

Raport 1

Martyna Bielec, Maciej Karczewski

2022-11-23

Wprowadzenie

Strona internetowa Lichess jest drugą co do wielkości platformą na świecie, która umożliwia grę w szachy online. Użytkownicy mogą grać z komputerem lub między sobą. Każdy gracz ma przypisany ranking, który jest ustalany na podstawie jego umiejętności. Jeśli użytkownik rozgrywa partię “rankingową”, w jej wyniku jego ranking rośnie lub maleje.

Dane, które zostaną poddane analizie, zawierają informacje o ponad 20 000 rozgrywkach szachowych ze strony Lichess Games. Dostępne są na witrynie kaggle. Składają się z następujących kolumn:

- id (zawiera elementy typu “character”) - numer ID rozgrywki
- rated (“character”) - wartość “True” lub “False” w zależności od tego, czy partia jest “rankingowa”
- created_at (“numeric”) - czas rozpoczęcia gry
- last_move_at (“numeric”) - czas zakończenia gry
- turns (“numeric”) - liczba rozegranych tur
- victory_status (“character”) - status zakończenia gry: “mate” dla zakończenia matem, “resign” dla zakończenia przez poddanie się jednej ze stron, “draw” dla remisu lub “outoftime” dla rozgrywki zakończonej z powodu upływu czasu
- winner (“character”) - zwycięzca gry: “white”, jeśli wygrał gracz biały, “black”, jeśli czarny lub “draw” w przypadku remisu
- increment_code (“character”) - składa się z dwóch liczb rozdzielonych znakiem “+”; pierwsza oznacza liczbę minut przypadających na gracza na wszystkie jego ruchy, druga liczbę sekund, o którą zwiększa się czas gracza po wykonaniu ruchu
- white_id (“character”) - ID gracza białego
- white_rating (“numeric”) - ranking gracza białego
- black_id (“character”) - ID gracza czarnego
- black_rating (“numeric”) - ranking gracza czarnego
- moves (“character”) - wszystkie ruchy zapisane w standardowej notacji szachowej
- opening_eco (“character”) - standaryzowany kod dla rozgranego otwarcia
- opening_name (“character”) - nazwa rozgranego otwarcia
- opening_ply (“character”) - liczba ruchów w fazie otwarcia

Celem analizy będzie odpowiedzenie na pytanie - jakie jest prawdopodobieństwo wygrania białego gracza? Będziemy rozważać ten problem pod kątem rozgranego otwarcia, rankingu białego gracza, różnicy w rankingu między graczem białym i czarnym oraz liczby tur w partii.

Obróbka danych

Analiza danych

Jak często wygrywa gracz biały?

W celu rozważenia od czego zależy prawdopodobieństwo wygrania białego gracza, najpierw oszacowaliśmy to prawdopodobieństwo w ogólności. Liczba partii zakończonych zwycięstwem jednej ze stron jest równa:

```
not_draw_count = sum(df$winner != "draw")
print(not_draw_count)
```

```
## [1] 19108
```

Z czego gracz biały wygrał następującą liczbę razy:

```
white_winner_count = sum(df$winner == "white")
print(white_winner_count)
```

```
## [1] 9998
```

Wynika z tego, że szacowane prawdopodobieństwo wygrania gracza białego wynosi w przybliżeniu 52,31% rozgrywek, a gracza czarnego 47,69%.

W dalszej części analizy sprawdzimy, jak to prawdopodobieństwo się zmienia, jeśli odpowiednio uwarunkujemy rozgrywkę.