

ANALÝZA PARAMETRŮ VÝTVARNÉHO STYLU U MAP PŘEDPOVĚDI POČASÍ METODAMI STROJOVÉHO UČENÍ

Mgr. Martin SADÍLEK

prof. RNDr. Vít Voženílek, CSc.

8. – 9. září 2022

GeoKARTO 2022



O práci

- diplomová práce obhájená 2020/2021
- hlavním cílem byla **identifikace výtvarných stylů** několika sad tematických map pomocí programu ORANGE technikami machine learning
- **analýza parametrů výtvarného stylu** (barevnost, rozlišení, grafická náplň...)



Pojmy

- **výtvarný styl:** provedení grafiky, kompozičních prvků a jejich parametrů (ukončení linií, jejich tloušťky, výběr fontu...), výběr barev, znakového klíče a provedení layoutu (*Vondráková, 2020*)
- **Orange:**
 - open-source software
 - určen především pro data mining a strojové učení
 - jednoduché, grafické provedení – uplatnění u začátečníků i profesionálů





sestavení datasetů
tematických map

kartografický popis
map

experimenty
*ořez, otočení, rozlišení,
barevnost, grafická
náplň mapy,
shlukovací metody*

uživatelské šetření

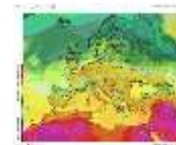
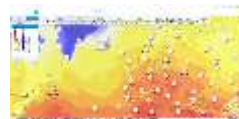
porovnání výsledků
a vyvození závěrů



Datasey tematických map

• FORECAST

- 25 map
- předpověď teploty vzduchu pro Evropu
- 11 zdrojů (Windy, Ventusky, Meteoblue...)



• TURIST

- 36 map (turistické + topografické)
- čtyři oblasti ČR
- 9 zdrojů



• TEKTO

- 25 map
- světová tektonika



Kartografický popis 86 map

- slovní a pomocí tabulky s koeficienty (přes 30 hodnocených prvků)
- pro snadné vyhodnocení – základní údaje o použitých mapách

ZDROJ:	AccuWeather	Dark Sky Map	Eldorado weather	CHMI	Meteoblue
Základní kompoziční prvky:					
legenda	1	0	1	1	1
měřítko číselné	0	0	0	0	0
měřítko grafické	0	0	0	0	0
tiráž	0,5	0,5	0,5	0,5	0,5
název: nadpis	1	0	1	1	1
podnadpis	0	0	0	0	0
mapové pole					
SOUČET_ZKP:	2,5	0,5	2,5	2,5	2,5
Nadstavbové kompoziční prvky:					
směrová růžice	0	0	0	0	0
zeměpisná síť	0	0	0	0	1
loga	1	0	0	1	0

