# Analysis of AlphaGo Artificial Intelligence Program and Implementation.

AIM: To analyze the working of Google's Deep Learning powerful AI i.e. AlphaGo and discovering its implementations. Also, study various latest implementations and derive a hybrid model named as 'MinAlphaGo' with a better and faster user interface along with the analysis of game state at any instant.

Mr. Devender Kumar (Project Supervisor) Department of Information Technology Netaji Subhas Institute Of Technology, Delhi.

By:-

Aditi Kumar (706/IT/14) Arushi Bhatt (718/IT/14) Diksha Sharma (731/IT/14)

# **OBJECTIVE**

- Studying AlphaGo (Google's Deep Learning powerful AI).
- Study of Google's research paper titled as: 1.Mastering the game of Go with deep neural networks and tree search
  - https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf
- Combining plus points of BetaGo and Leela Game Engine implementations and coming out with hybrid version named as 'MinAlphaGo'.
- Comparing the interface of Leela and MinAlphaGo.

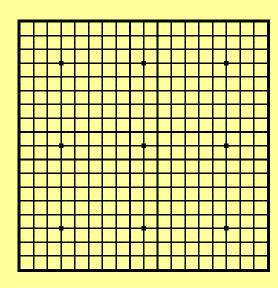
# WHAT IS GO GAME?

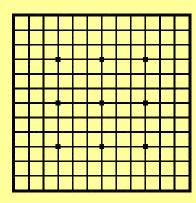
- -Go is an ancient(2500 yrs) Chinese two player abstract strategy board game(oldest board game).
- -Go has simpler rules than chess.

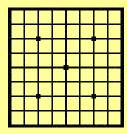


# More about go GAME:

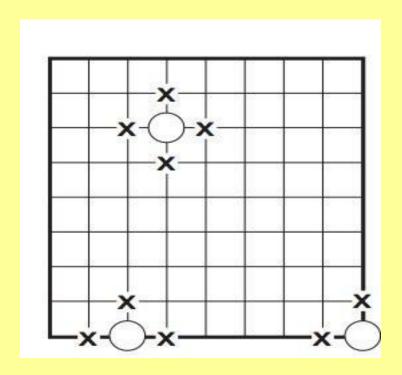
- Although the official size of a Go board is 19 x 19, but it can be played on a 13 x 13 board and 9 x 9 board.
- **Objective** is to capture more territories than the opponent and other players stone.

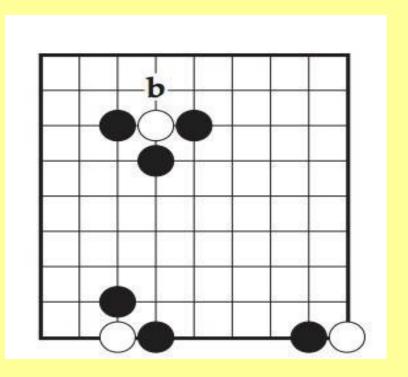




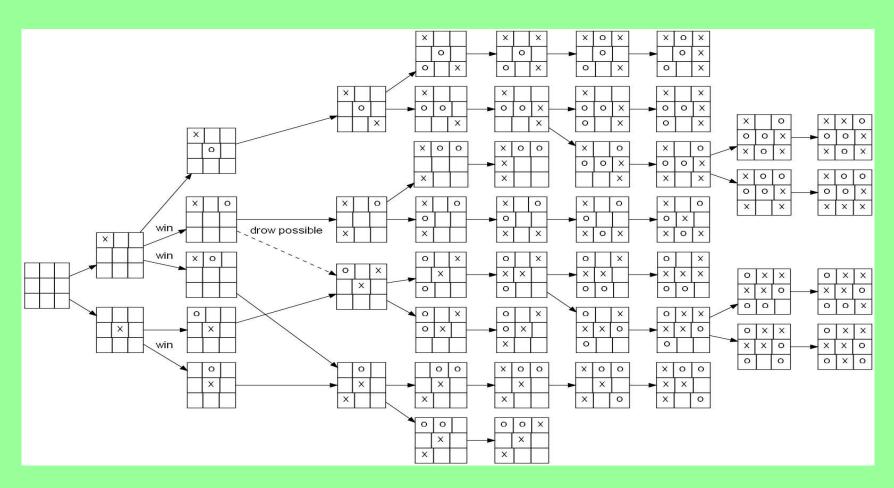


- A basic principle of Go is that a group of stones must have at least one liberty(open point bordering the group) to remain on the board.
- An enclosed liberty (or liberties) is called an eye, and a group of stones with two or more eyes is said to be unconditionally alive.
- open "point" (intersection) bordering the group
- Such groups cannot be captured, even if surrounded.
- A group with one eye or no eyes is "dead" and cannot resist eventual capture.





