a precise steady-state. A new RL problem solution is presented that enhances medical concept mining from a free text scenario via attributing associated observable indications to discover the most likely diagnoses. During retraining, the agent discovered a great strategy through iteratively searching for associating the most relevant medical ideas that quality products a definitive diagnosis.

Classifiers

Using diabetes information, KNN could be used to diagnose DT. In this example, the categorization was compelled by majority votes of neighbors, and a separation formula would be used to estimate its K closest neighbors. A DT is discriminate that partitions to occurrence space sequentially. It is made up of elements to create a tree, with the tree's basic framework being a component called "root" that has no international implications. Each of the other nodes has only one making initial.

$$IGain = Et(parent) - [W_{avg}] * Et(child) \quad (1)$$

$$Et = -Q(x_i) \log Q(x_i) \quad (2)$$

A collection of q-values was predicted using the Q-learning process by an entity using a set of events and behaviors. For each couple, the agent would receive a positive or negative incentive. Through acquiring effective policy choices for various situations, the agent enhanced this same positive or negative benefit as in the long-run operation. The amount of q table or q matrix remains zero at the start of the Q-learning process. The operator might run a series of instances to produce better appropriate q-values, that agent could use as a reference table to make the optimal option in a given situation. The owner's decision-making process can be divided into two main categories: exploitation and exploration. Exploitation refers to deciding on a course of conduct based on maximum q-value attainable. Exploration, on either extreme, would be a method of taking behavior at arbitrary rather than deciding depending on every strategy and exploring different realities that could be exploited. Both are regulated by epsilon (), which determines how often the agent would explore or utilize resource regions.

The primary repetitive operation of this technique was changing the Q value, this is a form of continuous compensation (R) or elements of subsequent government interference couple Q [Tu+1, Bu+1]. A q value update function is shown in Equation 3:

$$Q[Tu,Bu] = R[Tu,Bu] + \gamma * (S_u + \alpha * \max(Q[Tu+1, Bu+1]) - Q[Tu,Bu]) \quad (3)$$

A training error (α) in the above method changes the step length, which specifies the approval probability among new and old values. Gamma (γ) would be a novel value reduction element that is used to balance existing and future benefits. After numerous executions, the q-table has been modified to include more traditionally attributed, or the operator develops the best approach for making decisions in each phase.

Table 1: Parameter Levels

Stages	BMI	Glucose	Age
Stage 1	<25	<114	<22
Stage 2	>25.1 and 27	>115 and <129	>22 and <25
Stage 3	>27 and <28	>130 and <139	>26 and <30
Stage 4	>28 and <29	>140 and <150	>31 and <35
Stage 5	>29 and <30	>151	>36
Stage 6	>30 and <31	-	-
Stage 7	>31 and <32	-	-
Stage 8	>32 and <33	-	-
Stage 9	> 33 and <34	-	-
Stage 10	>34 and <35	-	-
Stage 11	>35 and <45	-	-

A created framework translates data into knowledge in a methodical manner, allowing us to develop forecast solutions. The Q-learning methodology is used on particular information set in this environment to enhance the precision of DT prediction, and the values are presented to other controlled ML techniques of KNN or DT. **Table 1** shows this same execution of a DT probability model predicated on Q-learning. After evaluating the association network, three parameters (hyperglycemia, BMI, and age) were chosen from the PIMA Indian Women Insulin Database. These characteristics have a considerable impact on the mellitus prediction outcome. Updated information was retrieved using selected three characteristics from collected data, which includes information on 768 subjects.

A data is then cleansed after the changed information has been reconstructed to eliminate noise. Numerous parameters have zero as their minimum amount. However, genetic influences such as hyperglycemia and body mass index (BMI) could have negative values; it appears that null values have been recorded as zeros. As a consequence, they deleted the impulse noise from our database and invented a separate one with 392 subjects' data for several attributes that were free of duplicated and inaccurate statements. A spatial structure of an individual was represented by a 2D matrix, with the x-axis denoting the combination of hypoglycemia level of students (56 = 30) or the y-axis denoting 11 BMI categories. The agents choose activities based on the action's maximal q-value or at arbitrary. Our approach uses the epsilon variable to manage the investigation (where q-agent acts arbitrarily) or domination (where q-agent acts depending on the greatest q-value). A two-dimensional database, initialized at zero, with some rows and columns as the q-model state vector.

