

DOKUMENTACJA APLIKACJI WSPOMAGAJĄCEJ WYBÓR KREDYTÓW MIESZKANIOWYCH

Przygotowali:

-Jakub Sarata
-Jan Szlagór
-Konrad Kropornicki
-Mateusz Zagórski
-Marcin Biela
-Michał Spinczyk
-Michał Święciło
-Grzegorz Penter

Link do repozytorium z kodem:

https://github.com/Marcin279/home_loans_choice

1. WPROWADZENIE

Głównym założeniem podczas pisania programu było stworzenie aplikacji wspomagającej decyzje wyboru konkretnego kredytu mieszkaniowego. Jej działanie sprowadza się do tworzenia rankingu kredytów na podstawie 2 lub 3 kryteriów. Użytkownik po uruchomieniu aplikacji przy pomocy specjalnych pól zaznacza jakie kryteria mają być brane pod uwagę oraz wybiera metodę, która ma być wykorzystana do stworzenia rankingu. Do wyboru ma 6 kryteriów oraz 4 metody. Po zakończeniu obliczeń program wyświetla utworzony ranking kredytów.

2. DANE

Do działania programu niezbędna jest baza danych. Stworzona jest ona w postaci arkusza programu Excel, w którym kolejne kolumny odpowiadają poszczególnym kryteriom decyzyjnym, a wiersze to kolejne oferty kredytowe. Kryteria decyzyjne zostały dobrane w taki sposób, aby nie było zbyt dużej korelacji między nimi. Użytkownik ma do wyboru 6 takich kryteriów:

- marża (dąży do minimum)
- prowizja (dąży do minimum)
- RRSO (dąży do minimum)
- koszt miesięczny (dąży do minimum)
- wkład własny (dąży do minimum)
- opinie (dąży do maximum)

Zaletą bazy danych w postaci pliku Excel jest to, że oferty kredytowe mogą być aktualizowane oraz można dodawać kolejne rekordy.

Punkt	Bank, data	Marża [%]	Prowizja [%]	RRSO [%]	Koszt mies	Wkład wła	Opinie[pkt. Max. 5]
N1	mBank, 12.2021	1,65	0,00	4,24	2066,00	20,00	4,90

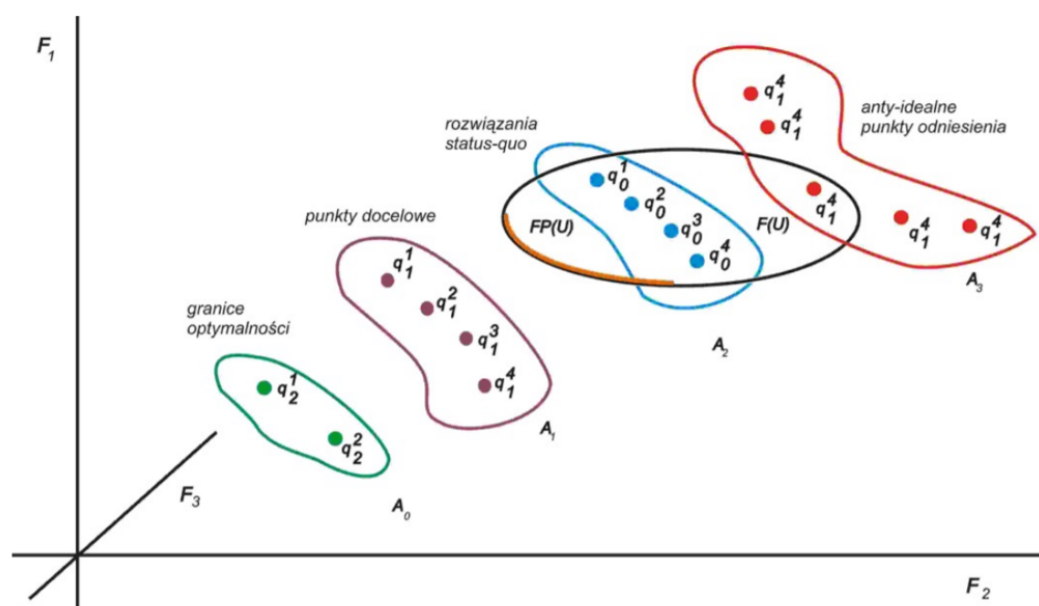
Rys.1. Pojedynczy rekord z bazy danych pokazujący kryteria.

