

Data oddania: _____

Ocena: _____

Radosław Grela 216769
Jakub Wachała 216914

Zadanie 2: Lingwistyczne podsumowania baz danych

1. Cel

2. Wprowadzenie

2.1. Funkcja trapezoidalna

Funkcja trapezoidalna przyjmuje 4 parametry a, b, c, d , dla których spełniony jest warunek $a \leq b \leq c \leq d$. Jej wzór jest następujący [1]:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & \text{gdy } x \in (a, b), \\ 1 & \text{gdy } x \in [b, c], \\ \frac{d-x}{d-c} & \text{gdy } x \in (c, d), \\ 0 & \text{w przeciwnym razie.} \end{cases} \quad (1)$$

2.2. Funkcja trójkątna

Funkcja trójkątna jest szczególnym przypadkiem funkcji trapezoidalnej. Przyjmuje ona trzy parametry a, b, c , dla których zachodzi warunek $a \leq b$

$\leq c$. Te parametry określają punkty „załamania” tej funkcji. Jej wzór jest następujący [4]:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & \text{gdy } x \in (a, b), \\ 1 & \text{gdy } x = b, \\ \frac{c-x}{c-b} & \text{gdy } x \in (b, c), \\ 0 & \text{w przeciwnym razie.} \end{cases} \quad (2)$$

2.3. Funkcja Gaussowska

Funkcja Gaussowska jest definiowana przez 2 parametry które określają środek funkcji oraz jej szerokość. Wzór jest następujący [3]:

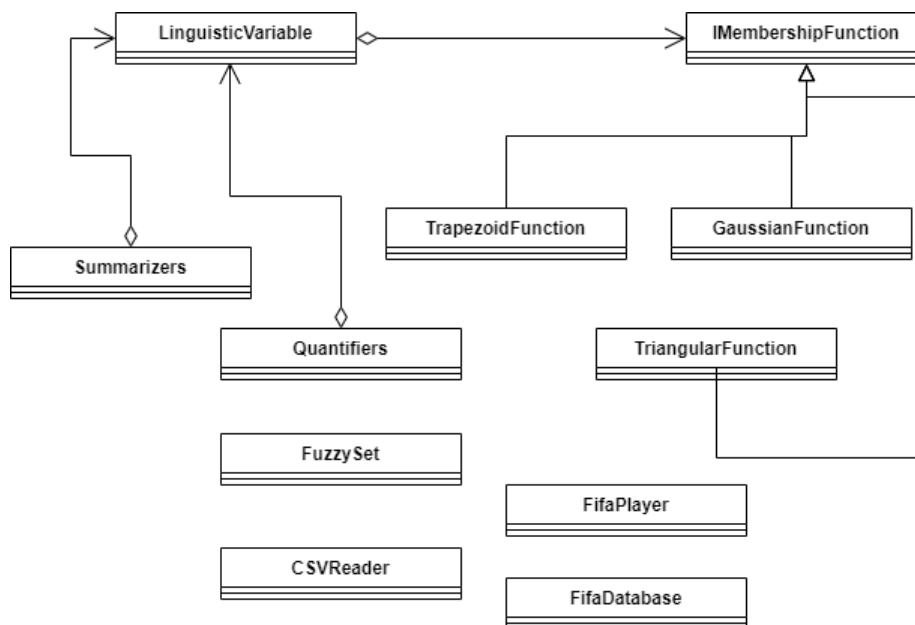
$$\mu_A(x) = e^{-(\frac{x-\bar{x}}{\sigma})^2} \quad (3)$$

gdzie

- \bar{x} jest środkiem funkcji,
- σ określa szerokość krzywej Gaussowskiej.

3. Opis implementacji

Program został stworzony w języku C#. Graficzny interfejs użytkownika został stworzony przy wykorzystaniu Windows Presentation Foundation. W programie wykorzystaliśmy bibliotekę AForge. Poniżej przedstawiamy uproszczony diagram UML naszego programu.



Rysunek 1. Diagram UML.

- Klasa Summarizers odpowiada za poszczególne sumaryzatory, np "młody", "wysoki"
- CSVReader odpowiada za wczytanie pliku csv z danymi do programu
- FIFADatabase odpowiada za bazę danych, czyli przechowywanie wszystkich rekordów
- FuzzySet to klasa odpowiadająca za zbiór rozmyty
- Klasy TrapezoidFunction, GaussianFunction, TriangularFunction odpowiadają za odpowiednie funkcje przynależności
- FIFAPlayer to klasa, która reprezentuje krotkę bazy danych
- Quantifiers jest klasą odpowiedzialną za kwantyfikatory
- LinguisticVariable to klasa reprezentująca zmienną lingwistyczną.

