

# Analiza zależności wyników drużyny piłkarskiej w lidze od wybranych statystyk drużynowych

Daniel Borkowski, 266593

5 czerwca 2023 r.

## Spis treści

<b>1</b>	<b>Wstęp</b>	<b>1</b>
1.1	Cel . . . . .	1
1.2	Plan . . . . .	2
<b>2</b>	<b>Dane</b>	<b>2</b>
2.1	Pozyskanie danych . . . . .	2
2.2	Analiza danych . . . . .	2
2.3	Przetwarzanie danych . . . . .	4
2.4	Analiza przetworzonych danych . . . . .	5
<b>3</b>	<b>Modele</b>	<b>7</b>
3.1	Narzędzia . . . . .	7
3.2	Dobór parametrów . . . . .	7
3.3	Wyniki . . . . .	7
3.4	Eksperyment . . . . .	8
3.4.1	Model trudny . . . . .	8
3.4.2	Model prosty . . . . .	8
3.4.3	Uwagi . . . . .	9
<b>4</b>	<b>Wnioski</b>	<b>10</b>
4.1	Model trudny . . . . .	10
4.2	Model prosty . . . . .	11
4.3	Uwagi końcowe . . . . .	11

## 1 Wstęp

### 1.1 Cel

Celem projektu jest **znalezienie, zbadanie i opisanie zależności** wyników drużyny piłkarskiej w lidze od wybranych mniej lub bardziej oczywistych statystyk drużynowych w celu **budowy modelu** przewidującego jak dobrze radzi sobie dana drużyna w lidze na podstawie tych statystyk. Ze względu na wybór czynników podzielimy to zadanie na dwie niezależne części:

- przewidywanie na podstawie bardziej oczywistych czynników (model *łatwy*):
  - ilości goli strzelonych
  - ilości goli straconych
- przewidywanie na podstawie mniej oczywistych czynników (model *trudny*):
  - udział młodych zawodników w grze zespołu
  - udział zagranicznych zawodników w grze zespołu

Czynnikiem w piłce nożnej, który określa wynik drużyny jest natomiast liczba zdobytych przez nią **punktów**. Definiujemy zatem nasz cel nieco bardziej szczegółowo, jako analizę zależności między **bramkami** strzelonymi i straconymi przez drużynę, a jej punktami w sezonie oraz analizę zależności między **strukturą składu drużyny**, a jej punktami w sezonie.

## 1.2 Plan

Skupimy się na naszym krajowym podwórku i rozgrywkach Ekstraklasy [1]. Za zgodą administratorów portalu EkstraStats [2] wykorzystamy dostępne tam dane na temat rozgrywek w latach 2015-2022. Nie wykorzystamy danych z sezonu 2014/2015 oraz 2022/2023 ze względu na ich niekompletność. Pozyskamy następujące metryki dla każdej drużyny w każdym sezonie:

1. liczba rozegranych meczów
2. liczba straconych bramek
3. liczba strzelonych bramek
4. liczba zdobytych punktów
5. liczba minut rozegranych przez młodzieżowców
6. liczba goli zdobytych przez młodzieżowców
7. liczba minut rozegranych przez zagranicznych zawodników
8. liczba goli zdobytych przez zagranicznych zawodników

Umożliwi nam to realizację celu na podstawie 114 rekordów występów drużyn w rozgrywkach ligowych.

## 2 Dane

### 2.1 Pozyskanie danych

Strona [2] udostępnia wiele ciekawych danych na temat rozgrywek ligowych, zawierających się w kilkudziesięciu tabelach dla każdego sezonu. Ograniczymy się do interesujących nas informacji: tabel ligowych, tabel młodzieżowców i tabel zawodników zagranicznych.

Używając skryptu `scrapping.py` dokonujemy zatem ekstrakowania interesujących nas tabel ze strony. Dane są zbierane hobbystów i na tym etapie natrafiamy na pewne problemy, jakim jest niespójność w ich zapisie: *klucze*, które zaraz będą potrzebne do łączenia danych nie są bowiem jednolite nawet dla różnych tabel w tym samym sezonie. Co więcej, ta niespójność nie jest konsekwentna i tak jak raz tę samą drużynę określa się jako *Wisła Kraków* i *Wisła K.*, tak czasem *Termalicę* w innej tabeli reprezentuje *Bruk-Bet*. Nie uciekniemy tu zatem od ręcznej korekty kluczy w pobranych plikach `.csv`.

Dopiero po tym kroku możemy przejść do dalszych działań, jakimi są połączenie tabel dla danych sezonów i odrzucenie kolumn, którymi nie jesteśmy zainteresowani, jak na przykład *liczba porażek*, czy też *procent goli drużyny zdobytych przez zagranicznych zawodników*, którą to metrykę możemy obliczyć. Na koniec otrzymujemy 7 tabel zawierających w sumie 114 rekordów ośmiu metryk wymienionych powyżej oraz kolumnę *KLUB* będącą kluczem używanym w celu łączenia danych i kolumnę *season* informującą o sezonie, w którym odbyły się rozgrywki.

