

Analiza porównawcza DEA-CCR w przemyśle młynarskim



Analiza porównawcza DEA-CCR (ang. *Data Envelopment Analysis – Charnes, Cooper, Rhodes*) to metoda stosowana do oceny efektywności jednostek decyzyjnych (ang. *Decision-Making Units, DMUs*), takich jak przedsiębiorstwa, instytucje czy oddziały, które przetwarzają wejścia (np. zasoby, nakłady) na wyjścia (np. produkty, usługi).

DEA-CCR pozwala na ocenę względnej efektywności jednostek decyzyjnych poprzez określenie granicy efektywności, na której znajdują się najbardziej efektywne jednostki. Jednostki poniżej tej granicy są uważane za mniej efektywne. Metoda ta identyfikuje również stopień nieefektywności i sugeruje możliwości poprawy poprzez zmniejszenie nakładów lub zwiększenie wyników przy niezmiennych zasobach.

DEA-CCR jest szczególnie użyteczna w sytuacjach, gdzie trudno jest porównać różne jednostki na podstawie tradycyjnych wskaźników efektywności,

umożliwiając ocenę efektywności w warunkach wielowymiarowych danych wejściowych i wyjściowych.

Porównać nieporównywalne

Załóżmy, że jesteśmy właścicielami kilkudziesięciu młynów i musimy stworzyć system premiowy dla pracowników obsługujących wagi. Przed każdym młynem znajduje się waga, ale czas ważenia różni się w zależności od infrastruktury, wielkości ważonych pojazdów, a nawet rodzaju ważonego surowca. Na niektórych wagach formują się kolejki, a inne stoją puste. Właściciel sieci młynów stoi przed bardzo trudnym zadaniem: jak sprawiedliwie przydzielić premie? Jakie kryteria zastosować? Przecież wagi są różne, nie można porównywać gruszek z jabłkami.

Rozważmy, że nasze młyny znajdują się w różnych miejscach w kraju. W jednym regionie mamy wielkie latyfundia, a w innym – liczne, małe gospodarstwa.

Każdy młyn ma swoją unikalną historię i wynikającą z tego odmienną technologię i różną liczbę pracowników biurowych. Niektóre urządzenia psują się częściej, inne są nowe, ale ich leasing kosztuje dużo. W jaki sposób ocenić, który młyn działa efektywnie, a który nie, skoro każdy jest inny? Czy możemy porównywać jabłka z pomarańczami?

Metoda porównawcza „jabłek z gruszkami”?

Na metodę DEA-CCR natknąłem się kilka lat temu i od razu zrobiła na mnie ogromne wrażenie. Pracowałem przez 15 lat jako analityk finansowy w dużych organizacjach. Jednym z moich obowiązków było porównywanie efektywności różnych jednostek, których mogło być kilkadziesiąt lub kilkanaście. Były to hotele, sklepy, bazy logistyczne, spedycje czy zakłady produkcyjne. Każda jednostka była inna, każda działała w odmiennych warunkach.



Aby je porównać musiałem posługiwać się standardowymi wskaźnikami oraz przypisywać do nich odpowiednie wagi. Np., gdy porównywałem dwie bazy logistyczne, z których jedna miała kilkanaście samochodów, a druga kilkadziesiąt, jasne było, że mniejsza baza miała większy udział kosztów stałych i wyższy procentowy udział kosztów administracji. Starałem się wyrównywać obie jednostki poprzez stosowanie wag, które sam określałem. Jednak te wagi były wyznaczane subiektywnie, co sprawiało, że moja metoda porównawcza miała wady. Mimo to trzeba było pracować i przedstawiać rankingi efektywności, nawet jeśli miałem świadomość, że metoda była niedoskonała.

Dopiero długo po zakończeniu pracy jako analityk finansowy odkryłem metodę DEA.

Metoda **Data Envelopment Analysis (DEA)** powstała, aby zaspokoić podstawowe potrzeby pomiaru efektywności organizacji, w których nakłady i efekty mają wiele wymiarów, i jest wykorzystywana do optymalizacji działań, identyfikacji liderów efektywności oraz wskazywania potencjalnych ulepszeń w mniej efektywnych jednostkach.

Wiele organizacji (np. szkoły, szpitale, instytucje publiczne) działa w warunkach, gdzie tradycyjne miary efektywności są mało adekwatne. DEA pozwalało na ocenę tych jednostek bez potrzeby ustalania z góry wag dla każdej zmiennej. Pomiar efektywności względnej DEA umożliwia ocenę jednostek względem siebie, wyznaczając jednostki efektywne jako odniesienie dla innych. Daje to możliwość identyfikacji jednostek, które są najlepsze (na granicy efektywności) i tych, które mają potencjał do

poprawy. Metoda pozwala na uwzględnienie wielu nakładów i efektów bez potrzeby redukcji ich do pojedynczych wskaźników, co czyni ją przydatną w wielu różnych kontekstach.

