

Data oddania: _____

Ocena: _____

Radosław Grela 216769
Jakub Wachała 216914

Zadanie 2: Lingwistyczne podsumowania baz danych

1. Cel

Celem zadania jest stworzenie aplikacji desktopowej, która ma za zadanie generować pewną ilość podsumowań lingwistycznych dla danej bazy danych. Aplikacja musiała umożliwić automatyczne generowanie podsumowań lingwistycznych służących do tworzenia wiadomości tekstowych na podstawie dużej bazy danych (ponad 10 tys rekordów). [4]

2. Wprowadzenie

W ramach projektu zajmowaliśmy się analizą działania lingwistycznych podsumowań baz danych na zbiorach rozmytych. Definicja zbioru rozmytego jest istotna w naszym zadaniu i brzmi następująco:

Definicja 1. Niech \mathcal{X} będzie zbiorem, który jest "zbiorem zwykłym". Wówczas zbiorem rozmytym opisanym na \mathcal{X} nazywamy każdy taki zbiór postaci:

$$A = \{\langle x, \mu_A(x) \rangle : x \in \mathcal{X}\}, \quad (1)$$

gdzie $\mu_A : \mathcal{X} \rightarrow [0, 1]$ nazywamy funkcją przynależności do zbioru rozmytego A . Jego wartość $x \in \mathcal{X}$ jest określany jako stopień przynależności x do A .

Funkcja przynależności określa, w jakim stopniu element x przynależy do zbioru. W zbiorach rozmytych są to wartości z przedziału $[0, 1]$. W naszym programie wykorzystujemy 3 funkcje przynależności: trapezoidalna, trójkątna oraz gaussowska.

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]:

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

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 [7]:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & \text{gdy } x \in (a, b), \\ 1 & \text{gdy } x = b, \\ \frac{c-x}{c-b} & \text{gdy } x \in (b, c), \\ 0 & \text{w przeciwnym razie.} \end{cases} \quad (3)$$

2.3. Funkcja Gaussowska

Funkcja Gaussowska jest definiowana przez 2 parametry które określają środek funkcji oraz jej szerokość. Wzór jest następujący [6]:

$$\mu_A(x) = e^{-(\frac{x-\bar{x}}{\sigma})^2} \quad (4)$$

gdzie

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

3. Miary jakości

3.1. Miary jakości dla zdań jednopodmiotowych

3.1.1. Degree of truth

Degree of truth to suma przynależności wszystkich rozważanych krotek do podsumowania lingwistycznego. Dla zdań jednopodmiotowych bez kwalifikatora oraz kwantyfikatorów relatywnych wzór wygląda następująco:

$$T_1 = \mu_Q\left(\frac{r}{m}\right) \quad (5)$$

natomiast zdań jednopodmiotowych bez kwalifikatora oraz kwantyfikatorów absolutnych:

$$T_1 = \mu_Q(r) \quad (6)$$

gdzie

$$r = \sum_{i=1}^m \mu_S(d_i) \quad (7)$$

a m to liczba krotek w bazie danych.

Dla zdań jednopodmiotowych z kwalifikatorem kwantyfikator może być tylko względny, a wzór wygląda następująco:

$$T_1 = \mu_Q(r) \quad (8)$$

gdzie

$$r = \frac{\sum_{i=1}^m (\mu_S(d_i) \wedge \mu_W(d_i))}{\sum_{i=1}^m \mu_W(d_i)} \quad (9)$$

3.1.2. Degree of imprecision

Degree of imprecision określa stopień precyzyjności sumaryzatora. Dany jest wzorem:

$$T_2 = 1 - \left(\prod_{j=1}^n \text{in}(S_j) \right)^{1/n} \quad (10)$$

gdzie $\text{in}(S_j)$ to stopień rozmycia wyrażony wzorem $\text{in}(s_j) = \frac{|\text{supp}(S_j)|}{|X|}$ a z kolei $\text{supp}(\cdot)$ oznacza nośnik zbioru rozmytego.