4. Materiały i metody

4.1. Baza danych

Do przeprowadzania badań oraz do generowania podsumowań wykorzystaliśmy bazę danych dotyczącą piłkarzy z gry FIFA 20. Pochodzi ona ze źródła [2]. Składa się ona z 18278 rekordów posiadających 104 atrybuty. Do naszego projektu skorzystamy z 11. Są to następujące atrybuty:

1. Wiek - age - wartość z przedziału [16, 42]

2. Wzrost (w cm) - *height_cm* - wartość z przedziału [156, 205]
3. Waga (w kg) - *weight_kg* - wartość z przedziału [50, 110]
4. Ocena ogólna - *overall* - wartość z przedziału [48, 94]
5. Wykończenie - *attacking_finishing* - wartość z przedziału [2, 95]
6. Dribbling - *skill_dribbling* - wartość z przedziału [4, 97]
7. Podkręcenie piłki - *skill_curve* - wartość z przedziału [6, 94]
8. Długie podania - *skill_long_passing* - wartość z przedziału [8, 92]
9. Sprint - *movement_sprint_speed* - wartość z przedziału [11, 96]
10. Siła strzału - *power_shot_power* - wartość z przedziału [14, 95]

Każda z kolumn jest typu całkowitego.

4.2. Zmienne lingwistyczne

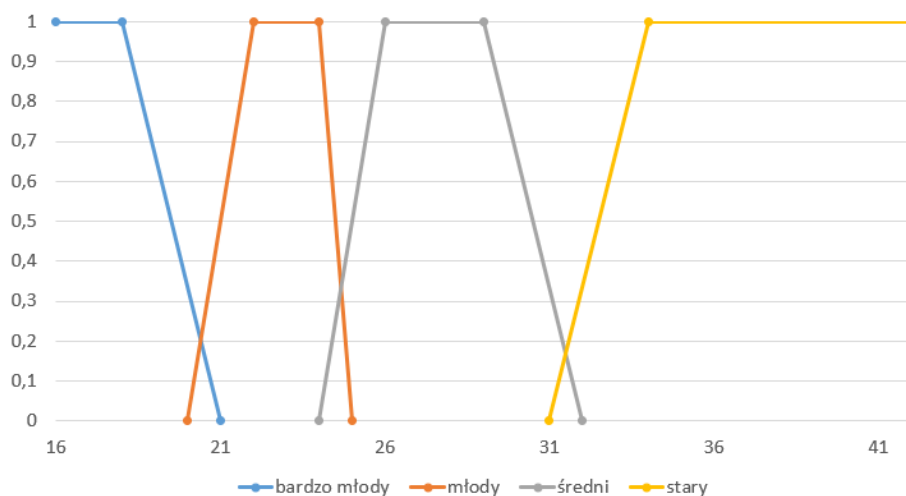
4.2.1. Wiek

Należy zauważyć, że wiek w przypadku zawodnika piłki nożnej oceniany jest w inny sposób niż wiek przeciętnego człowieka.

- (16-21) *bardzo młody*
- (20-25) *młody*
- (24-32) *średni*
- (31-42) *stary*

| Etykieta | a | b | c | d |
|--------------|----|----|----|----|
| bardzo młody | 16 | 16 | 18 | 21 |
| młody | 20 | 22 | 24 | 25 |
| średni | 24 | 26 | 29 | 32 |
| stary | 31 | 34 | 42 | 42 |

Tabela 1. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Wiek.



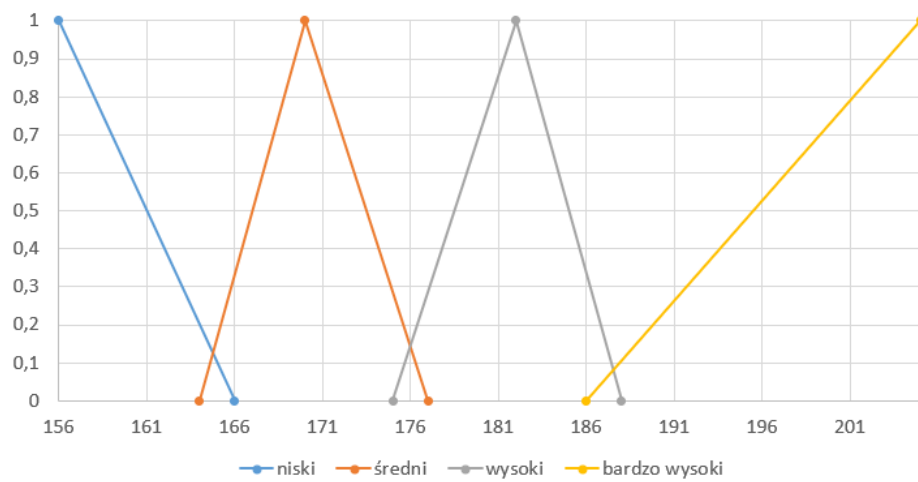
Rysunek 2. Funkcja przynależności (trapezoidalna) dla atrybutu Wiek.

