

**UNIWERSYTET RZESZOWSKI**

**Kolegium Nauk Przyrodniczych**

Jakub Fałowski

Nr albumu: 113605

Kierunek: Informatyka

**Serwis web gromadzący, weryfikujący i analizujący statystyki piłkarskie**

Praca inżynierska

Praca wykonana pod kierunkiem

dr inż. Marcin Ochab

Rzeszów 2023

Spis treści

**1 Wstęp 5**

* 1. Problematyka 5
  2. Cel i założenia 5
  3. Diagram przypadków użycia 6
  4. Użyte technologie 7
  5. Użyte API/Serwisy 7
     1. API Flashscore 7
     2. Forebet.com 8
     3. FIFA – futhead.com 8
     4. Football Manager – fmdataba.com 8
     5. API Sofascore 8

1. **Moduł zawodników 9**
   1. Zasada działania 9
      1. FIFA 9
      2. Football Manager 12
      3. Sofascore – rzeczywiste statystyki 14
   2. Przykładowe widoki 16
   3. Możliwe rozwinięcia 19
   4. Implementacja systemu 19
2. **Moduł klubów 21**
   1. Zasada działania 21
      1. System na podstawie ostatnich rezultatów 21
         1. Wyznaczenie zwycięzcy 23
         2. Wyznaczenie różnicy bramkowej 28
         3. Wyznaczenie dokładnego wyniku 30
      2. Kursy bukmacherów 30
      3. Forebet 31
   2. Wyświetlanie meczów 31
   3. Przykładowe widoki 32
   4. Możliwe rozwinięcia 35
   5. Implementacja systemu 35
3. **Testy aplikacji 39**
   1. Typowane wyniki a rzeczywistość 39
   2. Zależność pomiędzy statystykami Sofascore i danymi z gier piłkarskich 48
   3. Testy manualne 50

**Podsumowanie i wnioski 55**

**Bibliografia i netografia 57**

**Wykaz rysunków 59**

**Wykaz tabel 60**

**Streszczenie w języku polskim 61**

**1 Wstęp**

* 1. **Problematyka**

Piłka nożna jest jednym z najbardziej popularnych sportów na całym świecie. Jeden mecz potrafi przyciągnąć na stadion kilkadziesiąt tysięcy fanów, którzy są gotowi opłacić wysokie ceny biletów, tylko po to by obejrzeć 90 minut spotkania swoich ulubieńców [1]. Podobnie sprawa ma się z transmisjami telewizyjnymi, które potrafi obejrzeć nawet kilkanaście milionów Polaków [2]. Według zestawienia „Sponsor Monitoring” za 2022 rok, piłka nożna jest najpopularniejszym sportem w Polsce. Wskaźnik zainteresowania tą dyscypliną jest równy 71% [3]. Co za tym idzie, jest to zarobkowa szansa dla wielkich firm. Potencjał medialny, który generują najlepsze mecze, niekoniecznie wpływa na wiedze kibiców. Według badania CBOS z 2016 roku, tylko 15% korespondentów, przyznało, że w większym stopniu interesuje się tym sportem [4]. Jest to sporo mniejsza liczba i niekoniecznie świadcząca o fachowości, bądź większym zainteresowaniu tematem od strony teoretycznej. W sieci brakuje wielu merytorycznych źródeł, które umożliwiałyby zdobycie większej informacji na ten temat. Bardziej popularnym źródłem informacji są subiektywne opinie dziennikarzy.

* 1. **Cel i założenia**

Niniejsza praca inżynierska stara się rozwiązać problem łatwej analizy danych piłkarskich poprzez utworzenie aplikacji webowej. Aplikacja ta gromadzi dane z różnych popularnych źródeł i statystyk dotyczących piłki nożnej. Dzięki niej, każdy zainteresowany tym tematem, może na podstawie obiektywnych danych, wyciągać własne wnioski. Pomaga to zapobiegać przekazywaniu niesprawdzonych informacji, oraz zwiększyć świadomość i głębsze zainteresowanie tą dyscypliną od strony analitycznej. Pozwala użytkownikowi na przeprowadzanie poniższych akcji:

* porównywanie statystyk zawodników w oparciu o trzy źródła,
* zobrazowanie średnich wartości atrybutów oceny piłkarzy, porównanie najlepszych klubów pod względem atrybutów w oparciu o najlepszych zawodników na poszczególnych pozycjach, system zliczający rezultaty meczy na podstawie ostatnich spotkań,
* porównywanie różnych źródeł szacujących wynik spotkania.

Poprzez porównywanie atrybutów zawodników z trzech źródeł, można obiektywniej spojrzeć na umiejętności danego zawodnika. Jeśli wartości jednego ze źródeł znacząco wyróżnia się na tle reszty, bądź widoczne są zależności pomiędzy dwoma źródłami, można badać przyczyny takiej sytuacji. Dodatkowo trzecie źródło jakim są statystyki z meczów w trakcie sezonu, można interpretować jako weryfikacja słuszności danych z gier piłkarskich. Podobnie działa funkcjonalność porównywania aplikacji zajmującej się wynikami meczów. Trzy źródła mogą dawać trzy inne wyniki, lecz rzeczywistość po czasie weryfikuje, które z nich było najbliżej prawdy. Dodatkowo wiedząc w jaki sposób została stworzona jedna z aplikacji do oszacowania wyników, można obserwować wspólne zależności powodujące, że dana drużyna ma większe szanse na wygraną. Może to pomóc zrozumieć jak działają pozostałe źródła, z których pobierano informacje.

* 1. **Diagram przypadków użycia**

Obraz zawierający diagram

Opis wygenerowany automatycznie

Rysunek 1 - Diagram DPU

Diagramy przypadków użycia stanowią skuteczne narzędzie do opracowywania, prezentowania i udokumentowania działania jednostki. Dzięki nim, systemy, podsystemy oraz klasy stają się łatwiej zrozumiałe i dostępne. Diagramy te przedstawiają jednostkę z perspektywy zewnętrznej [5].

* 1. **Użyte technologie**

**Next.js** – framework React.js, który pozwala na renderowanie stron zarówno po stronie serwera, jak i klienta. Posiada wsparcie między innymi do arkuszy styli, czy typescript.js. Zawiera w sobie routing i umożliwia stosowanie metod DOM [20].

**Mantine –** biblioteka do React.js(w tym Next.js), zawierająca ponad 100 wbudowanych komponentów, oraz 40 hooków. Dzięki gotowym komponentom interfejsu, takim jak przyciski, pola tekstowe i inne elementy interaktywne, pozwala twórcom szybko i efektywnie tworzyć spójne i atrakcyjne interfejsy użytkownika.

**SASS –** język skryptowy opierający się na CSSie, pozwalający między innymi na tworzenie zmiennych i funkcji. Pomaga uporządkować i ulepszać kod CSS, co poprawia jego przejrzystość i rozwijalność.

**Cheerio.js –** biblioteka służąca do parsowania kodu i manipulowania danymi z HTML. Jest szybsza oraz łatwiejsza do użycia w porównaniu do jQuery. Działa na takiej zasadzie, że wczytuje dane, a następnie umożliwia wyszukiwanie i manipulowanie nimi za pomocą selektorów i funkcji, podobnych do jQuery.

**Axios –** klient HTTP umożliwiający tworzenie zapytań w przeglądarce. Pozwala na wysyłanie zapytań takich jak: GET, POST, PUT, DELETE oraz łatwą konfiguracje nagłówków i parametrów.

**Recharts –** biblioteka React.js służąca do tworzenia wizualizacji danych w aplikacjach internetowych. Zawiera komponenty ułatwiające tworzenie wykresów, dostosowanych do indywidualnych potrzeb.

**Cypress –** narzędzie do tworzenia do aplikacji internetowych. Pozwala na zdefiniowanie i wykonanie testów, które sprawdzają, czy aplikacja jest zgodna z oczekiwaniami testera. Cypress oferuje wiele przydatnych narzędzi i funkcji, takich jak śledzenie i debugowanie, które ułatwiają tworzenie i utrzymywanie testów.

* 1. **Użyte API/Serwisy**
     1. **API Flashscore**

Flashscore API to interfejs programowania aplikacji (API), który umożliwia dostęp do danych dotyczących wyników meczów i statystyk sportowych ze strony flashscore.com. API umożliwia przetwarzanie danych w czasie rzeczywistym oraz wyświetlanie wyników i statystyk meczów na różnych platformach. Zapewnia dostęp do danych z wielu dyscyplin sportowych, w tym piłki nożnej, koszykówki, hokeja czy tenisa. Umożliwia również wyszukiwanie i filtrowanie danych według różnych kryteriów, takich jak drużyna, liga, a nawet indywidualni zawodnicy. Jego zalety to:

* dostęp do danych w czasie rzeczywistym - Flashscore API umożliwia dostęp do danych w czasie rzeczywistym, co oznacza, że wyniki meczów i statystyki są aktualizowane na bieżąco,
* dostęp do danych z wielu dyscyplin - Flashscore API umożliwia dostęp do danych z wielu dyscyplin sportowych, w tym piłki nożnej, koszykówki, hokeja i wielu innych,
* możliwość wyszukiwania i filtrowania danych.
  + 1. **Forebet.com**

Forebet.com to strona internetowa, która oferuje prognozy i analizy dotyczące wyników meczów piłkarskich z całego świata. Strona zawiera informacje o meczach, drużynach, ligach i zawodnikach, a także statystyki i inne dane, które pomagają w typowaniu wyników meczów. Forebet.com jest popularną stroną internetową wśród miłośników piłki nożnej, którzy szukają informacji i analiz dotyczących wyników meczów. Dane pobierane na bieżąco z powyższej strony są trzecim źródłem analizującym wyniki meczów.

* + 1. **FIFA- futhead.com**

FIFA to seria gier piłkarskich wydawana przez firmę Electronic Arts (EA). Gra została pierwotnie wydana w 1993 roku i od tego czasu wydano wiele kolejnych wersji z ulepszeniami i dodatkami. Gra pozwala na rozgrywanie meczów piłkarskich z różnymi drużynami i ligami z całego świata. FIFA jest jedną z najpopularniejszych gier piłkarskich na świecie i cieszy się dużą popularnością wśród miłośników tego sportu. Gra jest regularnie aktualizowana i ulepszana przez twórców, aby zapewnić jak najlepszą rozgrywkę i realizm. Dane dotyczące statystyk piłkarzy, które pojawiają się w grze pobrano z portalu futhead.com, które zostały użyte jako jedno ze źródeł.

* + 1. **Football Manager – fmdataba.com**

Football Manager to seria gier menedżerskich, w których gracze mogą zarządzać drużyną piłkarską i prowadzić ją przez kolejne sezony. Gra została wydana pierwotnie w 1992 roku i od tego czasu wydano wiele kolejnych wersji z ulepszeniami i dodatkami. Pozwala na rozgrywanie meczów piłkarskich i prowadzenie drużyny w trakcie sezonu. Gracze mogą wybierać spośród setek licencjonowanych drużyn i lig z całego świata, a także zarządzać składem drużyny, taktyką i budżetem. Football Manager oferuje również szczegółowy system transferów, w którym gracze mogą negocjować kontrakty z piłkarzami. W aplikacji atrybuty zawodników z Football Managera zostały użyte jako źródło porównawcze ze statystykami z FIFY i API Sofascore

* + 1. **API Sofascore**

Sofascore to internetowa platforma, która oferuje aktualne wyniki sportowe oraz statystyki z wielu różnych dyscyplin. Platforma ta prócz informacji o rezultatach, udostępnia szczegółowe statystyki dotyczące poszczególnych zawodników i drużyn .Z API Sofascore można pobrać tak szczegółowe dane jak np. liczba odbiorów, wślizgów, procent celnych podań i strzałów. Te statystyki idealnie nadają się do porównania względem atrybutów z gier.

1. **Moduł zawodników**

Celem tego modułu jest możliwość weryfikacji umiejętności zawodników. Wysoka ocena wystawiona przez researcherów niekoniecznie musi się pokrywać z rzeczywistością. Wpływ na wartość oceny ogólnej może mieć również popularność, bądź subiektywna ocena. Teoretycznie każdy piłkarz ma swoje mocne i słabe strony, które można szczegółowo ocenić na podstawie jego ostatnich występów w drużynie.. .Jednak różnice w ocenie atrybutów mogą prowadzić do sytuacji, w której w jednej z gier zawodnik jest znacznie lepszy niż konkurenci. Dyspozycja zawodnika podczas rozgrywki może zależeć od kilku czynników, takich jak szybkość dla zawodników ofensywnych podczas gry z kontratakiem, a technika podczas wolnego rozgrywania akcji [6]. Aby umożliwić porównanie sytuacji, w których wartości atrybutów różnią się znacząco, powstał moduł ułatwiający analizę umiejętności piłkarzy. Posiadając dane zebrane z jednego i drugiego źródła można je wykorzystać do porównywania. Łączna liczba piłkarzy w źródłach różni się z różnych względów, przede wszystkim z powodu szerokiej kadry klubów, w skład której wchodzą zawodnicy grający na co dzień w pierwszym zespole oraz druga drużyna i juniorzy uzupełniający skład w razie kontuzji, zawieszenia lub woli trenera.. Dodatkowo, Football Manager zwraca dużą uwagę na młodszych zawodników, a dla wielu graczy obserwowanie rozwoju nastoletniego zawodnika w dłuższym okresie czasu potrafi być satysfakcjonujące. Aplikacja wyświetla danego zawodnika tylko w momencie, gdy pojawi się on w obydwu grach, a do analizy dołączają się również dane statystyczne z rozgrywek pozyskane z Sofascore.