3.1.3. Degree of covering

Degree of covering reprezentuje, stopień, w jakim nośnik sumaryzatora pokrywa się z nośnikiem kwalifikatora. Dany jest wzorem:

$$T_3 = \frac{\sum_{i=1}^m t_i}{\sum_{i=1}^m h_i} \quad (11)$$

gdzie dla zdań z kwalifikatorem:

$$t_i = \begin{cases} 1 & \text{gdy } \mu_S(d_i) > 0 \wedge \mu_W(d_i) > 0 \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

$$h_i = \begin{cases} 1 & \text{gdy } \mu_W(d_i) > 0 \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

a dla zdań bez kwalifikatora:

$$t_i = \begin{cases} 1 & \text{gdy } \mu_S(d_i) > 0 \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

$$h_i = 1$$

3.1.4. Degree of appropriateness

Degree of appropriateness definiuje, jak dużo krotek przynależy do sumaryzatora, czyli czy określone podsumowanie jest odpowiednie dla zestawu danych. Dany jest wzorem:

$$T_4 = \left| \prod_{j=1}^n r_j - T_3 \right| \quad (12)$$

gdzie

$$r_j = \frac{\sum_{i=1}^m g_{ij}}{m} \quad (13)$$

natomiast $g_{ij} = \begin{cases} 1 & \text{gd}y \mu_{S_j}(d_i) > 0 \\ 0 & \text{w przeciwnym wypadku.} \end{cases}$

3.1.5. Length of a summary

Length of a summary określa jakość podsumowania na podstawie złożoności sumaryzatora, czyli im więcej składowych sumaryzatora złożonego, tym niższa wartość tej miary. Dany jest wzorem:

$$T_5 = 2 \cdot \left(\frac{1}{2}\right)^{|S|} \quad (14)$$

gdzie $|S|$ to liczba zbiorów rozmytych z jakich złożony jest sumaryzator.

3.1.6. Degree of quantifier imprecision

Degree of quantifier imprecision przedstawia w jakim stopniu precyzyjny jest kwantyfiktor. Im mniejszy nośnik zbioru rozmytego tym wyższa jest jego precyzja. Dany jest wzorem:

$$T_6 = 1 - in(Q) = 1 - \frac{|supp(Q)|}{|\mathcal{X}_Q|} \quad (15)$$

gdzie $|\mathcal{X}_Q| = 1$ dla kwantyfikatora relatywnego, natomiast dla kwantyfikatora absolutnego $|\mathcal{X}_Q| = m$, czyli liczba krotek w bazie danych.

3.1.7. Degree of quantifier cardinality

Degree of quantifier cardinality opisuje stopień precyzji kwantyfikatora, im większa kardynalność kwantyfikatora tym jest on mniej precyzyjny. Dany jest wzorem:

$$T_7 = 1 - \frac{|Q|}{|\mathcal{X}_Q|} \quad (16)$$

gdzie $|\cdot| = clm(\cdot)$ - całka z funkcji przynależności zbioru rozmytego (czyli pole pod jego wykresem).

3.1.8. Degree of summarizer cardinality

Degree of summarizer cardinality opisuje stopień precyzji sumaryzatora, im mniejsza kardynalność sumaryzatora tym jest on bardziej precyzyjny. Dany jest wzorem:

$$T_8 = 1 - \left(\prod_{j=1}^n \frac{|S_j|}{|\mathcal{X}_j|} \right)^{\frac{1}{n}} \quad (17)$$

gdzie n to liczba zbiorów rozmytych z jakich stworzony jest sumaryzator.

3.1.9. Degree of qualifier imprecision

Degree of qualifier imprecision określa, w jakim stopniu precyzyjny jest kwalifikator. Im szerszy nośnik zbioru rozmytego tym niższa jest jego precyzja. Dany jest wzorem:

$$T_9 = 1 - \left(\prod_{j=1}^x in(W_{gj}) \right)^{\frac{1}{x}} \quad (18)$$

gdzie $in(W_{gj})$ to stopień rozmycia zbioru rozmytego W_{gj} .

3.1.10. Degree of qualifier cardinality

Degree of qualifier cardinality opisuje stopień precyzji kwalifikatora, im większa jest kardynalność kwalifikatora, tym jest on mniej precyzyjny. Dany jest wzorem:

$$T_{10} = 1 - \left(\prod_{j=1}^x \frac{|W_{gj}|}{|\mathcal{X}_{gj}|} \right)^{\frac{1}{x}} \quad (19)$$

3.1.11. Length of qualifier

Length of qualifier wyznacza jakość podsumowania na podstawie złożoności kwalifikatora. Im bardziej złożony kwalifikator, tym jakość podsumowania jest gorsza. Dany jest wzorem:

$$T_{11} = 2 \cdot \left(\frac{1}{2} \right)^{|W|} \quad (20)$$

gdzie $|W|$ to liczba zbiorów rozmytych, z jakich stworzony jest kwalifikator.