- To understand how Als are capable of playing games such as chess and Go, we have to understand what a game tree is.
- A GAME TREE represents game states (positions) as nodes in the tree, and possible actions as edges.
- The root of the tree represents the state at the beginning of the game. The next level represents the possible states after the first move, etc... For simple games such as tic-tac-toe, it is possible to represent all possible game states (the complete game tree) visually:



### **GO VS CHESS:**

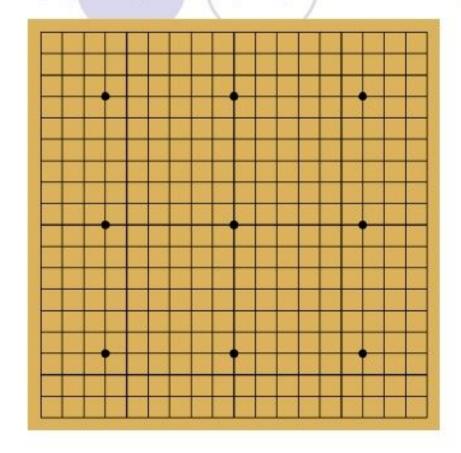
- Similar in some respects: both are played by two players taking turns, and there is no random element involved (no dice rolling, like in backgammon).
- In 1997, Garry Kasparov was defeated by **Deep Blue**, a computer program written by IBM, running on a supercomputer, Superficially, AlphaGo's win against Lee Sedol can be compared to Deep Blue's win against Gary Kasparov.

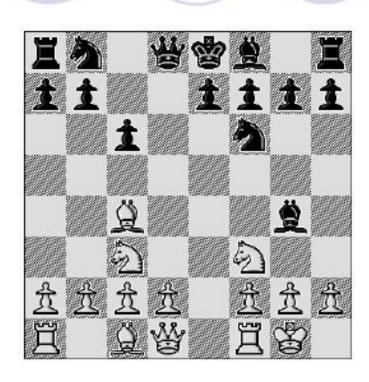
- In chess,
  - each player begins with 16 pieces of six different types.
    - Each piece type moves differently.
    - The goal of the game is to capture the opponent's king.
- Go starts with an empty board.
  - At each turn, a player places a stone (the equivalent of a piece in chess) on the board.
  - Stones all obey the same rules.
  - The goal of the game is to capture as much territory as possible. It can therefore be argued that **Go has** *simpler* **rules than chess.**
- In spite of the fact that the rules of Go might appear simpler than the rules of chess,
   the complexity of Go is higher. At each game state, a player is faced with a choice of a greater number of possible moves compared to chess (about 250 in Go vs. 35 in chess).

Because of this, the total number of possible games of Go has been estimated at 10<sup>761</sup>, compared to 10<sup>120</sup> for chess.

### **GO VS CHESS:**

# Game Board





19X19 lines vs. 64 squares (8 rows and 8 columns)

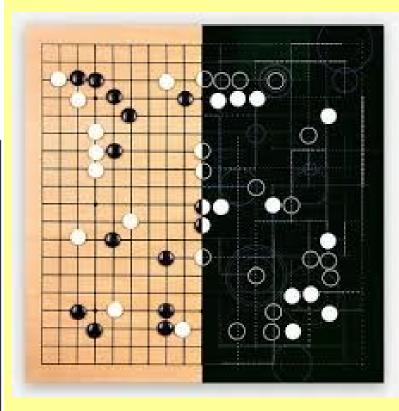
# What's interesting about Go?



- -Has been viewed as the most challenging of classic games for Artificial Intelligence owing to its enormous search space and the difficulty of evaluating board positions and moves.
- -Despite its relatively simple rules, Go is very complex, even more so than chess, and possesses more possibilities than the total number of atoms in the visible universe.
- -Number of atoms=10^80;
- -Number of possible games of chess=10^120;
- -Number of possible games of Go game= 10^761;







### What Is ALPHAGO?



- AlphaGo, an Al computer program which was developed by Alphabet Inc.'s
   Google DeepMind in London in October 2015, became the first Computer Go
   program to beat a human professional Go player on a official-sized 19×19 board.
- Instead of using brute force, alphago used reinforcement learning and Neural Network to mimic the leaning process of human mind.
- By combining deep learning and reinforcement learning in a series of artificial neural networks, AlphaGo first learned human expert-level play in Go from 30 million moves from human games.

