

# Komputerowe systemy rozpoznawania

## 2024/2025

### Projekt 2. Podsumowania lingwistyczne relacyjnych baz danych

Dominik Gałkowski, 247659  
Jan Śladowski, 247806  
Prowadzący: dr inż. Marcin Kacprowicz

2 czerwca 2025

## 1 Cel

Celem projektu jest stworzenie aplikacji, której główną funkcjonalnością jest lingwistyczna agregacja zawartości wybranego zbioru danych. Ma ona za zadanie generowanie podsumowań lingwistycznych dla wybranych przez użytkownika kwantyfikatorów, sumaryzatorów i kwalifikatorów dla różnych atrybutów. Analiza otrzymanych wyników polega na określeniu, znaczenia wybranych kwantyfikatorów, sumaryzatorów, kwalifikatorów oraz miar ich jakości dla wiarygodności i jakości otrzymanych podsumowań lingwistycznych. Przykładowe podsumowanie to, np. większość pomiarów ma wysokie ciśnienie.

## 2 Baza danych, zmienne lingwistyczne, kwantyfikatory lingwistyczne

### 2.1 Charakterystyka podsumowywanej bazy danych

W tym projekcie został wykorzystany zbiór danych zapisany w pliku w formacie .csv, na podstawie którego utworzono bazę danych - PostgreSQL. Baza danych o nazwie World Weather Repository zawiera różnego rodzaju pomiary danych atmosferycznych, np. temperatura lub prędkość wiatru. [1] Użyteczność bazy jest określona na stronie Kaggle jako 10.0, a dane z niej są wykorzystywane do prognozowania pogody oraz analizy klimatu na różnych kontynentach. Baza jest namiętnie aktualizowana, natomiast na dzień 18.05.2025r. składa się z 70292 rekordów.

Zmiennym lingwistycznym przypisuje się znaczenie ze względu na potrzebę lepszej interpretowalności danych przez użytkowników. Ludzie rzadko reagują

na dokładne wartości (np. 1033.8 mb ciśnienia), natomiast określenie „wysokie ciśnienie” pozwala im intuicyjnie rozumieć sytuację pogodową. Stąd istnieje zapotrzebowanie na „przekładanie” danych formalnych na język naturalny. Podmiotem podsumowań jest pomiar atmosferyczny, z bazy World Weather Repository wybrano 10 atrybutów, które zostaną rozmyte, są to następujące kolumny:

1. last\_updated - data przeprowadzenia pomiarów, z tego atrybutu zostanie wykorzystana godzina w celu określenia pory dnia, zakres [0 - 24].
2. temperature\_celsius - temperatura wyrażona w stopniach Celsjusza w zakresie [-25, 50].
3. wind\_kph - prędkość wiatru wyrażona w kilometrach na godzinę w zakresie [3, 92].
4. pressure\_mb - ciśnienie powietrza wyrażone w milibarach w zakresie [947 - 1052].
5. humidity - wilgotność w zakresie [2 - 100%].
6. visibility\_km - widoczność wyrażona w kilometrach w zakresie [0, 32].
7. uv\_index - wartość promieniowania słonecznego UVI w zakresie [0, 16].
8. air\_quality\_Carbon\_Monoxide - pomiar jakości powietrza ze względu na stężenie tlenku węgla, wyrażony w ppm (liczba cząstek CO2 w milionie cząstek powietrza), w zakresie [0 - 3000].
9. air\_quality\_Nitrogen\_dioxide - pomiar jakości powietrza ze względu na stężenie dwutlenku azotu, wyrażony w ppm (liczba cząstek NO2 w milionie cząstek powietrza) w zakresie [0 - 428].
10. air\_quality\_gb-defra-index - skala określająca poziomy zanieczyszczenia w powietrzu w zakresie [1 - 10].

## 2.2 Zmienne lingwistyczne (atrybuty/własności obiektów)

Poniżej zostały zaprezentowane zmienne lingwistyczne dla atrybutów opisanych w sekcji 2.1, ich wzory analityczne oraz wykresy funkcji przynależności wraz z dopasowanymi do nich etykietami. W każdym z poniższych wzorów  $L_x$  to zmienna lingwistyczna,  $\mathcal{L}_x$  - nazwa zmiennej lingwistycznej,  $H_x$  zbiór możliwych przyjmowanych wartości,  $X_x$  - przestrzeń rozważań,  $x$  - numer kolejnej zmiennej lingwistycznej.

1. last\_updated

$$L_1 = \langle \mathcal{L}_1, H_1, \mathcal{X}_1 \rangle \quad (1)$$

gdzie:  $\mathcal{L}_1$  – pora dnia,  $H_1 = \{\text{nocna, poranna, południowa, popołudniowa, wieczorna}\}$ ,  $\mathcal{X}_1 = [0, 24]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

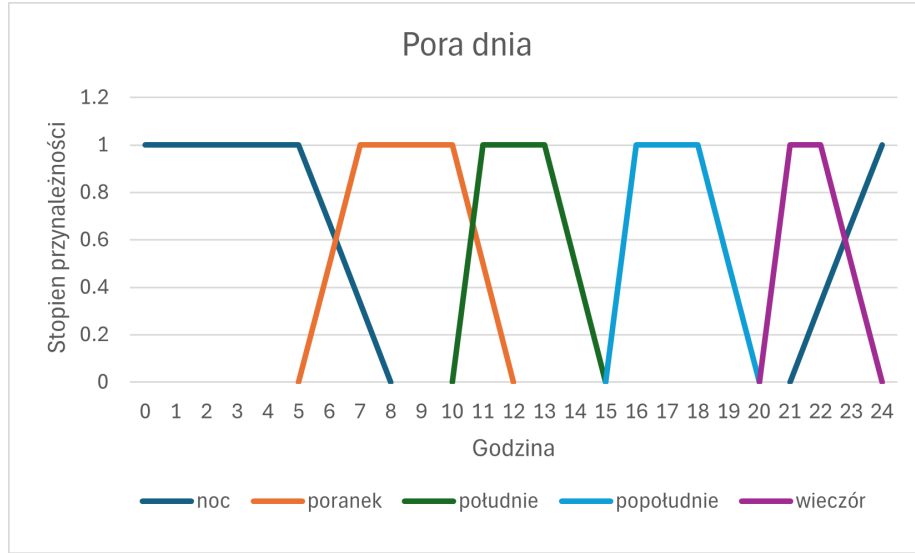
$$\mu_{\text{nocna}}(x) = \begin{cases} \frac{7-x}{3}, & x \in (4, 7) \\ 1, & x \in [0, 4] \\ \frac{x-21}{3}, & x \in [21, 24) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (2)$$