### 2.2 Analiza danych

Pozyskane dane dla różnych sezonów łączymy i zaczynamy analizę. Wykorzystamy narzędzia oferowane przez `y_data_profiling` oraz przyjrzymy się ręcznie pozyskanym tabelom. Poniżej w tabeli 1 pięć początkowych rekordów i kolejno ich indeks, liczba rozegranych meczów, zdobytych punktów, strzelonych bramek, straconych bramek, minut i strzelonych bramek obcokrajowców, minut i strzelonych bramek młodzieżowców oraz sezonu.

Tabela 1: Efekt wywołania metody `head` na zebranych danych

id	played	points	gf	ga	fm	fg	ym	yg	season
0	37	73	70	32	14846	51	2324	2	2015/2016
1	37	69	60	45	20666	40	3270	1	2015/2016
2	37	60	55	42	13202	9	5995	13	2015/2016
3	37	58	66	50	12329	36	5326	7	2015/2016
4	37	51	53	44	13296	19	5992	5	2015/2016

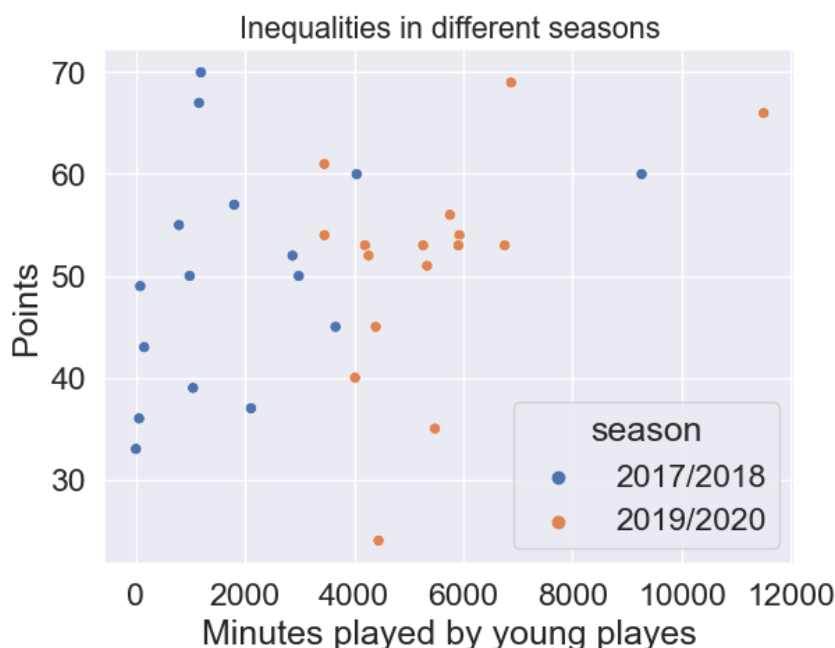
Ze względu na sposób i źródło pozyskanych danych w tabeli nie ma brakujących wartości.

Tabela 2: Mediany wartości dla poszczególnych sezonów

season	played	points	gf	ga	fm	fg	ym	yg
2015/2016	37.0	48.0	44.0	46.5	12765.5	15.0	3702.5	3.5
2016/2017	37.0	46.5	49.0	54.0	12894.5	18.5	5128.5	8.0
2017/2018	37.0	50.0	49.5	50.0	18306.5	26.0	1173.0	1.0
2018/2019	37.0	50.5	49.5	51.5	16298.5	31.5	2741.5	2.0
2019/2020	37.0	53.0	48.0	49.0	18275.0	25.0	5298.0	4.5
2020/2021	30.0	39.0	37.5	35.0	13658.5	18.5	3833.5	2.5
2021/2022	34.0	44.5	42.5	49.0	15354.5	20.5	4080.0	4.0

Jednakże, dokonując grupowania natrafiamy na inny problem potencjalnie rzutujący na późniejsze wyniki. W różnych sezonach, co widzimy na tabeli 2, różnie kształtowały się mediany poszczególnych wartości. Przyczyny tego zjawiska korzystając z własnej wiedzy na temat rozgrywek Ekstraklasy i dodatkowych materiałów [3] diagnozujemy następująco:

- **zmiana formuły rozgrywek** - co widoczne po różnej liczbie rozegranych meczów, rozgrywki w badanym okresie kilkakrotnie zmieniały format
- **zmiana innych przepisów** - w 2019 roku wprowadzono obowiązek gry młodzieżowca w każdym meczu
- **zmiana znaczenia wartości kolumn** - nierozwiązane dotychczas wątpliwości na temat znaczenia słowa *młodzieżowiec* warto rozwiązać - na przestrzeni różnych sezonów zmieniał się sposób pojmowania takich zawodników, najpierw jako piłkarzy poniżej 24. roku życia, a potem 21.