4.2.2. Wzrost

- (156-166) *niski*
- (164-177) *średni*
- (175-188) *wysoki*
- (186-205) *bardzo wysoki*

| Etykieta | a | b | c |
|---------------|-----|-----|-----|
| niski | 156 | 156 | 166 |
| średni | 164 | 170 | 177 |
| wysoki | 175 | 182 | 188 |
| bardzo wysoki | 186 | 205 | 205 |

Tabela 2. Przyporządkowane parametry funkcji trójkątnej dla atrybutu Wzrost.



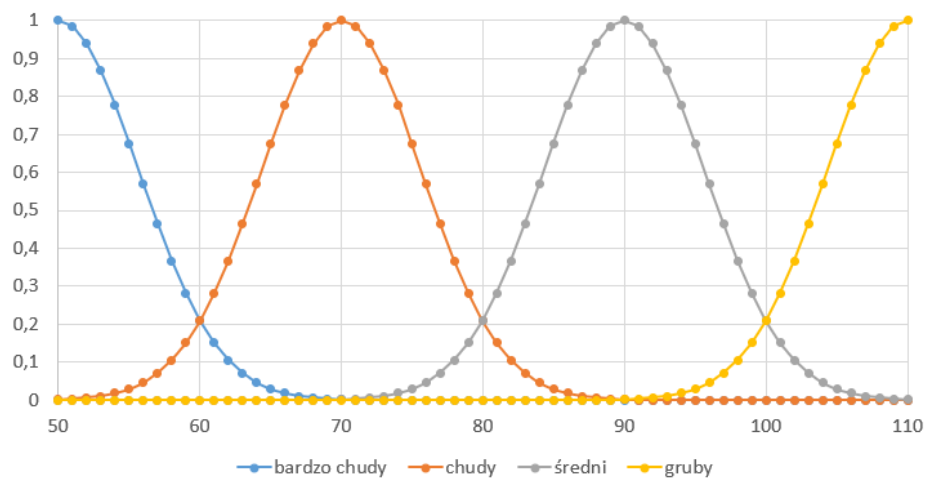
Rysunek 3. Funkcja przynależności (trapezoidalna) dla atrybutu Wzrost.

4.2.3. Waga

- (50-65) *bardzo chudy*
- (55-85) *chudy*
- (75-105) *średni*
- (95-110) *gruby*

| Etykieta | \bar{x} | σ |
|--------------|-----------|----------|
| bardzo chudy | 50 | 8 |
| chudy | 70 | 8 |
| średni | 90 | 8 |
| gruby | 110 | 8 |

Tabela 3. Przyporządkowane parametry funkcji gaussowskiej dla atrybutu Waga.



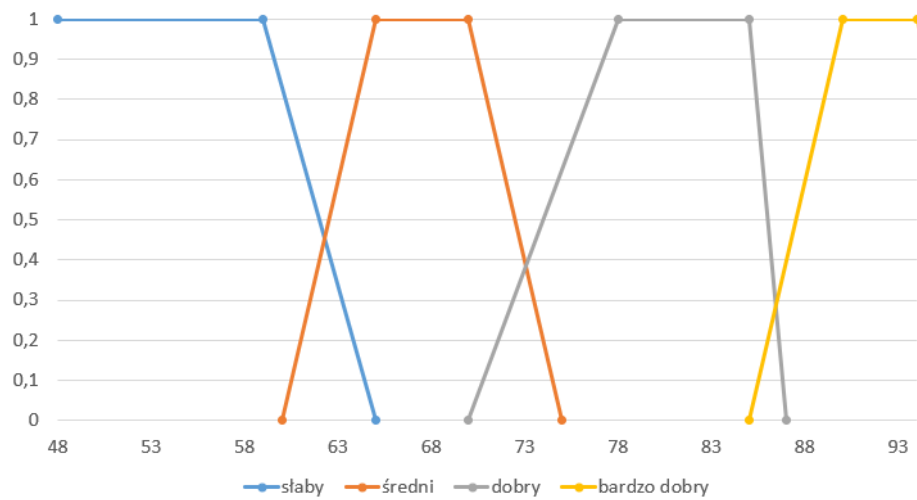
Rysunek 4. Funkcja przynależności (gaussowska) dla atrybutu Waga.

