

Verkefnið mitt fjallar um hvernig hægt er að nota vélrænan lærdóm til að greina stöður á skákborði.

Ég hef gaman af skák og því finnst mér áhugavert að geta nýtt það sem ég hef lært í þessu námskeiði í tengingu við áhugamál mín.

Það eru um  $10^{40}$  löglegar stöður á skákborði svo það þarf mjög mikið að þjálfunargögnum til að geta gert gott módel. Það gæti því verið hindrun að þjálfunin tæki ofboðslega mikinn tíma og tölvuafli til að fá gott módel.

Til eru ýmis gagnasett sem hægt væri að nota, t.d. er hægt að sækja skrár sem innihalda upplýsingar um alla þá 4,107,353,953 leiki sem spilaðir hafa verið á skáksíðunni lichess.org af heimasíðu þeirra. Til að þjálfna módelið þarf gögn um stöðu á skákborði og tilsvareandi greiningu á þeirri stöðu. Það gagnasett sem ég tel henta best er gagnasett af Kaggle.com sem innigeldur aðeins upplýsingar um stöður og tilsvareandi greiningar á þeim. Það býður upp á 12,954,834 stöður með greiningum á csv formi.

Dæmi um gildi úr því gagnasetti eru:

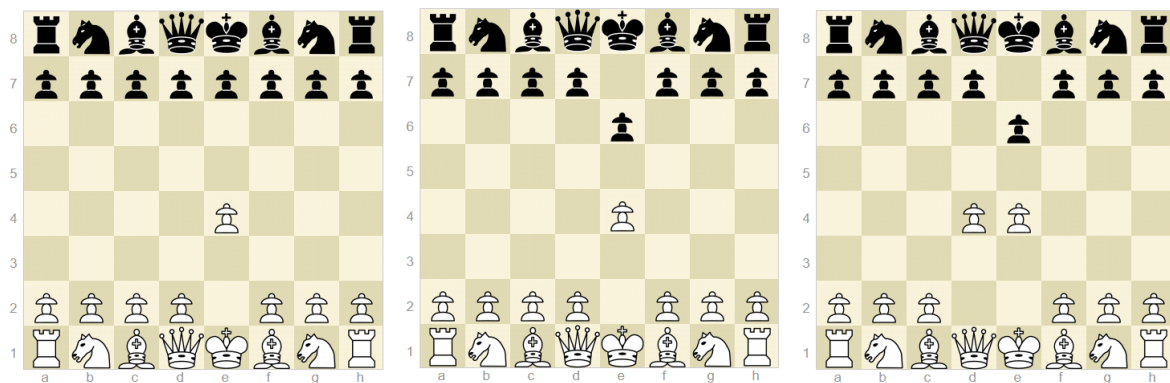
FEN,Evaluation

rnbqkbnr/pppppppp/8/8/4P3/8/PPPP1PPP/RNBQKBNR b KQkq - 0 1,-10

rnbqkbnr/pppp1ppp/4p3/8/4P3/8/PPPP1PPP/RNBQKBNR w KQkq - 0 2,+56

rnbqkbnr/pppp1ppp/4p3/8/3PP3/8/PPP2PPP/RNBQKBNR b KQkq - 0 2,-9

Þar sem FEN inniheldur upplýsingar um stöðuna á skákborðinu og Evaluation er mat sem skáktölva hefur gefið á stöðuna. Þetta eru fyrstu þrjár færslur í einhverjum leik og líta þær svona út:



Flestar af nútíma skáktölvum notast við tauganet og því tel ég að það sé hentugt módel fyrir mitt verkefni. TensorFlow er notað fyrir hönnun á tauganetum og myndi ég því nýta mér það.

Ég er ekki kominn með neinn ákveðinn lista af heimildum sem ég ætla að nota til að vinna verkefnið.

Til að meta niðurstöðurnar er hægt að reikna prófunarskekkju fyrir módelið þar sem þekkt mat á stöðu er borið saman við mat módelins.