$$\mu_{\text{poranna}}(x) = \begin{cases} \frac{x-4}{3}, & x \in (4, 7) \\ 1, & x \in [7, 9] \\ \frac{12-x}{3}, & x \in (9, 12) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (3)$$

$$\mu_{\text{południowa}}(x) = \begin{cases} \frac{x-9}{3}, & x \in (9, 12) \\ 1, & x \in [12, 13] \\ \frac{16-x}{3}, & x \in (13, 16) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (4)$$

$$\mu_{\text{popołudniowa}}(x) = \begin{cases} \frac{x-13}{3}, & x \in (13, 16) \\ 1, & x \in [16, 17] \\ \frac{20-x}{3}, & x \in (17, 20) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (5)$$

$$\mu_{\text{wieczorna}}(x) = \begin{cases} \frac{x-17}{3}, & x \in (17, 20) \\ 1, & x \in [20, 21] \\ \frac{24-x}{3}, & x \in (21, 24) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (6)$$



Rysunek 1: Wykres funkcji przynależności dla pory dnia.

2. temperature\_celsius

$$L_2 = \langle \mathcal{L}_2, H_2, \mathcal{X}_2 \rangle \quad (7)$$

gdzie:  $\mathcal{L}_2$  – temperatura,  $H_2 = \{\text{bardzo zimna, zimna, umiarkowana, ciepła, gorąca}\}$ ,  $\mathcal{X}_2 = [-25, 50]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

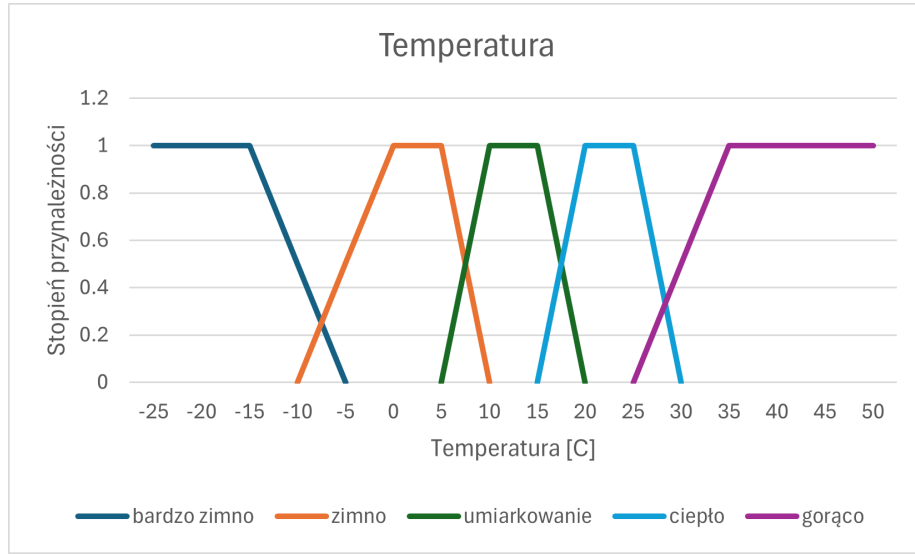
$$\mu_{\text{bardzo\_zimna}}(x) = \begin{cases} 1, & x \in [-25, -15] \\ \frac{-5-x}{10}, & x \in (-15, -5) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (8)$$

$$\mu_{\text{zimna}}(x) = \begin{cases} \frac{x+10}{10}, & x \in (-15, -5) \\ 1, & x \in [-5, 0] \\ \frac{10-x}{10}, & x \in (0, 10) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (9)$$

$$\mu_{\text{umiarkowana}}(x) = \begin{cases} \frac{x}{10}, & x \in (0, 10) \\ 1, & x \in [10, 15] \\ \frac{20-x}{10}, & x \in (15, 25) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (10)$$

$$\mu_{\text{ciepła}}(x) = \begin{cases} \frac{x-15}{10}, & x \in (15, 25) \\ 1, & x \in [25, 30] \\ \frac{30-x}{10}, & x \in (30, 40) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (11)$$

$$\mu_{\text{gorąca}}(x) = \begin{cases} \frac{x-30}{10}, & x \in (30, 40] \\ 1, & x \in [40, 50] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (12)$$



Rysunek 2: Wykres funkcji przynależności dla temperatury.

3. wind\_kph

$$L_3 = \langle \mathcal{L}_3, H_3, \mathcal{X}_3 \rangle \quad (13)$$

gdzie:  $\mathcal{L}_3$  – wiatr,  $H_3$  – {słaby, umiarkowany, silny, bardzo silny, gwałtowny},  $\mathcal{X}_3 = [3, 151]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