**2.1 Zasada działania**

**2.1.1 FIFA**

Witryna futhead stworzona dla graczy FIFY, posiada informacje o atrybutach zawodników, którzy pojawiają się w tej grze. Odpowiednik tej witryny istnieje również w Football Managerze - fmdataba. Ze względu na brak API, dane zostały pobrane z powyższych witryn.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 2 – Widok strony futhead – 11.11.2022 [7].

Jak widać na rysunku powyżej (Rysunek 2), obok nazwiska widnieją takie informacje jak:

* klub, w którym obecnie się znajduje (np. Manchester City),
* wartość oceny ogólnej(RAT- rating),
* szybkości (PAC – pace),
* strzałów (SHO – shotting),
* podań (PAS – passing),
* dryblingu (DRI – dribbling),
* obrony (DEF – defensive),
* fizyczności (PHY – physical).

Aby porównywane dane były w pełni rzetelne, baza piłkarzy ograniczona została do graczy z pola z angielskiej Premier League. Jak widać na zdjęciu poniżej (Rysunek 3) jest to najsilniejsza liga według współczynnika UEFA z 4 listopada 2022 roku. W aplikacji porównywane są takie statystyki jak liczba bramek. O wiele łatwiejszym zadaniem jest strzelenie dużej ilości bramek, gdy gra się przeciwko zawodnikom o słabszych umiejętnościach. Ciężki do określenia byłby również adekwatny stopień trudności danej ligi. Pomimo jasnego współczynnika UEFY, trudno przypuszczać, że zawodnik, który zdobędzie 20 bramek w Premier League, potroiłby swój dorobek, w lidze mającej trzy razy niższy współczynnik. Obraz zawierający stół

Opis wygenerowany automatycznie

Rysunek 3 - Współczynnik UEFA – 04.11.2022 [8]

Dane są pobierane z wszystkich podstron i zapisywane w postaci pliku .json do API. Każdy obiekt ma wartość name z imieniem i nazwiskiem piłkarza, rating z oceną ogólna oraz pozostałe potrzebne atrybuty.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 4 - Wartości wyświetlone w API pracy inżynierskiej

Zdarzały się pojedyncze przypadki, gdy nazwa piłkarza była tak długa, że na stronie futheadu nazwisko było niepełne i wypełnione trzema kropkami jak na zdjęciu poniżej (Rysunek 5). Wynikało to z takich danych po stronie serwera.



Rysunek 5 – Widok zawodnika na stronie futhead – 11.11.2022 [7]

W takich przypadkach nazwisko przed wrzuceniem obiektu tego piłkarza do API, zostało podmienione.

Obraz zawierający tekst, wewnątrz, pomarańczowy, zrzut ekranu

Opis wygenerowany automatycznie

Rysunek 6 – Informacje o zawodniku w API pracy inżynierskiej

Do prawidłowego wyświetlania informacji na stronie wymagane było ustawienie zezwolenia na CORSy po stronie serwera. CORS(Cross-Origin Resource Sharing) to mechanizm bezpieczeństwa stosowany w przeglądarkach internetowych

**2.1.2 Football Manager**

Kolejnym źródłem informacji są dane z portalu fmdataba. Ostatnio dostępne wartości dotyczą edycji z 2021 roku. Aby zachować spójność i realizm, wszystkie statystyki będą analizowane na podstawie danych z tego roku. Dodatkowo dane z obydwu gier będą porównywanie do statystyk z przyszłorocznych rozgrywek.

Obraz zawierający stół

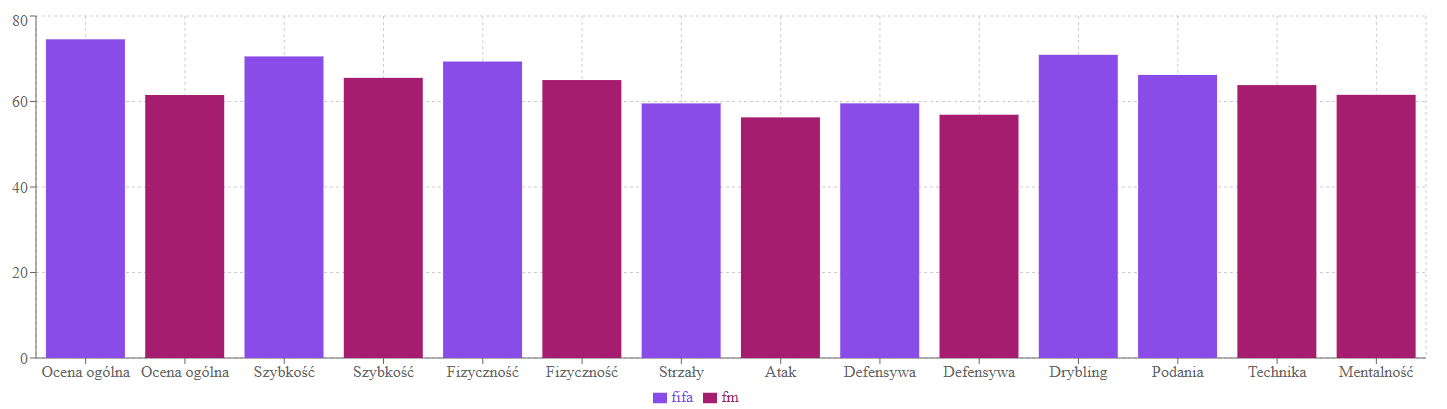
Opis wygenerowany automatycznie

Rysunek 7- Widok strony fmdataba – 11.11.2022 [9]

Dane te są niezmienne. Wartości atrybutów są aktualizowane z każdą nową edycją gry. W aplikacji analizowane są wartości z jednej, równoległej edycji. Na portalu z danymi Football Managera podobnie jak w przypadku witryny futhead, widnieje widok tabeli z nazwiskiem i atrybutami piłkarzy. Znajdują się tam informacje o:

* **klubie** (Club),
* **ocenie ogólnej**(Ove – overall), statystyka taka sama jak w FIFIE,
* **ataku**(Att – attacking), odpowiednik strzałów w FIFIE,
* **defensywie**(Def – defending), odpowiednik w FIFIE,
* **technice**(Tec – technical), składowa dryblingu i podań w FIFIE,
* **mentalności**(Men – mental), nie ma tego typu statystyki w FIFIE. Wynika to z tego faktu, że w Football Managerze gracz zostaje trenerem piłkarskim i zarządza swoimi zawodnikami, którzy mogą podejmować różne decyzje ze względu na swój charakter czy inteligencje boiskową. W FIFIE, to gracz podejmuje decyzje, ponieważ on bezpośrednio decyduje o zachowaniu na boisku za pomocą kontrolera, a nie tylko i wyłącznie na podstawie przydzielonej taktyki,
* **fizyczności**(Phy – physical), odpowiednik w FIFIE,
* **szybkości**(Spe – speed), odpowiednik w FIFIE.

Zakres atrybutów w Football Managerze jest inny niż w FIFIE. W menedżerze piłkarskim wynosi on 1-20, w przeciwieństwie do FIFY gdzie wartości znajdują się pomiędzy 1 a 100. Nie jest to jednak problemem, ponieważ podczas porównywania, w której grze zawodnik jest lepszy, zwracana jest uwaga, na miejsce w rankingu w stosunku do innych zawodników. Przykładowo Raheem Sterling zajmuje 1 miejsce pod względem szybkości w Football Managerze, a w FIFIE jest 4. Za pomocą takiej metody można wysunąć wnioski, że pod tym względem w pierwszej grze został bardziej doceniony. Można założyć również, że wartość 18 w Football Managerze jest równa 90 w FIFIE, ze względu na to, iż odpowiada ona 90% maksymalnego atrybutu.



Rysunek 8 - Porównanie średnich statystyk w pracy inżynierskiej

W tym założeniu mogłoby być problemem to, że średnie wartości różnią się w sposób znaczący. W FIFIE ta wartość wynosi 74.55, a w Football Managerze 61.52 (Rysunek 8). Wynika z tego, że wartości w FIFIE są zazwyczaj wyższe niż w Football Managerze. Jednak jest to sprawa sporna, czy ważniejszym aspektem jest to by zawodnik był jak najlepszy względem maksymalnej wartości czy lepszym względem innych piłkarzy. Dodatkowo kolejnym z kryteriów przy porównywaniu jest procent w stosunku do najlepszej oceny tego konkretnego atrybutu.

**2.1.3 Sofascore – rzeczywiste statystyki**

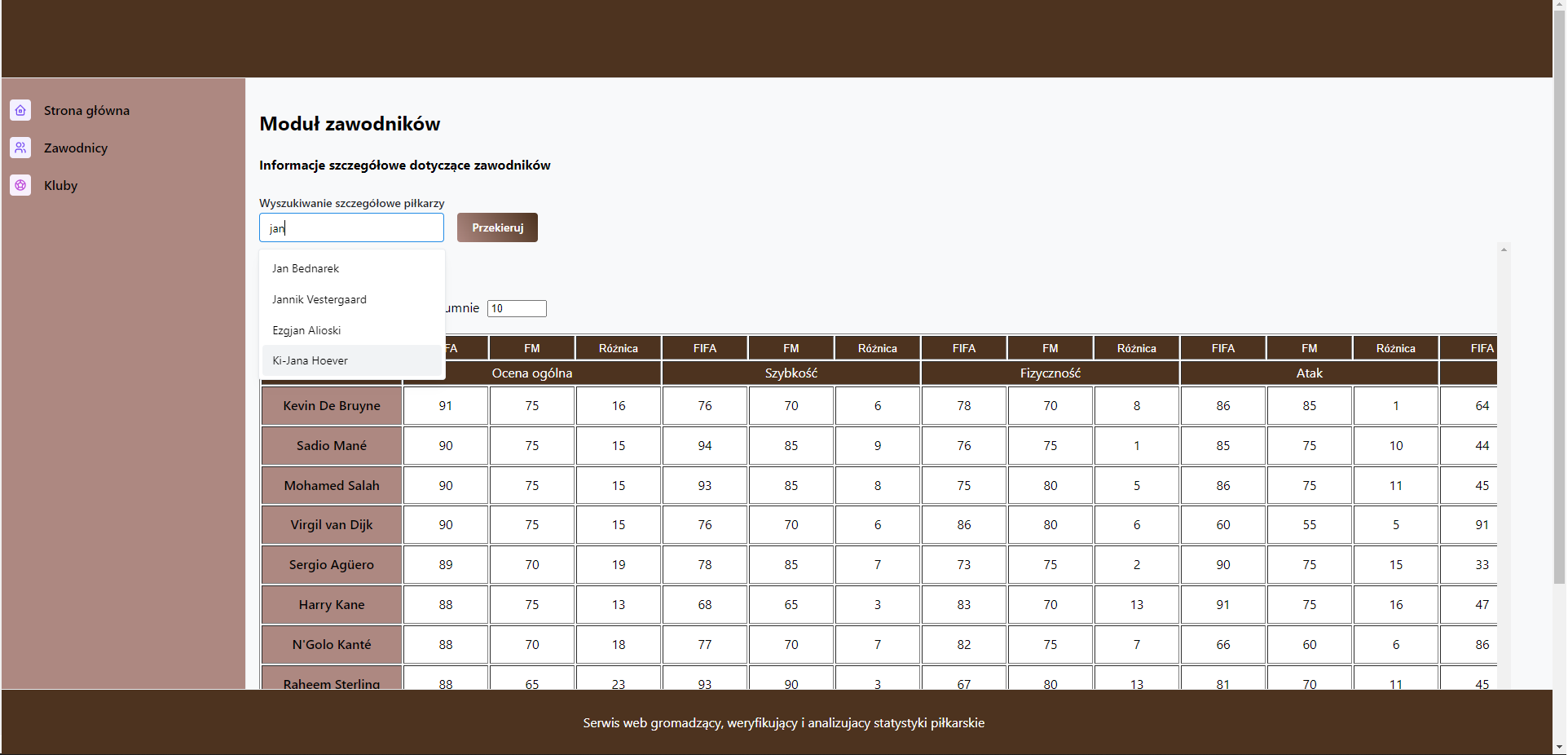
Dane z Sofascore służą głównie do własnej analizy i porównań względem gier. Niektóre statystyki takie jak średnia ocena są bezpośrednią informacją i dobrze odzwierciedlają formę piłkarza względem oceny ogólnej z gier. Jednak niektóre atrybuty takie jak „ilość fauli na danym zawodniku”(sklasyfikowane w kolumnie szybkość) nie świadczą tylko i wyłącznie o atrybucie, przy którym zostali przypisani. Jest to zależne od wielu czynników, takich jak pozycja na boisku, czy styl gry drużyny. W przypadku tej statystyki, o wiele częściej zatrzymywani nieprzepisowo są zawodnicy grający w ataku i na skrzydłach. Jednak szybkość jest wartością kluczową, gdyż przewaga w tej dziedzinie zmusza do błędów i konieczności uciekania się do nieprzepisowych zagrań. Mimo braku możliwości stworzenia sensownego i jednolitego rankingu, dane pozyskiwane mogą dać wiele cennych wskazówek przydatnych do analityki. Wartości poszczególnych danych można porównać procentowo w stosunku do najlepszej oceny. Jednakowe porównanie istnieje także przy pozostałych ocenach.