4.2.4. Ocena ogólna

- (48-65) *słaby*
- (60-75) *średni*
- (70-87) *dobry*
- (85-94) *bardzo dobry*

| Etykieta | a | b | c | d |
|--------------|----|----|----|----|
| słaby | 48 | 48 | 59 | 65 |
| średni | 60 | 65 | 70 | 75 |
| dobry | 70 | 78 | 85 | 87 |
| bardzo dobry | 85 | 90 | 94 | 94 |

Tabela 4. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Ocena ogólna.



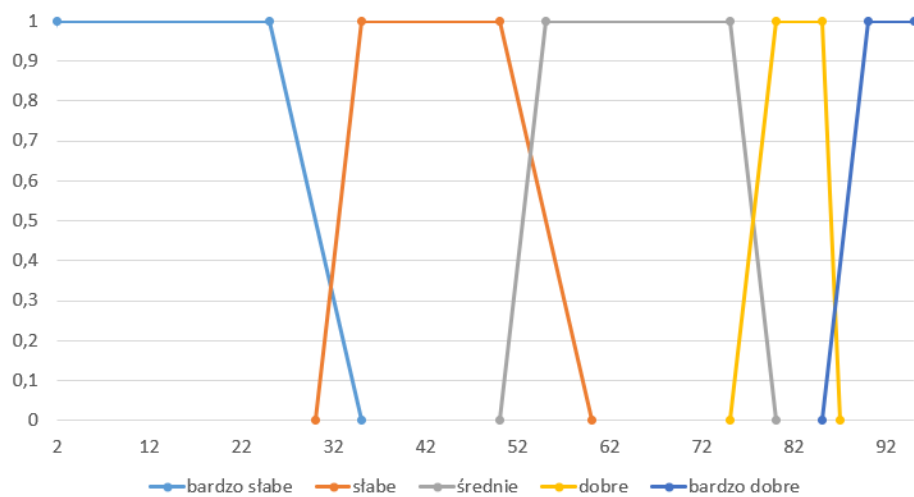
Rysunek 5. Funkcja przynależności (trapezoidalna) dla atrybutu Ocena ogólna.

4.2.5. Wykończenie

- (2-35) *bardzo słabe*
- (30-60) *słabe*
- (50-80) *średnie*
- (75-87) *dobre*
- (85-95) *bardzo dobre*

| Etykieta | a | b | c | d |
|--------------|----|----|----|----|
| bardzo słabe | 2 | 2 | 25 | 35 |
| słabe | 30 | 35 | 50 | 60 |
| średnie | 50 | 55 | 75 | 80 |
| dobre | 75 | 80 | 85 | 87 |
| bardzo dobre | 85 | 90 | 95 | 95 |

Tabela 5. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Wykończenie.



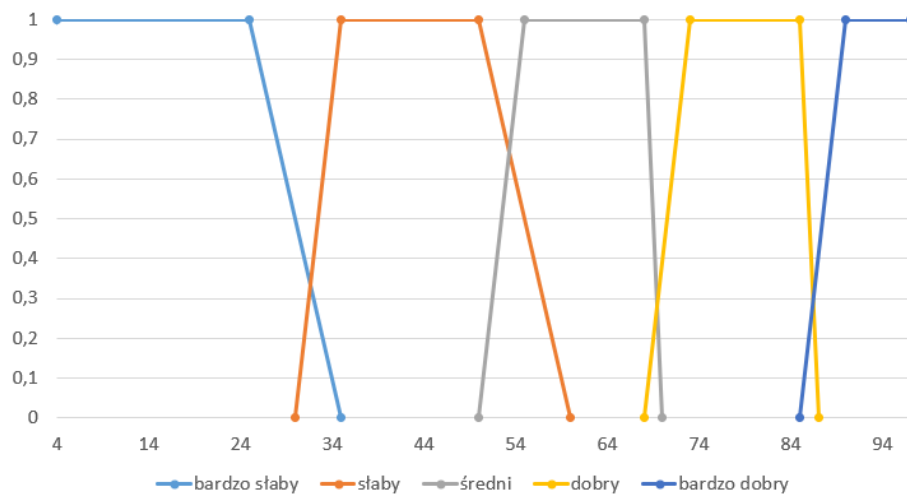
Rysunek 6. Funkcja przynależności (trapezoidalna) dla atrybutu Wykończenie.