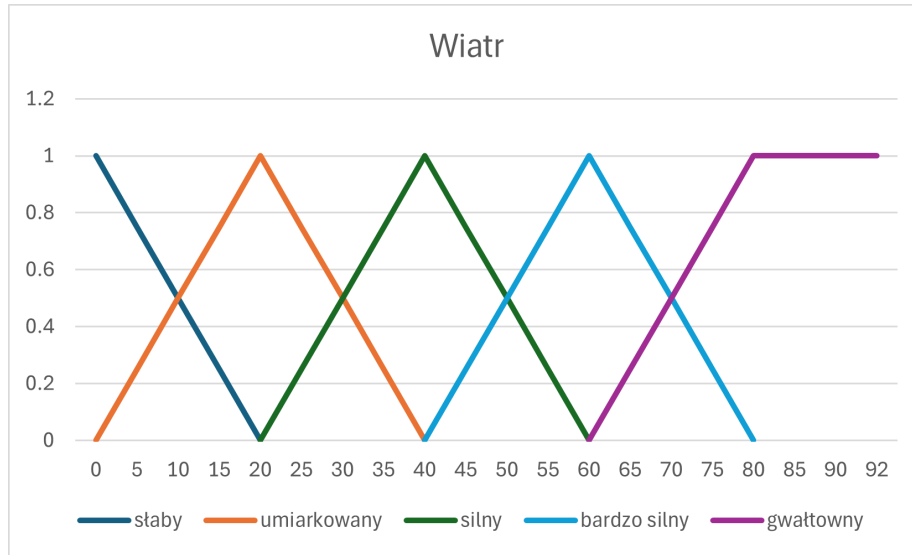
$$\mu_{\text{słaby}}(x) = \begin{cases} 1, & x = 0 \\ \frac{20-x}{20}, & x \in (0, 20) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (14)$$

$$\mu_{\text{umiarkowany}}(x) = \begin{cases} \frac{x}{20}, & x \in (0, 20) \\ 1, & x = 20 \\ \frac{40-x}{20}, & x \in (20, 40) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (15)$$

$$\mu_{\text{silny}}(x) = \begin{cases} \frac{x-20}{20}, & x \in (20, 40) \\ 1, & x = 40 \\ \frac{60-x}{20}, & x \in (40, 60) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (16)$$

$$\mu_{\text{bardzo\_silny}}(x) = \begin{cases} \frac{x-40}{20}, & x \in (40, 60) \\ 1, & x = 60 \\ \frac{80-x}{20}, & x \in (60, 80) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (17)$$

$$\mu_{\text{gwaltowny}}(x) = \begin{cases} \frac{x-60}{20}, & x \in (60, 80) \\ 1, & x \in [80, 92] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (18)$$



Rysunek 3: Wykres funkcji przynależności dla siły wiatru.

4. pressure\_mb

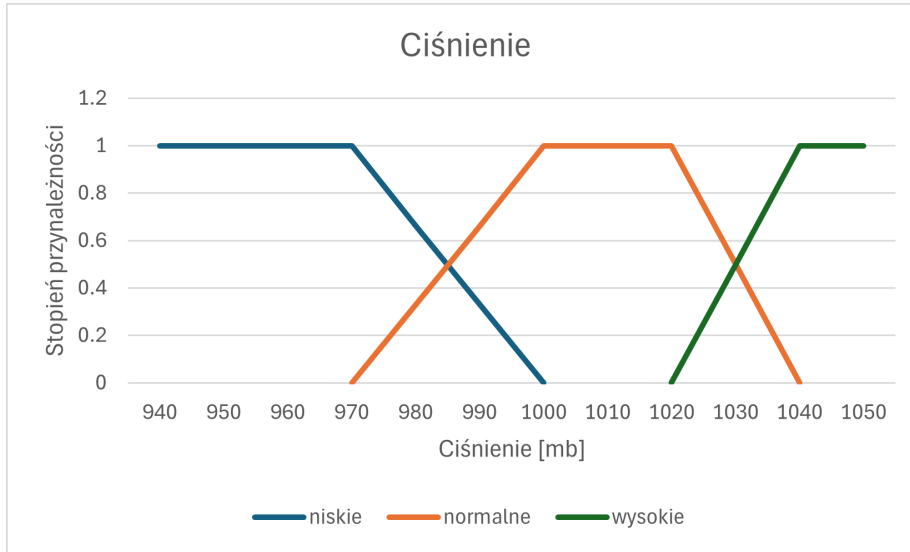
$$L_4 = \langle \mathcal{L}_4, H_4, \mathcal{X}_4 \rangle \quad (19)$$

gdzie:  $\mathcal{L}_4$  – ciśnienie,  $H_4 = \{\text{niskie, normalne, wysokie}\}$ ,  $\mathcal{X}_4 = [947, 1050]$ .  
Poniżej wzory dla wszystkich możliwych etykiet.

$$\mu_{\text{niskie}}(x) = \begin{cases} 1, & x \in [947, 970] \\ \frac{1000-x}{30}, & x \in (970, 1000] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (20)$$

$$\mu_{\text{normalne}}(x) = \begin{cases} \frac{x-970}{30}, & x \in (970, 1000] \\ 1, & x \in (1000, 1020] \\ \frac{1040-x}{20}, & x \in (1020, 1040] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (21)$$

$$\mu_{\text{wysokie}}(x) = \begin{cases} \frac{x-1020}{20}, & x \in (1020, 1040] \\ 1, & x \in (1040, 1052] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (22)$$



Rysunek 4: Wykres funkcji przynależności dla ciśnienia.

##### 5. humidity

$$L_5 = \langle \mathcal{L}_5, H_5, \mathcal{X}_5 \rangle \quad (23)$$

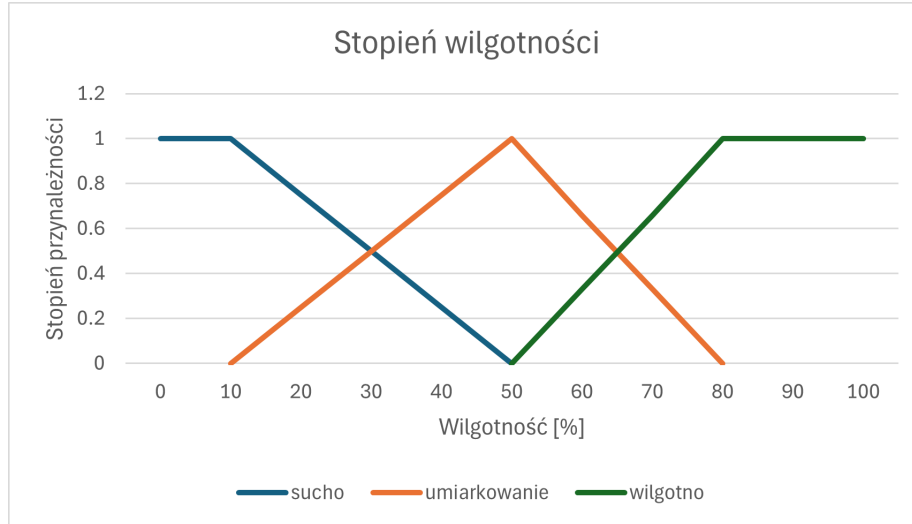
gdzie:  $\mathcal{L}_5$  – wilgotność powietrza,  $H_5 = \{\text{suche, umiarkowane, wilgotne}\}$ ,  
 $\mathcal{X}_5 = [2, 100]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