* **Ocena ogólna:**
* **średnia ocena** – domyślna ocena zawodnika przed meczem wynosi 6.6. W momencie, gdy zawodnik strzeli bramkę, wykona dokładne podanie, bądź wykona inne pozytywne zagranie, otrzymuje dodatnie punkty do początkowej średniej. To samo tyczy się ujemnych punktów w przypadku złych rozwiązań. Waga każdej statystyki jest inna. Po strzelonej bramce zawodnik otrzymuje większy bonus punktowy niż po dokładnym podaniu. Wartości oceny występu są popularnym źródłem piłkarskim, a lista z najlepszymi występami w danym tygodniu jest szeroko rozpowszechniana w mediach społecznościowych
* **Szybkość:**
* **minięcia zawodników na mecz** – ta statystyka mówi, ile razy zawodnik wyprzedził swojego przeciwnika z piłką przy nodze. Najlepsze wartości osiągają ofensywni, szybcy zawodnicy. Umiejętność szybkiego biegania, które przekłada się na wymijanie przeciwników jest bardzo ważnym elementem gry, ponieważ pomaga drużynie w efektywnym przedostaniu się pod bramkę przeciwnika oraz znalezieniu wolnego miejsca, które ułatwi skuteczne dośrodkowanie piłki w pole karne przeciwnika.
* **ilość fauli na danym zawodniku** – ten parametr wiąże się w dużym stopniu z powyższym, ponieważ do nieprzepisowego zagrania dochodzi zazwyczaj w momencie, gdy zawodnik wymknie się przeciwnikom spod kontroli. Gdy przeciwnicy nie potrafią nadążyć za zawodnikiem atakującym zmuszeni są do tak zwanego faulu taktycznego. Bardziej opłacalna jest sytuacja w której podczas rzutu wolnego piłka zatrzymuje się w miejscu przewinienia, niż taka gdy przeciwnik wykorzysta brak poprawnego ustawienia i zorganizowania w obronie. Istnieją także takie sytuacje w których zawodnik chce odebrać piłkę, ale ze względu na jego spóźnienie, trafia on w nogi przeciwnika.
* **Fizyczność:**
* **wygrane pojedynki powietrzne** – największy wpływ na to czy zawodnik wygra pozycje w momencie, gdy piłka jest w powietrzu ma wzrost, skoczność i siła. Są to atrybuty, które składają się na ocenę fizyczności.
* **wygrane pojedynki z piłką na trawie** – siła i zwinność to wartości, które mają duże znaczenie podczas pojedynków odbywających się na podłożu.
* **Atak:**
* **liczba bramek** – informacje o ilości zdobytych bramek przez danego zawodnika w przekroju całego sezonu
* **liczba strzałów** – ilość bezpośrednich prób zdobycia bramek
* **Defensywa:**
* **odbiory** – ilość sytuacji w których dany zawodnik przejmie piłkę przeciwnikowi i utrzyma się z nią, przez choćby krótki moment
* **wybicia** - ilość sytuacji w których dany zawodnik przeszkodził przeciwnikowi w utrzymaniu się przy piłce, lecz po wykonaniu manewru stracił posiadanie
* **wślizgi** – statystyka mówiąca o tym ile razy zawodnik próbuje odebrać piłkę za pomocą ruchu ślizgowego po ziemi
* **Technika:**
* **asysty** – liczba podań po których osoba z drużyny zdobędzie bramkę
* **kluczowe podania** – ilość podań, które pośrednio prowadzi do niebezpiecznej sytuacji
* **stworzone niebezpieczne sytuacje** – ilość ostatnich podań przed niebezpieczną sytuacją
* **procent celnych podań** – rozkład procentowy ilości podań do kolegi z drużyny w stosunku do wszystkich podań

Dodatkowo prócz stworzonej ogólnej tabelki porównawczej pomiędzy zawodnikami, istnieją wykresy w których można uzyskać bardziej zaawansowane informacje o atrybutach. W jednym z nich można porównywać średnią statystykę danego atrybutu pomiędzy FIFĄ, a Football Managerem. W pozostałych widnieje ranking według najlepszych klubów, gdzie siła drużyny równa się sile najlepszych 15 zawodników z danego klubu.

**2.2 Przykładowe widoki**

Dane dotyczące atrybutów w jednej i drugiej grze są porównywane w tabeli. Dodatkowo wyświetlona jest różnica pomiędzy obiema statystykami. Istnieje również możliwość sortowania piłkarzy na podstawie największych różnic różnicy w statystykach.



Rysunek 9 - Podstrona zawodników w pracy inżynierskiej

Na podstronie ze szczegółami, widnieją kolumny takie jak „nie gorszy niż” oraz „procent w stosunku do najlepszej oceny”. W pierwszej z nich sprawdzane jest miejsce, które zajmuje dany piłkarz w rankingu wszystkich zawodników. Miejsce to jest przyrównywane do liczby wszystkich graczy. Wzór na obliczenie:

**ngn = (1 - m/lp)\*100**, gdzie

**ngn** – nie gorszy niż,

**m** – miejsce w rankingu,

**lp** – liczba piłkarzy

Przykładowo, zawodnik będący 30 w rankingu, w którym znajduje się 100 piłkarzy, jest nie gorszy od 70% pozostałych graczy. Druga statystyka przyrównuje procentowo danego zawodnika, do najlepszego piłkarza w tym atrybucie. Wzór na obliczenie:

**sdno** = (op/on)\*100, gdzie

**sdno** – stosunek do najlepszej oceny,

**op** – ocena piłkarza,

**on** – ocena najlepszego

Przykładowo, gdy najszybszy zawodnik w całej grze, będzie miał wartość równą 90, to zawodnik mający 60 szybkości, w stosunku do najlepszej oceny będzie posiadał statystykę na poziomie 66,66%. W przypadku statystyki „technika”, która w grze FIFA jest rozdzielona na drybling i podania, brana pod uwagę jest średnia dwóch wyżej wymienionych wartości. Dodatkowo, dla ułatwienia analizy w interfejsie graficznym, porównywana jest wyższa ocena, biorąc pod uwagę kolumnę „nie gorszy niż”. Jeśli zawodnik pod kątem jednego z atrybutów jest lepiej oceniony w jednej z gier, wyświetlony jest ciemniejszy kolor wiersza.

Obraz zawierający stół

Opis wygenerowany automatycznie

Rysunek 10 - Podstrona zawodnika w pracy inżynierskiej- część 1

W przypadku wartości z realnie rozgrywanych rozgrywek, dane wskazują poglądowo formę danego piłkarza. Jest to jednak inny rodzaj przedstawienia atrybutów, niż w grach piłkarskich. W przypadku statystyki „podania”, twórcy gier oceniają całokształt zawodnika w tej dziedzinie, na podstawie obserwacji realnych meczy. Statystyki przedstawiające celność, bądź ilość kluczowych podań, są w stanie w dużym stopniu pomóc w analizie i nakreślić dyspozycje piłkarza, której niekoniecznie zauważy przeciętny kibic. Jeśli warunki, w których znajdują się dwaj piłkarze są podobne, przykładowo zawodnicy występujący na tych samych pozycjach w danym klubie rywalizują ze sobą o miejsce w składzie, to większa liczba stworzonych sytuacji podbramkowych przez jednego z nich może wskazać potencjalnego zwycięzcę. Aczkolwiek wpływ na celność podań, asyst i innych czynników pobranych ze źródła Sofascore, jest często zależna od klubu i pozycji, w którym gra zawodnik. Większe szanse na zdobycie dużej ilości bramek, ma napastnik, który dzięki dobrej grze drużyny, znacznie częściej znajduje się w dogodnych sytuacjach. Podobnie sytuacja wygląda z innymi wartościami, które są generowane na podstawie spotkań. Z tego względu wyznaczenie wartości w zakresie od 1 do 99, podobnie jak w przypadku pozostałych źródeł, byłoby ciężkie do zrealizowania w taki sposób, aby było to w pełni miarodajne.

Obraz zawierający stół

Opis wygenerowany automatycznie

Rysunek 11 – Podstrona zawodnika w pracy inżynierskiej – część 2

Na wykresach klubów przedstawiona jest suma ocen danej statystyki spośród 15 najlepszych zawodników, wybranego klubu. Wykresy gier porównują średnią wartość atrybutu w obu źródłach.

Obraz zawierający tekst, pianino, zrzut ekranu

Opis wygenerowany automatycznie

Rysunek 12 - Wykresy mocy drużyn w pracy inżynierskiej

**2.3 Możliwe rozwinięcia**

Aplikacja dość rozlegle pozwala na zrealizowanie tematu, jakim jest analiza umiejętności piłkarzy w danym sezonie. Możliwe jest jednak jej rozwinięcie. Obecny stan pobiera informacje dotyczące zawodników jedynie z Premier League i ich dyspozycji z sezonu w 2021 roku. W razie rozszerzenia danych o następne lata, możliwe stałoby się prognozowanie oceniania w kolejnych edycjach, na podstawie dostarczonych z API Sofascore statystyk, takich jak średnia ocena. Oprócz tego dodanie większej ilości lig, pozwoliłoby na zwiększeniu ilości interesariuszy w tym temacie.

**2.4 Implementacja systemu**

Webscrapping statystyk z FIFY za pomocą Cheerio i strony futhead

**Obraz zawierający tekst, wewnątrz, zrzut ekranu

Opis wygenerowany automatycznie**

Rysunek 13 – Pobranie statystyk z portalu futhead [7]

Wyświetlanie powyższych statystyk w formie pliku .json na serwerze.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 14 – Wyświetlanie statystyk w API

Podczas pobierania danych, zdarzały się przypadki, gdy ten sam zawodnik wystąpił dwa razy. By uniknąć takich błędów sprawdzane są imiona graczy, w raz z ich oceną ogólną. Kopia zawodnika nie jest dopuszczana do utworzonego API.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 15 – Sprawdzanie pobranych danych – część 1

Złączone zostają obydwie tablice w jedną, w celu łatwiejszego dostępu do atrybutów zawodników, którzy występują w obydwu grach piłkarskich.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 16 – Sprawdzanie pobranych danych – część 2

Pobieranie statystyk zawodnika po jego nazwie za pomocą Sofascore API.

Obraz zawierający tekst, zrzut ekranu, ekran

Opis wygenerowany automatycznie

Rysunek 17 – Pobranie statystyk zawodnika po jego nazwisku z API Sofascore

1. **Moduł klubów**

Kolejnym modułem w aplikacji jest szacowanie zwycięzców spotkań. Wynik meczu jest najważniejszym elementem w piłce nożnej. Każdy kibic lubi, jak jego drużyna gra efektownie i prezentuje ofensywny styl, aczkolwiek to zwycięstwa gwarantują lepszą przyszłość. Dzięki dobrym wynikom klub może awansować do wyższej ligi, co przekłada się na konkurowanie z lepszymi drużynami. Lepsze drużyny powodują zwiększenie poziomu, przyciągają więcej fanów na trybuny oraz gwarantują większe wpływy z telewizji, które można zainwestować na większy rozwój klubu. Dlatego też zainteresowanie dotyczące typowania wyników jest popularne, co wykorzystują firmy bukmacherskie. Szacowania podawane przez zakłady bukmacherskie w teorii powinny być najbardziej rzetelne i trafne. Wpływ na to mają duże pieniądze zarabiane na ofertach proponowanych swoim klientom. Aby pieniądze zwracały się w tym biznesie, najpierw trzeba zainwestować spore środki na system przewidujący. Można zauważyć, że często na przedmeczowe kursy ma wpływ nie tylko obecna dyspozycja i siła, lecz także reputacja danej drużyny. Dodatkowo, wpływ na kursy ma ilość pieniędzy zainwestowanych przez klientów. Ten czynnik jest mocno subiektywny[10] . Przykładem takiej sytuacji jest mecz Legii Warszawa z Radomiakiem Radom w 19 kolejce sezonu 2021/2022. Drużyna z Warszawy do momentu rozegrania meczu zdobyła 18 punktów w 18 meczach i plasowała się na 15 pozycji, a forma w jej ostatnich meczach nie zachwycała. W opozycji do jednego z najpopularniejszych klubów w Polsce, drużyna Radomiaka, dla której był to pierwszy sezon w najwyższej klasie rozgrywkowej. Mimo małego doświadczenia zajmowali oni 4 miejsce z 32 punktami na koncie. Firmy bukmacherskie oferowały 2 złote i 20 groszy za każdą postawioną złotówkę na Legie, a na Radomiaka 3.55 złotych. Oznacza, że dosyć dużym faworytem była Legia [10]. Mecz ostatecznie zakończył się wynikiem 3-0 dla zespołu z Radomia. Aby zbadać takie sytuacje i lepiej zrozumieć czynniki, które wpływają na zwycięstwo, został stworzony ten moduł.