Q-Agent communicates with the surroundings as in the proposed model, traveling from commencing to the desired condition and receiving positive or negative benefits based on the phase space. Positive DT receives a reward of +10, whereas all other phases receive a payout of -100. They separated our data set into two sections for forecasting DT using the Q-learning methodology: the classification model set (242 rows, 3 columns) or the testing set (242 rows, 4 columns) (150 rows, 3 columns). The Q-learning model selects the best DT forecasting strategy from patients' medical experience information during development operation. For each permutation of condition and movement, Q-agent evaluates or modifies the q-values table in each episode. For each conjunction of condition and movement, the Q-agent generates or maintained a q-values chart in each program. Investigating the phase regions begins with S (1, 1, 1) and ends with S(11, 5, 6) to create a single program. To develop an appropriate or trustworthy DT forecast consequence while limiting error rate and generalization, they used 20000 complete episodes in our procedure.

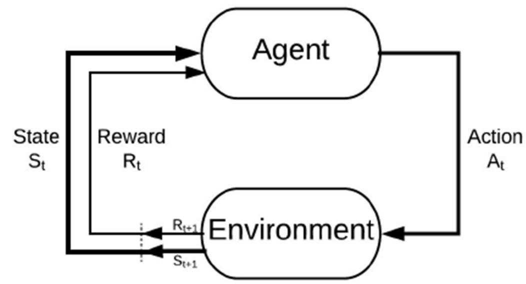

Figure 1: Q-learning design

The proposed methodology was evaluated using 150 test results from PIMA Indian women mellitus database. To forecast DT, consumers first estimate the patient's condition or then discover q-table values for that condition. A patient is said to have a favorable DT if they discover q-value to comparable in +10. Otherwise, the patient was diagnosed as having negative DT. The National Academy of Diabetes and Digestive and Kidney Diseases (NIDDK) produced the Pima Indian Women Insulin data set, which comprises diagnostic medical data processing components as well as Diabetes Mellitus (0 or 1) as an endogenous construct. Age, insulin level, hypertension, BIM, skin thickness, blood sugar, mellitus pedigree performance, or outcome are among the 9 variables in the data, which comprises 769 female individuals.

Table 2: Q-Learning Results

BMI level	Glucose	Age	Actual	Q – Agent	Remarks
41	188	39	1	1	True positive
36	176	35	1	1	True positive
25	199	24	1	-1	False-negative
30	187	30.5	1	1	True positive
31	133	30	-1	1	False-positive
39	187	33	1	1	True positive
. .	. .	. .	. .	. .	. .
25	88	25	-1	-1	True negative
27	97	26	-1	-1	True negative

RESULTS AND DISCUSSION

Q-learning approach is validated to use data from 150 Indian women with diabetes from the PIMA database, of whom 59 had positive DT, as well as 91, have negative DT. They used the similar lifestyle variables of an individual that was used in the simulation construction phase to evaluate the procedure's achievement: hyperglycemia, weight, or age. **Table 2** displays the weight, Hyperglycemia, Age, Real Outcomes, and Q-Agent Consequence of the 150 test group. The confusion matrix of a binary classifier would be a two-by-two chart which counts the true positive (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) explaining the reliable positive criterion, erroneous favorable predictor, correct negative predictor, and erroneous negative predictor, respectively. This work's ambiguity matrix was developed employing 150 test data. There were no FP errors, and categorization was 35 TP, 24 FN, or 91 TN.

Figure 2 shows the ROC curve for the Q-learning algorithm. The importance of a subject's BMI, sugar levels, and age in building accurate DT forecasting models cannot be overstated. In this study, they created an off-policy rewarded environment and our learning agent discover the best policy for identifying patients with DT predicated on those three critical variables. They created a subject's information in such a manner that the subject might be in any of 330 potential states (amount of blood glucose, amount of weight levels, and quantity of age categories = 5 11 6). A design correctly detected DT to an efficiency of 84%, 100% consistency, 59 percent memory, and 74% f-estimate. **Table 3** summarizes of results in machine learning techniques.