- But then it started playing against itself, using the outcome of each game to relentlessly refine its decisions about the best move in each board position.
- Two networks are used,
  - A value network learned which is used to predict the likely outcome given any position,
  - while a **policy network** learned the best action to take in each situation.
- Also much of AlphaGo's power is based on a technique called back-propagation learning that helps it correct errors
- Alphago is made up of number of relatively standard techniques:-
  - Convolution Neural Network
  - Supervised Learning
  - Reinforcement learning
  - Monte Carlo Tree Search
  - Value Network
  - Policy Network and Fast Policy Network









# Study of the Google's Official Paper...



- 1.Mastering the game of Go with deep neural networks and tree search <a href="https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf">https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf</a>
- 2. Google Research Blog:

https://research.googleblog.com/2016/01/alphago-mastering-ancient-game-of-go.html

- 3. And also various articles and Blogs.
- 4. Relevant Courses on:
- Machine Learning Specialization by University of Washington.
- Deep Learning and Neural Networks by Co-founder of Coursera.
- Game Theory by David Silver.



## Deep Mind's AlphaGo Official Site

Nature Paper published on 28 Jan 2016:

https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf

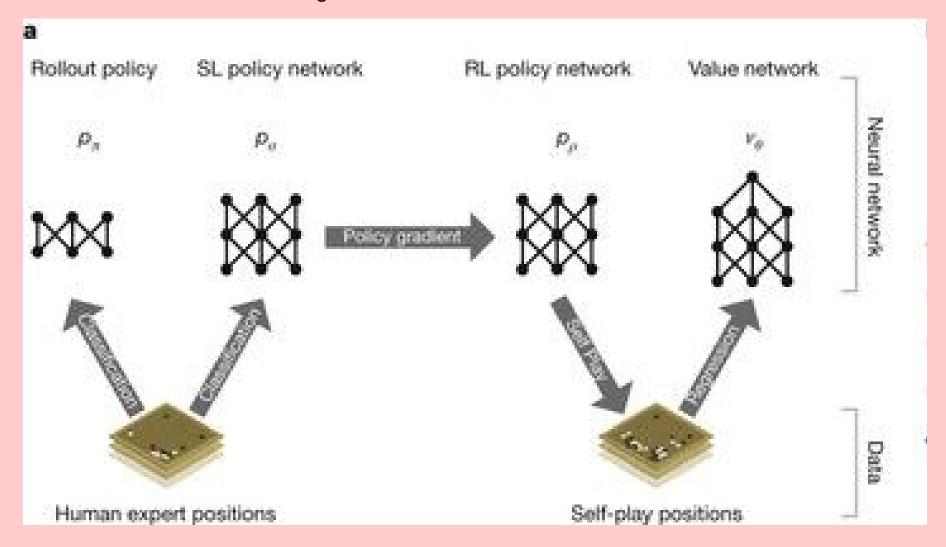
# Summary of the Research paper:

- New approach: uses 'Value Networks' to evaluate board positions and 'Policy Networks' to select moves.
- Combination of Supervised Learning and Reinforcement Learning
- New search algo: combines Monte Carlo simulation with value and policy networks (lookahead search)
- Using this search algo, AlphaGo achieved 99.8% winning rate.

- AlphaGo uses a SL policy to initialize the learning of an RL policy that gets perfected with self play, which they then estimate a value function from, which then plugs into MCTS that uses the SL policy to sample rollouts.
- Optimal Function: determines the outcomes of the game from every board position or state s.
- This optimal value function can be solved by traversing a search tree containing approx. b^d sequence.
   Where b-game's breadth (no. of legal moves per position) d-depth (game length)
- AlphaGo combines Monte Carlo simulations with value and policy networks called "Asynchronous policy and value MCTS" (APV-MCTS).

# **Training Pipeline**

- Supervised Learning of policy network
- Reinforcement Learning of policy network
- Reinforcement Learning of value network



- SL policy performed better in APV-MCTS than the stronger RL policy because SL policy better reflect the diverse beam of promising moves humans select, whereas RL optimizes for a single best move.
- In case of value function, the value function derived from the stronger RL policy network perform better than the value function derived from SL policy network.

