



Kierunek studiów: Geoinformacja

Nr albumu: 000000

Tomasz Matuszek

**Ocena wpływu zastosowania kanału termalnego
Landsat na wyniki nadzorowanej klasyfikacji
pokrycia terenu**

*Measuring impact of adding Landsat 8 thermal band on
supervised land cover classification*

Praca inżynierska napisana
w Instytucie Geoekologii i Geoinformacji
pod kierunkiem
dr. hab. Jakuba Nowosada

Poznań, 2022

Streszczenie

Abstrakt

Streszczenie powinno przedstawiać skrótowo główny problem pracy i jego rozwiązanie. Możliwa struktura streszczenia to: (1) 1-3 zdania wstępu do problemu (czym się zajmujemy, dlaczego jest to ważne, jakie są problemy/luki do wypełnienia), (2) 1 zdanie opisujące cel pracy, (3) 1-3 zdania przedstawiające użyte materiały (dane) i metody (techniki, narzędzia), (4) 1-3 zdania obrazujące główne wyniki pracy, (5) 1-2 zdania podsumowujące; możliwe jest też określenie dalszych kroków/planów.

Słowa kluczowe: (4-6 słów /zwrotów opisujących treść pracy, które nie wystąpiły w tytule)

Abstract

The abstract must be consistent with the above text.

Keywords: (as stated before)

Spis treści

Streszczenie	3
1 Wprowadzenie	5
2 Przegląd literatury	7
3 Dane źródłowe	9
3.1 Dane satelitarne	9
3.2 Dane o pokryciu terenu	9
4 Metody	13
4.1 Środowisko języka R	13
4.2 Metody uczenia maszynowego	13
5 Wyniki	15
5.1 Wynikowa mapa pokrycia terenu	15
5.2 Ocena jakości modelu klasyfikacyjnego	15
5.3 Ocena wpływu kanału termalnego na wyniki klasyfikacji	15
6 Podsumowanie	17

Rozdział 1

Wprowadzenie

- zastosowania i istotność map pokrycia terenu
- uczenie maszynowe i klasyfikacja nadzorowana obrazów satelitarnych jako narzędzie do tworzenia map pokrycia terenu
- zwrócenie uwagi na częste pomijanie kanału termalnego w modelach klasyfikacyjnych, nie do końca jasny wpływ informacji termalnej na wyniki klasyfikacji
- celem pracy jest stworzenie mapy pokrycia terenu metropolii poznańskiej oraz ocena wpływu kanału termalnego na wyniki modelu

Wprowadzenie powinno mieć charakter opisu od ogółu do szczegółu (np. trzy-pięć paragrafów). Pierwszy paragraf powinien być najbardziej ogólny, a kolejne powinny przybliżać czytelnika do problemu. Przedostatni paragraf powinien określić jaki jest problem (są problemy), który praca ma rozwiązać i dlaczego jest to (są one) ważne.

Wprowadzenie powinno być zakończone stwierdzeniem celu pracy. Dodatkowo tutaj może znaleźć się również krótki opis co zostało zrealizowane w pracy.

Rozdział 2

Przegląd literatury

Ten rozdział zawiera wyjaśnienie kontekstu pracy.

Pisząc ten rozdział proszę pomyśleć o osobach, które zupełnie nie znają opisywanej tematyki. Należy tutaj krok po kroku wyjaśnić podstawowe koncepcje, istotność problemu, wyniki poprzednich podobnych badań, itd. Ten rozdział obejmuje tylko kwestie, które już zostały wykonane przez inne osoby - nowe wyniki mają swoje miejsce w rozdziale 5.

Każda kwestia opisana w tym rozdziale powinna być cytowana. Dodatkowo cytowania odbywa się poprzez uzupełnienie pliku `thesis.bib` zapisem w formacie BibTeX, a następnie dodanie nazwy referencji poprzedzonej znakiem @. Przykładowo, zacytowanie książki *Geocomputation with R* odbywa się poprzez (Lovelace et al., [2019](#)).