Table 3 shows that the KNN classifier accomplished the normal and precision to 74.49 percent and reliability of 74.49 percent (64.29 percent). The DT extractor, on either extreme, fared well with a

75.51 percent prediction performance, 54.55 percent recognition accuracy, and an f-measure of 1.2121, but less accuracy (66.67 percent). Compared to this ML category, the suggested Lee proposed surpasses them, with an efficiency of 84 percent, specificity of 100 percent, and recollection of 59.3 percent. **Figure 3** and **Figure 4** illustrate the ROC curves for KNN and DT, correspondingly.

However, there has been a resurgence of interest in employing artificial attempting to learn techniques that can determine EHRs to diagnose disease and build individualized treatment plans in current history. It has marked the start of a new era or new treatments. This research hopes to be one to use an RL model to forecast DT. To begin, this study used hyperglycemia, weight, and age information variables which have a substantial impact on DT outcomes. These three crucial environmental characteristics, which were not previously examined as in research, were used to construct an effective and easiest DT diagnostic framework.

This research can help identify hypoglycemic patients as well as those who have DT. Personal q-values for each of 330 constructed asserts are engendered in this research, and a subject is deemed to have DT or unless the patient's province's corresponding q-value was +10. They could also identify hyperglycemia patients using the generated q-value chart.

If a patient receives a positive but less than +10 q-values for his entered condition, consumers could assume that subject has acquired hyperglycemia, which is the first stage of DT. Finally, the Q-learning system outperformed the supervised ML techniques in terms of accuracy (84 percent). Eventually, this design estimates DT to required accuracy (100%), implying that the approach is always correct in its correctly predicted. There was no one subject's information that was genuinely negative during the model evaluation process, although the model projected positive. Reducing the FP rate in medical forecasting was critical since it can lead to a false diagnosis.

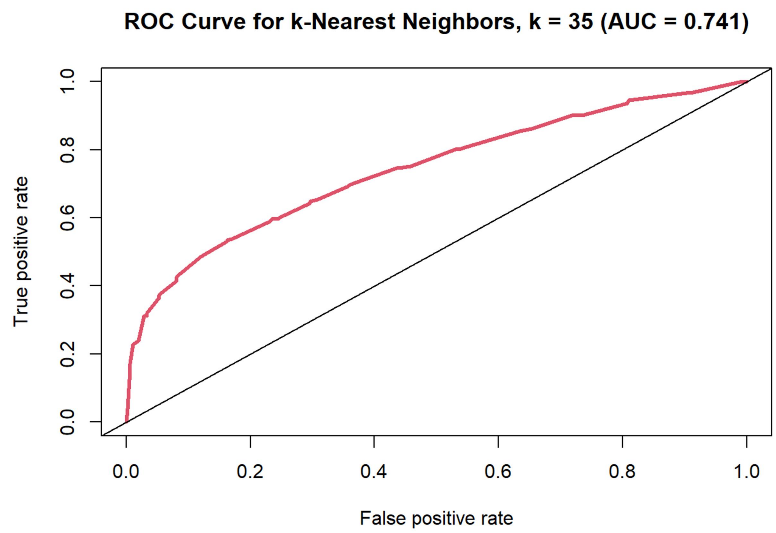

Figure 2: ROC diagram for Model

Table 3: Analysis of various methods

Classification	Accuracy	Precision	Recall	F-Measure	AuC
K-nearest neighbor	0.9743	0.7565	0.6645	1.1678	0.8923
Decision tree	0.8834	0.7434	0.6422	1.1221	0.8744
Q-learning	0.9899	1.021	0.5643	0.6755	0.9833