**3.1 Zasada działania**

**3.1.1 System na podstawie ostatnich rezultatów**

W celu zoptymalizowania szacowania do konkretnych wyników, system skupia się na jednej konkretnej lidze. Według statystyk, takie czynniki jak ilość remisów czy bramek, są zależne od kraju i poziomu rozgrywkowego [11]. Na ilość remisów wpływa w dużej mierze ilość bramek. Jeżeli wartość ta znajduję się w zakresie od 0 do 2 bramek to dwa z możliwych sześciu wyników oznacza podział punktów. W przypadku od 3 do 5 goli, zaledwie jeden z piętnastu rezultatów jest remisowych.

0: **0-0**

1: 1-0, 0-1

2: **1-1**, 0-2, 2-0

3: 3-0, 0-3, 2-1, 1-2

4: 4-0, 0-4, 3-1, 1-3**, 2-2**

5: 5-0, 0-5, 4-1, 1-4, 3-2, 2-3

W niektórych ligach, częściej można spotkać ofensywny styl i piłkarzom zależy na dużej zdobyczy bramkowej, gdyż to wpływa na atrakcyjność meczy i przyciąga więcej kibiców. W innych, częściej gra się bardziej zachowawczo, ponieważ to często pomaga w uzyskiwaniu dobrych wyników, co się przekłada na miejsce w tabeli. W dodatku niektóre narodowości słyną z tego, że są nadprzeciętnie dobre w danej dziedzinie. Przykładowo zawodnik z Brazylii kojarzy się z dobrym dryblingiem, szybkością i wieloma walorami ofensywnymi, a włoskie kluby z przykuwaniem dużej uwagi do taktyki.

Obraz zawierający stół

Opis wygenerowany automatycznie

*Rysunek 18 – Statystyka remisów w zależności od ligi – 2.12.2022 [11].*

Według danych portalu eprogresja.pl z dnia 2 grudnia 2022 roku, [11] najczęściej remis pada w bahrańskiej Premier League, a procent takich wyników wynosi 40%. Po drugiej stronie Indyjska Super League, gdzie ten współczynnik jest równy 9.52%. Można zauważyć, że wpływ ligi na ilość remisów jest ogromny.

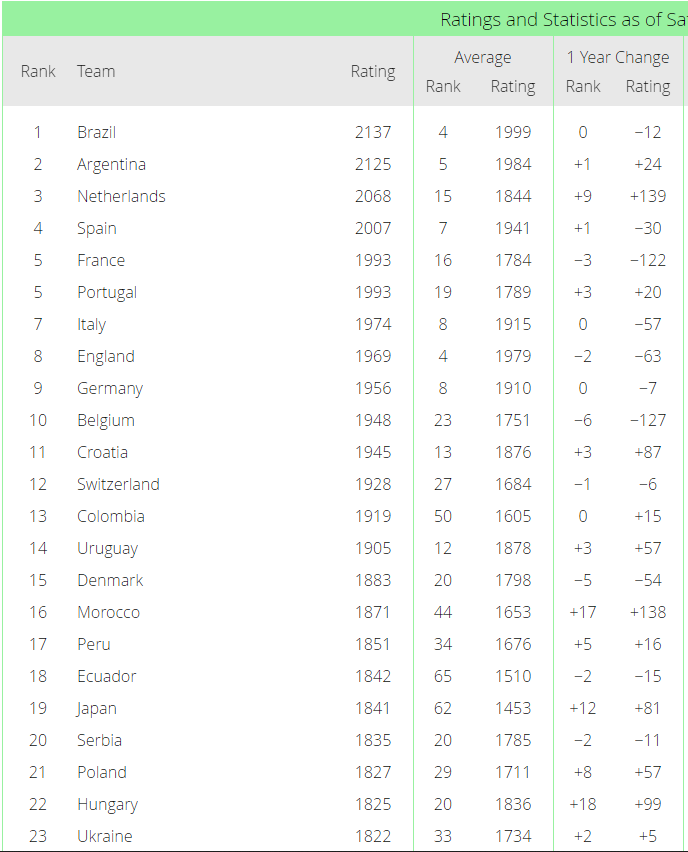
**Wartości najbardziej popularnych lig**:

* **angielska Premier League**: 23.29%,
* **hiszpańska La Liga**: 23.57%,
* **niemiecka Bundesliga** 22.96%
* **francuska Ligue 1**: 26.67%
* **włoska Serie A**: 24%
* **polska Ekstraklasa**: 23.03%
* **polska 1 liga**: 31.48%

Do analizy wybrana została polska Ekstraklasa, ze względu na największą ilość dostępnych, znanych artykułów, które są pomocne podczas szacowania wyników.

**3.1.1.1 Wyznaczenie zwycięzcy**

Pierwszym zadaniem aplikacji jest pobranie informacji dotyczących ostatnich meczy w Ekstraklasie. Do uzyskania ich użyte zostało API Flashscore. Za jego pomocą tworzone zostają oceny siły drużyn, w podobny sposób do rankingu ELO. Ranking ten jest stosowany do określenia mocy reprezentacji narodowych w piłce nożnej. Oficjalny ranking federacji piłki kobiecej opiera się na tym systemie. W przypadku reprezentacji narodowych stosuje się go na podstawie trzydziestu lub większej ilości spotkań. Taka pula meczy wynika z dużej liczby narodowości mogących grać między sobą i możliwości gry przez dłuższy okres czasu wyłącznie ze słabszymi zespołami.



Rysunek 19 - Ranking ELO męskich drużyn narodowych – 3.12.2022 [12]

Podczas rozgrywek ligowych w trakcie jednego sezonu trwającego około rok czasu, drużyny mierzą się z wszystkimi pozostałymi przeciwnikami, a ich liczba wynosi zazwyczaj od 12 do 20. Często zdarza się, że dyspozycja danej drużyny zmienia się bardzo szybko i wyniki sprzed roku nie są adekwatne do obecnej sytuacji w klubie. Takim przykładem jest forma Wisły Płock w pierwszych spotkaniach rundy jesiennej oraz wiosennej sezonu 2022/2023. Po szóstej kolejce, dzięki dobrej dyspozycji, zostali liderem tabeli. Jednak po przerwie zimowej, ich forma znacząco uległa pogorszeniu.

Obraz zawierający stół

Opis wygenerowany automatycznie*Rysunek 20 – Tabela Ekstraklasy za okres 15.07.2022 – 22.08.2022 [13]*

Obraz zawierający tekst

Opis wygenerowany automatycznie*Rysunek 21 – Tabela Ekstraklasy za okres 27.01.2023 – 7.03.2023 [13]*

Wpływ na tak gwałtowną zmianę pozycji mają w dużym stopniu roszady kadrowe. Dobre rezultaty zespołu wynikają nierzadko z nadprzeciętnej dyspozycji jednego czy dwóch piłkarzy, którzy w trakcie okienka transferowego przechodzą do innych, lepszych klubów. Skutkiem tego w tej metodzie stosowany jest nacisk przy szacowaniu wyników na bieżącą formę. Do punktacji pod uwagę brany jest przekrój ostatnich 15 meczy. Jest to reprezentatywna ilość danych, gdyż zespół mierzy się przez ten okres z praktycznie wszystkimi drużynami w lidze. Rywalizuje zarówno z lepszymi i słabszymi przeciwnikami i nie ma mowy o tym, że dobra dyspozycja była spowodowana jedynie sprzyjającym terminarzem. Z drugiej strony mecz, który odbywał się rok czy dwa lata temu nie powinien mieć wpływu na wynik z obecnego okresu.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 22 - Metoda obliczania punktów ELO – 18.11.2022 [14]

W rankingu ELO indeksem ważności oznaczana jest ranga meczu. W przypadku spotkania podczas mistrzostw świata, waga jest sporo większa niż w meczu towarzyskim. Jeśli chodzi o spotkania ligowe, każdy mecz jest równie ważny, ponieważ niezależnie od przeciwnika zdobycz punktowa w razie zwycięstwa jest taka sama. W tabeli, która weryfikuje siłę drużyn, liczy się suma wszystkich punktów i to od niej zależy, miejsce które drużyna zajmie w tabeli. To przekłada się na utrzymanie, bądź awans. Aplikacja weryfikuje formę drużyny, a w ocenieniu tego pomagają ostatnie mecze, szczególnie te które odbywały się w mniejszym odstępie czasowym od obecnej daty. System, który przy testowaniu dawał najlepsze efekty wygląda następująco:

**K = 2 – (X/7.5),** gdzie:

**K** – indeks ważności meczu,

**X** – numer meczu.

Podczas 15 spotkań drużyny rywalizują zarówno z lepszymi i słabszymi drużynami, stąd brak konieczności dodatkowego punktowania w zależności od siły przeciwników. Parametr, który dodatkowo znajduje się w punktacji, jest teren rozgrywania spotkania. W analizie statycznej Ekstraklasy z sezonu 2021/2022 można znaleźć informacje pokazujące, jak ważnym czynnikiem jest to, czy drużyna gra u siebie czy na wyjeździe.

*„W 306 meczach zdobyto 818 goli, co daje średnią 2,673 goli na mecz. Gospodarze zdobyli 459 goli (56%), a goście 359 goli (44%) (...) Gospodarze wygrali 135 meczów (44%), padły 84 remisy (27%), a goście wygrali 87 meczów (28%).”*

*[15]*

Stosunek zwycięstw gości do gospodarzy wynosi 64,(4)%. W lidze każda drużyna rozgrywa ze sobą po dwa mecze, zarówno na swoim jak i wyjazdowym terenie. Widoczna jest spora zależność wyniku od miejsca w którym spotkanie jest rozgrywane. Około jednej trzeciej meczów kończy się triumfem gospodarzy, a decydującym czynnikiem jest ich boisko, licznie zgromadzeni kibice oraz brak potrzeby męczących podróży przed spotkaniem. Jednakże ta sytuacja stanowi różnej wielkości atut w zależności od drużyny. W sezonie 2021/2022 aż 15 na 18 zespołów zdobyło więcej punktów u siebie względem meczów na wyjeździe. U niektórych zespołów widać było dużą dysproporcje pomiędzy tymi dwoma statystykami. Dobrym przykładem jest 40 zdobytych punktów Lechii Gdańsk u siebie i tylko 17 na obcym boisku. Podobnie Wisła z dorobkiem 35 punktów w Płocku i 13 na wyjazdach. Zdarzają się jednak sytuacje, w których drużynom lepiej się gra obcym boisku. W takiej sytuacji jest Warta Poznań, która rozgrywa swoje spotkania domowe w Grodzisku Wielkopolskim, ze względu na brak własnego stadionu, spełniającego wymogi licencyjne na najwyższym poziomie rozgrywkowym. W sezonie 2021/2022 stosunek punktów domowych do wyjazdowych Warty wynosił 18:24. Po rundzie jesiennej kolejnego sezonu, czyli po rozegraniu połowy spotkań, ta różnica jest sporo większa i wynosi 6:19. Mając na uwadze powyższe statystyki, system prócz punktów mających określić formę w ostatnich 15 meczach, do rankingu klubu wliczana jest forma zespołu na terenie, w którym będzie się odbywał mecz. Jeśli drużyna gra następny mecz u siebie, 1/3 punktów w rankingu będzie zależna od zdobyczy w ostatnich 5 domowych spotkaniach. Analogicznie w przypadku spotkań wyjazdowych. Ogólny wzór na obliczanie punktów w rankingu, który będzie miał zakres procentowy, czyli wartości od 0 do 100 jest następujący:

**PP = ((2\*(PO/45)+(PT/15))/3)\*100**, gdzie:

**PP** – punkty procentowo,

**PO** – punkty ogólne, czyli liczba punktów w ostatnich 15 meczach pomnożona razy indeks ważność meczu k(wzór powyżej, uśredniona wartość równa 1), 45 to maksymalna wartość po,

**PT** – punkty terenowe, czyli liczba punktów w ostatnich 5 meczach u siebie, bądź na wyjeździe w zależności od tego, gdzie dana drużyna gra, 15 to maksymalna wartość pt.