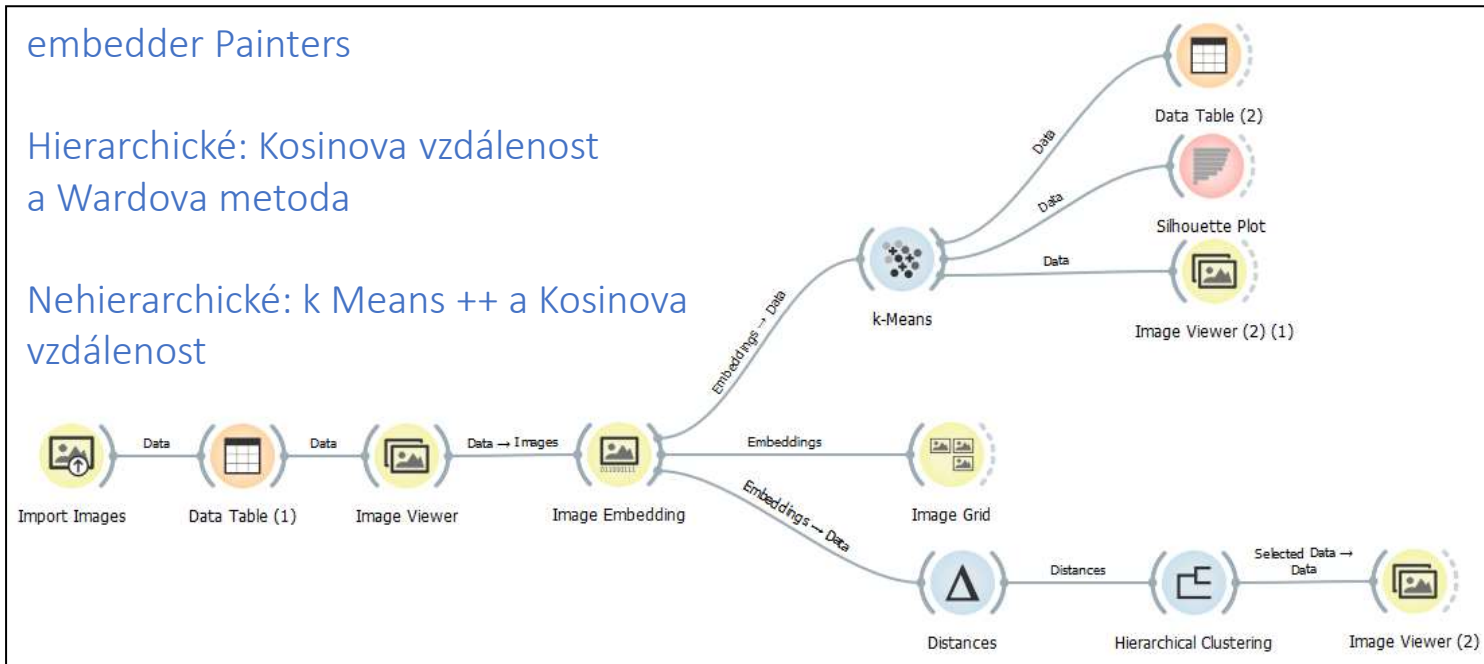


Workflow v Orange

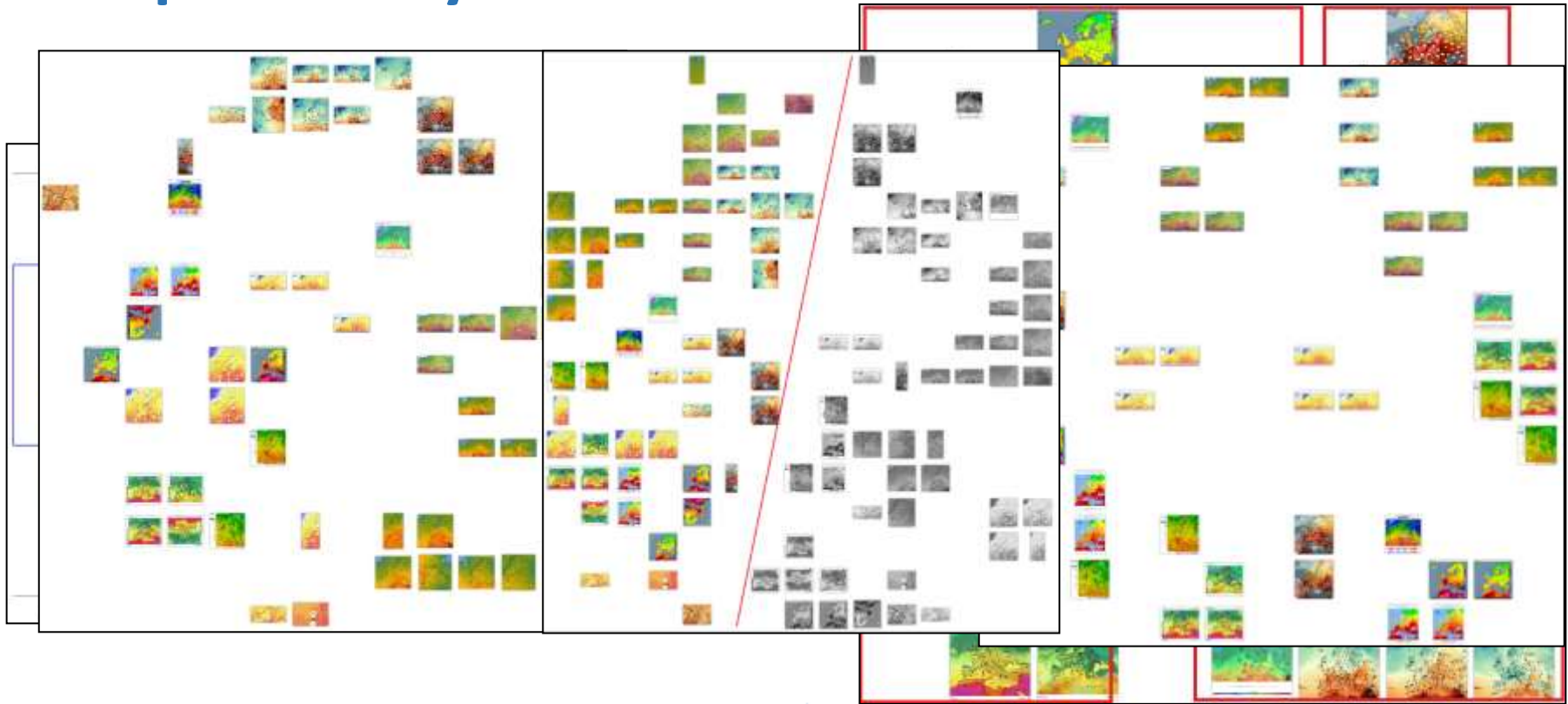
embedder Painters

Hierarchické: Kosinova vzdálenost
a Wardova metoda

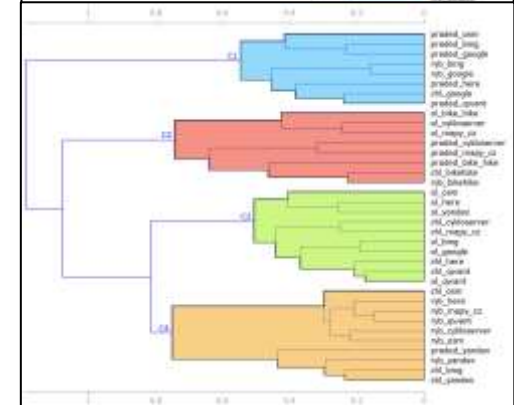
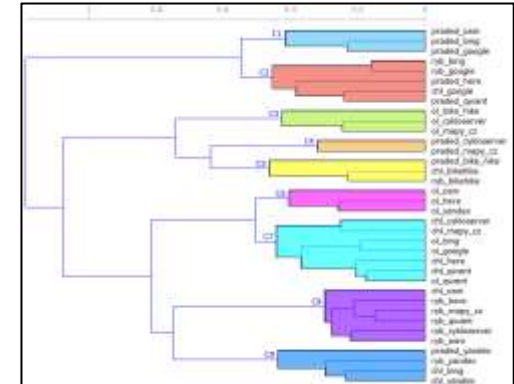
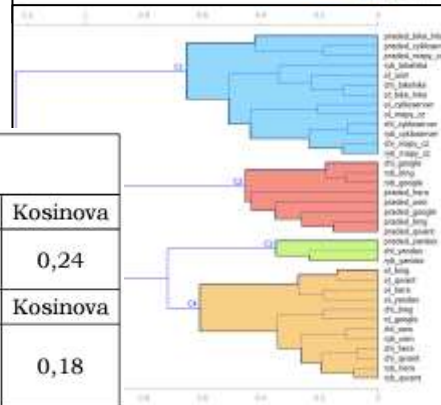
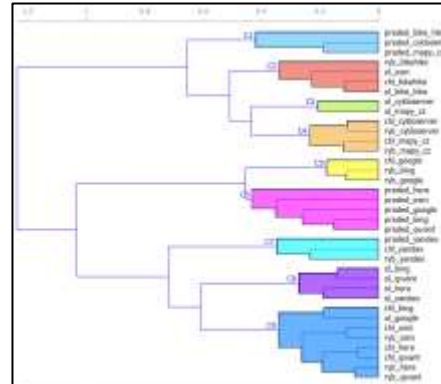
Nehierarchické: k Means ++ a Kosinova
vzdálenost



Experimenty 1 až 4: dataset FORECAST



Experiment 5: dataset TURIST



inicializační metoda				
k-Means ++	vzdálenostní metrika	Euklidovská	Manhattanská	Kosinova
	průměrná hodnota silhouette score	0,13	0,14	0,24
náhodná inicializace	vzdálenostní metrika	Euklidovská	Manhattanská	Kosinova
	průměrná hodnota silhouette score	0,1	0,12	0,18