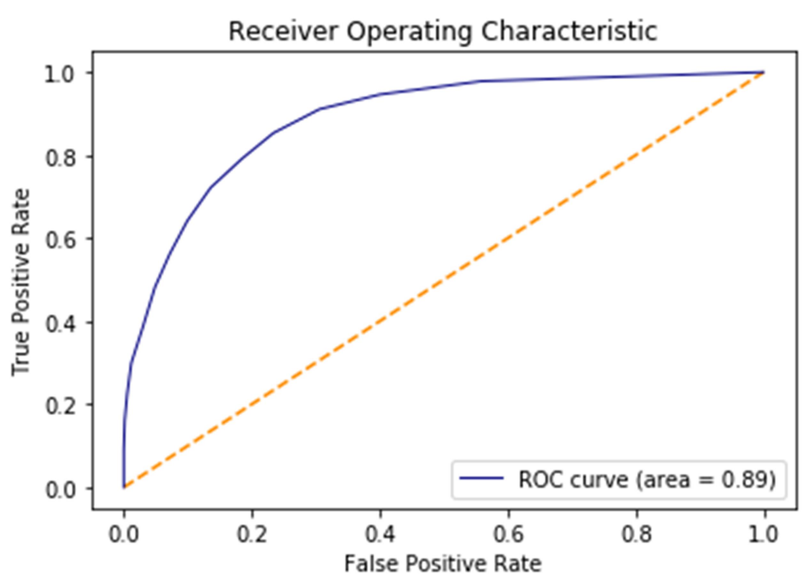

Figure 3: KNN ROC graphic

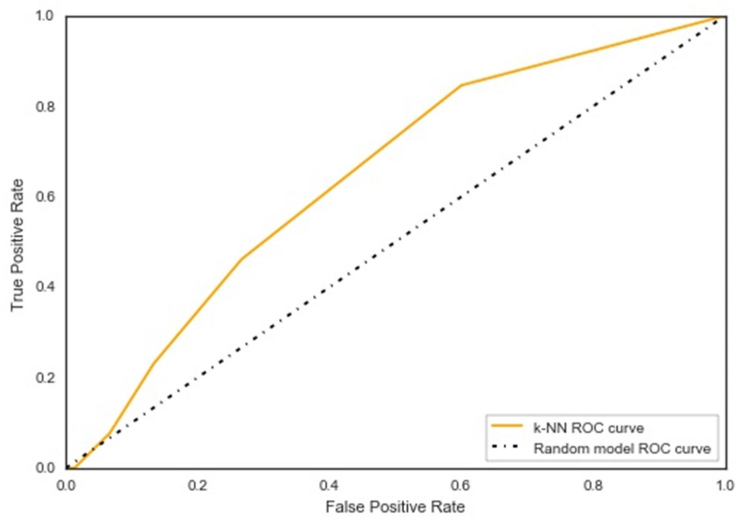

Figure 4: Shows a decision tree ROC chart

CONCLUSIONS

If DT was identified early enough, complications with the circulatory system, kidneys, brain, retina, nerves, and skin, as well as pregnancy, sleeping, hearing, and external genitalia, could be postponed or avoided. Machine intelligence technologies are frequently expected to have a significant impact on the healthcare industry and supervised learning has demonstrated living person or indeed outcome vector or dependability of world to clinical information processing. This model can help specialist doctors in reducing treatment effectiveness time by identifying DT with a 0% false-positive error using only some crucial behavioral data of a subject. This investigation could determine the hyperglycemic status of patients who are extremely close to acquiring positive DT, in addition to DT prediction. Future research could expand this approach to include additional characteristics that cause diabetes, such as hormone, diabetic pedigree performance, and hypertension, as well as bigger information. This approach could also be utilized to detect various types of disease utilizing EHRs by successfully altering the space vector specification.

REFERENCES

[1] Nordhaug Myhre J, Tejedor M, Kalervo Launonen I, El Fathi A, Godtliebsen F. In-Silico Evaluation of Glucose Regulation Using Policy Gradient Reinforcement Learning for Patients with Type 1 Diabetes Mellitus. Applied Sciences. 2020 Jan;10(18):6350.