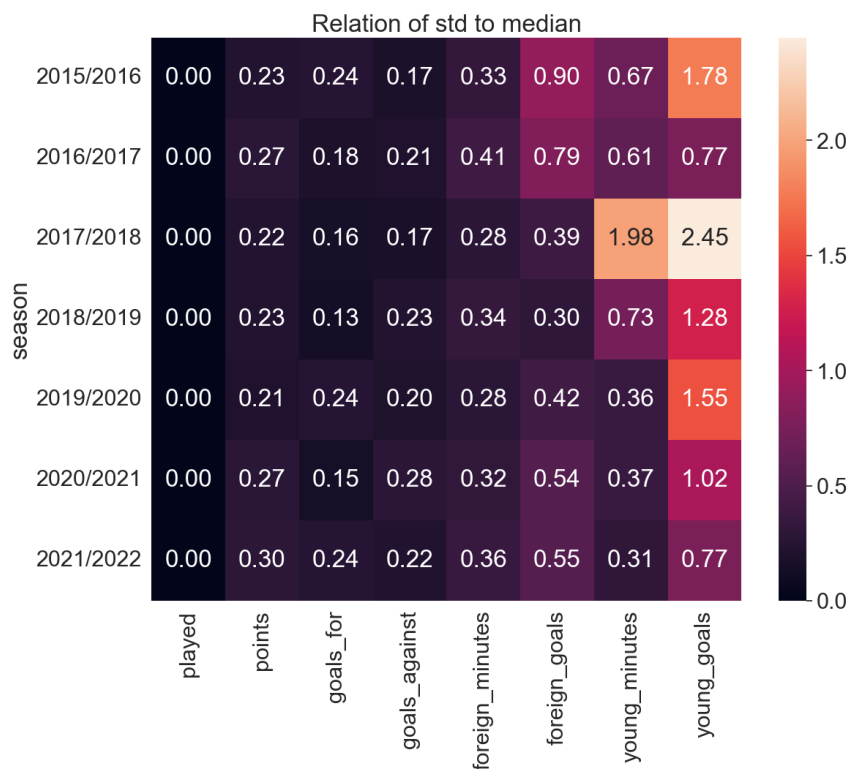
Rysunek 1: Zależność liczby punktów zdobytych przez drużynę od liczby minut rozegranych przez młodzieżowców.

Warto więc odwołać się do czynników, które się nie zmieniają. Oprócz tego, że parafrazując legendarnego polskiego trenera Kazimierza Górskiego, *piłka zawsze była jedna, a bramki zawsze były dwie*, nie zmieniły się najważniejsze zasady meczów piłkarskich. W każdym z rozgrywanych meczów do zdobycia były trzy punkty za zwycięstwo i jeden punkt za remis, a każda bramka liczyła się tak samo. Odpowiada nam to na zmianę formuły rozgrywek, ale nie rozwiązuje kwestii zmiany przepisów i zmiany znaczenia wartości kolumn. Nie chcąc rezygnować z pozyskanych metryk (wobec braku dostępnych lepszych danych) proponujemy standaryzację wartości dla poszczególnych sezonów.

Po zbadaniu odchylenia standardowego na tabeli 3 i jego stosunku do mediany na grafice 2 podejmujemy decyzję o zastosowaniu standaryzacji względem mediany, gdyż jest ona mniej podatna

Tabela 3: Odchylenia standardowe wartości dla poszczególnych sezonów

season	points	gf	ga	fm	fg	ym	yg
2015/2016	10.81	10.65	8.02	4201.37	13.55	2489.86	6.22
2016/2017	12.68	8.92	11.28	5226.91	14.69	3120.85	6.13
2017/2018	11.00	7.84	8.58	5097.70	10.23	2325.39	2.45
2018/2019	11.38	6.66	11.63	5531.65	9.48	2011.74	2.56
2019/2020	11.11	11.62	9.75	5027.71	10.40	1933.05	6.98
2020/2021	10.45	5.55	9.86	4350.18	10.03	1424.45	2.56
2021/2022	13.15	10.31	10.66	5489.81	11.27	1278.29	3.07



Rysunek 2: Stosunek odchylenia standardowego do mediany dla poszczególnych sezonów

na skrajne wartości, których, zwłaszcza w kolumnach dotyczących młodych graczy, nie brakuje, co widać na grafice 1, gdzie widoczne jest także wprowadzenie obowiązku gry młodzieżowca w sezonie 2019/2020.

Proponujemy zatem uzależnienie wszystkich wartości od liczby rozegranych meczów w sezonie, uzależnienie zdobywanych przez zawodników goli od goli drużyny oraz wstępną standaryzację wartości dotyczących piłkarzy młodzieżowych i zagranicznych, tak aby otrzymane wskaźniki wyrażały ogólnie *czy klub stawia na młodzieżowców, czy bramki zespołu są zdobywane przez zagranicznych zawodników* i były niezależne od trendu w danym sezonie spowodowanego przepisami lub innymi czynnikami zewnętrznymi.

Przetworzone w ten sposób dane poddamy właściwej analizie w celu stworzenia modelu.

## 2.3 Przetwarzanie danych

Dokonyjemy zatem transformacji danych, chcąc uzyskać następujące wskaźniki:

- **liczba punktów drużyny na mecz** - główny wskaźnik sukcesu bądź porażki drużyny, będziemy go przewidywać
- **liczba goli zdobytych i straconych na mecz** - podstawowe wskaźniki, z użyciem których będziemy przewidywać liczbę punktów na mecz
- **wskaźnik minut młodzieżowców** - oddanie skłonności drużyny do grania młodymi zawodnikami