$$\mu_{\text{sucho}}(x) = \begin{cases} 1, & x \in (0, 10] \\ \frac{50-x}{40}, & x \in (10, 50) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (24)$$

$$\mu_{\text{umiarkowane}}(x) = \begin{cases} \frac{x-10}{40}, & x \in (10, 50) \\ 1, & x = 50 \\ \frac{80-x}{30}, & x \in (50, 80) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (25)$$

$$\mu_{\text{wilgotne}}(x) = \begin{cases} \frac{x-50}{30}, & x \in (50, 80) \\ 1, & x \in [80, 100] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (26)$$



Rysunek 5: Wykres funkcji przynależności dla stopnia wilgotności.

6. visibility\_km

$$L_6 = \langle \mathcal{L}_6, H_6, \mathcal{X}_6 \rangle \quad (27)$$

gdzie:  $\mathcal{L}_6$  – stopień widoczności,  $H_6 = \{\text{słaba, umiarkowana, dobra, bardzo dobra}\}$ ,  $\mathcal{X}_6 = [0, 32]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

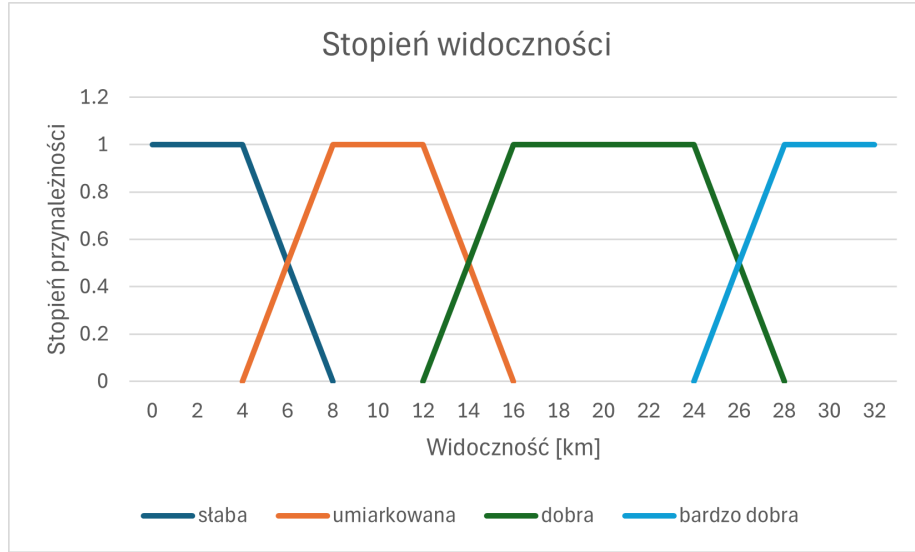
$$\mu_{\text{słaba}}(x) = \begin{cases} 1, & x \in [0, 4] \\ \frac{8-x}{4}, & x \in (4, 8] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (28)$$



$$\mu_{\text{umiarkowana}}(x) = \begin{cases} \frac{x-4}{4}, & x \in (4, 8] \\ 1, & x \in (8, 12] \\ \frac{16-x}{4}, & x \in (12, 16] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (29)$$

$$\mu_{\text{dobra}}(x) = \begin{cases} \frac{x-12}{4}, & x \in (12, 16] \\ 1, & x \in (16, 24] \\ \frac{28-x}{4}, & x \in (24, 28] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (30)$$

$$\mu_{\text{bardzo\_dobra}}(x) = \begin{cases} \frac{x-24}{4}, & x \in (24, 28] \\ 1, & x \in (28, 32] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (31)$$



Rysunek 6: Wykres funkcji przynależności dla stopnia widoczności.

7. uv\_index

$$L_7 = \langle \mathcal{L}_7, H_7, \mathcal{X}_7 \rangle \quad (32)$$

gdzie:  $\mathcal{L}_7$  – promieniowanie UV,  $H_1 = \{\text{niskie, umiarkowane, wysokie, bardzo wysokie, ekstremalne}\}$ ,  $\mathcal{X}_7 = [0, 16]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

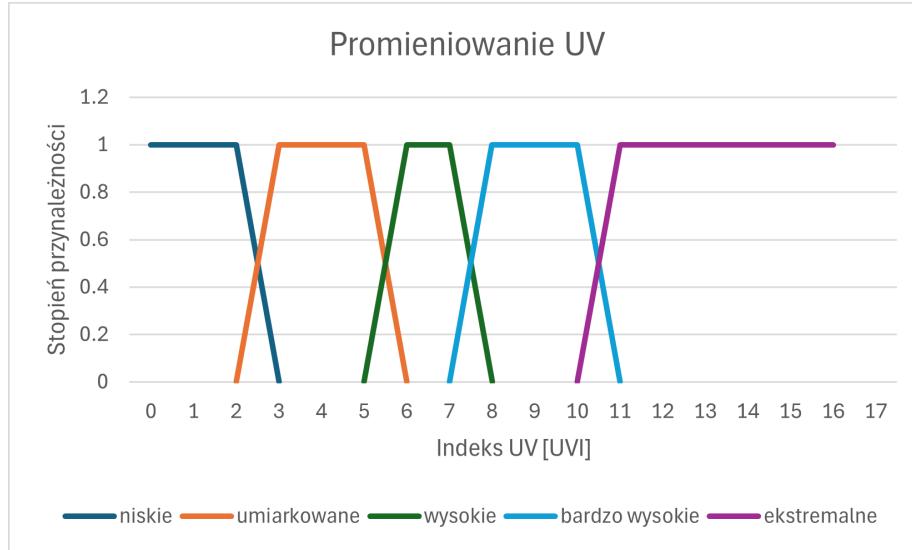
$$\mu_{\text{niskie}}(x) = \begin{cases} 1, & x \in [0, 2] \\ \frac{3-x}{1}, & x \in (2, 3] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (33)$$