W przypadku, gdy cytowanie zostało poprawnie wpisane oraz istnieje w pliku `thesis.bib` to bibliografia powinna się automatycznie wygenerować na końcu pracy.

W przypadku, gdy praca dyplomowa opisuje konkretny obszar to można po tym rozdziale stworzyć kolejny rozdział opisujący "obszar badań".

Ten i kolejne rozdziały mogą mieć także podrozdziały. Tworzenie podrozdziałów polega na stworzeniu nowej linii rozpoczynającej się od znaków ## a następnie tytułu podrozdziału. Dodatkowo w postaci {#sec - } można dodać skrót nazwy rozdziału/podrozdziału umożliwiający odnoszenie się do niego używając operatora [-@sec].

Rozdział 3

Dane źródłowe

Do przeprowadzenia analiz wykorzystano dwa typy danych:

- rastrowe dane satelitarne Landsat pobrane przy pomocy narzędzia GLAD ARD
- punktowe dane o pokryciu terenu pozyskane w wyniku programu LUCAS

3.1 Dane satelitarne

- opis procesu obróbki danych Landsat wykonany przez zespół z Uniwersytetu Maryland
- proces pobierania danych
- struktura danych GLAD ARD

3.2 Dane o pokryciu terenu

- czym jest program LUCAS i jak powstają dane w ramach jego działania
- LUCAS Grid i LUCAS Micro Data
- struktura danych punktowych, opis dodatkowych kolumn poza kategorią pokrycia terenu



Rycina 3.1: *Moja druga rycina*

- obróbka danych w celu wybrania odpowiednich punktów treningowych i reklasyfikacji do wspólnego podziału na kategorie pokrycia terenu

Celem tego rozdziału jest przedstawienie użytych w pracy danych. Należy wyjaśnić jakie dane zostały użyte, jakiego są one rodzaju, dla jakiego okresu zostały pobrane/stworzone, co one zawierają, etc.

W tym rozdziale warto dodać ryciny i tabele przedstawiające użyte dane.

Zwróć uwagę, że poniższe bloki kodu mają parametr `#| echo: false`; oznacza to, że będą one niewidoczne w wynikowym pliku PDF. Każdy z bloków kodu musi mieć unikalną nazwę; w przypadku rycin powinna się ona rozpoczynać od prefiksu `fig-`. Dodanie podpisu pod rycinę odbywa się używając parametru `#| fig-cap:`. Następnie do tej ryciny można się odnieść używając operatora `[-@]`.

Podobnie wygląda odnoszenie się do plików graficznych. Tutaj wewnątrz bloku kodu należy użyć funkcji `knitr::include_graphics()` (Rycina 3.1). Dodatkowo możliwa jest zmiana rozmiaru obrazka używając parametrów takich jak `#| out-width: i #| out-height:`.

Tabela 3.1: *Moja pierwsza tabela*

	a	b
1	a	
2	b	
3	c	
4	d	
5	e	

Odnoszenie się do tabel odbywa się poprzez operator [-@] wraz z prefiksem `tbl-`. Natomiast tworzenie podpisu nad tabelą ma miejsce używając parametru `#| tbl-cap:`. Dodatkowo możliwe jest użycie pakietu **kableExtra** (Zhu, [2021](#)) do określenia szerokości kolumn (Tabela 3.1).

Rozdział 4

Metody

4.1 Środowisko języka R

Krótki opis języka R oraz RStudio. Wymienienie wykorzystanych pakietów/bibliotek.

4.2 Metody uczenia maszynowego

- czym jest uczenie maszynowe i do czego się je wykorzystuje
- klasyfikacja a regresja
- klasyfikacja nadzorowana a nienadzorowana

4.2.1 Algorytm lasów losowych (ang.*random forest*)

- czym jest drzewo decyzyjne
- jak działają lasy losowe

4.2.2 Metody oceny jakości modelu

- idea resamplingu
- miary jakości modelu klasyfikacyjnego

4.2.3 Tuning parametrów modelu

- czym jest tuning parametrów?
 - resampling zagnieżdżony
-

Rozdział **Metody** zawiera opis użytych metod (np. statystycznych czy geostatystycznych) oraz technologii (np. pakiety R). Opis każdej z metod czy technologii powinien być zwarty i zawierać tylko najważniejsze informacje z punktu widzenia pracy dyplomowej.