# Our Contribution...

# Latest Works:

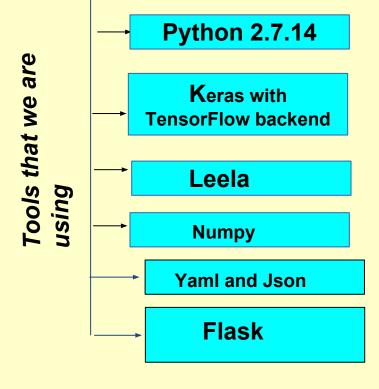
- BetaGO
- Leela Game Engine
- Rochester AlphaGo

We took 'BetaGo' AND 'Leela Game Engine' into consideration and have joined the plus points of the implementation and have come out with our newer version named as 'MinAlphaGo'

# PHASE 1

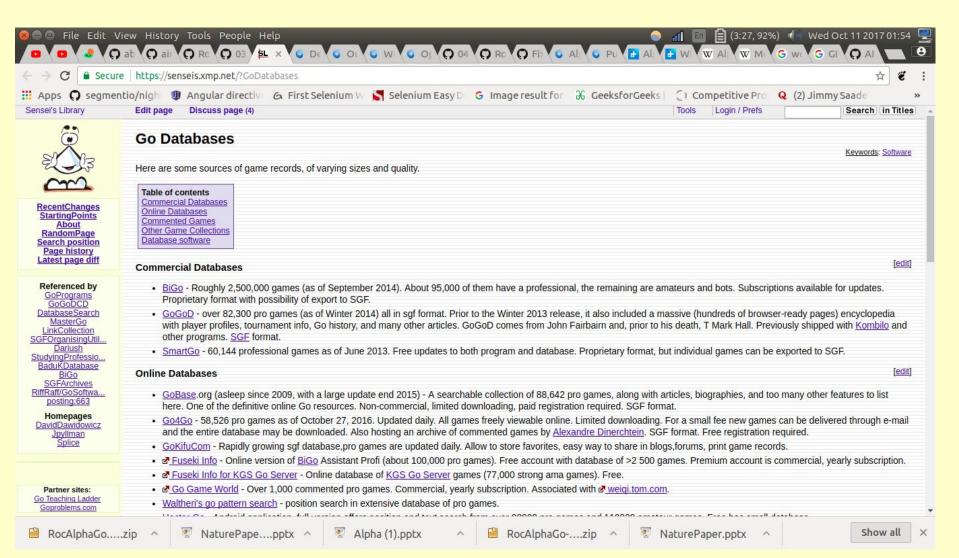
Contexual Learning Method and Exploration of Al WORLD