3. INTERFEJS UŻYTKOWNIKA

W celu ułatwienia korzystania z aplikacji została ona wyposażona w intuicyjny interfejs. Pozwala on na wybór rozważanych kryteriów oraz algorytmu, na podstawie którego ma zostać stworzony ranking. Możliwość wyboru kryteriów została zablokowana do trzech ponieważ dla wyższej ilości kryteriów nie można zwizualizować wyników. Wiąże się to z faktem, że każde kolejne kryterium to dodatkowy wymiar na wykresie, stąd wybór ograniczono do wykresów trójwymiarowych.

4. WYZNACZENIE KLAS PUNKTÓW I MACIERZY POTENCJALNYCH DECYZJI

Dla metody topsis utworzono algorytm klasyfikujący punkty odniesienia do zbiorów A_0 , A_1 , A_2 , A_3 oraz zbioru alternatyw do utworzenia ranking. Początkowo wyznaczono zbiory A_0 i A_3 jako zbiory punktów niezdominowanych: A_0 - najlepszych, A_3 - najgorszych. Kolejne zbiory A_1 i A_2 wyznaczono jako zbiory punktów nieporównywalnych w stosunku do preferencji podanych przez użytkownika. Zbiór A_1 został wyznaczony na podstawie preferencji będącej punktem docelowym lecz niemożliwym do osiągnięcia dla klienta. Z kolei zbiór A_2 na podstawie preferencji będącej punktem statusu quo dla decydenta.



Wykres 1.: Wykres obrazujący ide i sposób tworzenia klas punktów.

Następnie sprawdzono warunki zewnętrznej i wewnętrznej niesprzeczności dla zbiorów A_0 , A_1 , A_2 , A_3 i poprawiono je tak aby zostały one spełnione:

- warunek wewnętrznej niesprzeczności:

$\forall q_1, q_2 \in A_i$ q_1 oraz q_2 są nieporównywalne,

- warunek zewnętrznej niesprzeczności:

$$\forall x \in A_j \exists y \in A_{j+1} : x \leq_0 y$$

$$\forall y \in A_{j+1} \exists x \in A_j : x \leq_0 y$$

Ostatnim etapem było wyznaczenie zbioru potencjalnych wyborów znajdujących się pomiędzy zbiorami A_1 i A_2 .

5. WYKORZYSTANE ALGORYTMY WSPOMAGAJĄCE DECYZJE

Do wyznaczania rankingów aplikacja używa 4 algorytmów na poszukiwanie rozwiązań problemu optymalizacji wielokryterialnej. Są to metody SP-CS, RSM, TOPSIS i UTA Star.

a. SAFETY PRINCIPLES (SP-CS)

Metoda ta do działania wykorzystuje krzywe Woronoja (krzywe szkieletowe). Przy pomocy krzywych łączy się górne i dolne punkty odniesienia. Punkty załamania na krzywej wyznaczone są na podstawie współrzędnych punktów odniesienia, dla których rysowana jest dana krzywa. Następnie rozważany punkt u rzutujemy na każdą z krzywych na najkrótszej możliwej drodze, a współczynnik skoringowy jest sumą współczynników rzutów na wszystkie krzywe. Współczynniki na krzywych Woronoja normalizujemy w taki sposób, aby dla punktu ze zbioru A_1 (dolny punkt odniesienia) wartość wynosiła 0, a dla punktu ze zbioru A_2 (górny punkt odniesienia) wartość była równa 1.

b. METODA ZBIORÓW ODNIESIENIA (RSM)

W metodzie RSM tworzy się **prostokąty** (dla dwóch kryteriów), **prostopadłościany** (dla 3 kryteriów), których jeden wierzchołek leży w zbiorze punktów docelowych A_1 , a drugi w zbiorze A_2 . Pole tak powstałego prostokąta/prostopadłościanu jest współczynnikiem wagowym. Tworzymy tyle takich prostokątów ile jest możliwych kombinacji punktów z A_1 z punktami z A_2 .

