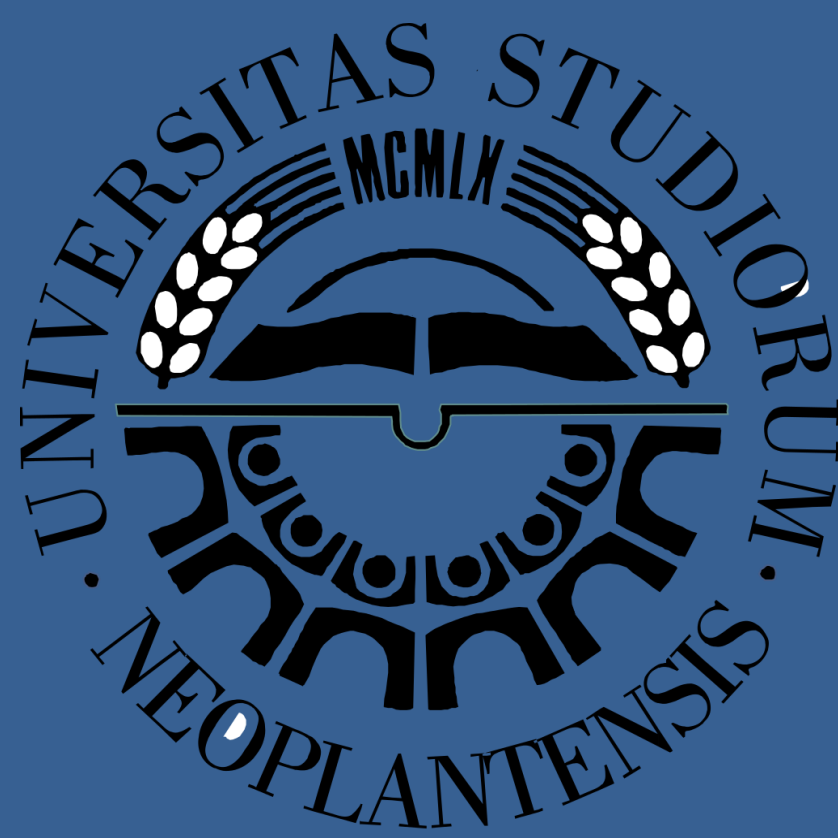




# Šah bot zasnovan na MCTS algoritmu i neuronskim mrežama

Nedeljko Vignjević i Dalibor Malić  
Softversko inženjerstvo i informacione tehnologije  
Asistent: Milica Škipina



## Apstrakt

Šah je strateška igra koja se već godinama proučava u oblasti veštačke inteligencije.

Najbolji šah botovi su do pre par godina bili zasnovani na kombinaciji sofiticiranih tehnika pretraživanja, prilagođavanja specifičnih za domen i ručno kreiranih funkcija procene.

Nasuprot tom pristupu cilj ovog projekta je napraviti šah bota koji ne koristi apsolutno nikakve karakteristike kreirane od strane čoveka.

## Uvod

Chessmc je šah bot koji koristi *Monte-Carlo Tree Search* (MCTS) algoritam kako bi odredio sledeći potez za igru.

Ideja je da bude sličan *AlphaZero* botu u smislu da ne koristi *random rollout* kod MCTS algoritma. *Random rollout* je zamenjen neuronskom mrežom. Neuronska mreža je obučavana tako da reprodukuje evalucionu funkciju *Stockfish* bota.

Rezultat procesa evaluacije je vrednost koja se naziva *centipawn* (cp). *Centipawn* odgovara 1/100 pijuna i najkorišćenija je metoda za evaluaciju stanja na šahovskoj tabli.