Każda użyta metoda i technologia powinna być zacytowana. W przypadku pakietów R, wystarczy wypełnić poniższy blok kodu (zwróć uwagę, że ten blok kodu ma parametr `echo: false`; oznacza to, że będzie on niewidoczny w wynikowym pliku PDF)...

... a następnie zacytować pakiet używając znaku @, po którym podać nazwę pakietu rozpoczynającą się od prefiksu R-. Przykładowe cytowanie języka R bez nawiasu to R Core Team (2021), a pakietu **kableExtra** w nawiasie to (Zhu, 2021). Więcej przykładów cytowania można znaleźć na stronie https://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html#citations.

W przypadkach, gdy cytowanie istnieje, ale nie jest pakietem R to należy dodać je do pliku `thesis.bib` i użyć powyższej składni ze znakiem @. W ostateczności, gdy dana technologia nie posiada cytowania, należy podać jej adres internetowy.

Rozdział 5

Wyniki

5.1 Wynikowa mapa pokrycia terenu

- mapa pokrycia terenu aglomeracji poznańskiej
- mapa prawdopodobieństwa/pewności wybrania danej klasy

5.2 Ocena jakości modelu klasyfikacyjnego

Tabela ze wskaźnikami:

- skuteczność (OA)
- błąd klasyfikacji (CE)
- skuteczność producenta i użytkownika (PA, UA)
- współczynnik Kappa

5.3 Ocena wpływu kanału termalnego na wyniki klasyfikacji

- porównanie wskaźników z i bez kanału termalnego
- średnia temperatura dla każdej klasy pokrycia terenu

- wykresy uśredniające wpływ zmiennych na model, profile zmiennych
- mapa wpływu kanału termalnego na model (metoda agregacji rastra oraz interpolacji z punktów LUCAS)
- średni wpływ kanału termalnego na każdą klasę pokrycia terenu
- mapa różnic pomiędzy modelem pokrycia terenu z i bez kanału termalnego, macierz zmian

Część **Wyniki** może składać się z jednego lub więcej rozdziałów. Każdy z tych rozdziałów powinien mieć tytuł adekwatny do swojej treści.

Rozdziały wynikowe powinny korzystać z wiedzy opisanej w poprzednich rozdziałach (Rozdziały 2, 3, 4). W przypadku prac analitycznych, ich treść powinna przedstawiać kolejne etapy eksploracji i analizy danych. W przypadku prac technicznych, treść tych rozdziałów powinna opisywać stworzone narzędzia, a następnie pokazywać ich zastosowanie/a.

W przypadku prac technicznych warto pokazywać fragmenty napisanego rozwiązania lub jego wywołania używając bloków kodu.

```
moja_funkcja = function(x){  
  cat(x, "rządzi!")  
}  
moja_funkcja("Autor tej pracy")
```

Autor tej pracy rządzi!

Rozdział 6

Podsumowanie

Podsumowanie pracy jest w pewnym sensie znacznie rozbudowanym abstraktem. Należy wyliczyć i opisać osiągnięcia uzyskane w pracy dyplomowej. Tutaj jednak (w przeciwieństwie do np. rozdziału 1) należy przechodzić od szczegółu do ogółu - co zostało stworzone/określone, jak zostało to zrobione, jakie ma to konsekwencje, itd.

Ten rozdział powinien też zawierać opis kwestii, których nie udało się rozwiązać w pracy dyplomowej (i dlaczego się nie udało) oraz pomysły na przyszłe ulepszenie uzyskanych wyników lub dalsze badania.

Bibliografia

Lovelace, R, J Nowosad, and J Muenchow (2019). *Geocomputation with R*. CRC Press.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.

Zhu, H (2021). *kableExtra: Construct Complex Table with kable and Pipe Syntax*. R package version 1.3.4. <https://CRAN.R-project.org/package=kableExtra>.