Zastosowanie analizy DEA-CCR w ocenie efektywności sieci młynów

Branża młynarska składa się zarówno z młynów, jak i elewatorów zbożowych, spedycji itp. W Polsce istnieją firmy, które zarządzają siecią młynów, podobnych do siebie jednostek, ponieważ jednostki te muszą działać w rygorze szeregu regulacji, w tym przepisami fitosanitarnymi, kodeksem pracy, kodeksem spółek handlowych oraz dobrymi praktykami obowiązującymi w branży młynarskiej.

Aby ocenić efektywność sieci 20 młynów można zastosować metodę analizy DEA-CCR. Pozwala ona na porównanie efektywności jednostek przy użyciu różnych nakładów i efektów. Poniżej znajduje się przykład nakładów wejściowych i efektów, które mogą być wykorzystane w takiej analizie.

Nakłady wejściowe:

1. Koszty stałe związane z utrzymaniem technicznym młyna

2. Koszty stałe związane z administracją młyna
3. Ilość samochodów rozładujących się w miesiącu
4. Ilość m² powierzchni produkcyjnej
5. Procentowy udział kosztów administracji w kosztach ogółem
6. Średni czas w minutach rozładunku jednej tony zboża

Wektor efektów:

1. Łączna ilość ton zmielonego ziarna w miesiącu
2. Zysk miesięczny ze sprzedaży mąki
3. Ilość rozładowanych samochodów lub wagonów ze zbożem

Poniżej przedstawiono przykład obliczeniowy, który pokaże, jak można przeprowadzić analizę DEA-CCR dla sieci 20 młynów przy użyciu języka Python.

Krok 1. Uruchamiam biblioteki programistyczne

```
import numpy as np
import pandas as pd
```

```
np.random.seed(148)
```

```
# Liczba młynów
num_mills = 20
```

Krok 2. Tworzę macierz fikcyjnych danych wejściowych (nakładów), które wymienię wcześniej

```
# Generowanie danych wejściowych
data_inputs = pd.DataFrame({
    "Maintenance_Cost": np.round(np.random.uniform(100000, 500000, num_mills)).astype(int),
    # Koszty stałe utrzymania technicznego
    "Admin_Cost": np.round(np.random.uniform(50000, 200000, num_mills)).astype(int),
    # Koszty administracji
    "Trucks_Unloading": np.random.randint(50, 300, num_mills),
    # Ilość samochodów/miesiąc
    "Production_Area": np.round(np.random.uniform(1000, 5000, num_mills)).astype(int),
    # Powierzchnia produkcyjna (m2)
    "Admin_Cost_Percentage": np.round(np.random.uniform(5, 20, num_mills)).astype(int),
    # Procentowy udział kosztów administracji
    "Avg_Unloading_Time": np.round(np.random.uniform(10, 60, num_mills)).astype(int),
    # Średni czas rozładunku (min/tona)
})
```

Krok 3. Tworzę fikcyjną macierz danych wyjściowych (efektów).

```
# Generowanie danych wyjściowych
data_outputs = pd.DataFrame({
    "Milled_Grain_Tons": np.random.uniform(500, 3000, num_mills).astype(int),
    # Łączna ilość ton zmielonego ziarna/miesiąc
    "Monthly_Profit": np.random.uniform(50000, 300000, num_mills).astype(int),
    # Zysk miesięczny
    "Unloaded_Vehicles": np.random.randint(40, 300, num_mills).astype(int),
    # Ilość rozładowanych samochodów lub wagonów
})
```

ane dla 20 młynów można wygenerować „z głowy” lub dostać podobne, prawdziwych jednostek, z działu księgowości.
dla pierwszych 4 młynów z 20 będziemy mieli takie dane:

Dane wejściowe (nakłady)

	Maintenance_Cost	Admin_Cost	Trucks_Unloading	Production_Area	Admin_Cost_Percentage	Avg_Unloading_Time
0	174751	175299	195	4267	13	60
1	490658	131695	139	3160	15	11
2	392405	179033	176	2511	6	54
3	435183	56044	254	3531	6	23

Dane wyjściowe (efekty)

	Milled_Grain_Tons	Monthly_Profit	Unloaded_Vehicles
0	2466	219872	233
1	2722	275694	253
2	1662	188888	281
3	2582	296249	207

Dane te mogą być trochę zniekształcone w stosunku do realiów pracy młyna. Najważniejsze jest tutaj, aby pokazać różne dane, które są niezbędne do przeprowadzenia analizy.