Experiment 6: dataset TEKTO

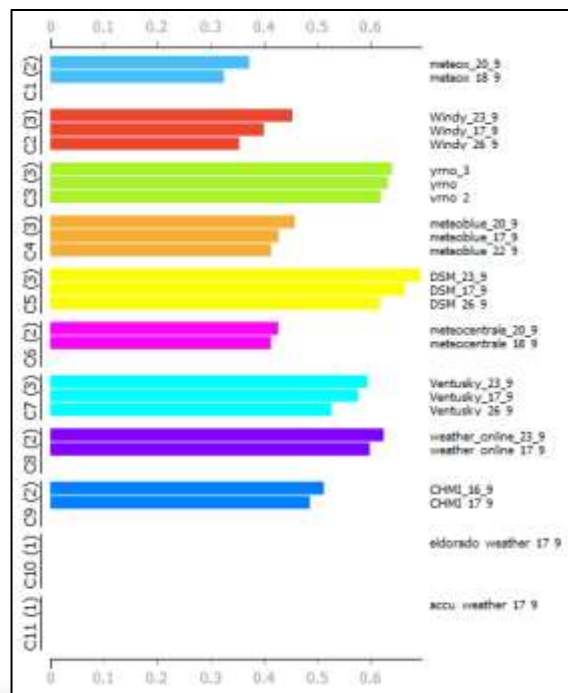
- zjištění vlivu grafické náplně na shlukování
- vyhodnocení pomocí ANOVY
- nemá vliv
- splněna podmínka normálního rozdělení datasetu?
- Conover-Imanův test
- nemá vliv

Anova: jeden faktor			
Faktor	<pre>> x <- c(8.4, 13.6, 20, 24.4) # shluk 1 > y <- c(17.7,17.9,18,20.8,22.5,22.5) # shluk 2 > z <- c(6.7,11.9,13.5,15,15,17.1,17.5,17.9,20.9,23.9,25.7) # shluk 3 > w <- c(11.6,12.1,16.3,19.2) # shluk 4 > conover.test(x=list(x,y,z,w))</pre>		
Sloupec 1	Kruskal-wallis rank sum test		
Sloupec 2			
Sloupec 3			
Sloupec 4	data: x and group Kruskal-wallis chi-squared = 3.8213, df = 3, p-value = 0.28		
ANOVA	Comparison of x by group (No adjustment)		
Zdroj v	Col Mean- Row Mean	1	2
Mezi výběry			3
Všechny vý	2	-1.038483 0.1554	
	3	0.145760 0.4428	1.488503 0.0757
Celkem	4	0.784552 0.2207	1.897918 0.0358
			0.804380 0.2151
alpha = 0.05			
Reject Ho if p <= alpha/2			