Następnie dla każdego punktu u (znajdującego się pomiędzy zbiorami A_1 i A_2) obliczamy odległości od danych punktów będących wierzchołkami prostokąta i przemnażamy je przez współczynnik wagowy danego prostokąta będący znormalizowaną wartością jego pola powierzchni (objętości). Tak czynimy kolejno dla każdego prostokąta, a współczynnik skoringowy jest sumą wszystkich takich mnożeń. Jeżeli dany punkt nie zawiera się w prostokącie/prostopadłościanie to współczynnik wagowy przyjmujemy 0. Współczynniki wagowe dla danego punktu u muszą być znormalizowane tak, aby ich suma wynosiła 1.

c. METODA FUZZY-TOPSIS

Metoda topsis w pierwszej kolejności na podstawie zbioru potencjalnych decyzji, punktu idealnego ze zbioru A1 oraz punktu nadir skaluje wartości kryteriów według normy euklidesowej (wzór podany poniżej), aby były równoważne w percepcji decydenta.

$$\|x\|_2 \stackrel{df}{=} \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}}$$

Podczas skalowania uwzględnia wagi kryteriów (domyślnie wagi wynoszą 1 ale implementacja metody pozwala na wprowadzenie wag) i zamienia kryteria dążące do maksimum na takie które chcemy minimalizować - po skalowaniu wszystkie kryteria minimalizujemy.

Kolejnym krokiem jest obliczenie odległości od punktów reprezentujących oceniane obiekty do punktu idealnego ze zbioru A1 oraz nadir w normie euklidesowej (im mniejsza odległość od idealnego A1 i większa od nadir tym lepiej).

Na podstawie tych odległości tworzony jest końcowy ranking.

d. METODA UTA STAR

Metoda UTA Star służy porządkowaniu wariantów decyzyjnych z rozważanego zbioru wariantów (czyli rozważana jest problematyka – tworzenie rankingu). Model preferencji decydenta jest ponownie funkcją użyteczności (o postaci addytywnej), która umożliwia uporządkowanie zbioru od wariantu o największej użyteczności globalnej do wariantu o najmniejszej użyteczności globalnej. W UTA Star:

- Globalna użyteczność wariantu (wektor ocen na poszczególnych kryteriach), oznaczana przez $U(x)$, zależy od jego użyteczności częściowych $u_i(x_i)$, (gdzie n – liczba kryteriów). Konkretnie, jest ich sumą (model addytywny):

$$U(x) = \sum_{i=1}^n u_i(x_i)$$

- Funkcje wyznaczane są niezależnie dla poszczególnych kryteriów w oparciu o przyjmowane założenie o niezależności kryteriów w sensie preferencji

$$g_i, \quad i = 1, \dots, n,$$

- Globalna funkcja użyteczności jest znormalizowana tak, by przyjmować wartości z przedziału od 0 do 1.

- Wartość funkcji $U(x)$ (x należy do A) określa pozycję wariantu decyzyjnego w rankingu końcowym na zbiorze A . I tak, wariant decyzyjny z najwyższą użytecznością globalną zajmuje miejsce pierwsze (najwyższe), ..., wariant decyzyjny z najniższą użytecznością globalną zajmuje miejsce ostatnie (najniższe).

- Dla najgorszej wartości kryterium zachodzi:

$$u_i(g_{i*}) = 0$$

Dzięki temu, wariant antyidealny ma użyteczność globalną równą 0 (suma samych 0).