Krok 4. Budujemy główny „silnik” obliczeniowy. Najważniejsze, aby najpierw przeprowadzić standaryzację danych.

```
def calculate_dea_ccr(inputs, outputs):

    from sklearn.preprocessing import MinMaxScaler
    scaler = MinMaxScaler()
    scaled_inputs = scaler.fit_transform(inputs)
    scaled_outputs = scaler.fit_transform(outputs)

    # Ocena efektywności (dla uproszczenia używamy sumarycznej miary efektywności)
    efficiency_scores = np.sum(scaled_outputs, axis=1) / np.sum(scaled_inputs, axis=1)
    return efficiency_scores
```

Krok 5. Obliczanie wyników efektywności

```
inputs = data_inputs.values
outputs = data_outputs.values

efficiency_scores = calculate_dea_ccr(inputs, outputs)

# Wyświetlanie wyników
result_df.set_index("Mill_ID", inplace=True)
print(result_df)
```

W rezultacie tych kilku linijek kodu otrzymujemy zestawienie, pokazujące liderów i maruderów efektywności.

Mill_ID	Efficiency_Score
4	0.979135
6	0.910291
2	0.876303
18	0.787929
19	0.771081
17	0.636867
5	0.610127
3	0.597836
1	0.579244
12	0.463178
15	0.408164
13	0.391725
10	0.388735
7	0.384174
9	0.341174
14	0.339898
11	0.292286
16	0.262455
20	0.238177
8	0.099524

Interpretacja wyników analizy 20 młynów

Młyny o najwyższych wskaźnikach efektywności to Młyn 4 (0,979) i Młyn 6 (0,910). Te jednostki są bardzo blisko granicy efektywności (1,0), co oznacza, że wykorzystują swoje zasoby niemal optymalnie w stosunku do osiąganych efektów.

Jednostki średnio efektywne to młyny: 2 (0,876), 18 (0,788), 19 (0,771). Osiągają

one stosunkowo wysokie wskaźniki efektywności, ale mają jeszcze pewne pole do poprawy, aby dorównać liderom. Te jednostki są blisko efektywności, ale mogą skorzystać z porównania się do liderów, aby dostosować swoje działania.

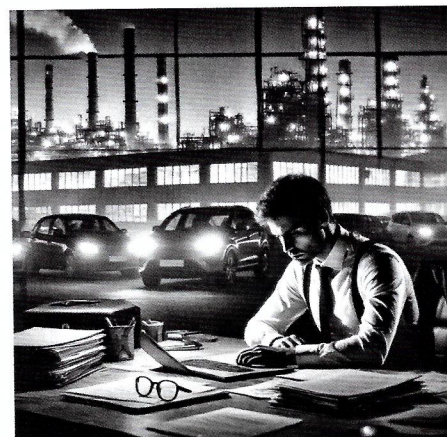
Młyny: 17 (0,637), 5 (0,610), 3 (0,598), 1 (0,579) wykazują się słabą efektywnością, uzyskały wskaźniki efektywności poniżej 0,7. To oznacza, że istnieje znaczna przestrzeń do poprawy ich działań, zarówno w zakresie

ograniczania kosztów, jak i zwiększania osiąganych efektów.

Najgorzej wypadły młyny: 8 (0,099) i 20 (0,238). Te jednostki działają znacznie poniżej poziomu efektywności i mogą wymagać bardziej kompleksowych zmian, takich jak poprawa zarządzania, optymalizacja nakładów lub zwiększenie produkcji i zysków.

Podsumowanie

Wyniki analizy DEA-CCR pozwalają zorientować się w realnej efektywności jednostek działających w podobnych warunkach prawnych, proceduralnych i ekonomicznych. Chyba najgorszą rzeczą jest niesprawiedliwe ocenienie jednostki. Ci, którzy mają „pod górkę” z rynkiem, starą technologią czy fatalnym położeniem, mimo że starają się wielkim wysiłkiem, aby osiągać najwyższą możliwą efektywność będą przegrywać z zakładem, który mimo doskonałej lokalizacji czy uwarunkowaniom ekonomicznym osiąga słabą efektywność krańcowo, a i który tylko i wyłącznie z powodu lokalizacji osiąga znacznie lepsze efekty i wygrywa w rankingu.



Metoda DEA wyłapuje te niesprawiedliwości i wskazuje kto jest prawdziwym zwycięzcą wyścigu o najwyższą efektywność.

Wojciech Moszczyński



Wojciech Moszczyński – absolwent Katedry Ekonometrii i Statystyki Uniwersytetu Mikołaja Kopernika w Toruniu, specjalista z zakresu ekonometrii, finansów, data science oraz rachunkowości zarządczej. Specjalizuje się w optymalizacji procesów produkcyjnych i logistycznych. Prowadzi badania w obszarze rozwoju i zastosowania sztucznej inteligencji. Od lat zaangażowany w popularyzację machine learning oraz data science w środowiskach biznesowych.