Celem aplikacji nie jest tylko i wyłącznie wygenerowanie siły drużyny, lecz wynik w spotkaniu z drużyną przeciwną. Aby móc typować jakim wynikiem zakończy się spotkanie potrzebne są informacje o tym, która drużyna zwycięży, jaka będzie różnica bramkowa i ile bramek padnie. W piłce nożnej dosyć często dochodzi do remisów i trzeba je też brać pod uwagę przy typowaniach. Przy obliczaniu tylu działań rzadko zdarzają się sytuacje, kiedy siła obu drużyn będzie identyczna. W celu rozwiązania tego problemu, nastąpiła komunikacja z mgr. Inż. [Piotrem Klimkiem](https://twitter.com/pklimek99) [16], który specjalizuje się w matematycznych obliczeniach i wykresach dotyczących piłki nożnej. Posiada on liczną grupę obserwujących na serwisie społecznościowym Twitter i jest uważany za eksperta w swojej dziedzinie. Dodatkowo brane są pod uwagę statystyki z poprzedniego sezonu. W ostatnim sezonie liczba remisów wynosiła 23,03%. Wzory, które po konsultacjach i analizie okazały się najlepsze do podziału sił drużyn A i B na procenty A, B i remis wynoszą następująco:

* **Jeśli H >= 60:**
* HW = 50 + H-60,
* DW = 30 - (H-60)/2,
* AW = 20 - (H-60)/2.
* **Jeśli H > 50 i H < 60:**
* HW = 30 + (H-50)\*2,
* DW = 40 - (H-50),
* AW = 30 - (H-50).
* **Jeśli H <= 50 i H > 40:**
* HW = 30 - (A-50),
* DW = 40 – (A-50),
* AW = 30 + (A-50)\*2.
* **Jeśli H <= 40:**
* HW = 20 – (A-60)/2,
* DW = 30 - (A-60)/2,
* AW = 50 + A-60.
* Gdzie

H – procentowy stosunek siły drużyny gospodarzy do siły obu drużyn, H równe 100-A,

A – procentowy stosunek siły drużyny gości do siły obu drużyn, A jest równe 100 – H,

HW – procent szacowania na zakończenie spotkania zwycięstwem gospodarzy,

DW – procent szacowania na zakończenie spotkania remisem,

AW – procent szacowania na zakończenie spotkania zwycięstwem gości,

* **Przykładowe rozkłady:**
* **55-45**: 40-35-25,
* **59-41**: 48-31-21,
* **70-30**: 60-25-15.

**3.1.1.2 Wyznaczenie różnicy bramkowej**

Posiadając parametry HW(procent na wygraną gospodarzy), DW(procent na remis) i AW(procent na wygraną gości) można określić potencjalnego zwycięzcę spotkania. Wiadomą sprawą jest, że czym większa wartość danej statystyki tym większa szansa na zwycięstwo przez daną drużynę. Wynika to z różnicy siły pomiędzy drużynami. Skoro drużyna A jest sporo lepsza od drużyny B, to można przypuszczać, że różnica bramkowa będzie trochę większa, niż w przypadku zbliżonych wartości. Z analizy statystycznej Ekstraklasy wynika, że: „

***Częstość wyników:*** *1-1 - 46 razy, 1-0 - 35 razy, 2-1 - 25 razy, 0-1 - 24 razy, 1-2 - 20 razy, 0-0 - 19 razy, 2-0 - 17 razy, 3-1 - 16 razy, 2-2 - 15 razy, 3-0 - 13 razy, 0-2 - 13 razy, 1-3 - 10 razy, 4-1 - 8 razy, 0-3 - 5 razy, 0-4 - 5 razy, 3-2 - 5 razy, 4-0 - 5 razy, 2-3 - 4 razy, 3-3 - 4 razy, 4-2 - 4 razy, 5-0 - 3 razy, 3-4 - 2 razy, 4-3 - 2 razy, 2-4 - 2 razy, 5-1 - 1 raz, 1-6 - 1 raz, 0-5 - 1 raz, 5-3 - 1 raz*

***Najwięcej goli (8) padło w meczu:***[*Legia Warszawa 5-3 Górnik Zabrze*](http://www.90minut.pl/mecz.php?id_mecz=1677841)

***Najwyższe zwycięstwo gospodarzy padło w meczach:***[*Lech Poznań 5-0 Wisła Kraków*](http://www.90minut.pl/mecz.php?id_mecz=1677625)*,*[*Raków Częstochowa 5-0 Jagiellonia Białystok*](http://www.90minut.pl/mecz.php?id_mecz=1677575)*,*[*Lech Poznań 5-0 Bruk-Bet Termalica Nieciecza*](http://www.90minut.pl/mecz.php?id_mecz=1677743)

***Najwyższy remis padł w meczach:***[*Cracovia 3-3 Stal Mielec*](http://www.90minut.pl/mecz.php?id_mecz=1677642)*,*[*Jagiellonia Białystok 3-3 Piast Gliwice*](http://www.90minut.pl/mecz.php?id_mecz=1677672)*,*[*Stal Mielec 3-3 Lechia Gdańsk*](http://www.90minut.pl/mecz.php?id_mecz=1677693)*,*[*Cracovia 3-3 Lech Poznań*](http://www.90minut.pl/mecz.php?id_mecz=1677733)

***Najwyższe zwycięstwo gości padło w meczach:***[*Wisła Kraków 0-5 Śląsk Wrocław*](http://www.90minut.pl/mecz.php?id_mecz=1677667)*,*[*Radomiak Radom 1-6 Zagłębie Lubin*](http://www.90minut.pl/mecz.php?id_mecz=1677843)*”*

*[15]*

Można zauważyć, że spośród 222 meczy, które nie zakończyły się remisem różnica bramkowa była następująca:

* **wygrane 1 bramką**: 117 (52,7%),
* **wygrane 2 bramkami**: 63 (28,4%),
* **wygrane 3 bramkami**: 26 (11,7%),
* **wygrane 4 bramkami**: 11 (4,9%),
* **wygrane 5 bramkami**: 5 (2,3%).

Dopasowując to do trzytorowych procentów na dany wynik (HW – wygrana gospodarzy, DW – remis, AW – wygrana gości),w praktyce najsensowniejszy podział był w momencie, gdy zależności procentów do różnicy bramkowej były następujące:

* WP < DW+10 – **różnica 1 bramki**,
* WP >= DW+10 i WP < DW+25 – **różnica 2 bramek**,
* WP >= DW+25 i WP < DW+40 – **różnica 3 bramek**,
* WP >= DW+40 – **różnica 4 bramek**, gdzie:
* **WP** – procent na drużynę, która według trzytorowych typowań ma największe szanse na wygraną,
* **DW** – procent na remis w tym meczu.

**3.1.1.3 Wyznaczenie dokładnego wyniku**

Następnie obliczana jest ilość strzelonych bramek przez oba zespoły w ostatnich 15 kolejkach. Suma wszystkich goli wpływa na stopień bramkostrzelności obu zespołów:

* **Stopień 1** – poniżej 2 goli,
* **Stopień 2** – 2.0-2.33,
* **Stopień 3** – 2.34-2.66,
* **Stopień 4** – 2.67-2.99,
* **Stopień 5** – powyżej lub równe 3 gole.

W zależności od stopnia bramkostrzelności obu ekip i zwróconej różnicy pomiędzy zespołami typowany wynik jest podawany przez aplikacje. Algorytmy dokładnych wyników zostały utworzone na podstawie cytowanych wyżej częstości wyników z portalu 90minut. Zależność jest następująca:

* Remis + stopień 1 = **0-0**,
* Remis + stopień 2 lub stopień 3 = **1-1**,
* Remis + stopień 4 lub stopień 5 = **2-2**,
* Wygrana 1 bramką + stopień 1 lub stopień 2 = **1-0**,
* Wygrana 1 bramką + stopień 3 lub stopień 4 = **2-1**,
* Wygrana 1 bramką + stopień 5 = **3-2**,
* Wygrana 2 bramkami + stopień 1 lub stopień 2 = **2-0**,
* Wygrana 2 bramkami + stopień 3 lub stopień 4 = **3-1**,
* Wygrana 2 bramkami + stopień 5 = **4-2**,
* Wygrana 3 bramkami + stopień 1 lub stopień 2 lub stopień 3 = **3-0**,
* Wygrana 3 bramkami + stopień 4 = **4-1**,
* Wygrana 3 bramkami + stopień 5 = **5-2**,
* Wygrana 4 bramkami = **4-0**.

**3.1.2 Kursy bukmacherów**

Kolejnym źródłem są kursy od bukmacherów. Jeśli kurs wynosi 2.0, to za każdą postawioną złotówkę, klient podwoi swoją kwotę. Z matematycznego punktu widzenia, szansa na wygraną w takich sytuacjach wynosi 50%. W momencie gdy kurs jest równy i 4.0, wskazuje to na szanse 25%. Wzór na obliczanie szans wynosi 1/(wysokość kursu) [10]. Zazwyczaj suma szans na zwycięstwo gospodarzy, gości oraz remis wynosi 105-110%. Wynika to z tego, że większy procent szans według tej metody oznacza mniejszą możliwą wygraną dla klienta. Zakłady bukmacherskie są konstruowane w taki sposób, aby móc zarabiać na swoich klientach. Aby móc czytelniej porównywać dane, powyższe procenty zostały podzielone w taki sposób by całość wynosiła 100%. Zakłady bukmacherskie w Polsce obejmuje 12% podatek. Żeby móc sprawdzić jakie minimalne szacowanie powinno być by na powyższy mecz opłacane było stawianie swoich pieniędzy, stworzona została statystyka „opłacalne szanse na wynik”. Jest to wysokość szacowania na dany wynik pomnożony przez 88% (kwota, która nie jest objęta podatkiem). Ze względu na to, że od kursów zależny jest zarobek bukmacherów, to źródło powinno dawać najlepsze wyniki. Jednak ewentualne znalezienie alternatywnego źródła o podobnej skuteczności, byłoby dużym odkryciem i zagrożeniem dla firm bukmacherskich.

**3.1.3 Forebet**

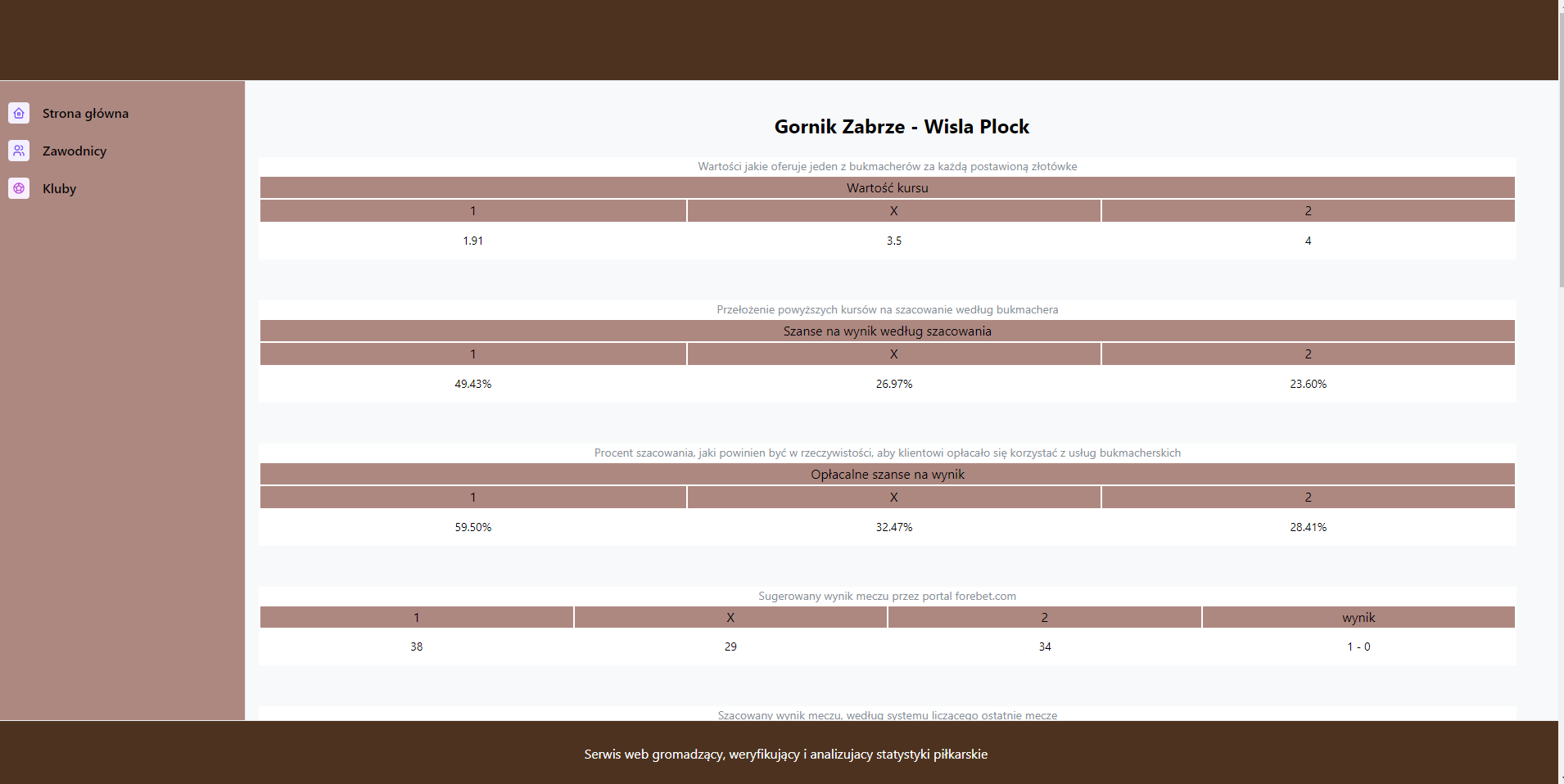
Ostatnim źródłem są dane z witryny forebet.com. Ta strona oferuje nie tylko procent szans na zwycięstwo danej drużyny, ale również zwraca sugerowany wynik. Stanowi dobrą alternatywę do pozostałych analiz. Podobnie jak system szacujący wynik spotkania, opiera ona się na matematyce i dotychczasowych meczach. Dokładny sposób liczenia procentów i wyników przez tą witrynę nie jest publicznie ujawniony. Pokazuje on czasami inne wyniki od pozostałych źródeł, lecz w pozyskanych wartościach widoczna jest zależność.