$$\mu_{\text{umiarkowane}}(x) = \begin{cases} \frac{x-2}{1}, & x \in (2, 3] \\ 1, & x \in (3, 5] \\ \frac{6-x}{1}, & x \in (5, 6) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (34)$$

$$\mu_{\text{wysokie}}(x) = \begin{cases} \frac{x-5}{1}, & x \in (5, 6] \\ 1, & x \in (6, 7] \\ \frac{8-x}{1}, & x \in (7, 8) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (35)$$

$$\mu_{\text{bardzo\_wysokie}}(x) = \begin{cases} \frac{x-7}{1}, & x \in (7, 8] \\ 1, & x \in (8, 10] \\ \frac{11-x}{1}, & x \in (10, 11) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (36)$$

$$\mu_{\text{ekstremalne}}(x) = \begin{cases} \frac{x-10}{1}, & x \in (10, 11] \\ 1, & x \in (11, 16] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (37)$$



Rysunek 7: Wykres funkcji przynależności dla promieniowania UV.

8. air\_quality\_Carbon\_Monoxide

$$L_8 = \langle \mathcal{L}_8, H_8, \mathcal{X}_8 \rangle \quad (38)$$

gdzie:  $\mathcal{L}_8$  – zanieczyszczenie CO<sub>2</sub>,  $H_8$  – {normalne, niezdrowe, niebezpieczne},  $\mathcal{X}_8 = [0, 2220]$ .

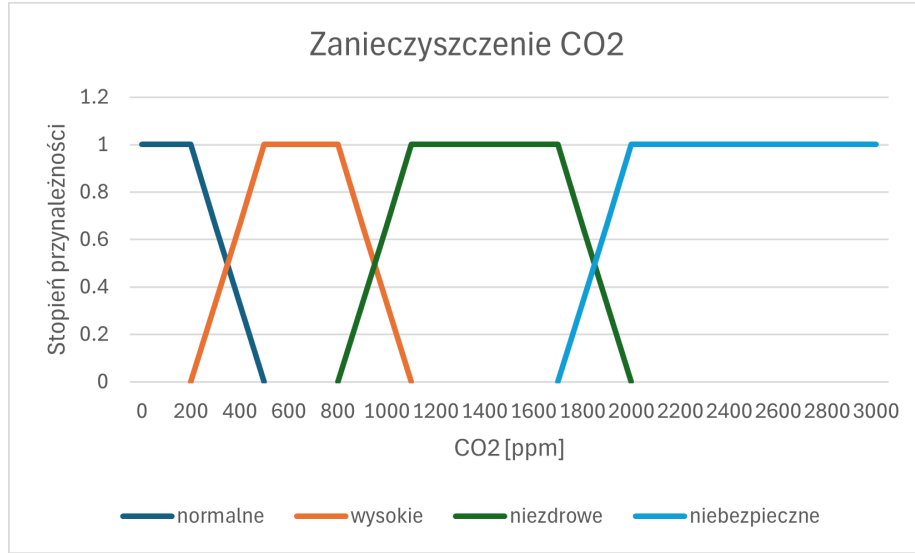
Poniżej wzory dla wszystkich możliwych etykiet.

$$\mu_{\text{normalne}}(x) = \begin{cases} 1, & x \in [0, 200] \\ \frac{500-x}{300}, & x \in (200, 500) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (39)$$

$$\mu_{\text{wysokie}}(x) = \begin{cases} \frac{x-200}{300}, & x \in (200, 500) \\ 1, & x \in [500, 800] \\ \frac{900-x}{300}, & x \in (800, 1100) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (40)$$

$$\mu_{\text{niezdrowe}}(x) = \begin{cases} \frac{x-800}{300}, & x \in (800, 1100) \\ 1, & x \in [1100, 1700] \\ \frac{2000-x}{300}, & x \in (1700, 2000) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (41)$$

$$\mu_{\text{niebezpieczne}}(x) = \begin{cases} \frac{x-1700}{300}, & x \in (1700, 2000) \\ 1, & x \in [2000, 3000] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (42)$$



Rysunek 8: Wykres funkcji przynależności dla zanieczyszczenia CO2.

#### 9. air\_quality\_Nitrogen\_dioxide

$$L_9 = \langle \mathcal{L}_9, H_9, \mathcal{X}_9 \rangle \quad (43)$$

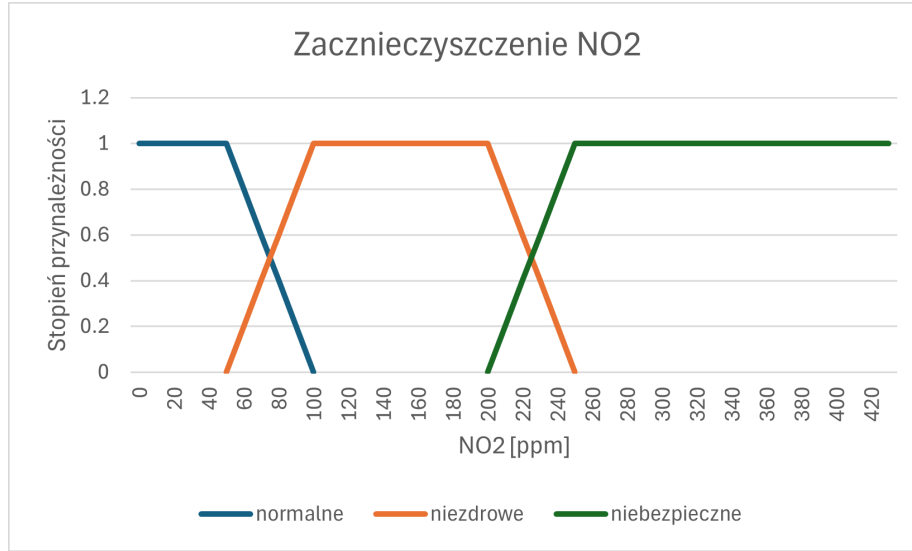
gdzie:  $\mathcal{L}_9$  – zanieczyszczenie NO2,  $H_9$  – {normalne, niezdrowe, niebezpieczne},  $\mathcal{X}_9 = [0, 428]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