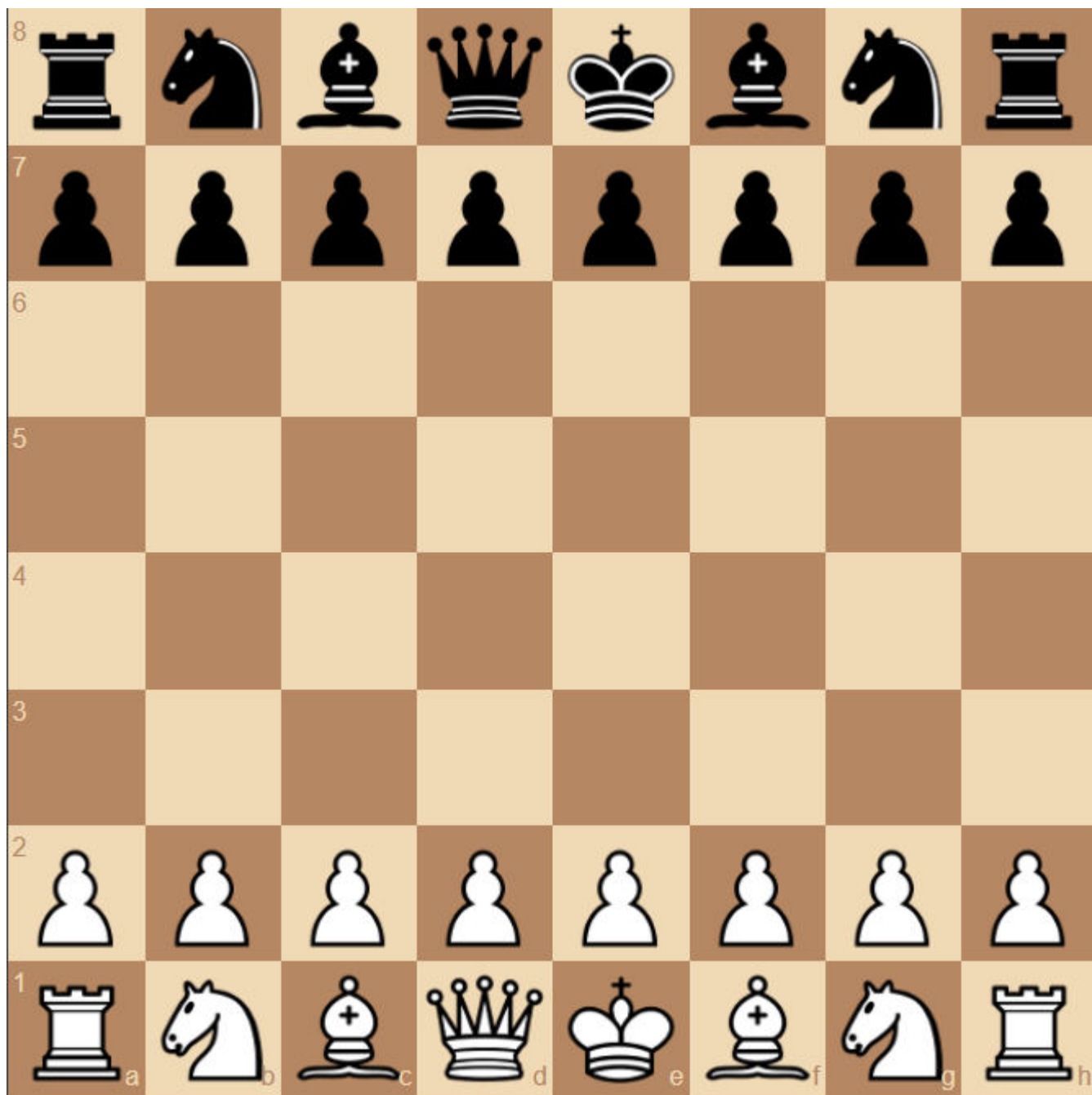
Svako stanje na tabli je labelirano kao pobeda, gubitak ili poraz u skladu sa cp vrednosti koju dodeljuje *Stockfish*. Pobeda je dodeljena ako je cp > 1.5, gubitak ako je cp < -1.5 i nerešeno ako je cp između ove 2 vrednosti.