4.2.6. Dribbling

- (4-35) *bardzo słaby*
- (30-60) *słaby*
- (50-70) *średni*
- (68-87) *dobry*
- (85-97) *bardzo dobry*

| Etykieta | a | b | c | d |
|--------------|----|----|----|----|
| bardzo słaby | 4 | 4 | 25 | 35 |
| słaby | 30 | 35 | 50 | 60 |
| średni | 50 | 55 | 68 | 70 |
| dobry | 68 | 73 | 85 | 87 |
| bardzo dobry | 85 | 90 | 97 | 97 |

Tabela 6. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Dribbling.



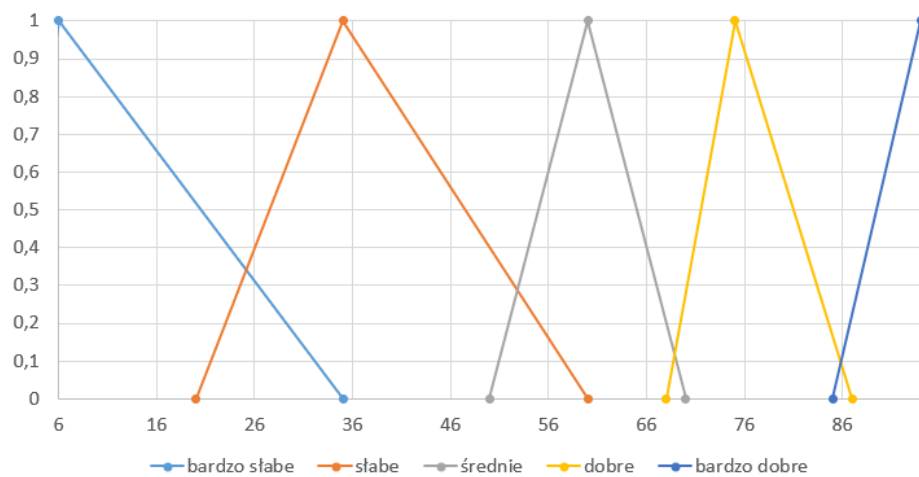
Rysunek 7. Funkcja przynależności (trapezoidalna) dla atrybutu Dribbling.

4.2.7. Podkręcenie piłki

- (6-35) *bardzo słabe*
- (30-60) *słabe*
- (50-70) *średnie*
- (68-87) *dobrze*
- (85-94) *bardzo dobre*

| Etykieta | a | b | c |
|--------------|----|----|----|
| bardzo słabe | 6 | 6 | 35 |
| słabe | 20 | 35 | 60 |
| średnie | 50 | 60 | 70 |
| dobrze | 68 | 75 | 87 |
| bardzo dobre | 85 | 94 | 94 |

Tabela 7. Przyporządkowane parametry funkcji trójkątnej dla atrybutu Podkręcenie piłki.



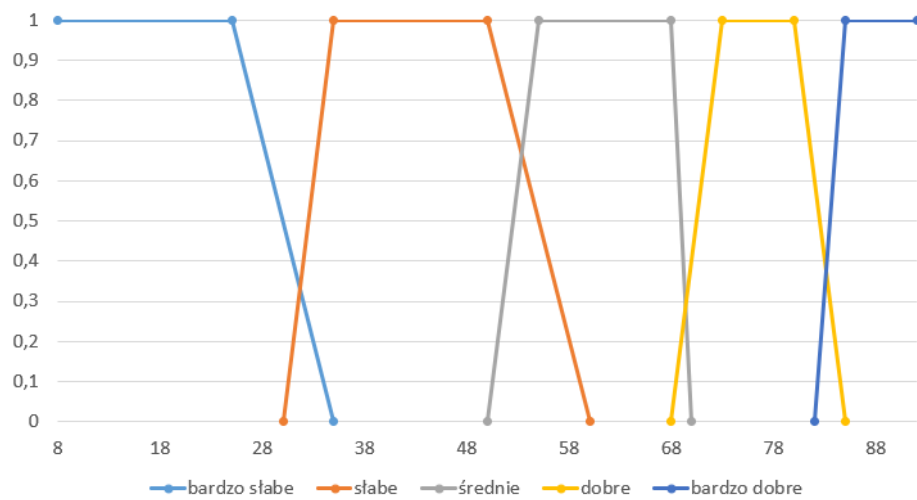
Rysunek 8. Funkcja przynależności (trójkątna) dla atrybutu Podkręcenie piłki.

