iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44208

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Playing Tetris with Reinforcement Learning

Dr. M. Shailaja[4], Nune Vinaya Reddy[1], Ambati Srujani[2], Cherukuthota Upeksha Reddy[3]
[1, 2, 3, 4]*Department of Electronics and Computer Engineering, Sreenidhi Institute of Science and Technology Yamnampet, Ghatkesar, Hyderabad, India*

Abstract: ***The essential inspiration for this uAndertaking was a pleasant utilization of AI. Tetris is a notable game that is cherished and loathed by a lot of people. Tetris game has a few qualities making it an intriguing issue for the field of ML. A total portrayal of the tetris issue incorporates tremendous number of states making a meaning of a non-learning procedure for all intents and purposes unthinkable. Late outcomes from the group at Google DeepMind have shown that support learning can have noteworthy execution at game playing, utilizing a negligible measure of earlier data about the game. We use support figuring out how to prepare an AI specialist to play tetris. Support learning permits the machine or programming specialist to gain proficiency with its conduct in light of the criticism that is gotten from the climate. The machine might adjust after some time or may advance once and proceed with that behavior. Tetris is played on a rectangular lattice divided into more modest square regions, regularly ten units wide by twenty units tall. The player controls the direction and even area of pieces that tumble from the highest point of the board to the base and procures focuses by framing total level lines, which are then eliminated from play, causing pieces put higher to move descending. The key speculation of this undertaking is that assuming that the focuses procured in Tetris are utilized as the prize capacity for an AI specialist, then that specialist ought to have the option to figure out how to play Tetris without other oversight.***

I. INTRODUCTION

The essential inspiration for this task was a fun ap-plication of brain organizations. Tetris is a notable game that is cherished and despised by a larger number of people. Late outcomes from the group at DeepMind have demonstrated the way that brain organizations can have momentous execution at game playing, utilizing a min-imal measure of earlier data about the game. Be that as it may, there has not yet been a fruitful endeavor to utilize a convo-lutional brain organization to figure out how to play Tetris. We desire to fill this hole.

Besides, examination into tackling control issues with brain organizations and negligible unequivocal featurization might have applications beyond computer games, like in self-driving vehicles or advanced mechanics.

Tetris is a game that includes putting pieces on top of one another so they fill in a rectangular framework. A typical human way to deal with this issue includes searching for open spaces that match the state of the ebb and flow piece in play. This is an eventually an undertaking in visual example identification and item acknowledgment, so we accept that a convolutional brain net-work is a characteristic way to deal with taking care of this issue. Tetris likewise includes a fair level of procedure, and late advances in profound support learning have shown that convolutional brain organizations can be prepared to learn technique. Machine Learning: Game AI has been filling over the most recent couple of years as a subject of interest both for the modern and scholastic communities. Most games have a choice parts of some sort, furthermore, AI procedures are progressively used to carry out them. Case based thinking (CBR) is predicted as a candidate approach as it would give a proficient mean to remember gaming encounters that could later be reused by some non-player characters (NPC). Anyway this acquaints a few difficulties with the current best in class of CBR. Cases generally address individual episodes of critical thinking occasions. As a large portion of the games are consecutive in nature, a simply wordy representation would just portray halfway circumstances. Additionally numerous strategic and receptive games force a few types of time dependent impediments on a game AI part. Not many drives have been taken to present time requirements in CBR thinking. Proficient administration of the case base would be one of the main pressing concerns to influence on the performance of CBR frameworks because of its closest neighbor style of recovery. The board arrangements would be expected to handle these issues. In this paper, we examine how much reinforcement learning (RL) can add to the administration of a heritage case base containing neighborhood designs for playing a sequential game. We utilize the round of Tetris as an application because of the dimensionality of the choice space which makes it a complex however manageable issue to tackle. Moreover this game present the fascinating property of being subject to time impediments as choices should be taken before a dropping figure arrives at the outer layer of the board. Our advantage for this work is on the assessment of earlier cases. We accept that cases are now given to the CBR framework either by an outer program or by some human player.

©IJRASET: All Rights are Reserved

Through disconnected preparing, we utilize the support values to survey the nature of the cases and to direct the forgetting of cases to control the size of the casebase. This contrasts from support learning endeavors in CBR expecting web based learning without earlier information (Sharma et al., 2007) (Gabel et al., 2005).

