

Metody stochastyczne

Tutaj kiedyś coś będzie

Filip Ręka
Daniel Kuc
Piotr Rzeźnik

A
G
H

22 kwietnia 2023

Wstęp

Migracja jest powszechna w dzisiejszym świecie. Ludzie migrują ze względu na pracę, szkołę, studia, rodzinę, a także przez wiele innych czynników. Migrację można podzielić ze względu na kilka aspektów, ale w tym przypadku najważniejszy jest podział na migrację wewnętrzną i zewnętrzną. Ta praca będzie opierała się na migracji wewnętrznej, czyli w obrębie granic jednego państwa. Stany Zjednoczone zostały wybrane ze względu na bardzo dużą mobilność ich ludności, czyli skłonność do zmiany miejsca zamieszkania.

Przegląd literatury

Temat migracji pojawia się w literaturze od wielu lat i jest przedmiotem badań w wielu obszarach naukowych. Michael Greenwood w swoim artykule rozważał trendy, kierunki i powody wewnętrznej migracji w Stanach Zjednoczonych w latach 1965-1983 [1]. Bardzo często badany jest wpływ zmiany klimatu i zagrożeń z niego płynących na migrację ludności [3]. Brany jest pod uwagę wzrost poziomu wody w morzach, powodzie, a także susze [4] [5] [6]. Często migracja jest zestawiana z rozprzestrzenianiem się chorób. Były przeprowadzone badania na temat wpływu migracji na rozprzestrzenianie się np. malarii lub koronawirusa [7] [8]. Do modelowania samej migracji wykorzystywane od dawna był wykorzystywany model grawitacyjny, który przewiduje stopień migracji pomiędzy dwoma miejscami i opiera się na prawie powszechnego ciążenia. Model ten jest nadal popularny w ostatnich latach [9]. Rozszerzeniem modelu grawitacyjnego jest model radiacyjny. Także bazuje on na zjawisku fizycznym, tym razem na rozchodzeniu się fal w próżni [10]. Może on zostać zgeneralizowany także na inne cechy dwóch miejsc np. ich poziom zurbanizowania [11]. W badaniach nad migracją ludności w sieci miejsc używano już łańcuchów Markowa do symulowania przepływu ludności między miastami [12]. Do przewidywania kierunków migracji ludności były użyte także metody uczenia maszynowego. W artykule Caleb Robinson i Bistra Dilkina zostały porównane ze sobą modele grawitacyjne, radiacyjne oraz modele uczenia maszynowego XGBoost i sieć neuronowa (MLP) [13].

Dane użyte przy realizacji projektu

Dane pochodzą z instytucji rządowych, przez co mamy pewność że są wiarygodne.

Opiszemy tutaj jakie parametry będą mieli ludzie jeśli już podejmiemy o nich decyzję.

Kroki przy realizacji projektu

Wybór cech wpływających na migrację w danym hrabstwie

Mając dane dotyczące ilości osób które migrują z danego hrabstwa do innych hrabstw, jesteśmy w stanie dowiedzieć się, które cechy z jaką siłą wpływają na migrację ludności. Oczywiście jest to, że w każdym hrabstwie będą to inne cechy, dlatego prezentujemy takie podejście do rozwiązania tego problemu.

Każde hrabstwo jest przedstawione przy pomocy wektora cech o stałym rozmiarze N . Mając dane dotyczące ilości osób, które z rozważanego hrabstwa wyemigrowały, jesteśmy w stanie policzyć siłę "przyciągania" w następujący sposób:

- h_A - macierz cech hrabstwa z którego emigracje rozważamy
- h_1, h_2, \dots, h_k - macierze cech hrabstw, do których mieszkańcy hrabstwa A emigrowali

Dla każdego hrabstwa od 1 do k posiadamy dane dotyczące ile osób z hrabstwa A wyemigrowało do każdego z pozostałych. Wpiszmy te dane do wektora d . Aby obliczyć prawdopodobieństwa stosunku migracji używamy funkcji softmax, która zamieni nasz wektor w wektor prawdopodobieństw.

$$t = \text{softmax}(d) \quad (1)$$

Jeżeli każde hrabstwo jest opisywane przez wektor o stałej długości, jesteśmy w stanie wszystkie hrabstwa nanieść na N -wymiarową przestrzeń. Osoba zamieszkująca dane hrabstwo jest opisywana przy pomocy takiej samej liczby cech co wektor dla hrabstwa. Osoby, które są bardziej skłonne do migracji, po naniesieniu ich wektora cech na przestrzeń, będą znajdowały się dalej od rodzinnego hrabstwa, odległość ta determinuje skłonność do migracji. Podejście to nie bierze pod uwagę czynników, które mogą wpłynąć na miejsce docelowe. Przykładowo osoba znajdująca się w hrabstwie o niskim PKB, będzie bardziej skłonna do emigracji do tych hrabstw, które również mają niskie PKB, co jest raczej nieintuicyjne, ponieważ będzie znajdowała się w przestrzeni bliżej nich. Dlatego chcemy znaleźć dla mieszkańca danego hrabstwa takie wagi odpowiadające za "siłę", które po wykonaniu produktu Hadamarda na wektorze cech osoby i wag, da nam lepsze umieszczenie w przestrzeni.

Dla każdego hrabstwa generujemy wektor, który reprezentuje hipotetyczne hrabstwo, które powstaje w następujący sposób:

$$h_{vA} = \sum_{i=1}^k t_i h_i \quad (2)$$

Aby dostać wektor reprezentujący wpływ danej cechy na migrację bierzemy różnicę wektora wirtualnego hrabstwa oraz badanego hrabstwa:

$$h_{wA} = 1 + \text{norm}(h_{vA} - h_A) \quad (3)$$

Różnica wektorów będzie poddana normalizacji, a następnie przesunięta o 1 tak aby zachować element neutralny skalowania.

Przebieg symulacji

1. Wylosowanie ustalonej liczby przedstawicieli z każdego hrabstwa
2. Wykonujemy produkt Hadamarda pomiędzy wektorem wag a pomiędzy każdym wylosowanym mieszkańcem
3. Dla każdego mieszkańca obliczmy jego “odległość” od hrabstwa rodzinnego i na podstawie tej odległości decydujemy czy będzie on migrował czy nie
4. Sprawdzamy ile osób wyemigrowało z danego hrabstwa i na tej podstawie losujemy bez zwracania właśnie taką liczbę emigrantów
5. Dla każdej wylosowanej osoby obliczamy odległość w przestrzeni do każdego innego hrabstwa i na podstawie tej odległości losujemy jedno hrabstwo do którego reprezentant wyemigruje
6. Ponownie losujemy wybrane zmienne cech (takie jak np. zarobki) inne pozostawimy bez zmian, a jeszcze inne ustawiamy na takie, jakie występują w obecnym miejscu zamieszkania
7. Przez ustaloną liczbę kroków symulacji wykonujemy polecenia zaczynając od numeru 2.

Bibliografia

- [1] Michael J. Greenwood. “Human migration: theory, models, and empirical studies.”
In: *Journal of regional science* 25 4 (1985), pp. 521–44.