Implementation of smaller
Version of AlphaGo as
MinAlphaGo

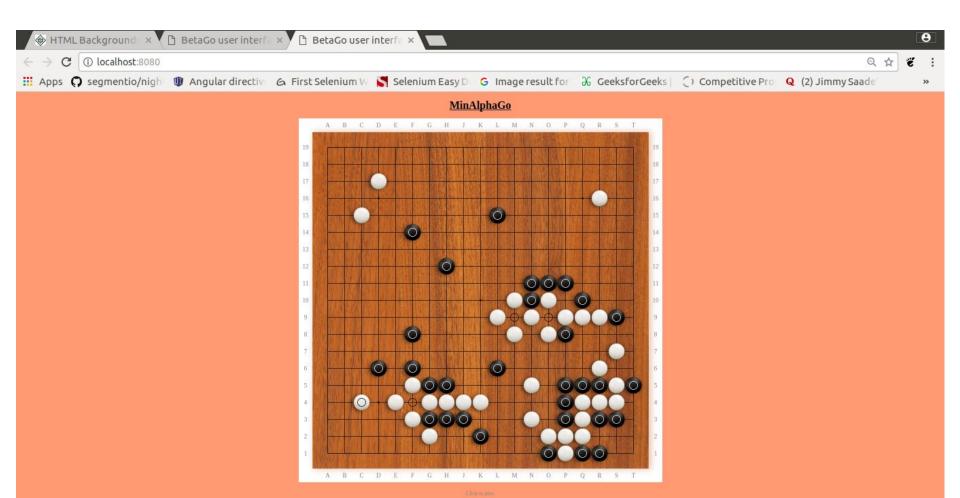


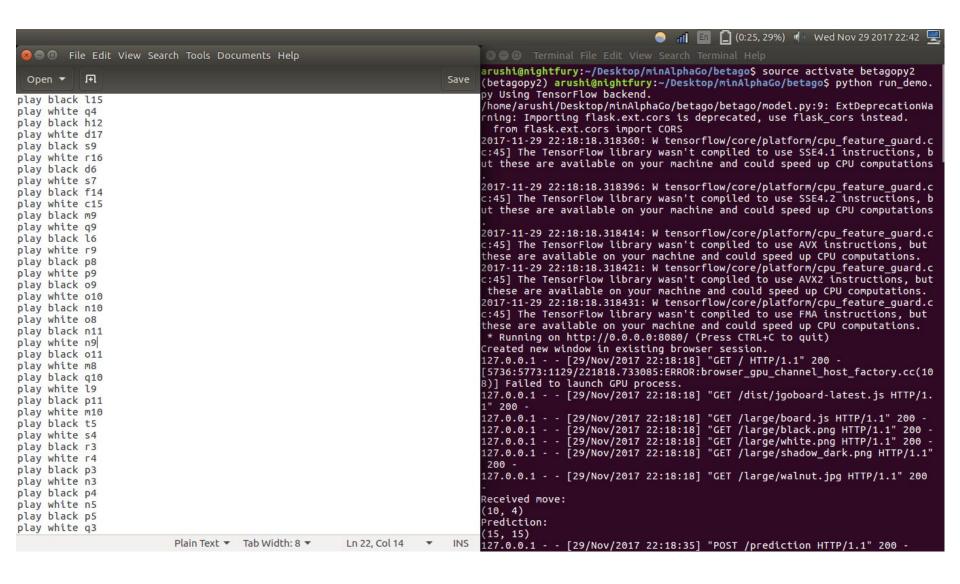
# **Data Source**

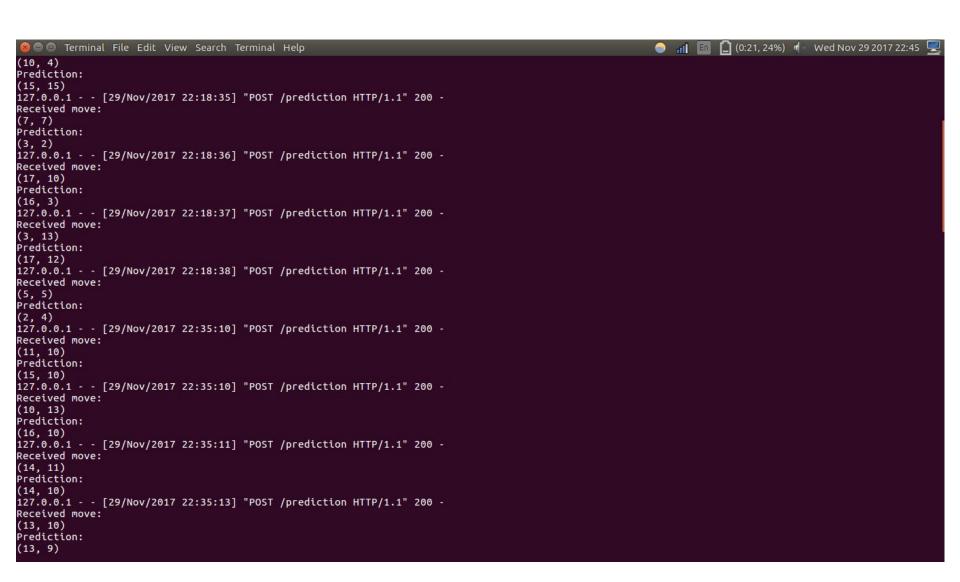
### https://senseis.xmp.net/?GoDatabases