In the following segments of this paper, we propose a representation for organizing neighborhood designs for playing Tetris with a CBR part. We then, at that point, survey the commitment of support figuring out how to assess the nature of the patterns. We contrast Q-learning and a straightforward transient difference definition without limiting to gauge the worth of every individual example. We at last play out some tests for neglecting cases and assessing what level of corruption can be seen by lessening the size of the case base involving support values as a pruning criterion.

A. Related Work

One of the first works in the field of significant help paper was DeepMind's 2013 paper, Playing Atari with Deep Reinforcement Learning . This paper and their 2015 resulting filled in as our fundamental inspiration, both for the general idea applying PC vision to help learning for PC games and for contemplations on the most capable technique to go on with our model plan. Their results are extraordinarily perfect and bleeding edge, with achieving divine execution on more than around half of the games attempted and coming up short essentially on games that require long stretch key arrangement and have outstandingly lacking awards, similar to Montezuma's Revenge. A couple of the techniques and our paper use have essentially longer stories. For example, experience replay, which stores the game-playing expert's experiences to be arranged long after the primary technique has been abandoned, was introduced in .RMSProp, an improvement strategy that typically reinforces learning rates down after some time and spotlights its learning rates on viewpoints with extra consistent and subsequently instructive slants, was first introduced proposed at this point didn't execute zeroed in on clearing as a technique for picking saved advisers for get ready on; we did this in our paper. Zeroed in on clearing, which readies the model on the models with the greatest mix- up, was introduced ;an even more computationally capable variation, which we didn't execute anyway would consider as a development, was introduced .Finally, one smoothing out strategy we investigated various roads in regards to yet didn't determine was Adadelta. Adadelta, which is less hyperparameter-subordinate than other improvement systems and reinforces convincing learning rates down considering how much a given perspective has proactively been revived, was introduced .Finally, it is vital that previous undertakings to manufacture experts fit for playing tetris all usage express featurization rather than going on from unrefined pixel data how we do. A couple of models are outlined .These previous techniques all usage features interfacing with the structure and level of the game board and the amount of openings present, so the heuristic limit that we used to help with setting up our association can be viewed as a superior on variation structure on the work in these paper.

II. LITERATURE SURVEY

Tetris is a very renowned game that was made in 1985 by Alexey Pajitnov and has been ported to basically every functioning structure and hardware stage in presence. The incomprehensible reputation and non-insignificant nature of the game have incited a ton of investigation, both into the math's enveloping the game and into the readiness of AI components to play it competently. The ease of showing Tetris, close by its unavoidable presence across systems, credits it to AI research. Support learning is a well established1 part of fake finding that isolates itself from various collections of man-made thinking in its accentuation on trial and error learning, and is a marvelously powerful area of assessment. It ensures a fair- minded expert fit for cultivating own procedures and recognizing designs are unquestionable in its relationship with the environment. This composing overview analyzes the pertinence of help sorting out some way to Tetris. A legitimate importance of Tetris is presented, close by a part of its intricacies. Existing man-made thinking Tetris players are furthermore explored, close by the viable applications and more significant attributes of help learning.

It will undeniably be entrancing to benchmark the expert against existing AI procedures, yet the obstructions constrained by the above subtleties will be waivered in confined assessments concerning the complete RL trained professional. The customary assurance of Tetris essentially reduces the complexity of the game away from the state of the art sign normally experienced by human players, and it will be interesting to see the expert being loosened up to adjust to a human level test. Support learning portrays a method for managing handling issues instead of deciding all of the intricacies drew in with dealing with the issue through to execution. It is described to the extent that an expert teaming up with an environment. The expert's perspective on the environment is encapsulated in a value limit, which vehicles the different states the environment can exist in, and accomplices an aggregate worth with each state. This value limit is revived subsequent to getting analysis, described by the award limit, from the environment.

This prize limit is statically articulated toward the start of an issue and is past the effect of the subject matter expert, and consequently directs the improvement of the value limit. It is crucial for observe that prizes can be either negative or positive, deflecting or enabling the expert suitably. The expert follows a methodology that guides states to exercises, and collaborates with the value limit in coordinating the approach to acting of the trained professional. The target of the expert is to intensify long stretch total honor. Its fundamental approach to acting is just trial and error driven, yet as the expert shapes an impression about the states, and their relative advantages, it ends up being logically huge for it to figure out some sort of agreement between the examination of new states which could give generally outrageous award, and the maltreatment of existing data. Support learning can be applied in non-deterministic circumstances, where taking a particular action inside the setting of a state doesn't expected lead to a comparable award or same state change. It does, regardless, expect that the environment be fixed and that the probabilities of getting a particular award or changing to a particular state go on as in the past.