- Suma użyteczności częściowych dla najlepszych wartości kryteriów jest równa 1:

$$\sum_{i=1}^n u_i(g_i^*) = 1$$

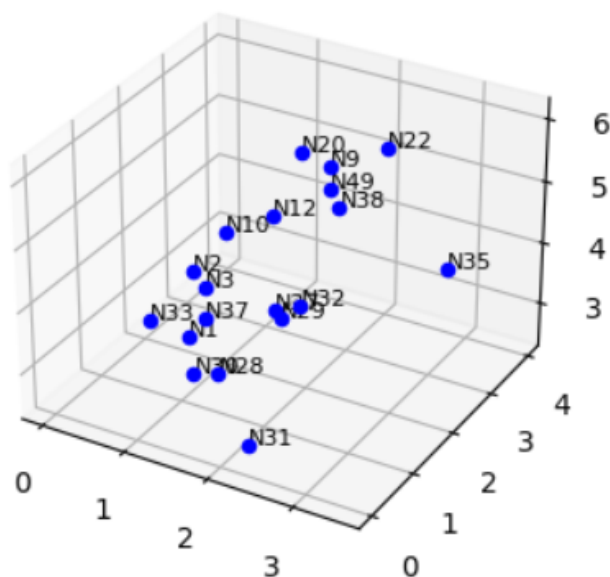
Dzięki temu, wariant idealny (sztuczny wariant o najlepszych możliwych wartościach na wszystkich kryteriach) ma użyteczność globalną równą 1.

- Funkcje użyteczności częściowych są odcinkami liniowe i monotoniczne (niemalejące dla kryteriów typu zysk, nierosnące dla kryteriów typu koszt). Pierwszy punkt charakterystyczny funkcji przypada w najmniejszej wartości kryterium (min), a ostatni punkt charakterystyczny w największej wartości kryterium (max).
- Metoda UTA Star wykorzystuje programowanie liniowe (PL) do wyznaczenia współrzędnych punktów charakterystycznych funkcji użyteczności częściowych.

6. PORÓWNANIE METOD

Dla 3 kryteriów (marży, prowizji i RRSO) utworzyliśmy rankingi, wykorzystując wszystkie zaimplementowane metody. Dane, które przyjmują algorytmy zaprezentowaliśmy na wykresach, a zwrócone wyniki w tabelach.

Metoda Fuzzy TOPSIS

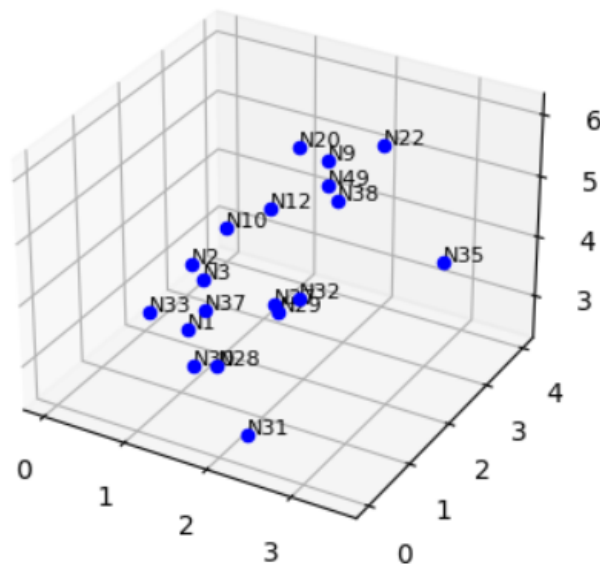


Rys.2. Wykres 3D prezentujący dane wejściowe metody Fuzzy TOPSIS.

Tabela 1. Ranking zwrócony przez metodę Fuzzy Topsis.

nr decyzji w bazie danych (od najlepszego)	Marża [%]	Prowizja [%]	RRSO [%]
12	1,70	1,8	5,11
32	1,99	1,89	3,7
27	1,89	1,5	3,82
29	1,95	1,5	3,72
37	1,85	0,0	4,62
3	1,84	0,0	5,09
2	1,70	0,0	5,28
28	1,99	0,0	3,79