**3.2 Wyświetlenie meczów**

Podstrona z analizami konkretnego meczu znajduje się po przejściu od terminarza w module klubów, gdzie można wybrać spotkanie, którym się jest zainteresowany. Dostępne są mecze, których termin rozpoczęcia, jest nie dłuższy niż dwa tygodnie od obecnej daty. Powodem tego jest fakt, że mecze odbywają się raz na tydzień. Ocena zwycięzcy meczu, który będzie się odbywał w odległym terminie mija się z celem, gdyż podczas rozegrania tego meczu, zespół może być w zupełnie innej formie. Z uwagi na to, kursy na mecz nie są dostępne u bukmacherów w dłuższym odstępie czasowym, podobnie jak z danymi z portalu forebet. Spotkania są wyświetlane w kolejności najbliższej daty. Problemem okazał się sposób zwracania terminu spotkania przez API Flashscore. Zamiast daty pojawiała się liczba, przykładowo 1654293600. Znając terminy realnych spotkań , można było porównywać do siebie dwie wartości. Wniosek obserwacji jest taki, że 3600 jest równy godzinie różnicy pomiędzy meczami. Dzięki takiej wiedzy, stworzony został kompilator liczby uzyskanej z API na datę.

**3.3 Przykładowe widoki**

Podstrona wyświetlająca szczegółowe analizy dotyczące konkretnego meczu



Rysunek 23 – Widok konkretnego meczu

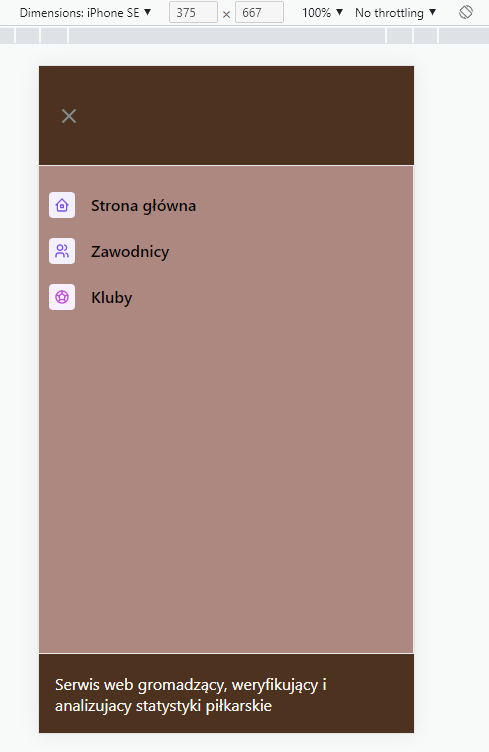
Podstrona wyświetlająca szczegółowe informacje o meczu w widoku mobilnym

Obraz zawierający stół

Opis wygenerowany automatycznie

Rysunek 24 – Widok konkretnego meczu w wersji mobilnej

Rozwinięty boczny panel w widoku mobilnym



Rysunek 25 – Widok panelu bocznego w wersji mobilnej

**3.4 Możliwe rozwinięcia**

Aplikacja pobiera większość informacji potrzebnych do generowania po stronie z API Flashscore. Obecnie nie jest to problemem. Lecz w razie rozwoju aplikacji i skierowania jej w większą liczbę odbiorców, dobrym rozwiązaniem byłoby uniezależnienie od niezawodności zewnętrznych źródeł. Dane z API można pobierać raz przed każdym cyklem spotkań (zazwyczaj odbywają się one od piątku do poniedziałku). W takiej sytuacji generowanie zapytania do API za każdym kliknięciem w podstronę nie jest potrzebne. Zamiast tego można skorzystać z bazy danych przy użyciu Redis. W razie takiej zmiany, możliwe stałoby się zapisywanie szacowań z wcześniejszych meczy. Umożliwiłoby to obliczanie skuteczności typowania każdego ze źródeł. Dodatkowo, możliwe jest rozszerzenie aplikacji o większa liczbę rozgrywek ligowych. Istnieje również możliwość włączenia statystyk z meczu (posiadanie piłki, strzały, niebezpieczne ataki) do skryptu typującego rezultat. Te dane dostępne są w API Flashscore i mogłyby ulepszyć tę funkcjonalność, aczkolwiek praca nad tym wymagałaby głębszej analizy i eksperymentów nad nowym skryptem.

**3.5 Implementacja systemu**

Funkcja pobierająca z API ostatnie mecze dwóch drużyn, których wynik meczu będzie liczony. Zmienna tablicowa „ekstraklasaTab” została stworzona z powodu wadliwości API. Z założenia w pliku .json każdy numer tablicy był przydzielony do osobnych rozgrywek

**Obraz zawierający tekst

Opis wygenerowany automatycznie**

Rysunek 26 – Pobranie ostatnich spotkań

Funkcja pobierająca z API wysokość kursów bukmacherskich

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 27 – Pobranie kursów bukmacherskich

Wyświetlanie danych z portalu Forebet

Obraz zawierający tekst, wewnątrz, ekran, zrzut ekranu

Opis wygenerowany automatycznie

Rysunek 28 – Wyświetlenie danych z portalu forebet

W API data przesyłana jest jako liczba, 1654293600 to 4 czerwca 2022 roku - północ. Zauważyć to można było na podstawie wcześniej rozgrywanych meczy. Kod przedstawia kompilator liczby na format daty.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 29 – Konwersja danych z API – część 1

Zwrócenie daty (ciąg dalszy Rysunek 29) w odpowiednim formacie

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 30 – Konwersja danych z API – część 2

**4 Testy**

**4.1 Typowane wyniki a rzeczywistość**

Jako przykład zostały wzięte pod uwagę kolejki 17-25 w sezonie 2022/2023 w Ekstraklasie. Kolejka oznacza wszystkie mecze odbywające się przez dany weekend w lidze. W niej zespoły rozgrywają po jedno spotkanie. Są to wszystkie mecze możliwe do sprawdzenia w trakcie jednego tygodnia, oraz takie które równomiernie testują wszystkie zespoły grające w tej lidze. Powyższy zakres to wszystkie mecze z cykli spotkań, które odbywały się od 11 listopada 2022 do 19 marca 2023. Spośród 72 rywalizacji, zwycięzcy zostali poprawnie wytypowani w 34 spotkaniach (47,2%) według danych od bukmacherów i portalu forebet. Niewiele mniej, bo 33 rozstrzygnięcia (45,8%) były zgodne z szacowaniem systemu obliczającego wynik meczu na podstawie ostatnich spotkań. Udane próby trafienia w zwycięzcę oraz różnice bramek wynosiły 21 (29,2%) w przypadku zarówno forebet i bukmachera, oraz 16 (22,2%) w przypadku systemu. Trafność dokładnych wyników, zgodna z danymi serwisu forebet oraz wskazaniami bukmachera, wyniosła dokładnie 9 razy (czyli 12,5%). Natomiast w przypadku systemu wynik sprawdził się 5 razy (czyli 6,9%).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Spotkanie** | **Rezultat** | **Bukmacher [18]** | **System** | **Forebet** |
| Górnik – Miedź | 0-3 | 1 / 1-0 | 4-2 | 2-0 |
| Jagiellonia – Lech | 1-2 | 2 / 1-1 | **1-2** | **1-2** |
| Korona – Widzew | 0-1 | 2 / 1-1 | 0-3 | 0-2 |
| Lechia – Piast | 1-3 | 1 / 1-1 | 1-0 | 0-0 |
| Radomiak – Pogoń | 1-2 | 2 / 1-1 | 2-2 | 2-2 |
| Śląsk – Legia | 0-0 | 2 / **1-1** | 1-2 | **0-0** |
| Warta – Stal | 1-1 | 1 / 1-0 | **1-1** | 0-2 |
| Wisła – Cracovia | 1-0 | 1 / 1-1 | 1-2 | 1-1 |
| Zagłębie – Raków | 1-2 | **2 / 0-1** | 0-2 | **1-2** |
| Cracovia – Górnik | 2-0 | 1 / 1-1 | 2-1 | 1-1 |
| Lechia – Wisła | 1-0 | 1 / 1-1 | **2-1** | 1-2 |
| Legia – Korona | 3-2 | 1 / 2-0 | 3-0 | 3-0 |
| Miedź – Radomiak | 0-0 | 1 / **1-1** | 1-2 | 1-2 |
| Piast – Jagiellonia | 1-1 | 1 / **1-1** | 0-1 | **1-1** |
| Stal – Lech | 0-0 | 2 / 0-1 | 0-1 | **1-1** |
| Śląsk – Zagłębie | 0-3 | 1 / 1-1 | 2-1 | 2-1 |
| Warta – Raków | 1-1 | 2 / 0-1 | 1-2 | 0-2 |
| Widzew – Pogoń | 3-3 | 1 / **1-1** | 2-1 | **1-1** |
| Górnik – Lechia | 1-1 | 1 / **1-1** | 1-2 | 1-2 |
| Korona – Cracovia | 2-1 | 2 / 1-1 | 0-3 | 0-1 |
| Lech – Miedź | 1-0 | **1 / 1-0** | 3-1 | **2-1** |
| Pogoń – Śląsk | 0-2 | 1 / 1-0 | 2-1 | 2-1 |
| Radomiak – Stal | 1-0 | 1 / 1-1 | 0-1 | 1-1 |
| Raków – Piast | 1-0 | **1 / 1-0** | 3-1 | **2-1** |
| Widzew – Jagiellonia | 1-1 | 1 / **1-1** | 3-1 | 2-1 |
| Wisła – Warta | 1-0 | 2 / 1-1 | 0-1 | 1-2 |
| Zagłębie – Piast | 0-2 | 2 / 1-1 | 1-2 | 1-1 |
| Cracovia – Stal | 2-1 | **1 / 1-0** | **1-0** | 1-1 |
| Korona – Lechia | 1-0 | **1** / 1-1 | 1-3 | 1-1 |
| Lech – Zagłębie | 1-2 | 1 / 1-1 | 3-1 | 1-0 |
| Miedź – Wisła | 2-1 | 1 / 1-1 | **2-1** | 0-2 |
| Piast – Legia | 0-1 | 2 / 1-1 | 1-3 | 0-2 |
| Pogoń – Warta | 3-1 | 1 / 1-0 | 1-2 | 1-1 |
| Radomiak – Jagiellonia | 0-0 | 1 / **1-1** | 1-0 | **2-2** |
| Raków – Górnik | 2-0 | 1 / 1-0 | **4-2** | **2-0** |
| Widzew – Śląsk | 1-0 | **1 / 1-0** | **1-0** | 3-1 |
| Jagiellonia – Raków | 1-2 | 2 / **0-1** | 1-3 | 0-2 |
| Lechia – Radomiak | 1-3 | 1 / 1-0 | 1-0 | 0-1 |
| Legia – Widzew | 2-2 | 1 / **1-1** | 2-1 | **1-1** |
| Piast – Cracovia | 2-1 | 1 / **1-0** | 0-2 | **2-1** |
| Stal – Górnik | 0-1 | 2 / 1-1 | 2-1 | 1-0 |
| Śląsk – Lech | 2-1 | 2 / 0-1 | 0-1 | 0-1 |
| Warta – Korona | 5-1 | 1 / 1-0 | 2-1 | 1-0 |
| Wisła – Pogoń | 0-1 | 2 / 1-1 | **1-2** | 2-1 |
| Zagłębie – Miedź | 2-1 | 1 / 1-1 | 1-2 | 1-1 |
| Cracovia – Śląsk | 1-1 | 1 / 1-0 | 1-0 | **1-1** |
| Górnik – Legia | 0-1 | 2 / 1-1 | 1-3 | 0-2 |
| Korona – Wisła | 1-0 | 1 / 1-1 | **2-1** | **2-1** |
| Lech – Lechia | 5-0 | **1 / 2-0** | **1-0** | **2-0** |
| Miedź – Jagiellonia | 1-1 | 1 / **1-1** | 1-0 | **1-1** |
| Pogoń – Raków | 0-2 | 2 / 1-1 | 2-3 | 0-1 |
| Radomiak – Zagłębie | 0-1 | 1 / 1-1 | 2-1 | 2-0 |
| Stal – Piast | 0-2 | 2 / 1-1 | 0-1 | 0-0 |
| Widzew – Warta | 0-2 | 1 / 1-1 | 1-1 | 1-0 |
| Jagiellonia – Górnik | 2-1 | 2 / 1-1 | 1-2 | 1-1 |
| Korona – Radomiak | 2-1 | 2 / 1-1 | 0-1 | 0-1 |
| Lechia – Miedź | 4-0 | 1 / 1-1 | 1-0 | 1-1 |
| Legia – Stal | 2-0 | 1 / 1-0 | **3-1** | 2-1 |
| Piast – Lech | 1-1 | 2 / **1-1** | 0-1 | 0-2 |
| Raków – Śląsk | 4-1 | 1 / 2-0 | 3-1 | 1-0 |
| Warta – Cracovia | 0-0 | 1 / **1-1** | 2-1 | **1-1** |
| Wisła – Widzew | 1-1 | 2 / **1-1** | 0-1 | **0-0** |
| Zagłębie – Pogoń | 0-1 | 2 / 1-1 | **2-3** | 0-2 |
| Górnik – Wisła | 3-2 | 1 / 1-1 | **2-1** | **1-0** |
| Jagiellonia – Zagłębie | 2-2 | 1 / **1-1** | 1-2 | 2-0 |
| Miedź – Piast | 0-1 | 2 / 1-1 | **0-1** | **2-3** |
| Pogoń – Korona | 0-0 | 1 / 1-0 | 2-1 | 2-1 |
| Radomiak – Legia | 0-2 | 2 / 1-1 | **1-3** | 0-1 |
| Raków – Śląsk | 4-1 | 1 / 2-0 | 3-1 | 1-1 |
| Śląsk – Stal | 1-1 | 1 / **1-1** | 2-0 | **1-1** |
| Warta – Lechia | 2-0 | 1 / 1-0 | 2-1 | 2-1 |
| Widzew – Lech | 1-2 | 2 / 1-1 | **0-1** | 1-1 |