Tetris has been the subject of much solicitation concerning its multifaceted design. It is known to be NP-Complete (Demaine, Hohenberger, Liben-Nowell, 2003), and that a long gathering of Z or S tetrominoes will cause any Tetris game to be lost (Burgiel, 1997). No victorious framework has been portrayed, yet every player most certainly has their own methodology. Many ventures exist to play Tetris, yet of explicit note is one made by Colin Fahey, who achieved standing in 2003 for making a very strong Tetris playing subject matter expert. Fahey did a version of Tetris considering the fundamental PC release, and made a design by which it might be worked by either a human player, or by an expert planned to simply be associated with the deliberate framework1

A changed version of Fahey‟s Tetris is used for this preliminary. Fahey‟s subject matter expert, given a board arrangement, considers possible board plans made by placing the right presently falling piece in every single under the sun region (known as translations) and headings (turns). Each made board is given out an authenticity considering characteristics including the general level of the pile of pieces, and the amount of covered, void openings in the load. These characteristics are weighted not completely settled by Fahey through a course of organized and randomized looking. Moreover using Fahey‟s framework is Pierre Dellacherie, whose expert correspondingly researches possible moves and surveys the resulting sheets considering a weighted plan of characteristics.As stood out from Fahey, Dellacherie made a couple of new estimations of evaluating a board plan, and picked his heaps by hand after genuinely contrasting them. While Dellacherie‟s agent‟s evaluation estimations are hand-picked, Colin Fahey saw that Roger Llima involved the AI methodology for genetic computations to further develop the heaps of the different appraisal estimations available to the experts in his framework (Fahey, 2003). This methodology for AI has as of late been applied to Tetris (Siegel, Chaffee, 1996), but had not as of late been used for figuring metric burdens.

III. METHODS USED

A. *Introduction To Uml*

This Uml empowers a computer programmer to verbalize a logical structure utilizing demonstrating documentation that follows a bunch of syntactic, semantic, and realistic rules. Five separate perspectives are utilized to show an UML organization, every one of which portrays the system from a more extensive perspective. An assortment of drawings, that are as following, portrays each perspective.

1) *Portrayal of the Simulation Framework:* As according to client's perspective, this is the manner by which the framework would look. According to the perspective of such gotten done, it insightful portrayal portrays genuine necessity.
2) *Image of the Tower Is planned:* Both data or capacities all through this worldview come from inside the establishment. The fixed parts are demonstrated consequently in displaying layer.
3) *Image of a Modeling Instrument:* It portrays the associations of social affair among stiffeners characterized inside the particular client and corroborative variable view, addressing the liquid of conduct as various subsystems.
4) *Point of view of the Methodology:* The framework's dynamic pieces of the product are portrayed as underlying parts in this outline.
5) *Picture of the Ecological Simulator:* The matrix's institutional and functional perspectives.

B. *Uses of UML*

A typical use for composing programming plans is the Universal Design Methodology (Extensible markup language). The Xml is a demonstrating language.

1) Imagining
2) Explaining
3) Laying out
4) Monitoring an application framework's curios.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

The Plm is a particular viewpoints that incorporates a word reference and rules for blending things from that word reference. A demonstrating language called that spotlights on the psychosomatic portrayal of a framework through its dictionary and rules. Reenactment gives knowledge into a cycle.

C. Building Blocks of the Uml

A Pcm language comprises of three fundamental of essential parts:

- Items
- Connections
- Patterns

Things will be the main occupants in a model; associations associate them; and portrayals orchestrate important groupings of items on the whole.

Things in the UML

With in Plm, there appear to be 4 sorts of items:

- Items with a construction
- Elements that influence your way of behaving
- Putting things in gatherings
- Items with explanations

These action words inside Uml graphs are building objects. The accompanying structure individuals were utilized for the arranging stage: First or premier, any subclass is a meaning of a gathering of items with comparable properties, strategies, associations, and implications. At long last, a utilization perfect representation of a rundown of steps performed by a cycle that guarantees in a recognizable result of significant worth to a specific player.

Investigations Thirdly, a group is a realtime fundamental piece that represents a registering limit, with at least restricted stockpiling and often strong processors.