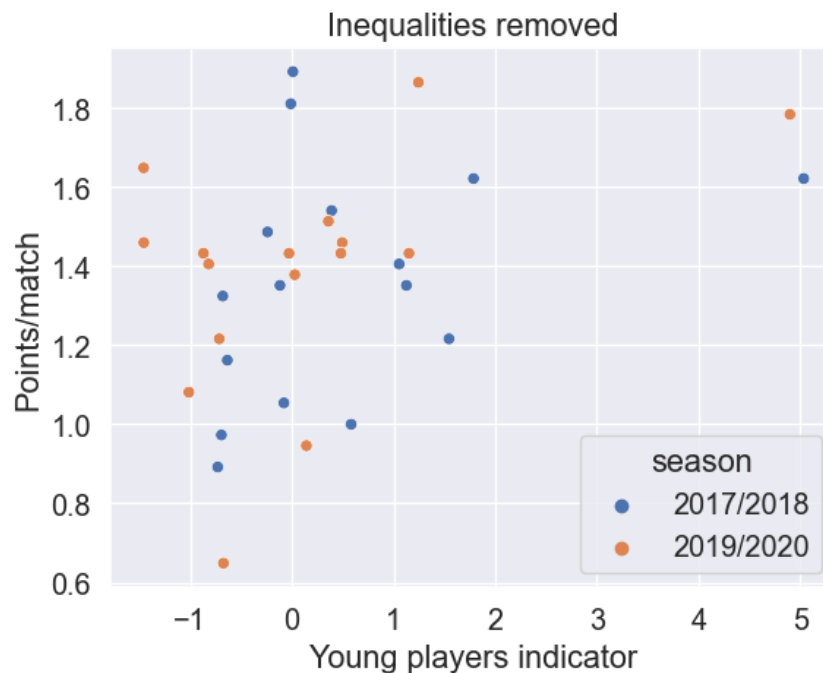
- **wskaźnik udziału młodzieżowców w golach zespołu** - oddanie rzeczywistego znaczenia młodych zawodników w drużynie
- **wskaźnik minut zawodników zagranicznych** - oddanie skłonności drużyny do grania zawodnikami zza granicy
- **wskaźnik udziału zawodników zagranicznych w golach zespołu** - oddanie rzeczywistego znaczenia zawodników zagranicznych w drużynie

Warto zaznaczyć, że używane słowo *wskaźnik*, nomen omen, *wskazuje* na wysoki poziom przetworzenia danych - na tym etapie nie mają już bezpośredniego przełożenia na rzeczywistość, jedynie oddają skłonności i trendy, na czym - w kontekście celu projektu - nam zależy.

Tabela 4: Wartości mediany dla poszczególnych sezonów po transformacji

season	points	gf	ga	fm	fg	ym	yg
2015/2016	1.30	1.19	1.26	-0.0	-0.0	0.0	0.0
2016/2017	1.26	1.32	1.46	0.0	0.0	0.0	0.0
2017/2018	1.35	1.34	1.35	-0.0	0.0	0.0	0.0
2018/2019	1.36	1.34	1.39	0.0	0.0	0.0	-0.0
2019/2020	1.43	1.30	1.32	0.0	0.0	0.0	0.0
2020/2021	1.30	1.25	1.17	0.0	0.0	0.0	-0.0
2021/2022	1.31	1.25	1.44	-0.0	0.0	0.0	0.0

Po przeprowadzeniu transformacji mediany nierówności zostały usunięte. Na rysunku 3 nie widzimy znaczących różnic pomiędzy danymi dla sezonów sprzed i po wprowadzeniu przepisu o obowiązku grania młodzieżowca w każdym meczu. Tak przetworzone dane możemy analizować pod kątem stworzenia modelu.



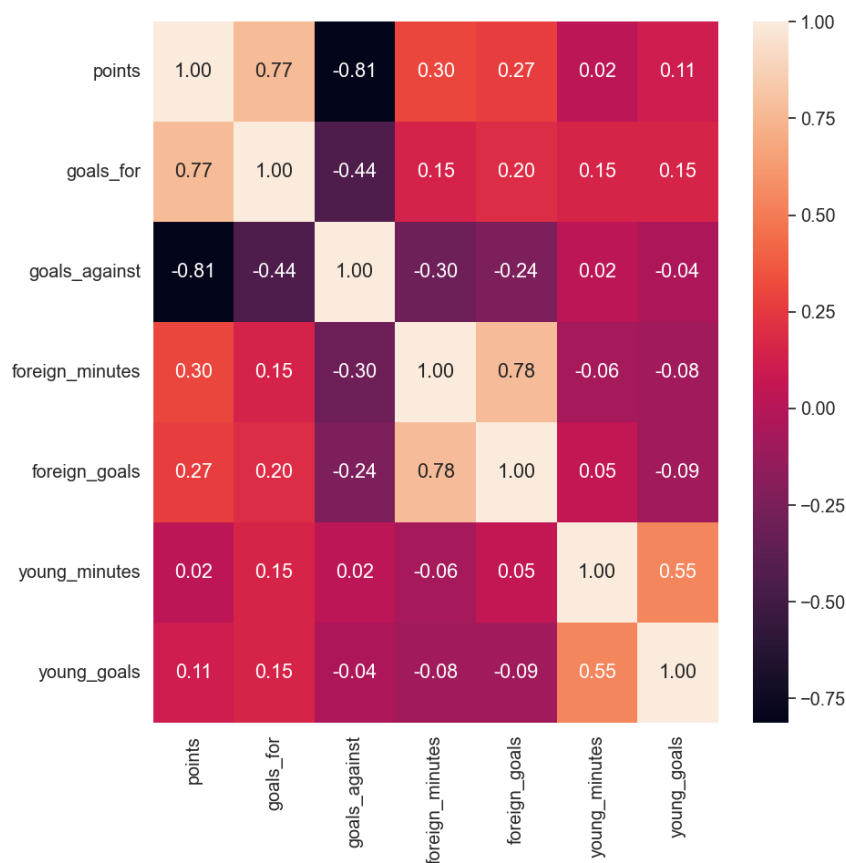
Rysunek 3: Zależność liczby punktów zdobytych przez drużynę na mecz od tego ile w zespole rozegrali młodzi zawodnicy.