## Metodologija i rezultati

Korišćen je *Monte-Carlo Tree Search* algoritam u kombinaciji sa *Multilayer Perceptron* (MLP) neuronskom mrežom. Neuronska mreža na ulazu dobija trenutno stanje na šahovskoj tabli. Stanje je predstavljeno nizom od 8x8x12 (768) elemenata, 8x8 za svako polje na šahovskoj tabli i 12 za svaku figuru na tabli (6 figura za igrača koji je na potezu i 6 za protivničkog igrača). Polje na kom se nalazi figura igrača koji je na potezu predstavljeno je brojem 1, polje na kom se ne nalazi ni jedna figura predstavljeno je brojem 0, a polje na kom se nalazi figura protivničkog igrača predstavljeno je brojem -1.

Neuronska mreža je obučavana tako da reprodukuje evalucionu funkciju jednog od najboljih svetskih šah programa – *Stockfish*, na način opisan u uvodu.

Najbolje obučeni model je trenutno MLP model sa 2 međusloja od po 256 i 64 neurona. Međuslojevi koriste *Leaky ReLU* kao aktivacionu funkciju. Model je obučavan *Adam* optimizatorom sa sledećim vrednostima parametara:  $\eta = 0.001$ ,  $\beta_1 = 0.90$ ,  $\beta_2 = 0.99$  i  $\epsilon = 1e-0.8$ . Kako bi se sprečilo preterano podudaranje s podacima za trening (*overfitting*) modela korišćena je regularizacija *dropout* metodom od 50% na svakom međusloju.



Slika 1. Izgled šahovske table korišćene u projektu

Model je obučavan 100 epoha, sa 128 batch veličina, nad 500.000 primera dobijenih iz skupa podataka koji je preuzet sa sajta pgnmentor [2]. Za validacioni skup kao i za testni skup podataka korišćeno je 50.000 (po 10%) primera. Dobijeni rezultati su sledeći:

- Training loss – 0.08
- Accuracy on test data – 76%

S obzirom da je MCTS algoritam (verzija sa UCB1, kao i verzija sa neuronskom mrežom) pokazao loše rezultate u igri šah, sledeće dve tabele predstavljaju rezultate pri igranju iks-oks igre sa minimax i random botom, u cilju validacije samog algoritma.

Tabela 1. Uspešnost različitih botova u iks-oks igri

MCTS	Pobeda	Nerešeno	Poraz
Minimax	0	67	33
Random	55	43	2

Minimax	Pobeda	Nerešeno	Poraz
Random	77	0	23

## Zaključak

Iako MCTS, zajedno sa UCB1, postiže dobre rezultate u drugim igrama, za ovaj algoritam se smatra da nije najpogodniji za igru kao što je šah [1].

Procene koje vidimo u radu [1] ukazuju na to da bez dobre i dovoljno tačne strategije izvođenja simulacija, dolazi do asimetričnog rasta stabla i nemogućnosti identifikovanja efikasnih rezultata pretrage.

U ovom projektu je korišćen MCTS (u kombinaciji sa UCB1 kao i sa neuronskim mrežama). Zbog hardverskih ograničenja, broj simulacija je veoma nizak za igru kao što je šah. Povećavanjem broja simulacija se produžava vreme izvršavanja poteza od strane bota, a rezultat malog broja simulacija je izuzetno loša igra bota.

Naredni korak, u cilju poboljšanja rada trenutnog bota, bio bi unapređenje trenutnog algoritma modifikacijama kao što su AMAF/RAVE ili *progressive bias*.

## Reference i pomoćni linkovi

1. [http://www.ke.tu-darmstadt.de/lehre/arbeiten/bachelor/2012/Arenz\\_Oleg.pdf](http://www.ke.tu-darmstadt.de/lehre/arbeiten/bachelor/2012/Arenz_Oleg.pdf)
2. <https://www.pgnmentor.com/files.html>
3. <https://arxiv.org/abs/1712.01815>
4. <https://stockfishchess.org/>