3.2. Miary jakości dla zdań wielopodmiotowych

Dla zdań wielopodmiotowych liczymy tylko stopień prawdziwości, czyli miarę T_1 .

3.2.1. Degree of truth dla zdań wielopodmiotowych w pierwszej formie [3]

$$T_{1_1} = \left(\frac{\frac{1}{M_1} \Sigma - count(S_{P_1})}{\frac{1}{M_1} \Sigma - count(S_{P_1}) + \frac{1}{M_2} \Sigma - count(S_{P_2})} \right) \quad (21)$$

gdzie

- $\Sigma - count(S_{P_1}) = \sum_{i=1}^m \{\mu_S(d_i) : d_i \in^* P_1\}$
- $\Sigma - count(S_{P_2}) = \sum_{i=1}^m \{\mu_S(d_i) : d_i \in^* P_2\}$
- M_1 to liczba krotek reprezentujących podmiot P_1
- M_2 to liczba krotek reprezentujących podmiot P_2

3.2.2. Degree of truth dla zdań wielopodmiotowych w drugiej formie [2]

$$T_{1_2} = \left(\frac{\frac{1}{M_1} \Sigma - count(S_{P_1})}{\frac{1}{M_1} \Sigma - count(S_{P_1}) + \frac{1}{M_2} \Sigma - count(S_{P_2} \cap W)} \right) \quad (22)$$

gdzie

$$— \Sigma - count(S_{P_2} \cap W) = \sum_{i=1}^m \{min(\mu_{S_1}(d_i), \mu_W(d_i)) : d_i \in^* P_2\}$$

3.2.3. Degree of truth dla zdań wielopodmiotowych w trzeciej formie [2]

$$T_{1_3} = \left(\frac{\frac{1}{M_1} \Sigma - count(S_{P_1} \cap W)}{\frac{1}{M_1} \Sigma - count(S_{P_1} \cap W) + \frac{1}{M_2} \Sigma - count(S_{P_2})} \right) \quad (23)$$

gdzie

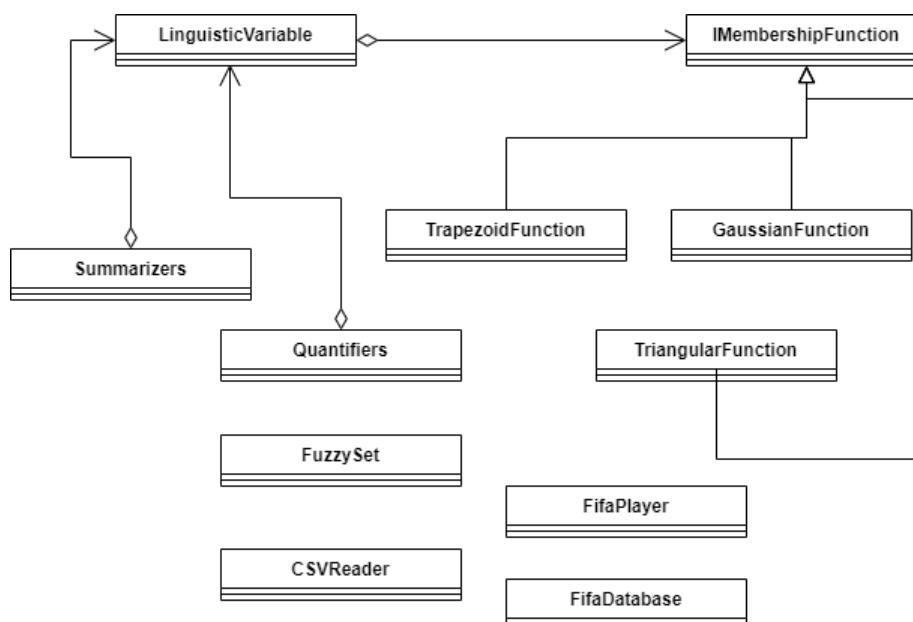
$$— \Sigma - count(S_{P_1} \cap W) = \sum_{i=1}^m \{min(\mu_{S_1}(d_i), \mu_W(d_i)) : d_i \in^* P_1\}$$

3.2.4. Degree of truth dla zdań wielopodmiotowych w czwartej formie [3]

$$T_{1_3} = \left(\frac{\Sigma - count(S_{P_1})}{\Sigma - count(S_{P_1}) + \Sigma - count(S_{P_2})} \right) \quad (24)$$

4. Opis implementacji

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



Rysunek 1. Diagram UML.