## 2.4 Analiza przetworzonych danych

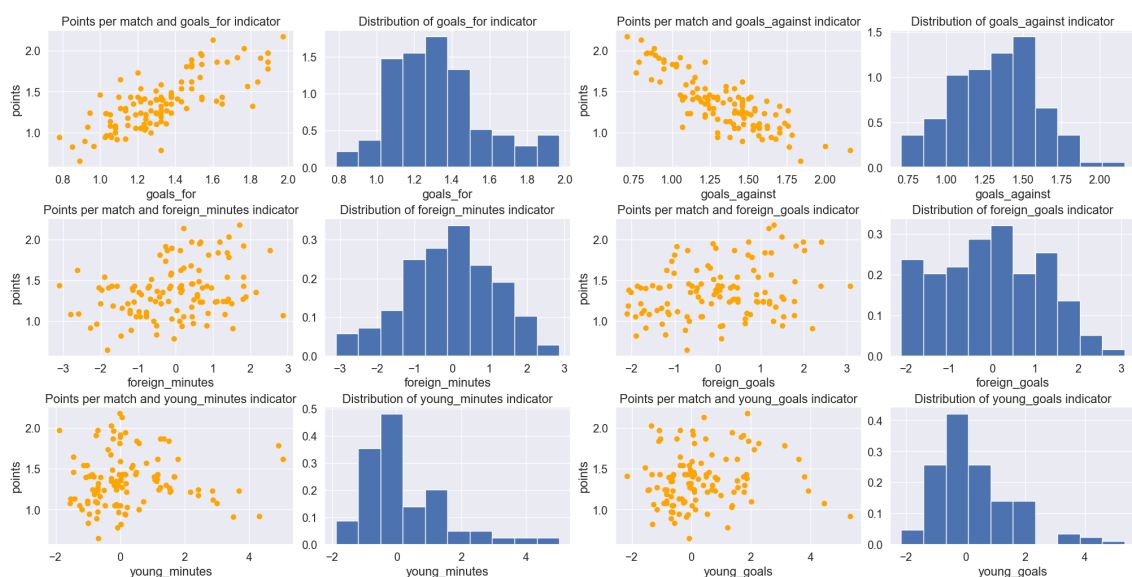
Poszukujemy zależności między poszczególnymi metrykami korzystając z macierzy korelacji na rysunku 4.

Skupiając się na celu przedsięwzięcia przyglądamy się poszczególnym metrykom i im zależnościom względem metryki *points*.

Na podstawie uzyskanych wykresów możemy wysnuć następujące wnioski:



Rysunek 4: Macierz korelacji.



Rysunek 5: Zależności między metrykami a punktami na mecz.

- mniej oczywiste metryki są słabo zależne względem punktów:
  - uzyskane wskaźniki dotyczące młodzieżowców są bardzo słabo zależne względem wszystkich innych metryk
  - uzyskane wskaźniki dotyczące młodzieżowców są silnie skupione wokół średnich wartości
  - uzyskane wskaźniki dotyczące zawodników zza granicy są słabo zależne względem punktów
- bardziej oczywiste metryki są silnie zależne względem punktów

Przeprowadzona dogłębna analiza posłuży nam jako uzasadnienie wniosków uzyskanych w eksperymencie.

## 3 Modele

### 3.1 Narzędzia

W budowie modelu będziemy wykorzystywać narzędzia dostarczane przez pakiet `scikit-learn`:

- model: maszyna wektorów nośnych do regresji liniowej
- metryka: średni błąd bezwzględny (o wzorze 2)
- preprocessing: `StandardScaler`
- inne narzędzia: `train_test_split`, `GridSearchCV`