$$\mu_{\text{normalne}}(x) = \begin{cases} 1, & x \in [0, 50] \\ \frac{100-x}{50}, & x \in (50, 100] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (44)$$

$$\mu_{\text{niezdrowe}}(x) = \begin{cases} \frac{x-50}{50}, & x \in (50, 100] \\ 1, & x \in (100, 200] \\ \frac{250-x}{50}, & x \in (200, 250) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (45)$$

$$\mu_{\text{niebezpieczne}}(x) = \begin{cases} \frac{x-200}{50}, & x \in (200, 250] \\ 1, & x \in (250, 428] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (46)$$



Rysunek 9: Wykres funkcji przynależności dla zanieczyszczenia no2.

#### 10. Indeks jakości powietrza

$$L_1 0 = \langle \mathcal{L}_{10}, H_{10}, \mathcal{X}_{10} \rangle \quad (47)$$

gdzie:  $\mathcal{L}_{10}$  – jakość powietrza,  $H_{10} = \{\text{bardzo dobra, dobra, umiarkowana, zła, bardzo zła}\}$ ,  $\mathcal{X}_{10} = [1, 10]$ .

Poniżej wzory dla wszystkich możliwych etykiet.

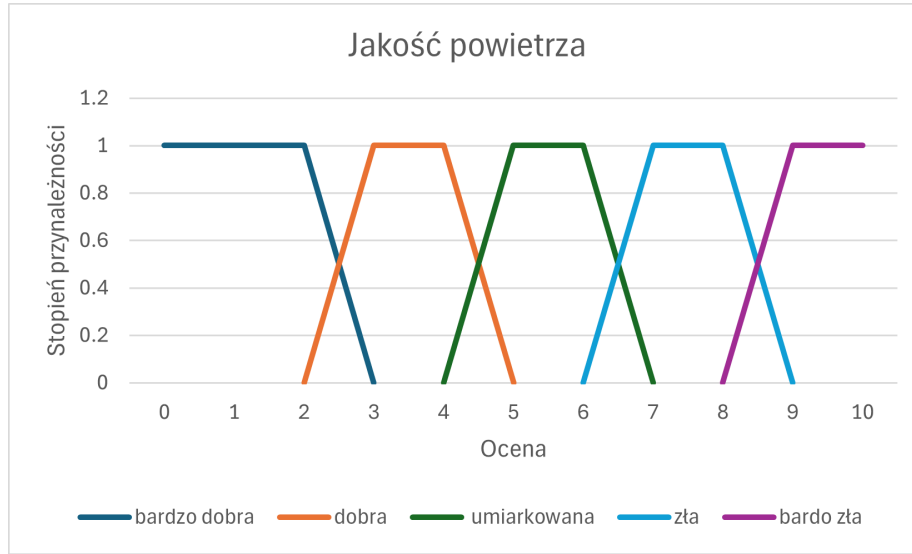
$$\mu_{\text{bardzo\_dobra}}(x) = \begin{cases} 1, & x \in [1, 2] \\ \frac{3-x}{1}, & x \in (2, 3] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (48)$$

$$\mu_{\text{dobra}}(x) = \begin{cases} \frac{x-2}{1}, & x \in (2, 3] \\ 1, & x \in (3, 4] \\ \frac{5-x}{1}, & x \in (4, 5] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (49)$$

$$\mu_{\text{umiarkowana}}(x) = \begin{cases} \frac{x-4}{1}, & x \in (4, 5] \\ 1, & x \in (5, 6] \\ \frac{7-x}{1}, & x \in (6, 7] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (50)$$

$$\mu_{\text{zła}}(x) = \begin{cases} \frac{x-6}{1}, & x \in (6, 7] \\ 1, & x \in (7, 8] \\ \frac{9-x}{1}, & x \in (8, 9) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (51)$$

$$\mu_{\text{bardzo zła}}(x) = \begin{cases} \frac{x-8}{1}, & x \in (8, 9] \\ 1, & x \in (9, 10] \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (52)$$



Rysunek 10: Wykres funkcji przynależności dla jakości powietrza.

### 2.3 Kwantyfikatory lingwistyczne (liczności obiektów)

Poniżej zostały zaprezentowane kwantyfikatory lingwistyczne wraz z ich wzorami analitycznymi oraz wykresem funkcji przynależności.

- prawie żaden

$$\mu_{\text{prawie\_żaden}}(x) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{x-0,0}{0,06}\right)^2\right), & x \in (0, 0.2) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (53)$$

- trochę

$$\mu_{\text{trochę}}(x) = \begin{cases} \frac{x-0.10}{0.15}, & x \in (0.10, 0.25) \\ 1, & x \in [0.25, 0.30] \\ \frac{0.45-x}{0.15}, & x \in (0.30, 0.45) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (54)$$