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

5. Materiały i metody

5.1. Baza danych

Do przeprowadzania badań oraz do generowania podsumowań wykorzystaliśmy bazę danych dotyczącą piłkarzy z gry FIFA 20. Pochodzi ona ze źródła [5]. 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.

5.2. Zmienne lingwistyczne

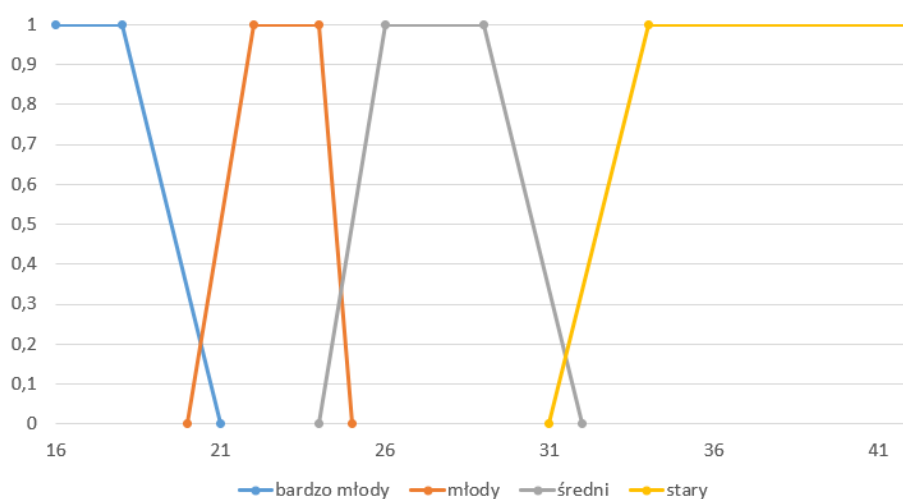
5.2.1. Wiek

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.



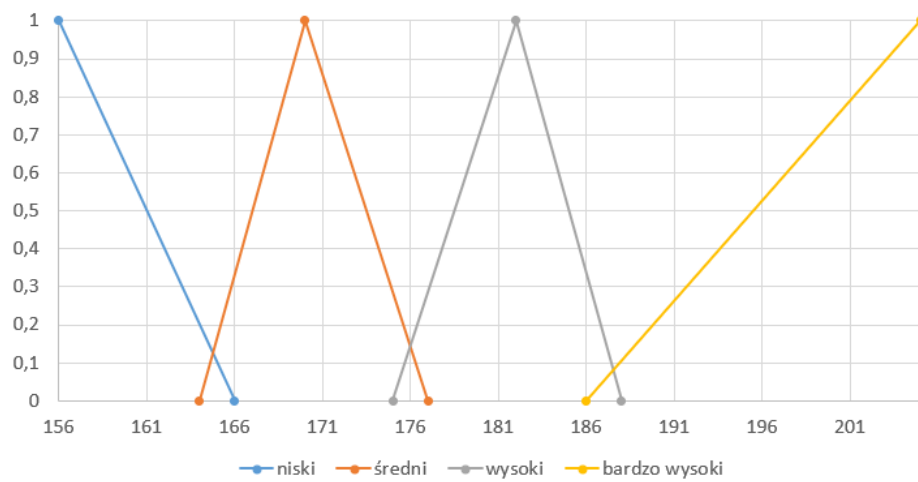
Rysunek 2. Funkcja przynależności (trapezoidalna) dla atrybutu Wiek.

5.2.2. 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.