Metoda SP-CS



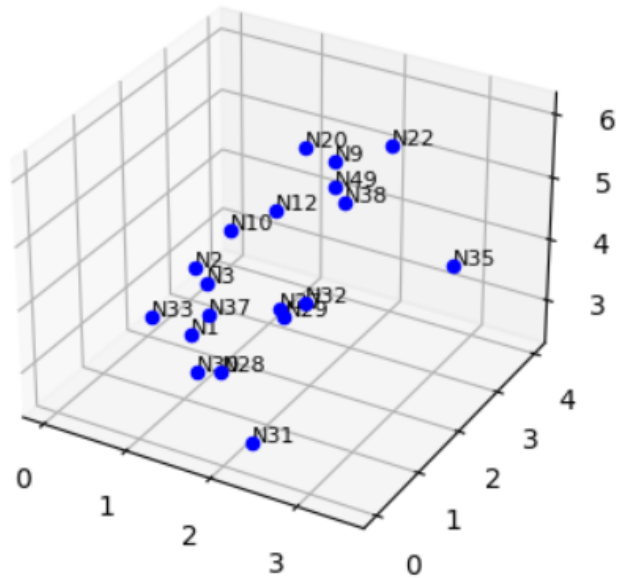
Rys.3. Wykres 3D prezentujący dane wejściowe metody SP-CS.

Tabela 2. Ranking zwrócony przez metodę SP-CS.

nr decyzji w bazie danych (od najlepszego)	Marża [%]	Prowizja [%]	RRSO [%]
33	0,00	2,1	2,54
31	2,33	0,0	2,8
30	1,70	0,0	3,68
1	1,65	0,0	4,24
27	1,89	1,5	3,82
29	1,95	1,5	3,72
37	1,85	0,0	4,62
3	1,84	0,0	5,09
2	1,70	0,0	5,28
32	1,99	1,89	3,7

28	1,99	0,0	3,79
12	1,70	1,8	5,11
49	2,29	2,0	5,61
38	2,39	2,0	5,38
35	3,50	2,29	4,66
22	2,00	4,0	5,05
20	1,94	2,0	6,08
10	2,10	0,0	6,02
9	2,19	2,2	5,83

Metoda RSM



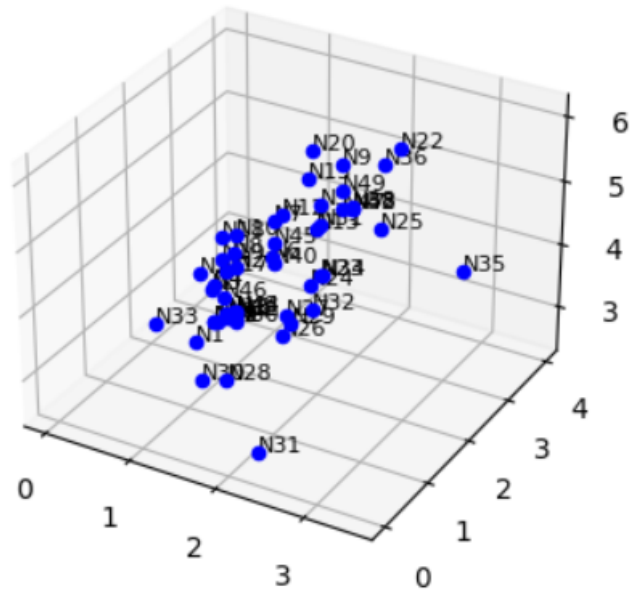
Rys.4. Wykres 3D prezentujący dane wejściowe metody RSM.

Tabela 3. Ranking zwrócony przez metodę RSM.

nr decyzji w bazie danych (od najlepszego)	Marża [%]	Prowizja [%]	RRSO [%]
33	0,00	2,1	2,54
31	2,33	0,0	2,8
30	1,70	0,0	3,68
1	1,65	0,0	4,24
37	1,85	0,0	4,62
32	1,99	1,89	3,7
29	1,95	1,5	3,72
27	1,89	1,5	3,82
12	1,70	1,8	5,11
28	1,99	0,0	3,79
3	1,84	0,0	5,09

2	1,70	0,0	5,28
49	2,29	2,0	5,61
38	2,39	2,0	5,38
35	3,50	2,29	4,66
22	2,00	4,0	5,05
20	1,94	2,0	6,08
10	2,10	0,0	6,02
9	2,19	2,2	5,83

Metoda UTA Star



Rys.5. Wykres 3D prezentujący dane wejściowe metody UTA Star.