4.2.8. Długie podania

- (8-35) *bardzo słabe*
- (30-60) *słabe*
- (50-70) *średnie*
- (68-85) *dobrze*
- (82-92) *bardzo dobre*

| Etykieta | a | b | c | d |
|--------------|----|----|----|----|
| bardzo słabe | 8 | 8 | 25 | 35 |
| słabe | 30 | 35 | 50 | 60 |
| średnie | 50 | 55 | 68 | 70 |
| dobrze | 68 | 73 | 80 | 85 |
| bardzo dobre | 82 | 85 | 92 | 92 |

Tabela 8. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Długie podania.



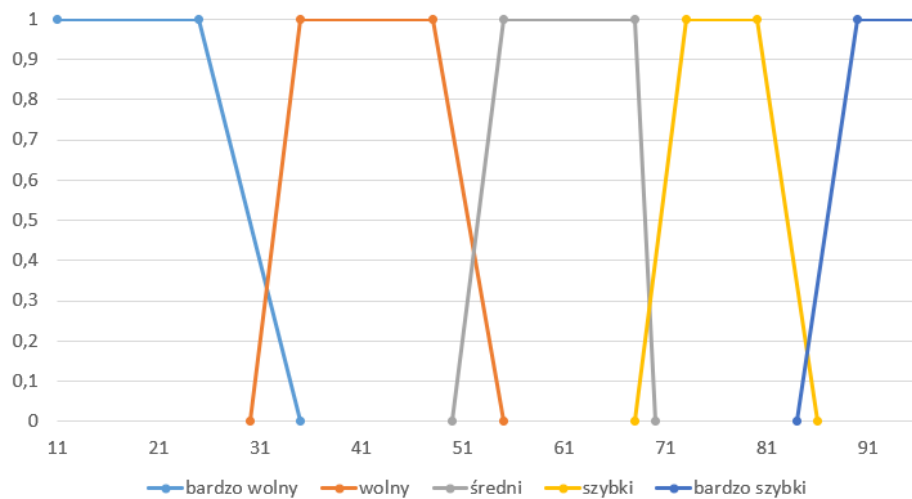
Rysunek 9. Funkcja przynależności (trapezoidalna) dla atrybutu Długie podania.

4.2.9. Sprint

- (11-35) *bardzo wolny*
- (30-55) *wolny*
- (50-70) *średni*
- (68-86) *szybki*
- (84-96) *bardzo szybki*

| Etykieta | a | b | c | d |
|---------------|----|----|----|----|
| bardzo wolny | 11 | 11 | 25 | 35 |
| wolny | 30 | 35 | 48 | 55 |
| średni | 50 | 55 | 68 | 70 |
| szybki | 68 | 73 | 80 | 86 |
| bardzo szybki | 84 | 90 | 96 | 96 |

Tabela 9. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Sprint.



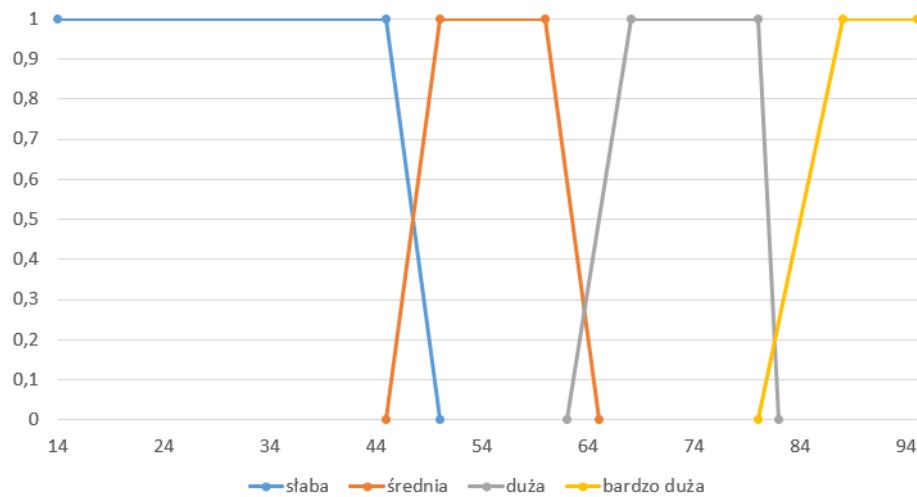
Rysunek 10. Funkcja przynależności (trapezoidalna) dla atrybutu Sprint.