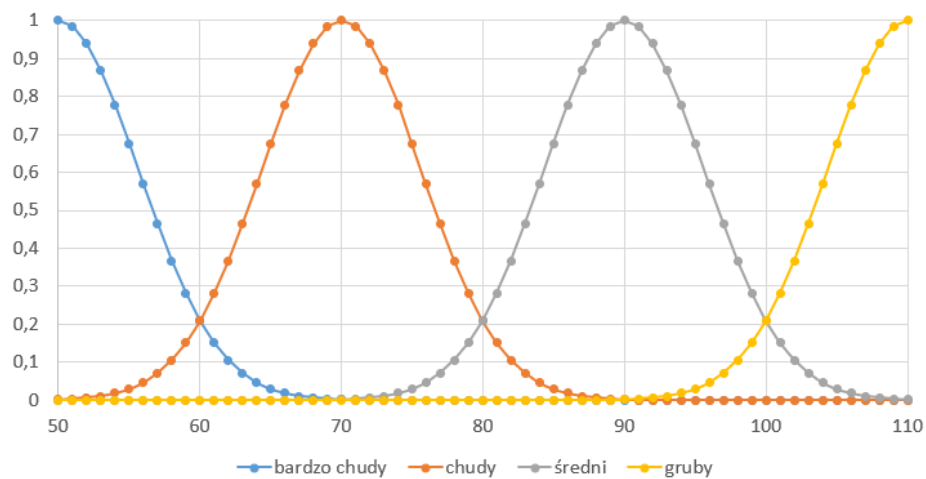
Rysunek 3. Funkcja przynależności (trapezoidalna) dla atrybutu Wzrost.

5.2.3. 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.



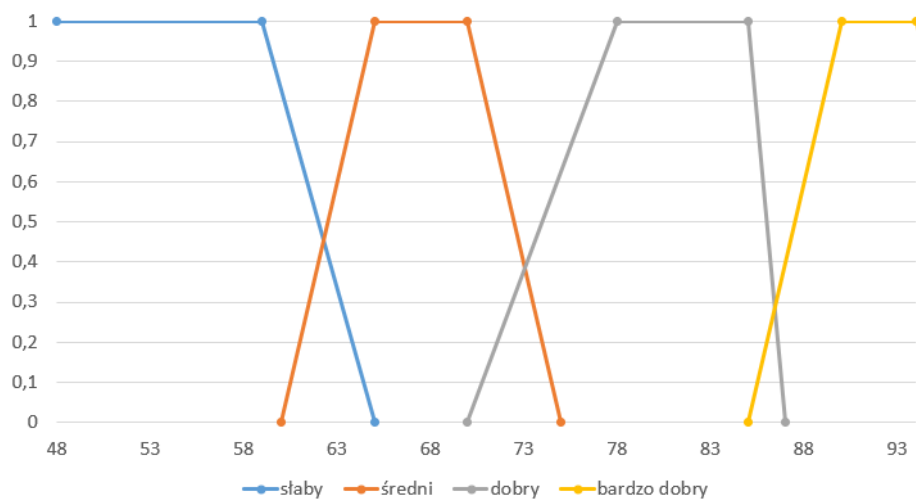
Rysunek 4. Funkcja przynależności (gaussowska) dla atrybutu Waga.

5.2.4. 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.



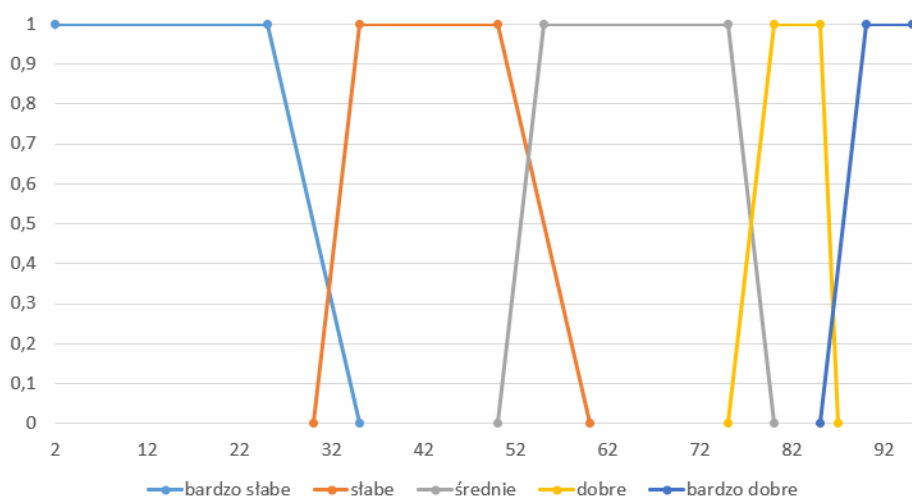
Rysunek 5. Funkcja przynależności (trapezoidalna) dla atrybutu Ocena ogólna.

