

Plan działania:

- Mapowanie abstraktu na wektory:
 - TF-idf
 - Ekstrakcja ficzerów z TF-idf
 - Wyrzucić te, które są wszędzie, albo prawie nigdzie
 - select k-best
 - Recursive feature elimination
- Dorzucenie własnych ficzerów:
 - Afiliacja - openalex
 - Usunięcie rzadkich i wrzucenie do jednej kategorii
 - Typ dokumentu - openalex
 - czy to jest artykuł, czy coś innego
 - "Ważność afiliacji" - własna zmienna
 - $\log\left(\frac{\#\{\text{ilość wszystkich paperów}\}}{\#\{\text{ilość paperów z którymi jest dana instytutacja powiązana}\}}\right)$
 - Powyższa wartość by była w one-hot wektorze zamiast 1
 - Liczba cytowań w innych paperach - altmetrics
 - Autor - openalex
 - Jeżeli wyjdzie coś sensownego, to coś analogicznego do ważności afiliacji
 - Kategoria MAG - openalex
 - Czy to otwarty dostęp - openalex