4.2.10. Siła strzału

- (14-50) *słaba*
- (45-65) *średnia*
- (62-82) *duża*
- (80-95) *bardzo duża*

| Etykieta | a | b | c | d |
|-------------|----|----|----|----|
| słaba | 14 | 14 | 45 | 50 |
| średnia | 45 | 50 | 60 | 65 |
| duża | 62 | 68 | 80 | 82 |
| bardzo duża | 80 | 88 | 95 | 95 |

Tabela 10. Przyporządkowane parametry funkcji trapezoidalnej dla atrybutu Siła strzału.



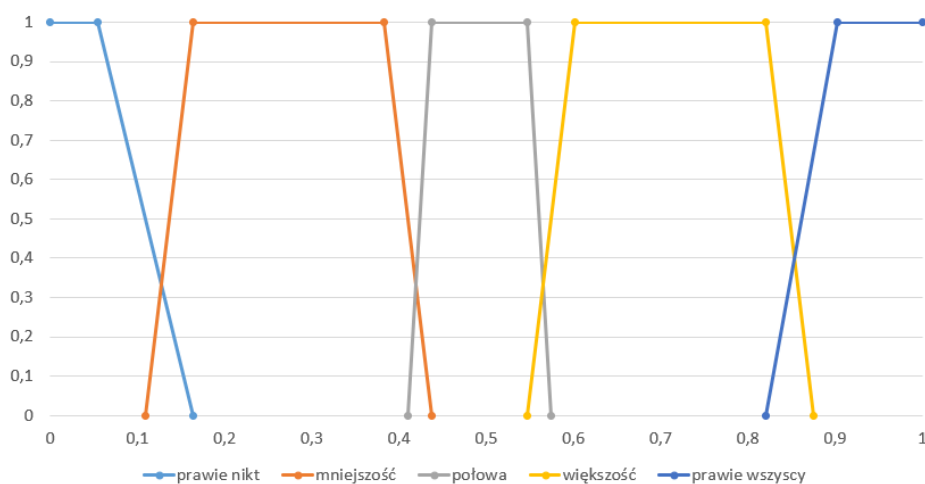
Rysunek 11. Funkcja przynależności (trapezoidalna) dla atrybutu Siła strzału.

4.3. Kwantyfikator względny

Poniżej przedstawiliśmy wartości parametrów oraz wykres funkcji przynależności dla kwantyfikatora względnego. Liczba rekordów w naszej bazie danych wynosi 18278, wykres zawiera się w wartościach $[0, 1]$.

| Etykieta | a | b | c | d |
|----------------|-------|-------|-------|-------|
| prawie nikt | 0,000 | 0,000 | 0,055 | 0,164 |
| mniejszość | 0,109 | 0,164 | 0,383 | 0,438 |
| połowa | 0,410 | 0,438 | 0,547 | 0,574 |
| większość | 0,547 | 0,602 | 0,821 | 0,875 |
| prawie wszyscy | 0,821 | 0,903 | 1,000 | 1,000 |