5.2.5. 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.



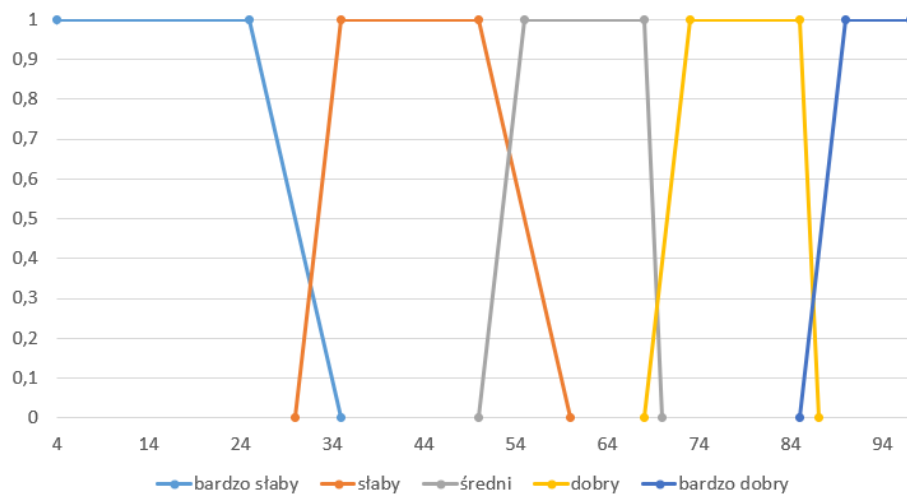
Rysunek 6. Funkcja przynależności (trapezoidalna) dla atrybutu Wykończenie.

5.2.6. 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.



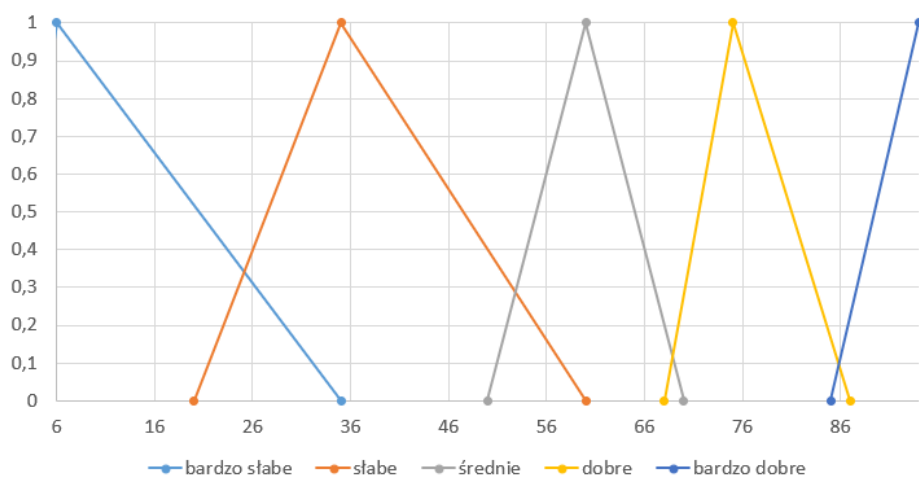
Rysunek 7. Funkcja przynależności (trapezoidalna) dla atrybutu Dribbling.

5.2.7. 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.