Tabela 1 – Porównanie wyników spotkań do innych źródeł

Podkreślenie: poprawny zwycięzca,

Pogrubienie: poprawna różnica bramek,

Zielony kolor: dokładny wynik.

* **Przykłady poprawnych szacowań systemu na podstawie ostatnich rezultatów**
* Widzew Łódź 1-0 Śląsk Wrocław 17.02.2023

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 31 – Przykład poprawnego szacowania – część 1 [17]

* Miedź Legnica 2-1 Wisła Płock – 19.02.2023

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 32 – Przykład poprawnego szacowania – część 2 [17]

* Wisła Płock 0-1 Lech Poznań – 11.02.2023

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 33 – Przykład poprawnego szacowania – część 3 [17]

* Zagłębie Lubin 1–2 Legia Warszawa – 04.02.2023

Obraz zawierający tekst

Opis wygenerowany automatycznie

Rysunek 34 – Przykład poprawnego szacowania – część 4 [17]

Wygrana gospodarzy oznacza 1, remis X, a zwycięstwo gości 2. Wyniki które zgadzają się z rzeczywistością, są podkreślone i pogrubione. Wskazania bukmachera opierają się na najmniejszym kursie w kategorii dokładnych wyników [18]. Zazwyczaj jest to wynik 1-0 bądź 1-1. Wynik 1 bramkowy pojawia się zazwyczaj w momencie, gdy jedna z drużyn jest wyraźnie lepsza od drugiej. Remis sugerowany jest, gdy spotkanie według tego źródła powinno być bardziej wyrównane. Zdarzają się takie sytuację, jak np. w meczu Lechia – Piast [Tabela 1], gdzie typowanym wynikiem według bukmachera jest wygrana Lechii, a dokładny wynik to remis 1-1. Fakt, że bukmacherzy szacują większość meczy głównie na te dwa wyniki to nie przypadek. W piłce nożnej statystycznie pada około 2,5 gola w zależności od ligi [11]. W momencie, gdy padnie dużo bramek, wynik jest trudniejszy do typowania, gdyż możliwych wyników z dużą ilością bramek jest sporo więcej. Można to zauważyć w podsumowaniu sezonu 2021/2022 prowadzonego przez portal 90 minut [15]. Wyniki te są najczęściej występującymi rezultatami. Spośród 306 meczy aż 105 zakończyło się wynikiem 1-0 bądź 1-1 [15].

**4.2 Zależność pomiędzy ocenami z Sofascore i danymi z gier piłkarskich**

Sprawdzane są oceny ogólne z gier najlepszych zawodników z Premier League i porównywane do jakości występów w przyszłym sezonie. Jakość występów jest definiowana jako średnią ocena z całego sezonu według Sofascore. Ta witryna wystawia oceny za każdy mecz na podstawie szczegółowych statystyk, takich jak ilość odbiorów czy celne podania . Można te dane analizować w szczegółowy sposób i każdy atrybut przyrównywać do pozostałych źródeł. Pomoże to w obserwacji oceniania w grach piłkarskich. Aby porównanie było jak najbardziej przejrzyste w każdej z kategorii, wartością obserwowaną będzie jakość piłkarza w stosunku do reszty, czyli procent zawodników od których jest lepszy bądź tak samo dobry. W tabeli porównywanych jest najlepszych 10 zawodników, którzy rozegrali minimum 450 minut w sezonie.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Imię i nazwisko** | **FIFA** | **FM\*** | **Sofascore** | **Źródło z bliższą liczbą** |
| Bernardo Silva | 98.17 | 100 | 73.74 | FIFA |
| Kevin De Bruyne | 100 | 99.77 | 100 | FIFA |
| Sadio Mane | 99.77 | 99.77 | 97.26 | brak |
| Mohamed Salah | 99.77 | 99.77 | 93.15 | brak |
| Harry Kane | 98.86 | 99.77 | 99.77 | FM |
| Andrew Robertson | 98.17 | 99.77 | 84.47 | FIFA |
| Bruno Fernandes | 98.17 | 99.77 | 99.54 | FM |
| Sergio Aguero | 99.09 | 98.17 | 71 | FM |
| Ngolo Kante | 98.86 | 98.17 | 78.77 | FM |
| Fabinho | 98.17 | 98.17 | 73.06 | Brak |
| Raheem Sterling | 98.86 | 87.21 | 60.27 | FM |

Tabela 2 – Porównanie statystyk zawodników – część 1, FM\* – Football Manager

Można zauważyć, że niektórzy zawodnicy, którzy mieli być tymi najlepszymi, spisywali się co najwyżej dobrze i analitycy gier niekoniecznie dobrze trafili z niektórymi ocenami. Na podstawie tak dużych różnic, tak wielu zawodników, nie można zaprzeczyć ,że niektórzy zawodnicy dostali wysoką ocenę w oparciu o dobrą reputacje. Reputacja nie jest przypadkowa, ponieważ żeby zawodnik był znany w pozytywny sposób, musi grać w dobrym klubie i na dobrym poziomie przez jakiś okres czasu. W tabeli przedstawieni są zawodnicy, którzy nie znaleźli się wśród najlepszej dziesiątki rankingu obu gier.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Imię i nazwisko** | **FIFA** | **FM\*** | **Sofascore** | **Źródło z bliższą liczbą** |
| Jack Grealish | 83.33 | 98.17 | 99.32 | FM |
| Gareth Bale | 91.10 | 87.21 | 99.09 | FIFA |
| James Rodriguez | 87.90 | 98.17 | 98.86 | FM |
| Mason Mount | 83.33 | 98.17 | 98.63 | FM |
| Riyad Mahrez | 95.66 | 98.17 | 98.17 | FM |
| Trend Alexander-Arnold | 98.17 | 98.17 | 97.95 | brak |
| Rodri | 95.66 | 87.21 | 97.72 | FIFA |
| Antonio Rudiger | 86.53 | 87.21 | 97.49 | FM |
| James Ward-Promise | 78.54 | 87.21 | 97.26 | FM |
| James Maddison | 86.53 | 87.21 | 96.80 | FM |
| John Stones | 83.33 | 98.17 | 96.57 | FM |
| Kalvin Phillips | 56.85 | 87.21 | 96.09 | FM |
| Phil Foden | 78.54 | 98.17 | 95.86 | FM |
| Matheus Pereira | 56.85 | 52.28 | 95.63 | FIFA |
| Luke Shaw | 86.53 | 87.21 | 95.30 | FM |
| Wilfred Ndidi | 92.92 | 87.21 | 94.84 | FIFA |
| Thiago Silva | 95.66 | 98.17 | 94.38 | FIFA |
| Pascal Gross | 67.12 | 87.21 | 94.15 | FM |
| Thiago Alcantara | 95.66 | 98.17 | 93.92 | FIFA |
| Kurt Zouma | 83.33 | 87.21 | 93.69 | FM |
| Youri Tielemans | 86.53 | 87.21 | 93.46 | FM |

*Tabela 3 – Porównanie statystyk zawodników – część 2, FM\* – Football Manager*

**4.3 Testy manualne**

Testy manualne są bardzo potrzebnym i używanym w nowoczesnej informatyce zagadnieniem. Zdarzają się takie sytuacje, gdy programista usunie część kodu, gdyż stwierdzi, że jest on obecnie niepotrzebny. Wtedy może się okazać, że jedna z funkcjonalności przestaje działać i po przejściu do kolejnego zadania, użytkownik nie jest tego świadomy. Dzięki testom, które mogą być uruchamiane przy każdej zmianie np. za pomocą takiego narzędzia jak Husky, do takiego zdarzenia nie ma prawa dojść. W momencie błędu, czyli niespełnienia oczekiwań zdefiniowanych w testach, zmiany nie zostaną wysłane do repozytorium. Wprowadzenie zorganizowanego procesu testowania może przynieść również inne znaczące korzyści. Jednym z istotnych aspektów, jest aktywny udział testera od etapu oszacowania i analizy. Dzięki takim działaniom błędy podczas wprowadzania zmian zdarzają się bardzo rzadko. Programiści podczas pracy są w stanie skoncentrować się na kluczowych aspektach, znając na wstępie zakres i cel testów [19]. W aplikacji za pomocą narzędzia Cypress dodane zostały testy sprawdzające tabele sortującą. W podstronie zawodników, po naciśnięciu przycisku danego atrybutu w danym źródle, witryna uporządkowuje graczy w kolejności od największej wartości tej statystyki. Pierwszy z testów sprawdza, czy tabela została wygenerowana i czy wszystkie kolumny znajdują się na odpowiednim miejscu. Drugi podpunkt testuje, czy wartości, które są zawartością tabeli wyświetlają się w odpowiedniej kolejności, od największej do najmniejszej.

**Obraz zawierający tekst, zrzut ekranu, monitor, ekran

Opis wygenerowany automatycznie**

Rysunek 35 - Zawartość testów

Naciśnięcie przycisku sortowania po różnicy wartości pomiędzy grami w ocenie ogólnej

**Obraz zawierający tekst, monitor, telefon komórkowy, ekran

Opis wygenerowany automatycznie**

Rysunek 36 – Wykonywanie testów – część 1

Odzwierciedlenie graficzne polecenia (Rysunek 36)

**Obraz zawierający stół

Opis wygenerowany automatycznie**

Rysunek 37 – Wykonywanie testów po stronie interfejsu graficznego – część 1

Przeszukiwanie wszystkich elementów tablicy

**Obraz zawierający tekst, ekran, czarny, zrzut ekranu

Opis wygenerowany automatycznie**

Rysunek 38 – Wykonywanie testów – część 2

Odzwierciedlenie graficzne polecenia (Rysunek 38)

**Obraz zawierający stół

Opis wygenerowany automatycznie**

Rysunek 39 – Wykonywanie testów po stronie interfejsu graficznego – część 2

Sprawdzenie czy wartości będące wyżej po sortowaniu (Rysunek 39) są większe bądź równe pozostałym komórkom

**Obraz zawierający tekst, monitor, ekran

Opis wygenerowany automatycznie**

Rysunek 40 – Wykonane testy

**Podsumowanie i wnioski**

Za pomocą aplikacji porównującej statystyki zawodników można zauważyć, że zarówno w jednej, jak i drugiej grze atrybuty oceny ogólnej, czyli najważniejszej statystyki wyglądają podobnie. Przykładowo trójka Sadio Mane, Mohamed Salah i Virgil van Dijk, która jest wśród najlepszych, jeśli chodzi o zawodników z Premier League, ma takie same wartości ocen. Virgil van Dijk złapał poważną kontuzje i grał bardzo mało w analizowanym sezonie, lecz pozostali zawodnicy w rzeczywistości spisywali się dobrze. Znalazło się paru piłkarzy, którzy byli niżej oceniani i zaliczyli lepszy sezon według Sofascore. Mimo wszystko jest to akceptowalna różnica. Mimo paru podobieństw można zauważyć jednak, że zdarzają się też takie sytuacje w których statystyki są lepsze w jednej z gier. Tak dzieje się w przypadku wartości zawodnika nazywającego się Bernardo Silva [Tabela 2]. W Football Managerze jest on uważany za najlepszego zawodnika w całej lidze, a w FIFIE zajmuje 9 miejsce w rankingu. W weryfikacji boiskowej okazało się, że to osoby odpowiedzialne za FIFE były bliżej prawdy według źródła Sofascore. Kevin De Bruyne, który był stawiany przez nich w roli lidera tych rozgrywek, rzeczywiście posiadał najlepszą średnią ocenę ogólną w serwisie Sofascore po całym sezonie. Dodatkowo Bernardo Silva, oceniony najlepiej w Football Managerze, zaliczył przeciętny sezon. Algorytmy portalu Sofascore oceniły jego postawę gorzej od ponad 26 procent piłkarzy. Można też dojść do takich wniosków, że szacowanie dokładnej dyspozycji piłkarza nie jest łatwą dziedziną. Prócz czołowej piątki, większość zawodników uważanych za jednych z najlepszych, zawodziła. Ich obecna wysoka ocena być może jest skutkiem świetnej gry w poprzednich sezonach. Mało który piłkarz potrafi utrzymać topową formę przez parę lat. Aczkolwiek duża ocena ogólna może wynikać częściowo z wyrobionej sobie reputacji na temat danego piłkarza. Obydwie gry mogą celowo wolno obniżać notowania zawodników. W razie nagłego powrotu do szczytowej formy, niska ocena w grze mogłaby powodować zła reakcję ze strony fanów. W razie niedocenienia zawodnika, którego dyspozycja do tej pory nie była na najwyższym poziomie, sprawa jest bardziej zrozumiała.