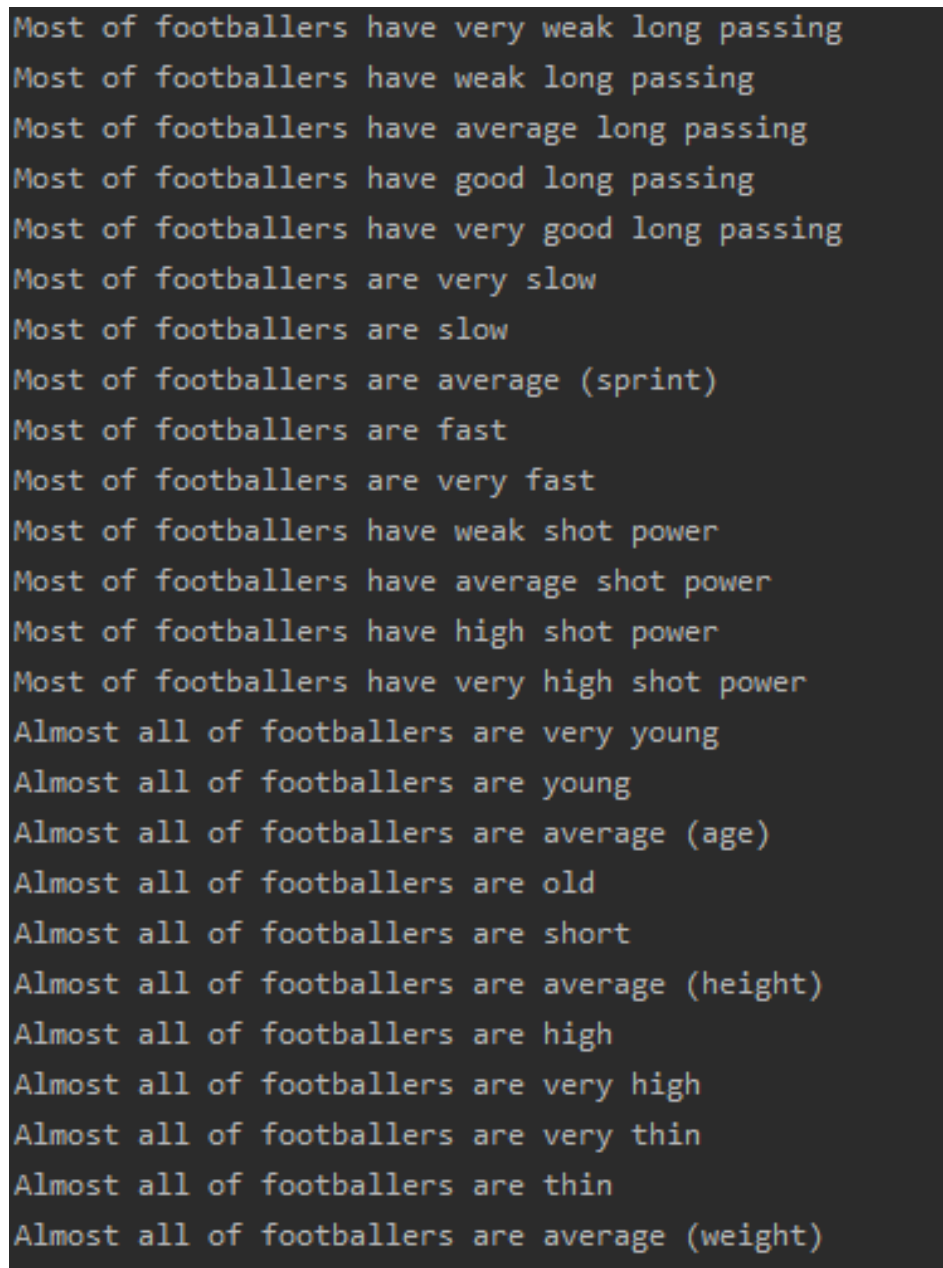
Tabela 11. Przyporządkowane parametry funkcji trapezoidalnej dla kwantyfikatora względnego.



Rysunek 12. Funkcja przynależności kwantyfikatora względnego.

5. Wyniki

Poniżej przedstawiamy przykładowe zdania podsumowujące bazę danych wygenerowane przez nas program.



```
Most of footballers have very weak long passing
Most of footballers have weak long passing
Most of footballers have average long passing
Most of footballers have good long passing
Most of footballers have very good long passing
Most of footballers are very slow
Most of footballers are slow
Most of footballers are average (sprint)
Most of footballers are fast
Most of footballers are very fast
Most of footballers have weak shot power
Most of footballers have average shot power
Most of footballers have high shot power
Most of footballers have very high shot power
Almost all of footballers are very young
Almost all of footballers are young
Almost all of footballers are average (age)
Almost all of footballers are old
Almost all of footballers are short
Almost all of footballers are average (height)
Almost all of footballers are high
Almost all of footballers are very high
Almost all of footballers are very thin
Almost all of footballers are thin
Almost all of footballers are average (weight)
```

Rysunek 13. Zdania wygenerowane przez program.

6. Dyskusja

7. Wnioski

Literatura

- [1] Niewiadomski, Adam. Methods for the Linguistic Summarization of Data: Applications of Fuzzy Sets and Their Extensions. Akademicka Oficyna Wydawnicza EXIT. Warszawa, 2008. ISBN 978-83-60434-40-6
- [2] <https://www.kaggle.com/stefanoleone992/fifa-20-complete-player-dataset>
- [3] https://pracownik.kul.pl/files/31717/public/Funkcje_przynaleznosci.pdf [dostęp 07.05.2020]
- [4] <http://ii.uwb.edu.pl/rudnicki/wp-content/uploads/2016/02/P07.pdf> [dostęp 08.05.2020]