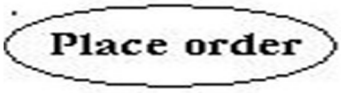


Fig 1: Analysis

Thirdly, a bunch is a realtime fundamental part that represents a figuring limit, with at least restricted stockpiling and as often as possible strong processors.

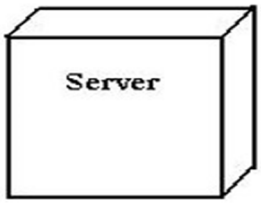


Fig 2: End points

The dynamical components of Modeling are conduct qualities. Coming up next is an illustration of a way of behaving:

Association:

An exchange is a kind of conduct where a bunch of texts is traded among a gathering of items in a restricted example to accomplish a particular objective. Both these parts of an interchange incorporate texts, battle scenes (the conduct set off by a text), and associations. Connections in the UML

Inside the Plm, there's numerous four sorts of associations:

- Reliance
- Association
- Clearing articulation
- Reality

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

A reliance is a lexical organization of associated substances wherein the translation of one thing can be impacted by the implications of the other (the reliant thing).

A bunch of associations is portrayed by a connection, which is a shape and volume that alludes to a bunch of connections. A hyperlink is a connection between things. Bunching is a connection between two that addresses a reliance structure among entire rather and its constituent pieces.

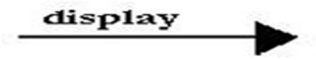


Fig 3: Messages

A speculation is a mastery association wherein the specific thing's (youngsters) things can be subbed with the extended thing's (father) things (the parent).

An acknowledgment is a calculated connection among students wherein 1 characterization gives a commitment that is destined to be satisfied by another grouping.

IV. PROPOSED SYSTEM

A. Issue Analysis

- Express The quantity of conditions of a Tetris board is an enormous on the off chance that we record whether the stones exist in each container. The quantity of conceivable outcomes is around 220*9 hence it is difficult to save the diagram of the board in the data set to choose the following ideal move of the AI. To stay away from the issue of memory use, we address the conditions of the game by various elements of the board:
- Total amount of the pinnacle of every section: Heuristically, this component of the board ought to be limited for forestalling losing because of the blocked board.
- Emptiness of the board: This element count the number of void blocks under a piece of filled block (stone). All in all, it is the complete amount of the pinnacle less the absolute amount of the filled block. This component ought to likewise be limited as the vacancy would impede the freedom of columns of blocks.
- Total number of cleared lines: Obviously and heuristically, this element can address how well players act in the game and it ought to be augmented all through the game.
- Surface harshness: This component evaluates the unpleasantness of the top surface of the blocks. Heuristically, smoother the top surface of the blocks is, simpler to get pushes free from stone it is. Additionally, this element ought to be limited for facilitating the filling of holes in the game. With these highlights, we can sum up the territories of Tetris to straight mix of factors, which is effectively to be learned for AI.
- Activity For each piece of stone, the decisions of activity are the last areas of the stone in the wake of dropping. We bunch the activity instead of recreate the keys played by human. Subsequently, we formalized the activities by addressing the bearing and number of times it moves and rotatio Tetriminos, at last immediately drop the Tetriminos without pausing.
- Assessment work Value of the board = $\sum$ weightingi x featurei Using the component recently referenced, we utilized a direct capacity to address the worth of the board, with the weightings for each element. For the explanation of investigation, the weightings are randomized at the beginning phase. Furthermore, in the later piece of the game, the weightings follow the upgraded assessment work in the support advancing piece of the AI.
- Ideal activity To decide each ideal development of the Tetriminos, we search the worth of the board involving the assessment work for all finals imaginable area for each new dropping stone.

The AI will compute the worth of the board subsequent to putting the stone in each conceivable loaction utilizing the assessment work above. Also, it acts the most ideal activity with the most ideal board esteem. As the following Tetriminos is given to the player, AI can streamline the move via scanning the worth of the ongoing board for current stone and next stone. Nonetheless, it is tedious for AI to look through every one of the outcomes because of the enormous number (>900) of conceivable outcomes. It prompts what is going on that the AI will invest energy to work out the ideal move while it is holding up in the game, which isn't effective. Thusly, we choose to separate three best maneuvers for flow stone and looking through their greatest worth in view of the following stone to choose the ideal development. It saves a huge piece of looking through time on the grounds that the vast majority of the choice isn't ideal and we disregard them in the thought. Thus, it can work on the looking through speed of the AI actually.