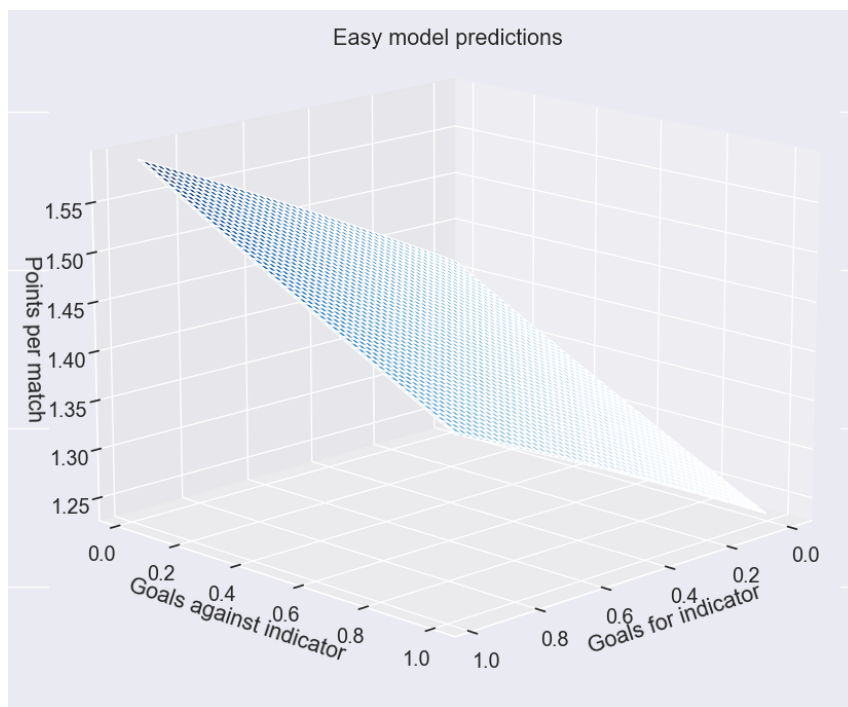
### 3.2 Dobór parametrów

Wykorzystując `GridSearchCV` na podstawie wartości `mean_absolute_error` i danych treningowych poszukujemy najlepszego modelu, spośród następujących wartości:

- **kernel**, jądro wskazujące zastosowany typ regresji liniowej - *linear*, *poly*, *rbf*
- **C**, wskaźnik regularyzacji (uproszczenia) - 1, 2, 5, 10, 15
- **gamma**, parametr jądra - 0.001, 0.005, 0.01, 0.05, 0.1, 1

Poszukiwania określają optymalne wartości tych parametrów jako:  $c=2$ ,  $\gamma=0.05$ ,  $\text{kernel}=\text{rbf}$  dla modelu przewidującego na podstawie danych dotyczących młodzieżowców i piłkarzy zza granicy oraz  $c=15$ ,  $\gamma=0.01$ ,  $\text{kernel}=\text{rbf}$  dla modelu przewidującego na podstawie goli strzelonych i straconych.

### 3.3 Wyniki



Rysunek 6: Wykres przewidywań dla różnych wartości czynników goli strzelonych i straconych (model *łatwy*)

Otrzymane wyniki będziemy odnosić do średniego odchylenia bezwzględnego (medianowego), ze względu na analogiczny sposób jego obliczenia do używanego w modelu średniego błędu bezwzględnego, gdzie występuje analogia między różnicą wartości i mediany, a różnicą wartości i wartości

przewidywanej przez model. Wartość średniego odchylenia bezwzględnego dla rubryki *punkty na mecz* to **0.25**.

$$MAD = \frac{1}{N} \sum_{i=1}^N |x_i - m(X)| \quad (1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| \quad (2)$$

Dla modelu *nieoczywistego* średni błąd bezwzględny wynosi **0.29**, natomiast dla modelu *oczywistego* - **0.10**.

Na tej podstawie należy stwierdzić, że **przewidywanie wyników drużyny na podstawie udziału w grze młodzieżowców i zawodników zza granicy jest obarczone dużym błędem**, a model *trudniejszy* radzi sobie znacznie gorzej niż *prostszy*.

Działanie *prostsze* z modeli widoczne jest na rysunku 6. Zgodnie z intuicją, drużyna radzi sobie lepiej jeśli traci mniej goli i strzela więcej goli.

### 3.4 Eksperyment

Żeby ostatecznie ocenić jakość otrzymanych modeli przeprowadzimy eksperyment na danych z aktualnego sezonu. Nie są one pełne, brakuje ostatniej serii gier, jednakże dla naszych modeli nie stanowi to żadnego problemu. Wyniki drużyn w tym sezonie widoczne są w tabeli 5.

Tabela 5: Tabela ligowa w sezonie 2022/2023 po 33 spotkaniach

id	klub	punkty
1	Raków	74
2	Legia	63
3	Lech	58
4	Pogoń	57
5	Piast	50
6	Górniki	47
7	Warta	45
8	Radomiak	44
9	Zagłębie	44
10	Cracovia	43
11	Widzew	41
12	Jagiellonia	41
13	Stal	40
14	Korona	38
15	Śląsk	38
16	Wisła	37
17	Lechia	30
18	Miedź	22

#### 3.4.1 Model trudny

W tabeli 6 widoczne są przewidywania modelu *trudnego*. Jest kilka sukcesów (ostatnia Miedź, czołowe trzy zespoły są wysoko) i kilka spektakularnych porażek (czwarta Pogoń druga od końca).

W sumie model pomylił się o 84 miejsca w tabeli (średnio 4.67 na drużynę) i 162 punkty (9 na drużynę). Na wykresach zamieszczonych na rysunku 7 widać pewne próby **dopasowania modelu do wskaźnika minut rozegranych przez piłkarzy zagranicznych, które nie zawsze przynosiły oczekiwany efekt**. Jest to przyczyną niedoszacowania Pogoni, która mniej stawiała na zagranicznych graczy od innych zespołów z czołówki.

Największą **premię** dostały Piast i Cracovia, zapewne w związku z dużą bramkostrzelnością graczy zagranicznych i, w odróżnieniu od Jagielloni, pewną bramkostrzelnością młodzieżowców.