- około połowa

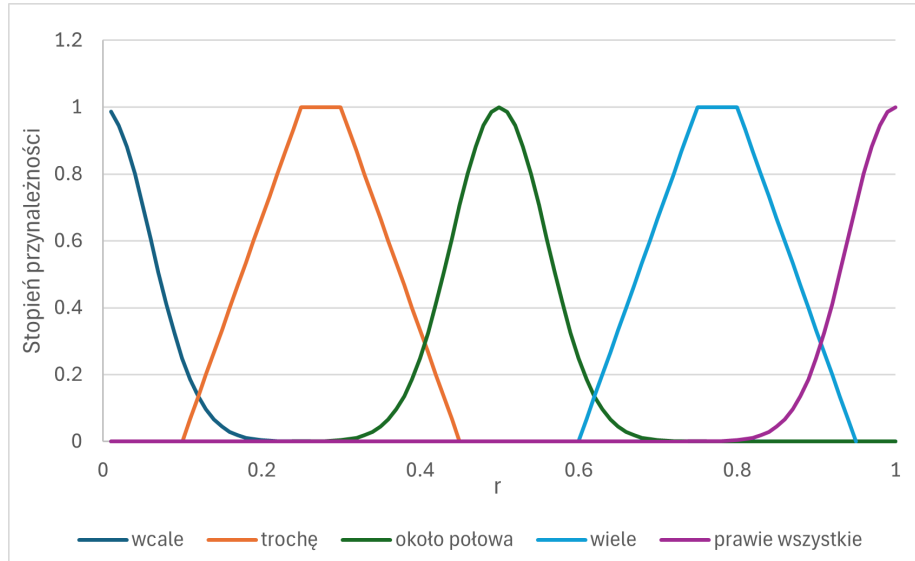
$$\mu_{\text{około połowa}}(x) = \begin{cases} \exp\left(-\frac{1}{2} \left(\frac{x-0.5}{0.06}\right)^2\right), & x \in (0.35, 0.75) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (55)$$

- wiele

$$\mu_{\text{wiele}}(x) = \begin{cases} \frac{x-0.60}{0.15}, & x \in (0.60, 0.75) \\ 1, & x \in [0.75, 0.80] \\ \frac{0.95-x}{0.15}, & x \in (0.80, 0.95) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (56)$$

- prawie wszystkie

$$\mu_{\text{prawie wszystkie}}(x) = \begin{cases} \exp\left(-\frac{1}{2} \left(\frac{x-1.0}{0.06}\right)^2\right), & x \in (0.8, 1.0) \\ 0, & \text{w przeciwnym razie} \end{cases} \quad (57)$$

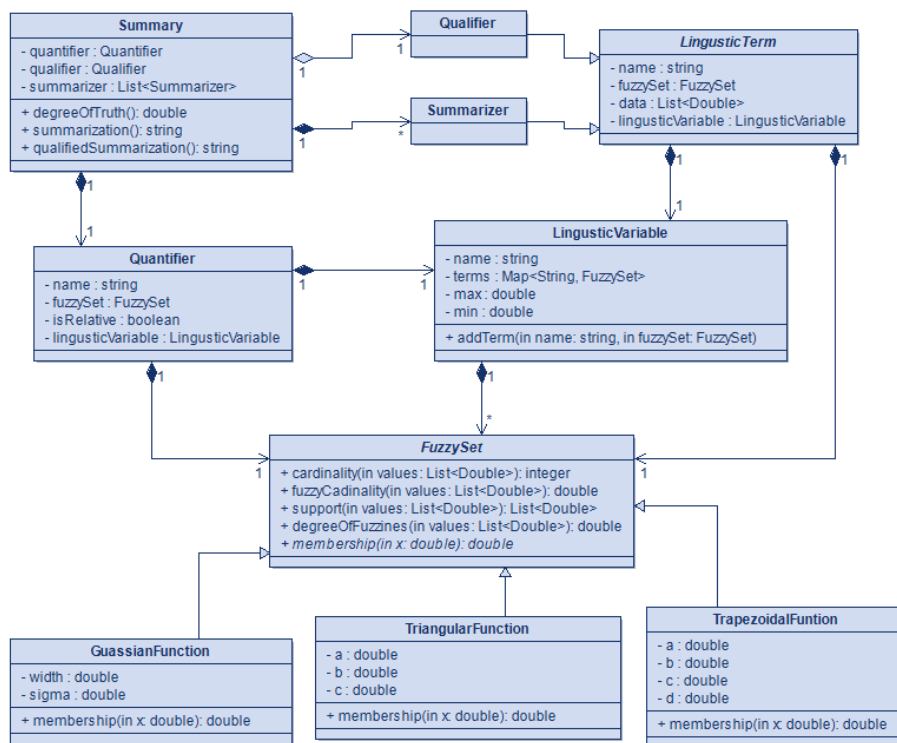


Rysunek 11: Funkcja przynależności dla wybranych kwantyfikatorów lingwistycznych.