[2] Logé F, Pennec EL, Amadou-Boubacar H. Challenging common bolus advisor for self-monitoring type-I diabetes patients using

Reinforcement Learning. arXiv preprint arXiv:2007.11880. 2020 Jul 23.

[3] Zheng H, Ryzhov IO, Xie W, Zhong J. Personalized Multimorbidity Management for Patients with Type 2 Diabetes Using Reinforcement Learning of Electronic Health Records. arXiv preprint arXiv:2011.02287. 2020 Oct 29.

[4] Kahkoska AR, Pokaprakarn T, Alexander GR, Crume TL, Dabelea D, Divers J, Dolan LM, Jensen ET, Lawrence JM, Marcovina S, Mottl AK. The Impact of Racial and Ethnic Health Disparities in Diabetes Management on Clinical Outcomes: A Reinforcement Learning Analysis of Health Inequity Among Youth and Young Adults in the SEARCH for Diabetes in Youth Study. Diabetes Care. 2022 Nov 1.

[5] Zhu T, Li K, Herrero P, Georgiou P. Basal glucose control in type 1 diabetes using deep reinforcement learning: an in silico validation. IEEE Journal of Biomedical and Health Informatics. 2020 Aug 5;25(4):1223-32.

[6] Liu Z, Ji L, Jiang X, Zhao W, Liao X, Zhao T, Liu S, Sun X, Hu G, Feng M, Xie G. A Deep Reinforcement Learning Approach for Type 2 Diabetes Mellitus Treatment. In2020 IEEE International Conference on Healthcare Informatics (ICHI) 2020 Nov 1 (pp. 1-9). IEEE.

[7] Zhu T, Li K, Georgiou P. Personalized dual-hormone control for type 1 diabetes using deep reinforcement learning. InExplainable AI in Healthcare and Medicine 2021 (pp. 45-53). Springer, Cham.

[8] Manoja, I., Sk, N. S., & Rani, D. R. (2017, March). Prevention of DDoS attacks in cloud environment. In 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC) (pp. 235-239). IEEE.

[9] Dr.P.Sivakumar an Mr.K Rajagobal, ‘Object Based Ring Routing Path Management Algorithm for Energy Efficient Nest Node of Sensor Network’, Journal of Computer and System Sciences, Volume 83, Issue 3, ISSN: 0022-0000, 2017, Pages 3-21. (Scopus Indexed))

[10] Dr.P. Sivakumar, ‘Efficient Job Scheduling of Genetic Algorithm with Tabu Search and Round Robin’, International Journal of

Printing, Packaging & Allied Sciences, (ISSN 2320-4387), vol. 4, no. 4, pp. 2864-2878,2016.

[11] Latchoumi, T. P., & Parthiban, L. (2021). Quasi Oppositional Dragonfly Algorithm for Load Balancing in Cloud Computing Environment.

[12] Chinnamahammad Bhasha, A., & Balamurugan, K. (2021). Studies on Mechanical properties of Al6061/RHC/TiC hybrid composite. International Journal of Lightweight Materials and Manufacture.

[13] Pavan, M. V., Balamurugan, K., Srinivasadesikan, V., & Lee, S. L. (2021). Impact and Shear Behavior of PLA/12% Cu Reinforced Composite Filament Printed at Different FDM Conditions. Arabian Journal for Science and Engineering, 1-12.

[14] Latchoumi, T. P., Vasanth, A. V., Bhavya, B., Viswanadapalli, A., & Jayanthiladevi, A. (2020, July). QoS parameters for Comparison and Performance Evaluation of Reactive protocols. In 2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE) (pp. 1-4). IEEE.

[15] Yamagata T, O'Kane A, Ayobi A, Katz D, Stawarz K, Marshall P, Flach P, Santos-Rodríguez R. Model-Based Reinforcement Learning for Type 1Diabetes Blood Glucose Control. arXiv preprint arXiv:2010.06266. 2020 Oct 13.

[16] Tejedor M, Woldaregay AZ, Godtliebsen F. Reinforcement learning application in diabetes blood glucose control: A systematic review. Artificial Intelligence in Medicine. 2020 Apr 1;104:101836.