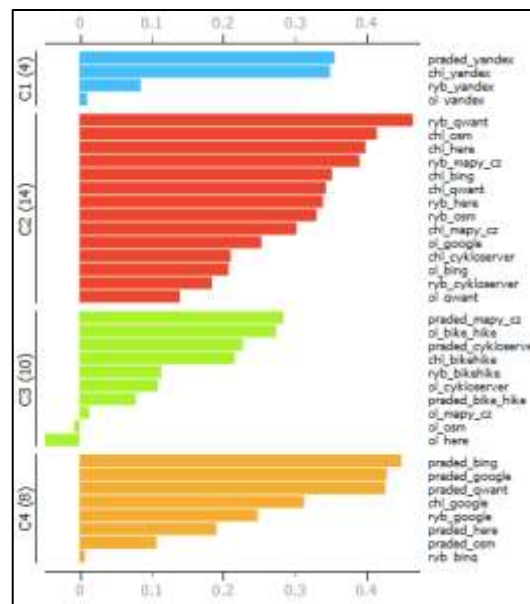


Porovnání datasetů podle Silhouette Score

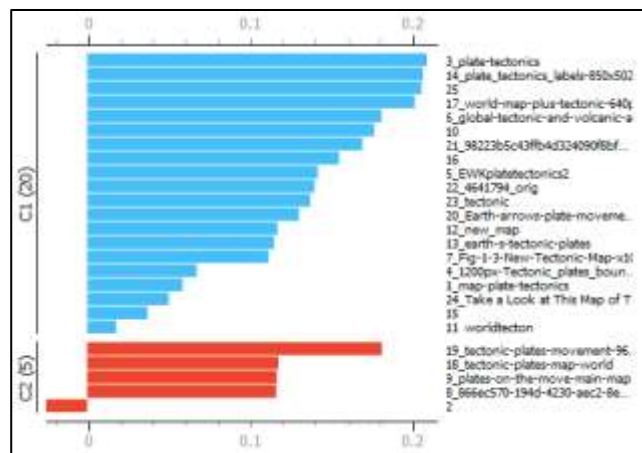
FORECAST 0,47



TURIST 0,24



TEKTO 0,13



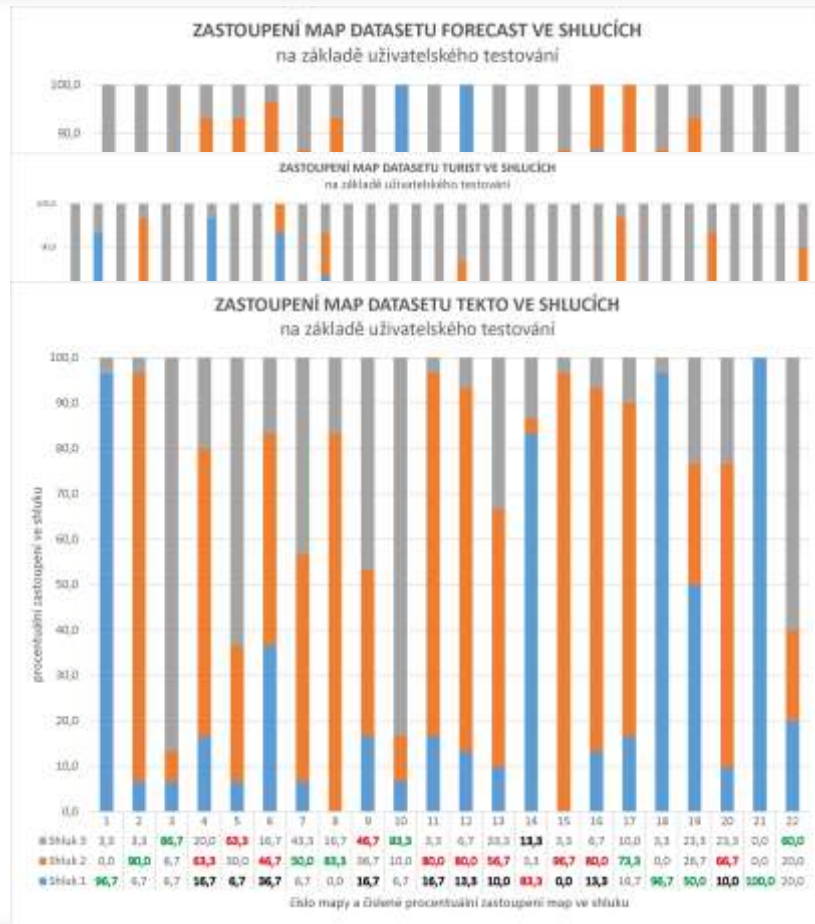
Slouží pro vyhodnocení kvality shlukování.

Uživatelské šetření

- 30 respondentů



- 26 respondentů vytvářelo shluky podle barevnosti
- dále zastoupena kompozice, popis, barvy linií a uspořádání legendy



Závěr

- zjištěny možnosti Orange pro vymezení výtvarných stylů
 - ☒ • největší vliv na shlukování má **barevnost** map *experiment 1*
 - ☒ • **změna rozlišení a rozdíly ve výřezech** mají vliv jen pokud jsou velmi výrazné; otočení ani grafická náplň mapy nemají vliv *experimenty 2 až 5*
 - ☐ • nevýrazné rozdíly mezi různými shlukovacími metodami *experiment 6*
- ↓
- ☒ • embeddery v Orange **nejsou vhodné** pro vymezení výtvarných stylů map
 - ☐ • „správné“ fungování pouze u map s výraznými barvami



Podněty pro navazující práce

- definice výtvarného stylu
- natrénování neuronové sítě na tuto problematiku
- kvantitativní vyjádření barevnosti map



Děkujeme za pozornost!

Martin SADÍLEK

Vít VOŽENÍLEK