Tabela 4. Ranking (Najlepsze 19 elementów) zwrócony przez metodę UTA Star.

nr decyzji w bazie danych (od najlepszego)	Marża [%]	Prowizja [%]	RRSO [%]
33	0,00	2,1	2,54
31	2,33	0,0	2,8
30	1,70	0,0	3,68
28	1,99	0,0	3,79

1	1,65	0,0	4,24
37	1,85	0,0	4,62
39	1,85	0,0	4,62
41	1,89	0,0	4,64
14	2,00	0,0	4,75
42	2,00	0,0	4,77
16	2,00	0,0	4,8
50	2,10	0,0	4,73
26	2,10	1,0	3,88
11	2,09	0,0	4,8
3	1,84	0,0	5,09
43	2,09	0,0	4,86
44	2,10	0,0	4,87
2	1,70	0,0	5,28
29	1,95	1,5	3,72

7. PORÓWNANIE METOD

Otrzymane wyniki przetworzyliśmy w Excel, tak aby móc porównać rankingi poszczególnych metod współczynnikiem korelacji rankingowej Spearmana według wzoru:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

gdzie d_i to różnica między parą punktów na pozycji rankingowej, a n to liczba pozycji w rankingu.

Wyliczone współczynniki wstawiliśmy do tabeli poniżej.

Tabela 5. Tabela współczynników korelacji rankingowej Spearmana.

	TOPSIS	SPCS	RSM	UTA Star
TOPSIS	1	-	-	-
SPCS	-0,44697	1	-	-
RSM	-0,15999	0,767994	1	-
UTA Star	-0,59908	-0,07701	-0,11248	1

Należy wyjaśnić, że zaimplementowane metody Topsis, SPCS i RSM działają na podstawie tworzenia punktów odniesienia i nie zwracają pełnego rankingu decyzji, tak jak robi to metoda UTA Star.

Ze względu na współczynniki korelacji można powiedzieć, że największym podobieństwem cechują się wyniki zwrócone przez metody RSM i SP-CS, zaś najmniejsza korelacja występuje między FUZZY-TOPSIS i UTA Star. Każda metoda ma swoją charakterystykę i działa na innej zasadzie dlatego otrzymane rankingi różnią się między sobą. Błędy mogą również wynikać z przybliżeń jakich dokonujemy po drodze.

8. Instrukcja instalacji

Instrukcja odnośnie instalacji i włączania aplikacji znajduje się w pliku README.md w repozytorium:

https://github.com/Marcin279/home_loans_choice/blob/main/README.md

7. LITERATURA I ŹRÓDŁA WIEDZY

1. Dokumentacja oraz instrukcje zawarte na kursie przedmiotu ze strony UPEL.
2. Skulimowski A.M.J. (1996) Decision Support Systems Based on Reference Sets. Wydawnictwa AGH, Monografie, Nr 40, s. 167
3. Skulimowski A.M.J. (2019). Selected methods, applications, and challenges of multicriteria optimization. Seria Monografie, t.19, Komitet Automatyki i Robotyki Polskiej Akademii Nauk, Wydawnictwa AGH, s. 380
4. Andrzej M.J. SKULIMOWSKI (1986). Foreseen Utility in Multi-Stage Multicriteria Optimization. Seminar on Nonconventional Problems of Optimization, Warszawa, May 9-11, 1984. W: J. Kacprzyk (ed.), Proceedings, Part III, s. 365-386.

Tabela 6. Tabela określająca udział każdego członka grupy w realizacji ćwiczenia

Wybór problemu	Cały zespół
Stworzenie bazy danych	Konrad Kropornicki
Utworzenie GUI	Marcin Biela, Michał Świącilo
Przygotowanie danych	Jan Szlagór, Mateusz Zagórski
Zaimplementowanie metody Fuzzy Topsis	Jan Szlagór, Mateusz Zagórski
Implementacja metody SP-CS	Jakub Sarata, Michał Spinczyk
Implementacja metody zbiorów odniesienia (RSM)	Jakub Sarata, Marcin Biela, Michał Spinczyk
Implementacja metody UTA Star	Konrad Kropornicki, Michał Świącilo, Jan Szlagór, Grzegorz Penter
Porównanie metod	Grzegorz Penter
Koordinacja, redakcja wniosków końcowych, edycja sprawozdania	Cały zespół