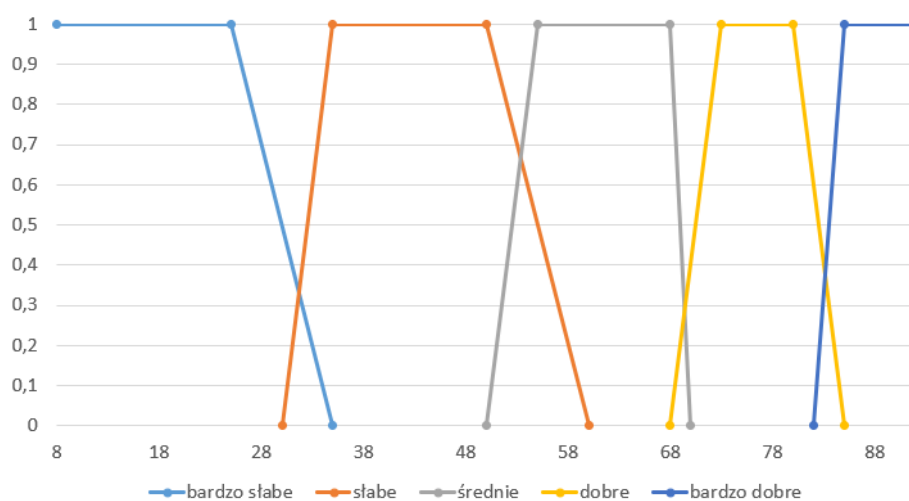
Rysunek 8. Funkcja przynależności (trójkątna) dla atrybutu Podkręcenie piłki.

5.2.8. 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.



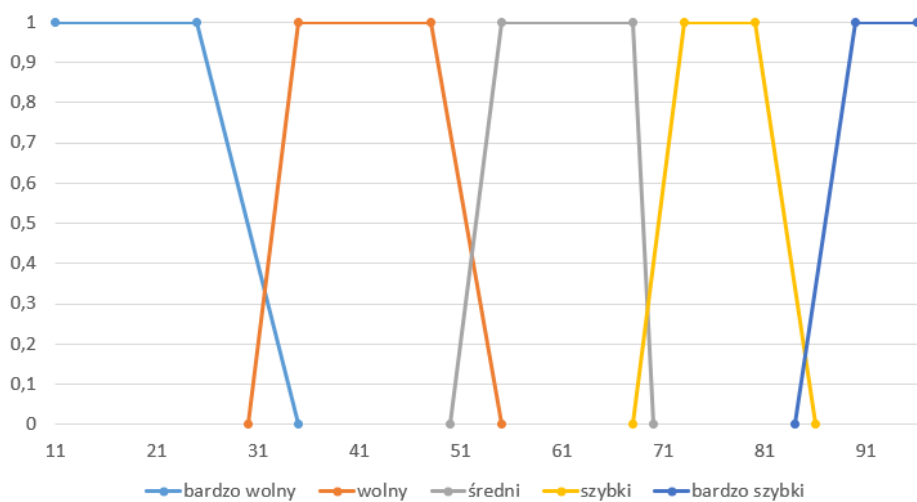
Rysunek 9. Funkcja przynależności (trapezoidalna) dla atrybutu Długie podania.

5.2.9. 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.



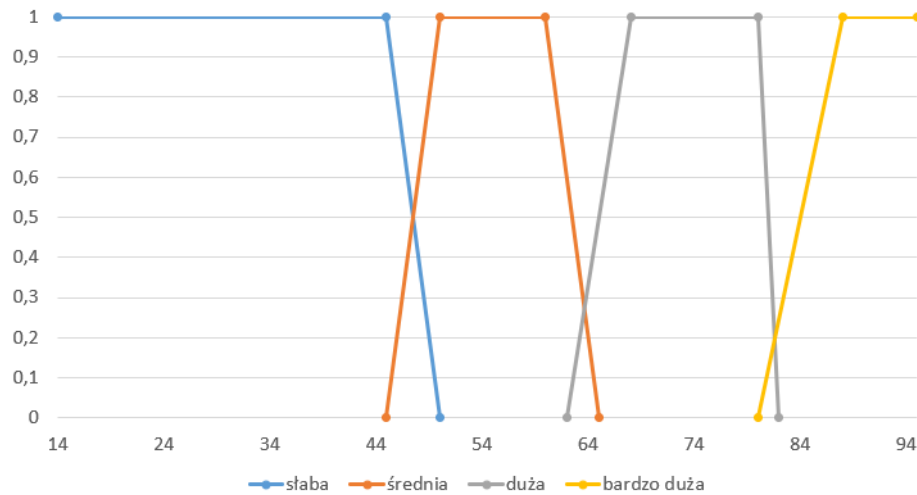
Rysunek 10. Funkcja przynależności (trapezoidalna) dla atrybutu Sprint.

5.2.10. 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.