©IJRASET: All Rights are Reserved

B. Preparing Agent using Reinforcement Learning

As recently said, there are various practical arrangements. The most well known, then again, permit the PC program to pick a movement that will enhance the extensive result . The extent of such techniques is known to be limitless. By and by, this is achieved through figuring out how to evaluate a state's worth. This assessment is refreshed over the course of time by spreading a part of the prize from the accompanying state. In the event that the states and acts are all attempted a sufficiently huge number of times, an ideal technique still up in the air; the activity that boost the value of another state is picked.

As often as possible, an office should contemplate a running cycle or has no clue about the number of exercises it that will be taken to wrap up. Limitless skyline challenges happen when the action might go on endlessly, or vague skyline blunders emerge when the office will at last stop yet doesn't see when. To address these situations, we add activities to the Dynamical framework. At each progression, the office picks which activity to do; the last not entirely settled by this choice. the earlier condition as well as the activity taken You would rather not ponder the utility right at end of an interaction in the event that it's actually running on the grounds that the specialist may in all likelihood never have one. An administrator can be given a progression of prizes. These prizes incorporate the expenses of the movement, as well as any prizes or fines that might be given. Sanctions are the name for negative motivators. An ending point can be utilized to demonstrate boundless reach concerns. At the point when a specialist is in a stopping or retention express, movements of every sort have no effect; that is, the point at which the office is in that country, movements of every kind get once again to that country with a 0 prize.

C. MARKOV Decision Process

A Markov decision processor a MDPconsists of

- S,a set of states of the world.
- A,a set of exercises.
- P:S×S×A→[0,1], which shows the components. This is made P(s'|s,a), where
- ∀ s ∈S ∀a ∈A ∑ P(s'|s,a) = 1. s'∈S
- In particular, P(s'|s,a) decides the probability of changing to state s' given that the expert is in state sand exercises a.
- R: :S×A × S→R, where R(s,a,s') gives the ordinary brief honor from doing movement aand changing to state s'from state s.

Both the components and the awards can be stochastic; there can be a few haphazardness in the ensuing state and prize, which is shown by having a dissemination over the ensuing state and by R giving the typical honor. The outcomes are stochastic when they depend upon unpredictable variables that are not shown in the MDP.

V. DESIGN FLOW

In like manner with decision associations, the fashioner also needs to consider what information is available to the expert when it picks what to do. There are two typical assortments:

In a totally detectable Markov decision cycle, the expert will see the current status while picking what to do.

A somewhat detectable Markov decision connection (POMDP) is a blend of a MDP and a mysterious Markov model.At each time point, the expert will make reference to a couple of genuine realities that depend upon the state. The expert simply moves toward the verifiable scenery of discernments and past exercises while seeking after a decision. It can't clearly see the current status.

To pick to do,the expert's opinion on different groupings of compensations. The most notable strategy for doing this is to change over a gathering of compensations into a number called the value or the joined award. To do this, the expert gets a brief honor together with various awards from this point forward. Accept the expert gets the gathering of compensations:

r1,r2,r3,r4,....

There are three notable approaches to merging pay into a value V: Complete Prize:

V = ∑ ∞ r. For this present circumstance, the value is how much the remunerations overall. This works
i=i

exactly when you can guarantee that the total is restricted; yet accepting the total is unending, it offers zero chance to break down which progression of compensations is ideal. For example, a gathering of $1 rewards has a comparable complete as a progression of $100 rewards (both are boundless). One circumstance where the award is restricted is when there is an ending state; when the expert by and large has a non-no probability of entering a stopping state, the total honor will be restricted.

Normal Award:

V = lim (r+⋯+r)/n. For this present circumstance, the expert's worth is the ordinary of its n→∞ 1 n rewards, showed up at the

midpoint of all through for each stretch of time. However lengthy the awards are restricted, this value will in like manner be restricted. Regardless, whenever the total award is restricted, the ordinary award is zero, in this way the typical honor will forget to allow the expert to pick among different exercises that each have a zero common award.

Under this standard, the primary thing that matters is where the expert breezes up. Any restricted progression of stunning demonstrations doesn't impact the limit. For example, getting $1,000,000 followed by compensations of $1 has a comparable commonplace award as getting $0 followed by compensations of $1 (both of them have an ordinary remuneration of $1).