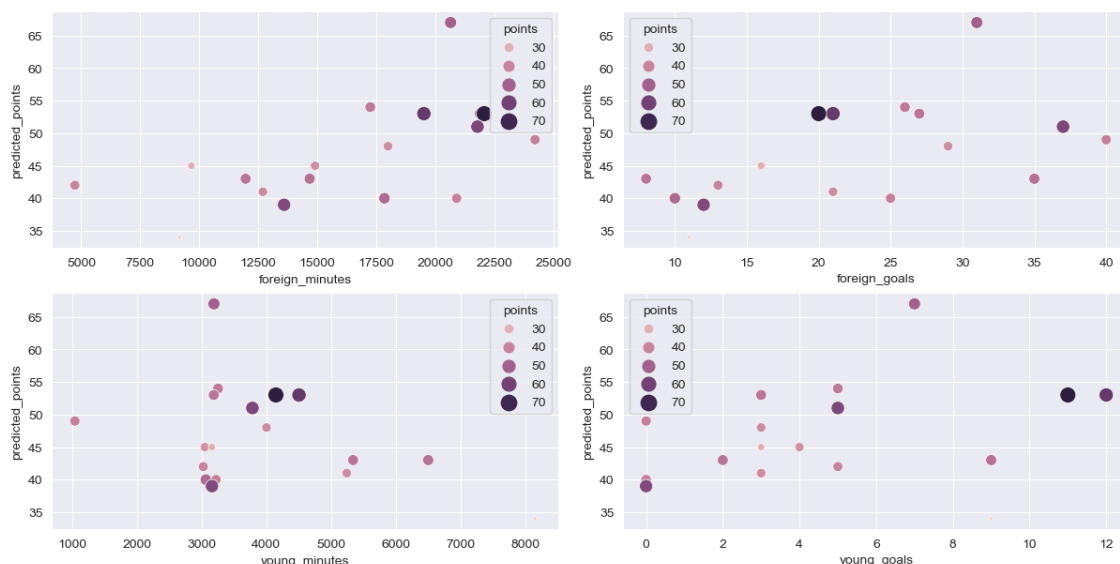
#### 3.4.2 Model prosty

Przewidywania wyników drużyny na podstawie bramek strzelonych i straconych widoczne są w tabeli 7.



Tabela 6: Predykcja punktów po 33 meczach na podstawie udziału w grze zespołu młodzieżowców i zawodników zza granicy w porównaniu do stanu faktycznego z tabeli 5

id	+/-	klub	punkty	+/-	fm	fg	ym	yg
1	4	Piast	67	17	20654	31	3191	7
2	8	Cracovia	54	11	17250	26	3255	5
3	-1	Legia	53	-10	19521	21	4506	12
4	4	Radomiak	53	9	21886	27	3188	3
5	-4	Raków	53	-21	22075	20	4149	11
6	-3	Lech	51	-7	21792	37	3786	5
7	5	Jagiellonia	49	8	24230	40	1042	0
8	6	Korona	48	10	18001	29	4004	3
9	7	Wisła	45	8	14908	16	3050	4
10	7	Lechia	45	15	9668	16	3160	3
11	-4	Warta	43	-2	14681	35	6503	9
12	-3	Zagłębie	43	-1	11966	8	5342	2
13	0	Stal	42	2	4731	13	3027	5
14	1	Śląsk	41	3	12698	21	5244	3
15	-9	Górniki	40	-7	17849	10	3066	0
16	-5	Widzew	40	-1	20909	25	3224	0
17	-13	Pogoń	39	-18	13597	12	3161	0
18	0	Miedź	34	12	9165	11	8154	9



Rysunek 7: Zależności przewidywań modelu *trudnego* od poszczególnych wskaźników

W sumie model pomylił się o 30 (robi wrażenie w porównaniu do poprzednich 84) miejsc w tabeli i 131 punktów uprzednio (162). Przyczyny takiego dysonansu można upatrywać w doborze metryki - *punktów na mecz* - jako pewnego wskaźnika sukcesu. Z punktu widzenia postawionych celów dobre umiejscowienie drużyn w tabeli jest zatem bardziej znaczące niż słabe przewidywania punktów na mecz, ze względu na ograniczenia tej rubryki.

Można powiedzieć, że model szukał balansu między ofensywą, a defensywą. Widoczne są silnie liniowe zależności, na co wskazywała już wcześniej na rysunku 4 macierz korelacji. Zależność ta jest zatem bliska prawdzie, co widoczne jest na wykresach na rysunku 8.