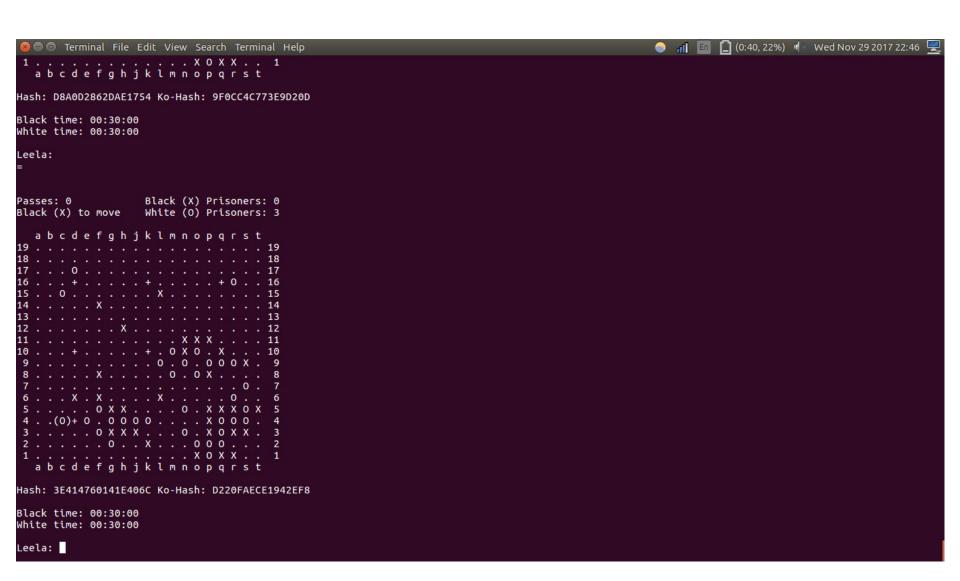


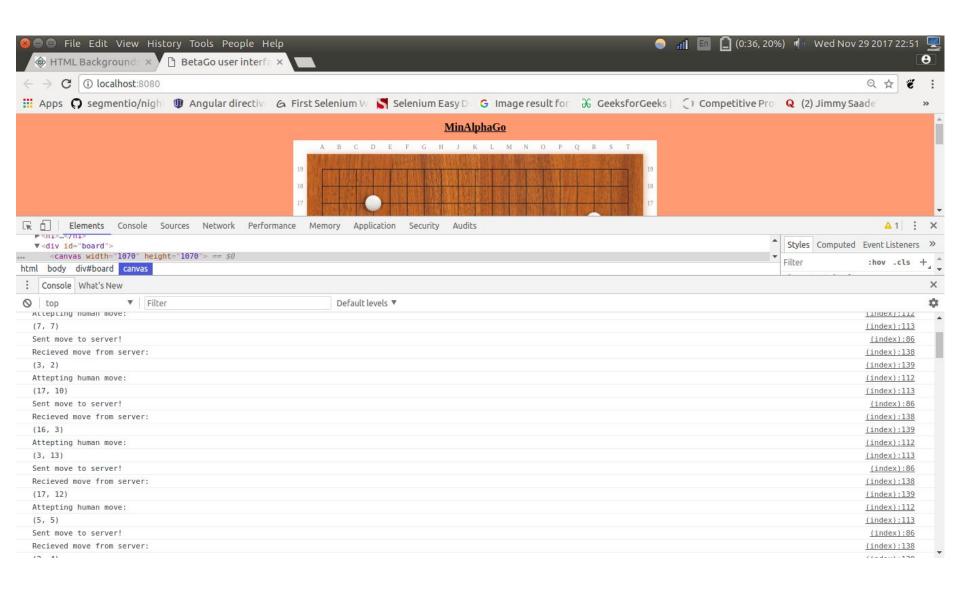
# ANALYSIS AND RESULTS



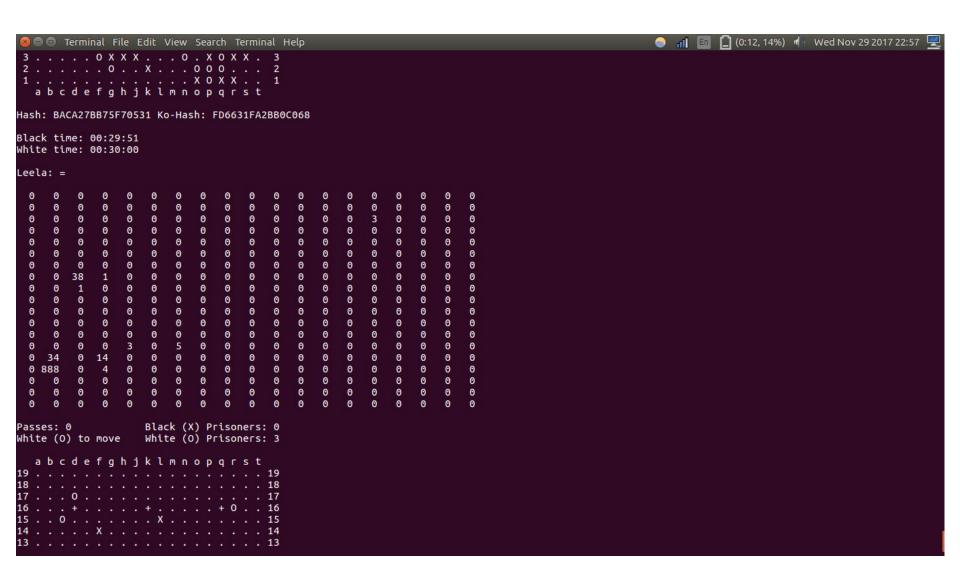


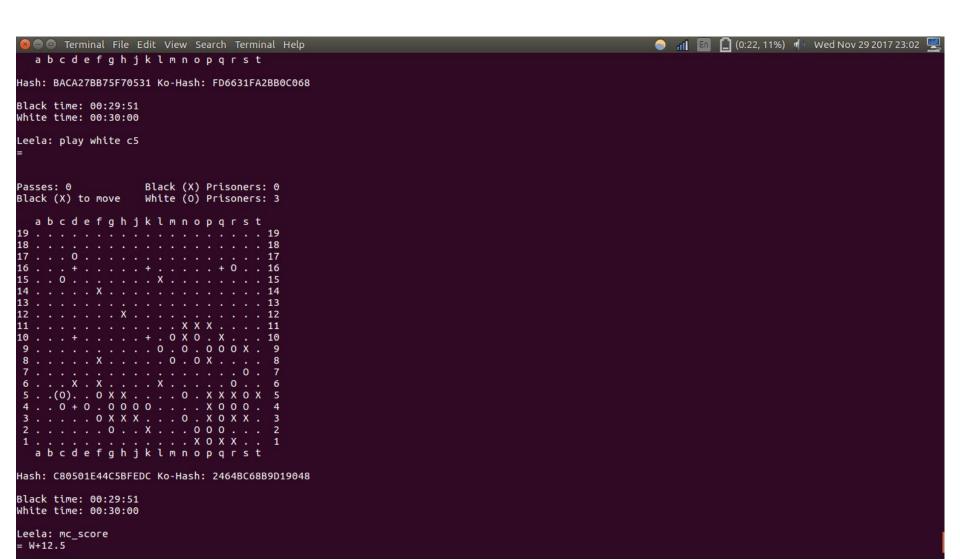


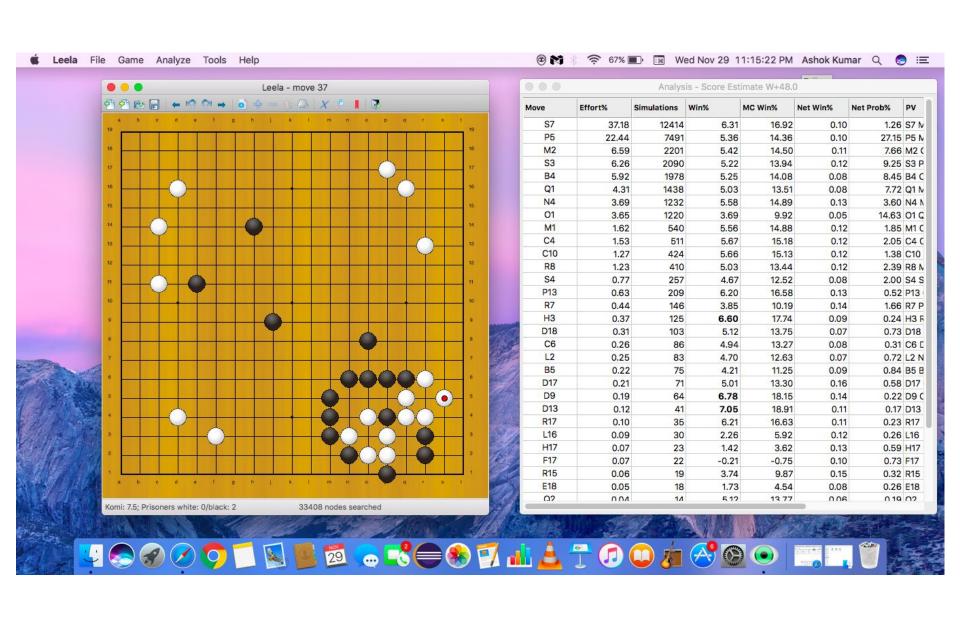




```
🔵 📶 🗈 🖺 (0:12, 14%) ゆ Wed Nov 29 2017 22:56 🖳
🔞 🖨 📵 Terminal File Edit View Search Terminal Help
6 . . . X . X . . . . X . . . . . . 6
5 . . . . . O X X . . . . O . X X X O X 5
4 . .(0)+ 0 . 0 0 0 0 . . . . X 0 0 0 . 4
3 . . . . . O X X X . . . O . X O X X . 3
1 . . . . . . . . . . . X O X X . . . 1
abcdefghjklmnopqrst
Hash: 3E414760141E406C Ko-Hash: D220FAECE1942EF8
Black time: 00:30:00
White time: 00:30:00
Leela: genmove black
MC winrate=0.000000, NN eval=0.000044, score=W+85.4
Nodes: 2961, Win: 1.26% (MC: 3.20%/VN: 0.00%), PV: H2 F2 L3 L4
Nodes: 6936, Win: 1.37% (MC: 3.46%/VN: 0.00%), PV: H2 F2 L3 L4 M2 J1,
Nodes: 10679, Win: 1.40% (MC: 3.55%/VN: 0.00%), PV: H2 F2 L3 L4 M2 J1 J2
Allowing early exit: score: 1.509197%
 C5 ->
                                                  60) (N: 23.6%) PV: C5 B4 C13 G17 P17 Q17 P16
         3190 (W: 1.51%) (U: 3.81%) (V: 0.00%:
 H2 ->
         2149 (W: 1.51%) (U: 3.81%) (V: 0.00%:
                                                  44) (N: 16.0%) PV: H2 F2 L3 L4 M2 J1 J2
C13 ->
         2009 (W: 1.43%) (U: 3.61%) (V: 0.00%:
                                                  38) (N: 16.6%) PV: C13 B6 F17 E15 F15 D13 C12