VI. RESULTS

Result will be displayed after an episode completes. Following are the parameters displayed after each episode.

1) Alpha for currently running episode
2) Score obtained in currently running
3) Number of lines cleared
4) CHANGE_IN_PILE_HEIGHT
5) CHANGE_IN_HOLES

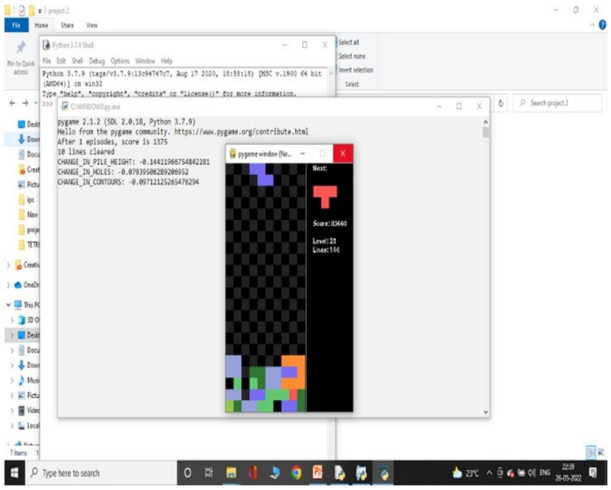


Fig 4: Result

VII. CONCLUSION

Support Learning is a discipline of Ai Technology which is an AI Method. It empowers robots other PC creatures that independently pick the best conduct in a provided circumstance to improve. Thusly for student to see new way of behaving, it needs just basic compensating input, which is called as the current year's building up message. These issue is tended to by an assortment of methods. As a general rule, a particular sort of issue characterizes Reinforcement Learning, a portion of its responses were named Reinforcement Learning calculations. Support profound learning used in an assortment of issues, and its expansiveness is wide enough where past endeavors to apply it to a circumstance had been inferable from execution blemishes. Prior disappointments uncovered the intricacy of taking care of Tetris, but they didn't demonstrate that support learning itself is unseemly or inadequate to dominate Tetris. Tetris is a reviving encounter which has ignited an arrangement of insightful, and it could benefit again from presence of an unlimited, conceivably capricious player. Support learning can be utilized in non-deterministic circumstances while making a remedial showing just in the system of a condition doesn't bring about a motivation or stage move. This may, in any case, require that one more environmental factors stay static so the possibilities getting a particular payout of changing to a particular condition exist.

VIII. FUTURE SCOPE

Request investigating controlling troubles with least evident objectives are achieved could of purposes past gamers, remembering for creative mind cars with robots.

Tetra is a tabletop game where the firm size could be changed.

Tetris game is Board game and size of the board can be modified. On scaling size up, states increment dramatically. This permits an improved execution utilizing brain organizations, profound learning and so on.

There are various potential enhancements in the space of this task. Learning can be extended to test numerous calculations and procedures. Tetris has a wide range of modes like Two- Piece Tetris (One piece look forward) Ill-disposed Tetris (Competitive 2-player Numerous line extra focuses Endurance Tetris (Last as far as might be feasible) Run Tetris (Clear 20 lines as quick as could really be expected)It would be fascinating to concentrate various adaptations of Tetris to check whether support learning is fit well for them or some other kind learning is best preferred.