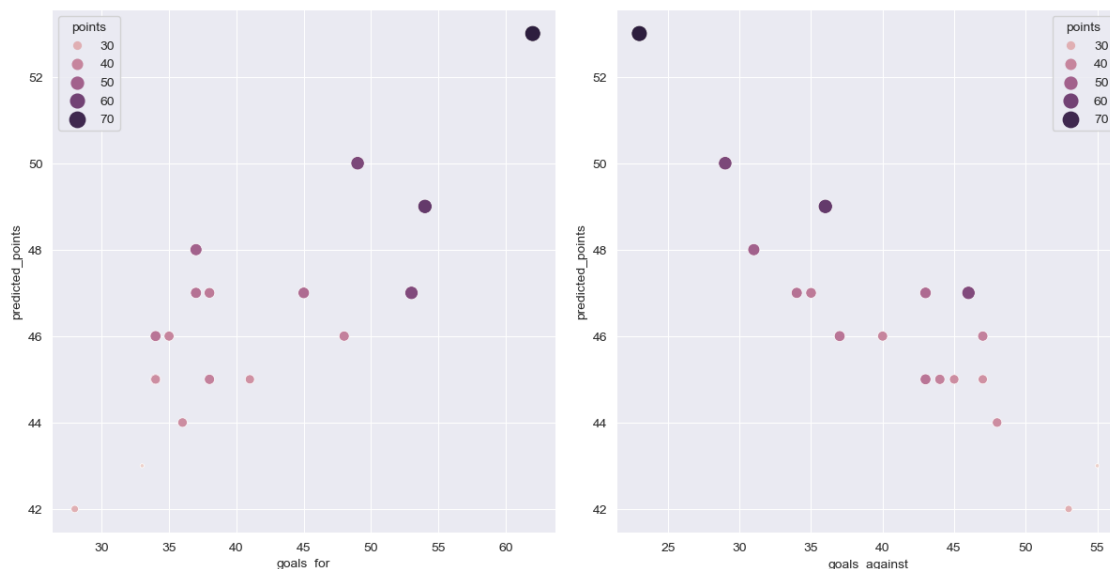
Co bardzo ciekawe, najmocniej niedoszacowana jest **ponownie** Pogoń, która traciła bardzo dużo goli, a dużą premię **ponownie** zgarnia Cracovia, która wyróżniała się od zespołów obok w *prawdziwej* tabeli, Górnika i Zagłębia, pozytywnym bilansem.

### 3.4.3 Uwagi

Powtarzające się w dwóch eksperymentach błędy modeli mogą być dosyć zdumiewające. Ich przyczyn możemy upatrywać w naturze piłki nożnej - nie zawsze lepsza drużyna wygrywa, czasem wynik jest *ponad stan*, a czasem drużyna wygrywa, mimo że przeciwnik grał lepiej. Na tej podstawie mo-

Tabela 7: Predykcja punktów po 33 meczach na podstawie strzelonych i straconych goli w porównaniu do stanu faktycznego z tabeli 5

id	+/-	klub	punkty	+/-	gf	ga
1	0	Raków	53	-21	62	23
2	1	Lech	50	-8	49	29
3	-1	Legia	49	-14	54	36
4	1	Piast	48	-2	37	31
5	2	Warta	47	2	37	34
6	4	Cracovia	47	4	38	35
7	-1	Górník	47	0	45	43
8	-4	Pogoń	47	-10	53	46
9	-1	Radomiak	46	2	34	37
10	2	Jagiellonia	46	5	48	47
11	2	Stal	46	6	35	40
12	-3	Zagłębie	45	1	34	43
13	-2	Widzew	45	4	38	44
14	1	Śląsk	45	7	34	45
15	1	Wisła	45	8	41	47
16	-2	Korona	44	6	36	48
17	1	Miedź	43	21	33	55
18	-1	Lechia	42	12	28	53



Rysunek 8: Zależności przewidywań modelu *łatwego* od poszczególnych wskaźników

żemy stwierdzić z dużą dozą prawdopodobieństwa, że **Cracovia grała lepiej niż wskazywałyby na to tabela**, a **wyniki Pogoni są nieco ponad stan**.

## 4 Wnioski

Przeprowadziliśmy analizę zależności wyników drużyny od pewnych mniej lub bardziej oczywistych metryk i zbudowaliśmy na tej podstawie modele.

### 4.1 Model trudny

Przewidywanie punktów na mecz na podstawie wskaźników udziału w grze zespołu zawodników młodych i obcokrajowców jest obciążone dużym błędem. Jako główną z przyczyn należy wskazać słabą zależność tych metryk od siebie nawzajem, zwłaszcza jeśli chodzi o te dotyczące młodych zawodników.

Wyniki drużyny są słabo zależne od udziału w grze drużyny młodzieżowców i graczy zza granicy.

## 4.2 Model prosty

Przewidywanie punktów na mecz na podstawie wskaźników goli strzelanych i traconych jest obarczone pewnym błędem, jednakże samo określenie wyników danej drużyny jest bliskie prawdzie.

Wyniki drużyny są silnie zależne od bramek strzelanych i traconych.

## 4.3 Uwagi końcowe

Ponadto należy zaznaczyć, że obydwie doświadczenia niedoszacowywały i przeszacowywały te same drużyny, co może wskazywać na rzeczywiste anomalie w stosunku do sposobu gry drużyny i doboru składu, a osiągnięte przez drużynę wyniki.

Analizę kończę zaznaczając swój osobisty sukces, jakim niewątpliwie jest budowa modelu, w którym Raków jest na pierwszym, jedynym słusznym dla tego klubu, miejscu.

## Źródła

- [1] Ekstraklasa, . URL <https://www.ekstraklasa.org>.
- [2] Ekstrastats, . URL <https://ekstrastats.pl>.
- [3] Sport.pl. Duża zmiana w ekstraklasie stała się faktem. pzn wprowadził obowiązek gry młodzieżowca!, 2018. URL <https://www.sport.pl/pilka/7,65039,24255185,rewolucja-w-ekstraklasie-stala-sie-faktem-pzn-wprowadzil-obowiazek.html>.