### 3 Narzędzia obliczeniowe: wybór/implementacja. Diagram UML klas do obliczeń rozmytych i generowania podsumowań

Program został napisany w języku Java w wersji JDK 24. Do obsługi relacyjnej bazy danych PostgreSQL wykorzystano framework Spring Boot, który umożliwia wygodną konfigurację połączenia z bazą danych. Kod źródłowy komponentu odpowiedzialnego za obliczenia rozmyte został umieszczony w pakiecie o nazwie **fuzzy**. Znajduje się w nim podstawowy model reprezentujący zbiory rozmyte, którego centralnym elementem jest abstrakcyjna klasa **FuzzySet**, która definiuje fundamentalne operacje na zbiorach rozmytych. Klasa ta udostępnia metody pozwalające na obliczanie podstawowych właściwości zbioru rozmytego, m.in. takich jak kardynalność oraz nośnik. Na bazie klasy **FuzzySet** zbudowane zostały trzy klasy reprezentujące konkretne typy funkcji przynależności: **GaussianFunction**, **TrapezoidalFunction** oraz **TriangularFunction**. Każda z tych klas implementuje matematyczną funkcję przynależności o charakterystycznym kształcie: odpowiednio gaussowskim, trapezoidalnym i trójkątnym. Dzięki temu możliwe jest precyzyjne modelowanie różnych pojęć językowych. Oprócz modelu zbiorów rozmytych, pakiet **fuzzy** zawiera również klasy odpowiedzialne za tworzenie i przetwarzanie podsumowań lingwistycznych. Klasa **LinguisticVariable** reprezentuje zmienną lingwistyczną, która zawiera nazwę, przestrzeń rozważań oraz zestaw powiązanych z nią terminów lingwistycznych, np. „bardzo zimno”, „ciepło”, „gorąco”, które są reprezentowane przez odpowiednie zbiory rozmyte. W strukturze podsumowań lingwistycznych wyróżniamy także kwalifikatory i sumaryzatory, które mają bardzo podobną strukturę, ponieważ są złożone z nazwy oraz przypisanego im zbioru rozmytego. Ich różnice wynikają jedynie z zastosowania, w jaki sposób są wykorzystywane w podsumowaniach lingwistycznych. W celu zachowania przejrzystości i uniknięcia duplikacji kodu, stworzona została wspólna klasa bazowa **LinguisticTerm**, z której dziedziczą klasy **Qualifier** oraz **Summarizer**. Pozwala to na zachowanie semantycznego rozróżnienia przy jednoczesnym wykorzystaniu wspólnej logiki przetwarzania. Klasa **Quantifier** odpowiada za reprezentację kwantyfikatorów, takich jak „około połowy” czy „prawie żaden”. Kwantyfikator opisany jest za pomocą nazwy, funkcji przynależności oraz typu — może być kwantyfikatorem absolutnym lub względnym. Ostatnim elementem pakietu jest klasa **Summary**. To ona odpowiada za generowanie podsumowań lingwistycznych na podstawie przekazanych obiektów: sumaryzatora, kwalifikatora oraz kwantyfikatora. Strukturę logiczną opisanych klas oraz zależności między nimi przedstawiono na poniższym diagramie UML.





Rysunek 12: Diagram UML pakietu fuzzy.

## 4 Jednopoziomowe podsumowania lingwistyczne. Miary jakości, podsumowanie optymalne

Wyniki kolejnych eksperymentów wg punktów 2.-4. opisu projektu 2. Listy podsumowań jednopoziomowych i tabele/rankingi podsumowań dla danych atrybutów obowiązkowe i dokładnie opisane w „captions” (tytułach), konieczny opis kolumn i wierszy tabel. Dla każdego podsumowania podane miary jakości oraz miara jakości podsumowania optymalnego. **Wzorów podsumowań ani miar nie należy przepisywać ani cytować, wystarczy podać literaturę, ale należy skomentować co oznaczają i jaką informację niosą wybrane miary w wybranych przypadkach.**

Sekcja uzupełniona jako efekt zadania Tydzień 11 wg Harmonogramu Zajęć na WIKAMP KSR.

## 5 Wielopodmiotowe podsumowania lingwistyczne i ich miary jakości

Wyniki kolejnych eksperymentów wg punktów 2.-4. opisu projektu 2. Uzasadnienie i metoda podziału zbioru danych na rozłączne podmioty. Listy podsumowań wielopodmiotowych i tabele/rankingi podsumowań dla danych atrybutów obowiązkowe i dokładnie opisane w „captions” (tytułach), konieczny opis kolumn i wierszy tabel. **Wzorów podsumowań ani miar nie należy przepisywać ani cytować, wystarczy podać literaturę, ale należy skomentować co oznaczają i jaką informacje niosą wybrane miary w wybranych przypadkach.** Konieczne uwzględnienie wszystkich 4-ch form podsumowań wielopodmiotowych.

\*\* Możliwe sformułowanie zagadnienia wielopodmiotowego podsumowania optymalnego \*\*.

\*\* Ewentualne wyniki realizacji punktu „na ocenę 5.0” wg opisu Projektu 2. i ich porównanie do wyników z części obowiązkowej \*\*.

**Sekcja uzupełniona jako efekt zadania Tydzień 12 wg Harmonogramu Zajęć na WIKAMP KSR.**

## 6 Dyskusja, wnioski

Dokładne interpretacje uzyskanych wyników w zależności od parametrów klasyfikacji opisanych w punktach 3.-4 opisu Projektu 2. Omówić i wyjaśnić napotkane problemy (jeśli były). Każdy wniosek/problem powinien mieć poparcie w przeprowadzonych eksperymentach (odwołania do konkretnych wyników: tabel i miar jakości). Ocena które podsumowania i dlaczego niosą najistotniejsze informacje i które ich miary jakości mają małe albo duże znaczenie dla wiarygodności i jakości otrzymanych agregacji/podsumowań.

Dla końcowej oceny jest to najważniejsza sekcja sprawozdania, gdyż prezentuje poziom zrozumienia rozwiązywanego problemu.

\*\* Możliwości kontynuacji prac w obszarze logiki rozmytej i wnioskowania rozmytego, zwłaszcza w kontekście pracy inżynierskiej, magisterskiej, naukowej, itp. \*\*

**Sekcja uzupełniona jako efekt zadań Tydzień 11 i Tydzień 12 wg Harmonogramu Zajęć na WIKAMP KSR.**

## 7 Braki w realizacji projektu 2.

Wymienić wg opisu Projektu 2. wszystkie niezrealizowane obowiązkowe elementy projektu, ewentualnie podać merytoryczne (ale nie czasowe) przyczyny

tych braków.

## Literatura

- [1] World Weather Repository - kaggle, <https://www.kaggle.com/datasets/nelgiriyeewithana/global-weather-repository?resource=download>. [dostęp 18.05.2025r.]
- [2] A. Niewiadomski, Zbiory rozmyte typu 2. Zastosowania w reprezentowaniu informacji. Seria „Problemy współczesnej informatyki” pod redakcją L. Rutkowskiego. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2019.
- [3] S. Zadrozny, Zapytania nieprecyzyjne i lingwistyczne podsumowania baz danych, EXIT, 2006, Warszawa
- [4] A. Niewiadomski, Methods for the Linguistic Summarization of Data: Applications of Fuzzy Sets and Their Extensions, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2008.

Literatura zawiera wyłącznie źródła recenzowane i/lub o potwierdzonej wiarygodności, możliwe do weryfikacji i cytowane w sprawozdaniu.