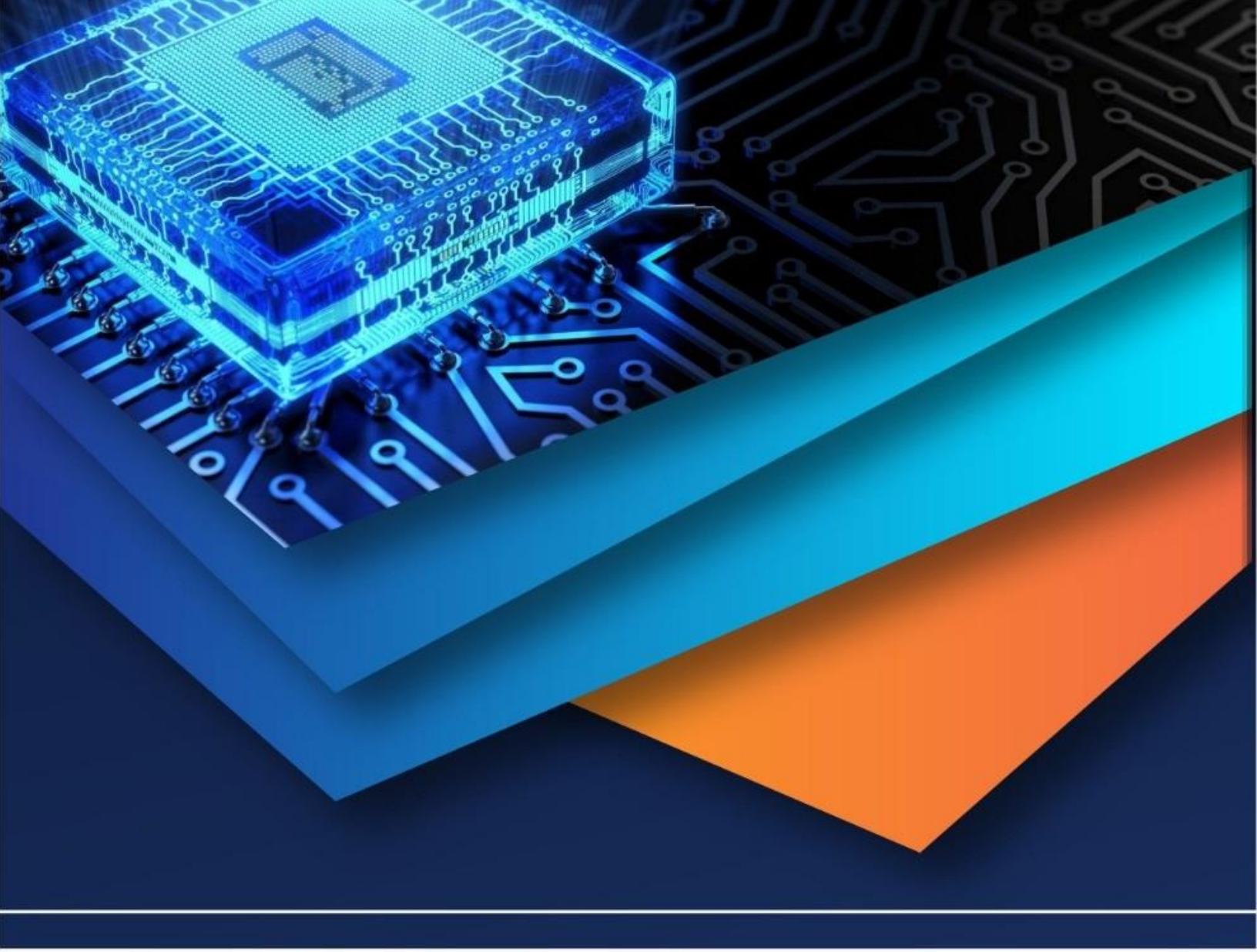
One more heading of development is speculation motor. This calculation can be upgraded to help gradual and social speculation motors. In the event that we have a specialist to play a game, applying this calculation to that specific game can be accomplished.

IX. ACKNOWLEDGEMENT

We'd like to express our deepest appreciation to everyone who has helped us create this project report. We'd like to thank Dr.M.Shailaja, an assistant professor in the department of electronics and computer engineering, for guiding us through the project. We appreciate your excellent advice and encouragement during the project, as well as your efforts to ensure that we function systematically. It is a wonderful pleasure for us to be able to collaborate with her.

REFERENCES

[1] GitHub. Matris – a clone of tetris made using pygame. https://github.com/smartViking/MaTris
[2] International Journal of Computational Geometry and Applications, vol 14
[3] L. Yiyuan. Tetris ai – the (near) perfect bot. https://codemyroad.wordpress.com/2013/04/14/tetris-ai-the-near-perfect-player/.
[4] D.Carr.Applying reinforcement learning to tetris. Department of Computer Science Rhodes University, 2005.
[5] Artificial Intelligence: foundatios of computational agents, Cambridge University 2017. http://artint.info/html/ArtInt_224.html#discounted-reward
[6] GitHub Matris - a clone of tetris made using pygame. https://github.com/SmartViking/MaTris
[7] GitHub : Q-learning for keras. https://github.com/farizrahman4u/qlearing4k
[8] Playing tetris with Deep Reinforcement Learning, M Stevens http://cs231n.stanford.edu/reports/2016/pdfs/121_Report.pdf



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)