Jeśli chodzi o porównywanie źródeł służących do typowania wyników, to w tym module również dane są w miarę podobne do siebie. Najbardziej sensownym źródłem, spośród trzech możliwości wyróżniają się kursy bukmacherskie. Działają one trochę inaczej, ponieważ biorą pod uwagę, takie zależności jak nieobecność zawodników pauzujących, indywidualne oceny ekspertów na podstawie oglądanych. Jednak system ten nie jest idealny i w niektórych sytuacjach może gorzej sprawdzać się od innych źródeł. Kolejnymi źródłami są witryna forebet oraz system do szacowania wyników na podstawie ostatnich meczy. Proponowane wyniki są często podobne i można zauważyć w nich jedną zależność. Rezultaty są szczególnie odmienne, gdy spotykają się drużyny, dla których miejsce rozgrywania meczu jest ważnym czynnikiem. Tak było w przypadku meczu Warta Poznań – Stal Mielec. Na tamten moment drużyna z Poznania bardzo słabo sobie radziła na własnym stadionie. Dla Stali statystycznie teren nie wpływał w dużym stopniu na wynik spotkania. Jeśli chodzi o ogólną formę to wyglądało to podobnie i oba zespoły były blisko siebie w tabeli. Algorytm oparty o formę działa w taki sposób, że 2/3 siły drużyny stanowi jej obecna forma w ostatnich 15 spotkaniach, a 1/3 to forma na terenie na jakim gra. Taka proporcja została stworzona na podstawie danych statystycznych z poprzedniego sezonu z portalu 90 minut [15]. W serwisie forebet czynnik formy w miejscu rozgrywania meczu prawdopodobnie jest większy. Dokładny wzór obliczania typowania w forebet nie jest znany, lecz na podstawie obserwacji można założyć, że wynosi około 50%. Powodem takich wniosków jest fakt, że zarówno Warta i Stal Mielec pod względem formy czy też innych czynników, które mogłyby być brane pod uwagę (np. ostatnie mecze między obiema drużynami), na moment przed rozegraniem spotkania, osiągały podobne rezultaty. Przedmeczowe dane witryny forebet wyświetlały 26% szans na zwycięstwo Warty, 28% na remis i 47% na zwycięstwo Stali. Daje to w przybliżeniu około 40-60 na korzyść Stali, gdyby mecz nie mógł zostać zakończony remisem. Stosunek ilości punktów Stali na wyjeździe w tym sezonie do sumy ilości punktów Stali na wyjeździe i punktów u siebie Warty wynosi 70%. Gdyby wyliczyć średnią 50-50 formy ogólnej i 70-30 formy zależnej od terenu taką samą wagą to wynik wynosiłby 60-40. Taka proporcja to przypuszczalna wartość sił obydwu drużyn przed rozkładem bez możliwości remisu. Analiza procentów witryny forebet w oparciu o tę tezę przy innych spotkaniach pozwala wysunąć podobne wnioski. Strona ta jest na pewno bardziej zaawansowana obliczeniowo od algorytmu opartego wyłącznie o ostatnią formę. Lecz ta druga metoda skupia się na jednej lidze, przez co może być bardziej dopracowana. Oczywiście żaden system nie jest w stanie idealnie typować wyników każdego meczu, lecz sprawdzane metody są użyteczne przy analizie

**Bibliografia i netografia**

[1] <https://sportowefakty.wp.pl/pilka-nozna/558431/200-tys-kibicow-na-meczu-pilkarskim-oto-spotkania-z-najwieksza-liczba-widzow> (dostęp: 09.03.2023)

[2] <https://transfery.info/aktualnosci/mistrzostwa-swiata-tvp-pochwalilo-sie-wynikami-ogladalnosci-meczu-polska-argentyna-sa-rekordy/176688> (dostęp: 09.03.2023)

[3] <https://arc.com.pl/juz-jest-najnowszy-raport-sponsoring-monitor/> (dostęp: 09.03.2023)

[4] https://cbos.pl/SPISKOM.POL/2016/K\_097\_16.PDF (dostęp: 09.03.2023)

[5] Grady Booch, James Rumbaugh, Ivar Jacobson „UML przewodnik użytkownika” tłum: Krzysztof Stencel, 2002

[6] Borys Struski, Piotr Sebastian „Encyklopedia Football Managera 2023”, 2022

[7] Źródło danych, z którego pozyskiwane są informacje o statystykach piłkarzy w FIFIE, <https://www.futhead.com/21/players/?level=all_nif&league=13&club=all&bin_platform=ps> (dostęp: 11.11.2022)

[8] Współczynnik UEFA, <https://www.transfermarkt.pl/statistik/5jahreswertung> (dostęp: 04.11.2022)

[9] Źródło danych, z którego dane są pozyskiwane informacje o statystykach piłkarzy w Football Managerze, <https://fmdataba.com/21/l/2613/premier-league/best-players/> (dostęp: 11.11.2022)

[10] Zasada działania kursów bukmacherskich, <https://legalbuk.pl/jak-dzialaja-kursy-bukmacherskie/> (dostęp: 09.03.2023)

[11] Współczynnik remisów w zależności od ligi, <https://www.eprogresja.pl/statystyki/pilkarskie/remisy_w_ligach/> (dostęp 02.12.2022)

[12] Klasyfikacja narodowości według rankingu ELO, <https://eloratings.net/> (dostęp 03.12.2022)

[13] Tabela Ekstraklasy według dat, <http://www.90minut.pl/stats.php?id=12330> (dostęp: 01.03.2023)

[14] Algorytm obliczania rankingu ELO, <https://pl.wikipedia.org/wiki/Ranking_Elo_(pi%C5%82ka_no%C5%BCna)> (dostęp 18.11.2022)

[15] Statystyki dotyczące wyników w Ekstraklasie w rundzie wiosennej sezonu 2021/2022, <http://img.90minut.pl/misc/stats-ekstraklasa-2021-22-wiosna.html> (dostęp: 01.03.2023)

[16] Prywatny kontakt ze specjalistą od matematycznych analiz piłkarskich, <https://twitter.com/pklimek99>

[17] Wyniki spotkań Ekstraklasy <https://www.flashscore.com/football/poland/ekstraklasa/results/> (dostęp: 09.03.2023)

[18] Serwis umożliwiający przeglądanie kursów bukmacherskich, również tych meczy, które już się odbyły, <https://www.oddsportal.com/soccer/poland/ekstraklasa/results/> (dostęp: 09.03.2023)

[19] Adam Roman, Karolina Zmitrowicz „Testowanie oprogramowania w praktyce: studium przypadków”, 2017

[20] Alex MacCaw „Javascript. Aplikacje WWW” Helion, 2021

**Wykaz rysunków**

Rysunek 1 – Diagram DPU 6 [Rysunek 2 – Widok strony futhead – 11.11.2022 10](#_Toc130892559) Rysunek 3 – Współczynnik UEFA - 04.11.2022 11 [Rysunek 4 – Wartości wyświetlone w API pracy inżynierskiej 12](#_Toc130892559) [Rysunek 5 – Widok zawodnika na stronie futhead – 11.11.2022 12](#_Toc130892559) [Rysunek 6 – Informacje o zawodniku w API pracy inżynierskiej 12](#_Toc130892559) Rysunek 7 – Widok strony fmdataba - 11.11.2022 13 [Rysunek 8 – Porównanie średnich statystyk w pracy inżynierskiej 14](#_Toc130892559) Rysunek 9 – Podstrona zawodników w pracy inżynierskiej 16 [Rysunek 10 – Podstrona zawodnika w pracy inżynierskiej – część 1 17](#_Toc130892559) Rysunek 11 – Podstrona zawodnika w pracy inżynierskiej - część 2 18 Rysunek 12 – Wykresy mocy drużyn w pracy inżynierskiej 19 [Rysunek 13 – Pobranie statystyk z portalu futhead 20](#_Toc130892559) [Rysunek 14 – Wyświetlanie statystyk w API 20](#_Toc130892559) [Rysunek 15 – Sprawdzanie pobranych danych – część 1 20](#_Toc130892559) [Rysunek 16 – Sprawdzanie pobranych danych – część 2 21](#_Toc130892559) [Rysunek 17 – Pobranie statystyk zawodnika po jego nazwisku z API Sofascore 21](#_Toc130892559) [Rysunek 18 – Statystyka remisów w zależności od ligi – 02.12.2022 23](#_Toc130892559) [Rysunek 19 – Ranking ELO męskich drużyn narodowych - 03.12.2022 25](#_Toc130892559) [Rysunek 20 – Tabela Ekstraklasy za okres 15.07.2022 - 22.08.2022 26](#_Toc130892559) [Rysunek 21 –](#_Toc130892559) Tabela Ekstraklasy za okres 27.01.2023 - 07.03.2023 26 [Rysunek 22 – Metoda obliczania punktów ELO 27](#_Toc130892559) [Rysunek 23 – Widok konkretnego meczu 33](#_Toc130892559) [Rysunek 24 – Widok konkretnego meczu w wersji mobilnej 34](#_Toc130892559) Rysunek 25 – Widok panelu bocznego w wersji mobilnej 35 [Rysunek 26 – Pobieranie ostatnich spotkań 37](#_Toc130892559) Rysunek 27 – Pobieranie kursów bukmacherskich 37 [Rysunek 28 – Wyświetlanie danych z portalu forebet 38](#_Toc130892559) [Rysunek 29 – Konwersja danych z API – część 1 38](#_Toc130892559) [Rysunek 30 – Konwersja danych z API – część 2 39](#_Toc130892559) [Rysunek 31 – Przykład poprawnego szacowania – część 1 45](#_Toc130892559) [Rysunek 32 – Przykład poprawnego szacowania – część 2 46](#_Toc130892559) [Rysunek 33 – Przykład poprawnego szacowania – część 3 47](#_Toc130892559) [Rysunek 34 – Przykład poprawnego szacowania – część 4 48](#_Toc130892559) [Rysunek 35 – Zawartość testów 52](#_Toc130892559) [Rysunek 36 – Wykonywanie testów – część 1 53](#_Toc130892559) [Rysunek 37 – Wykonywanie testów po stronie interfejsu graficznego – część 1 53](#_Toc130892559) [Rysunek 38 – Wykonywanie testów – część 2 54](#_Toc130892559) [Rysunek 39 – Wykoywanie testów po stronie interfejsu graficznego – część 2 54](#_Toc130892559) [Rysunek 40 – Wykonane testy 55](#_Toc130892559)

**Wykaz tabel**

Tabela 1 – Porównanie wyników spotkań do innych źródeł 44 [Tabela 2 – Porównanie statystyk zawodników – część 1 49](#_Toc130892559) Tabela 3 – Porównanie statystyk zawodników - część 2 51

**Streszczenie pracy w języku polskim**

**Temat pracy w języku polskim:** Serwis web gromadzący, weryfikujący i analizujący statystyki piłkarskie

**Temat pracy w języku angielskim:** A website that collects, verifies and calculates football statistics

**Streszczenie:**

W ramach pracy analizowane są wartości umiejętności zawodników i szacowania dotyczące meczów. Celem jest ułatwienie analizy danych związanych z piłką nożną. W sekcji poświęconej zawodnikom zgromadzone zostały oceny przyznane przez gry FIFA oraz Football Manager za sezon 2019/2020, a także wartości odzwierciedlające postawę w trakcie rzeczywistych meczów, nadane przez serwis Sofascore. W sekcji poświęconej klubom dokonywane są porównania szacowań dotyczących wyników meczów, które pochodzą z różnych źródeł, takich jak serwis forebet.com, kursy bukmacherskie oraz system analizujący wyniki poprzednich spotkań.