 E5 ->
         1423 (W: 1.42%) (U: 3.59%) (V: 0.00%:
                                                  31) (N: 11.8%) PV: E5 F4 C13 G17 C5
P17 ->
         1337 (W:
                 1.45%) (U: 3.67%) (V: 0.00%:
                                                 23) (N: 10.7%) PV: P17 P16 Q16 Q15 Q17
 J6 ->
          582 (W:
                 1.30%) (U: 3.28%) (V: 0.00%:
                                                 12) (N: 5.6%) PV: J6 P17 C13
                 1.79%) (U: 4.52%) (V: 0.00%:
F17 ->
          501 (W:
                                                 10) (N: 3.1%) PV: F17 P17 R14
                                                  3) (N: 2.3%) PV: P16 P17 017
                  1.26%) (U: 3.18%) (V: 0.00%:
P16 ->
          210 (W:
 L4 ->
          176 (W: 1.36%) (U: 3.42%) (V: 0.00%:
                                                  3) (N: 1.9%) PV: L4 K3
 D3 ->
          112 (W: 2.77%) (U: 7.00%) (V: 0.00%:
                                                  3) (N: 0.5%) PV: D3 C3
 M4 ->
           94 (W: -0.47%) (U: -1.18%) (V: 0.00%:
                                                  2) (N: 2.4%) PV: M4 M3
 F2 ->
           63 (W: 0.81%) (U: 2.05%) (V: 0.00%:
                                                  1) (N: 1.1%) PV: F2
______
3190 visits, score 1.51% (from 1.46%) PV: C5 B4 C13 G17 P17 Q17 P16
11907 visits, 3537 nodes, 11907 playouts, 1378 p/s
= C5
```