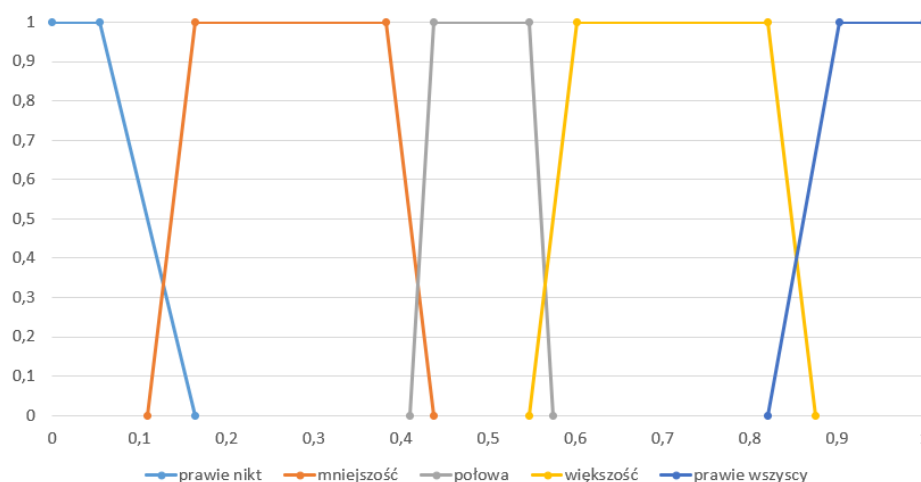
Rysunek 11. Funkcja przynależności (trapezoidalna) dla atrybutu Siła strzału.

5.3. 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.

5.4. Kwantyfikator absolutny

| Etykieta | a | b | c | d |
|-----------------|-------|-------|-------|-------|
| Około 1000 | 900 | 960 | 1040 | 1100 |
| Więcej niż 3000 | 0,109 | 0,164 | 0,383 | 0,438 |
| Mniej niż 3000 | 0 | 0 | 2990 | 3010 |
| Około 500 | 450 | 480 | 520 | 550 |
| Około 100 | 80 | 90 | 110 | 120 |

Tabela 12. Przyporządkowane parametry funkcji trapezoidalnej dla kwantyfikatora absolutnego.

5.5. Przeprowadzone eksperymenty

Eksperymenty, jakie zdecydowaliśmy przedstawić są następujące:

1. Porównanie miar jakości dla różnych kwantyfikatorów w zdaniach jednopodmiotowych.
2. Porównanie podsumowań z kwalifikatorem oraz bez (zdania jednopodmiotowe w formie pierwszej i drugiej).
3. Porównanie podsumowań z sumaryzatorem oraz złączeniem kilku sumaryzatorów
4. Porównanie miary degree of truth dla różnych typów zdań wielopodmiotowych.

6. Wyniki

6.1. Porównanie miar jakości dla różnych kwantyfikatorów

s

6.2. Porównanie podsumowań z kwalifikatorem oraz bez

s

6.3. Porównanie podsumowań z sumaryzatorem oraz złączeniem kilku sumaryzatorów

s

6.4. Porównanie miary degree of truth dla różnych typów zdań wielopodmiotowych

s

7. Dyskusja

8. 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] Niewiadomski, Adam. Zbiory rozmyte typu 2. Zastosowania w reprezentowaniu informacji. Akademicka Oficyna Wydawnicza EXIT. Warszawa, 2019. ISBN 978-83-7837-595-1
- [3] Pozyskiwanie wiedzy z relacyjnych baz danych: wielopodmiotowe podsumowania lingwistyczne. [online] [dostęp 04.06.2020] <http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-48b6e678-8bdb-4a2b-ab79-77a76631b2e5/c/30.pdf>

- [4] Treść zadania 2. [online] [dostęp 04.06.2020] https://ftims.edu.p.lodz.pl/pluginfile.php/132360/mod_resource/content/6/ksr-projekt2-2019.pdf
- [5] Źródło bazy danych zawodników z gry FIFA 20. [online] [dostęp 04.06.2020] <https://www.kaggle.com/stefanoleone992/fifa-20-complete-player-dataset>
- [6] https://pracownik.kul.pl/files/31717/public/Funkcje_przynaleznosci.pdf [dostęp 07.05.2020]
- [7] <http://ii.uwb.edu.pl/rudnicki/wp-content/uploads/2016/02/P07.pdf> [dostęp 08.05.2020]