# RESULTS

### MinAlphaGO:

- We were able to perform the analysis part successfully.
- It predicts move faster than leela.
- Improved User Interface.

### BetaGo:

- Model trained using Supervised Learning.
- UI was not implemented properly.
- No analysis part.
- Faster than Leela.

### Leela:

- Has a good User Interface.
- It is comparatively slower.

# SCOPE OF PROJECT

The learning done so far will be implemented In one of these areas ...

Climatic Modelling for Monitoring Environmental issues

**Space Explorations** 



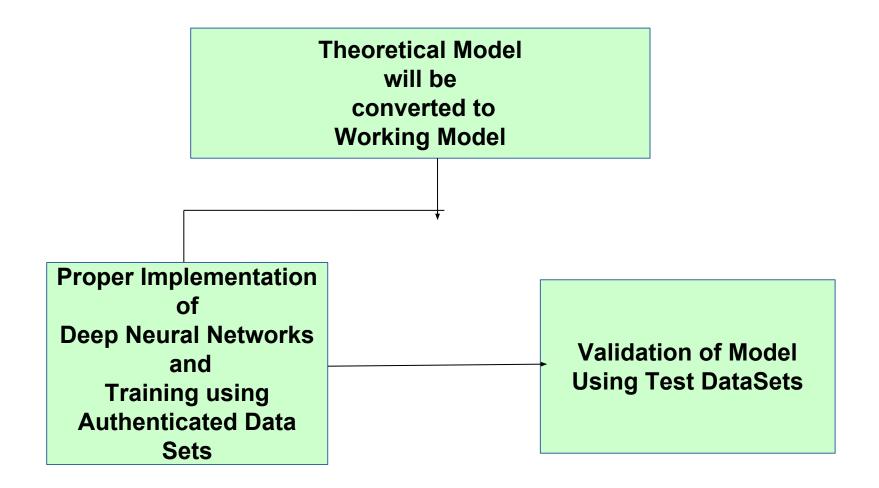
Complex Disease
Analysis
{Zika Virus}



# FUTURE SCOPE

- MCST can also be used to refine the training process.
- Reinforcement learning along with policy and value networks can be used to train the model more efficiently.
- Different bots can be implemented.

# PHASE 3



# Improvements and Suggestions are most welcomed ...

:)